Author(s): Tedja, Milly S.; Wojciechowski, Robert; Hysi, Pirro G.; Eriksson, Nicholas; Furlotte, Nicholas A.; Verhoeven, Virginie J. M.; Iglesias, Adriana I.; Meester-Smoor, Magda A.; Tompson, Stuart W.; Fan, Qiao; Khawaja, Anthony P.; Cheng, Ching-Yu; Höhn, René; Yamashiro, Kenji; Wenocur, Adam; Grazal, Clare; Haller, Toomas; Metspalu, Andres; Wedenoja, Juho; Jonas, Jost B.; Wang, Ya Xing; Xie, Jing;

Title: Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error

Year: 2018

Version: Accepted version (Final draft)

Please cite the original version:

Tedja, M. S., Wojciechowski, R., Hysi, P. G., Eriksson, N., Furlotte, N. A., Verhoeven, V. J. M., Iglesias, A. I., Meester-Smoor, M. A., Tompson, S. W., Fan, Q., Khawaja, A. P., Cheng, C.-Y., Höhn, R., Yamashiro, K., Wenocur, A., Grazal, C., Haller, T., Metspalu, A., Wedenoja, J., . . . Klaver, C. C. W. (2018). Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. Nature Genetics, 50(6), 834-848. https://doi.org/10.1038/s41588-018-0127-7

Large genome-wide meta-analysis highlights light-induced signaling as a driver for refractive error

Milly S. Tedja[1,2,76], Robert Wojciechowski[3,4,76], Pirro G. Hysi[5,76], Nicholas Eriksson[6,76], Nicholas A. Furlotte[6,76], Virginie J.M. Verhoeven[1,2,7,76], Stuart W.J. Tompson[8], Adriana I. Iglesias Gonzalez[2], Qiao Fan[9], Anthony P. Khawaja[10], Ching-Yu Cheng[9,11], René Höhn[12,13], Elisabeth M. van Leeuwen[1,2], Kenji Yamashiro[14], Adam Wenocur[15], Clare Grazal[15], Toomas Haller[16], Andres Metspalu[16], Juho Wedenoja[17,18], Jost B. Jonas[19,20], Ya Xing Wang[20], Jing Xie[21], Paul Mitchell[22], Paul J. Foster[23], Barbara E.K. Klein[8], Ronald Klein[8], Andrew D. Paterson[24], Mohsen S. Hosseini[24], Rupal L. Shah[25], Cathy Williams[26], Yik Ying Teo[27,28], Yih Chung Tham[11], Preeti Gupta[29], Wanting Zhao[30], Yuan Shi[30], Woei-Yuh Saw[31], E Shyong Tai[28], Xue Ling Sim[28], Jennifer E. Huffman[32], Ozren Polašek[33], Caroline Hayward[32], Goran Bencic[34], Igor Rudan[35], James F. Wilson[35], CREAM[#], 23andMe Research Team, Peter K. Joshi[35], Akitaka Tsujikawa[14], Fumihiko Matsuda[36], Kristina N. Whisenhunt[8], Tanja Zeller[37], Peter J. van der Spek[38], Roxanna Haak[38], Arthur A. B. Bergen[39,40], Hanne Meijers-Heijboer[40,41], Magda A. Meester-Smoor[1,2], Sudha K. Iyengar[42-44], Jonathan H. Lass[42,43], Albert Hofman[2,45,46], Fernando Rivadeneira[2,46,47], André G. Uitterlinden[2,46,47], Johannes R. Vingerling[1], Terho Lehtimäki[48,49], Olli T. Raitakari[50,51], Ginevra Biino[52], Maria Pina Concas[53], Tae Hwi Schwantes-An[4,54], Robert P. Igo, Jr.[42], Gabriel Cuellar[55], Nicholas G. Martin[56], Jamie E. Craig[57], Puya Gharahkhani[55], Katie M. Williams[5], Abhishek Nag[58], Jugnoo S. Rahi[59], Phillippa Cumberland[59], Cécile Delcourt[60], Céline Bellenguez[61-63], Janina S. Ried[64], Thomas Meitinger[65,66], Christian Gieger[64], Tien Yin Wong[67,68], Alex W. Hewitt[21,69], David A. Mackey[70], Claire L. Simpson[4,71], Norbert Pfeiffer[13], Olavi Pärssinen[72,73], Paul N. Baird[21], Veronique Vitart[32], Najaf Amin[2], Cornelia M. van Duijn[2], Joan E. Bailey-Wilson[4], Terri L. Young[8], Seang Mei Saw [28,74], Dwight Stambolian[15], Stuart MacGregor[55], Jeremy A. Guggenheim[25,77], Joyce Tung[6,77], Christopher J. Hammond[5,77], Caroline C.W. Klaver[1,2,75,77]

Affiliations

1. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands.
2. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.
3. Department of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.
4. Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
5. Department of Ophthalmology, King's College London, London, UK.
6. 23andMe, Inc., Mountain View, USA.
7. Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
8. Department of Ophthalmology and Visual Sciences, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA.
9. Ophthalmology & Visual Sciences Academic Clinical Program and Centre for Quantitative Medicine, DUKE-National University of Singapore, Singapore.
10. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
11. Ocular Epidemiology Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore.
12. Department of Ophthalmology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland.
13. Department of Ophthalmology, University Medical Center Mainz, Mainz, Germany.
14. Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan.
15. Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
16. Estonian Genome Center, University of Tartu, Tartu, Estonia.
17. Department of Ophthalmology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.
18. Department of Public Health, University of Helsinki, Helsinki, Finland.
19. Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University of Heidelberg, Mannheim, Germany.
20. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China.
21. Centre for Eye Research Australia, University of Melbourne, Melbourne, Australia.
22. Department of Ophthalmology, Centre for Vision Research, Westmead Institute for Medical Research, University of Sydney, Sydney, Australia.
23. NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.
24. Program in Genetics and Genome Biology, Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada.
25. School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
26. School of Social and Community Medicine, University of Bristol, Bristol, UK.
27. Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore.
28. Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore, Singapore.
29. Department of Health Service Research, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore.
30. Statistics Support Platform, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore.
31. Life Sciences Institute, National University of Singapore, Singapore, Singapore.

77	32.	Institute of Genetics and Molecular Medicine, Medical Research Council Human Genetics Unit,
78		University of Edinburgh, Edinburgh, UK.
79	33.	Faculty of Medicine, University of Split, Split, Croatia.
80	34.	Department of Ophthalmology, Sisters of Mercy University Hospital, Zagreb, Croatia.
81	35.	Centre for Global Health Research, Usher Institute for Population Health Sciences and
82		Informatics, University of Edinburgh, Edinburgh, UK.
83	36.	Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan.
84	37.	Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg,
85		Germany.
86	38.	Department of Bioinformatics, Erasmus Medical Center, Rotterdam, The Netherlands.
87	39.	Department of Ophthalmology, Academic Medical Center, Amsterdam, The Netherlands.
88	40.	Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands.
89	41.	Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands.
90	42.	Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland,
91		Ohio, USA.
92	43.	Department of Ophthalmology and Visual Sciences, Case Western Reserve University and
93		University Hospitals Eye Institute, Cleveland, Ohio, USA.
94	44.	Department of Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
95	45.	Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston,
96		Massachusetts, USA.
97	46.	Netherlands Consortium for Healthy Ageing, Netherlands Genomics Initiative, the Hague, the
98		Netherlands.
99	47.	Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands.
100	48.	Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland.
101	49.	Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere, Tampere,
102		Finland.
103	50.	Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku,
104		Turku, Finland.
105	51.	Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku,
106		Finland.
107	52.	Institute of Molecular Genetics, National Research Council of Italy, Sassari, Italy.
108	53.	Institute of Genetic and Biomedical Research, National Research Council of Italy, Sassari, Italy.
109	54.	Department of Medical and Molecular Genetics, Indiana University - Purdue University
110		Indianapolis, Indianapolis, Indiana, USA.
111	55.	Department of Statistical Genetics, Queensland Institute of Medical Research, Brisbane,
112		Australia.
113	56.	Department of Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane,
114		Australia.
115	57.	Department of Ophthalmology, Flinders University, Adelaide, Australia.
116	58.	Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
117	59.	Great Ormond Street Institute of Child Health, University College London, London, UK.
118	60.	Team LEHA, UMR 1219, Université de Bordeaux, Inserm, Bordeaux Population Health
119		Research Center, Bordeaux, France.
120	61.	Institut Pasteur de Lille, Lille, France.
121	62.	Université de Lille, U1167, Lille, France.
122	63.	Inserm, U1167, Lille, France.
123	64.	Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for
124		Environmental Health, Neuherberg, Germany.
125	65.	Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany.
126	66.	Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Munich,
127		Germany.

128 67. Academic Medicine Research Institute, Singapore.
129 68. Retino Center, Singapore National Eye Centre, Singapore, Singapore.
130 69. Department of Ophthalmology, Menzies Institute of Medical Research, University of Tasmania,
131 Hobart, Australia.
132 70. Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western
133 Australia, Perth, Australia.
134 71. Department of Genetics, Genomics and Informatics, University of Tennessee Health Sciences
135 Center, Memphis, Tenessee.
136 72. Department of Ophthalmology, Central Hospital of Central Finland, Jyväskylä, Finland.
137 73. Gerontology Research Center and Department of Health Sciences, University of Jyväskylä,
138 Jyväskylä, Finland.
139 74. Myopia Research Group, Singapore Eye Research Institute, Singapore National Eye Centre,
140 Singapore, Singapore.
141 75. Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands.
142 76. These authors contributed equally to this work.
143 77. These authors jointly directed this work.
144
145 [#] A full list of consortium members appears at the end of the paper.

146 **Corresponding author**

147 Prof. dr. Caroline C.W. Klaver, Erasmus Medical Center, room Na-2808, PO Box 2040, 3000 CA,

148 Rotterdam, the Netherlands, e-mail: c.c.w.klaver@erasmusmc.nl, phone number +31651934491, fax

149 number +31107044657

150

151

152

153

154

155

156

157
158

159

160

161

Abstract

Refractive errors, including myopia, are the most frequent eye disorders worldwide and an increasingly common cause of blindness. Through a genome-wide association meta-analysis in 160,420 participants of mixed ancestry from CREAM and 23andMe and replication in 95,505 participants from the UK Biobank, we increased the number of significant independent signals from 37 to 161 and found a high genetic correlation between Europeans and Asians (>0.78). Enrichment analysis identified retinal cell physiology and light processing as the most prominent mechanisms. Expression experiments and comprehensive *in silico* analyses of the novel genes showed functional contribution of all cell types in the neurosensory retina (*GNB3, DRD1, AKAP6, ZEB2, TFAP2B, CA8, EDN2*), the retinal pigment epithelium (*EFEMP1, ANO2*), vascular endothelium (*CD34, FLT1*), and extracellular matrix (*VIPR2, ANTXR2, TCF7L2, COL10A1*) to refractive error development. The newly identified genes also elicited novel mechanisms such as rod as well as cone bipolar synaptic neurotransmission (rod: *CLU*; cone: *GNB3*), anterior segment morphology (*TCF7L2, VIPR2, MAF*), and angiogenesis (*FLT1*). Twenty eight SNPs resided in or near DNA structures transcribing small RNAs (non-coding, tRNAs, snoRNas, rRNAs, miRNA), suggesting a role for post-transcriptional regulation. Our results support the notion that refractive errors are caused by a light-dependent retina-to-sclera signaling cascade, and delineate potential molecular drivers defining the pathobiology of refractive errors and myopia.

Introduction

Refractive errors are common optical aberrations determined by mismatches in the focusing power of the cornea, lens and axial length of the eye. Their distribution is rapidly shifting towards myopia, or nearsightedness, all over the world. The myopia boom is particularly prominent in urban East Asia where up to 95% of twenty-year-olds in cities such as Seoul and Singapore have this refractive error[1-4]. Myopia prevalence is also rising throughout Western Europe and the USA, affecting ~50% of young adults in these regions[5,6]. While refractive errors can be optically corrected, even at moderate values they carry a significant risk of ocular complications with a high economic burden[7-9]. One in three individuals with high myopia (-6 diopters or worse) will develop irreversible visual impairment or blindness, mostly due to myopic macular degeneration, retinal detachment, or glaucoma[10,11]. At the other extreme, high hyperopia predisposes to strabismus, amblyopia and angle-closure glaucoma[10,12].

Refractive errors result from a complex interplay of lifestyle and genetic factors. Most established lifestyle factors for myopia are high education, lack of outdoor exposure, and excessive near work[3]. Recent research has identified many genetic variants for refractive errors, myopia, and axial length[13-25]. Two large studies, the international Consortium for Refractive Error and Myopia (CREAM)[26] and the personal genomics company 23andMe, Inc.[17,27] have provided the most comprehensive results. Despite differences in design and methodology, 37 associated genetic loci were identified in common, and most strikingly, there was a near-linear relationship in genetic effect sizes of the associated variants[28].

Given that only 3.6% of the variance of the refractive error trait was explained by the identified genetic variants[26], we presumed a high missing heritability. We therefore combined CREAM and 23andMe, and expanded the study sample to 160,420 individuals from a mixed ancestry population with quantitative information on refraction for a genome-wide association (GWA) meta-analysis. Significant variants were tested for replication in an independent cohort consisting of 95,505 individuals from the UK Biobank. We conducted systematic comparisons to assess differences in genetic inheritance and

205 distribution of risk variants between Europeans and Asians. Polygenic risk analyses were performed to

206 evaluate the contribution of the identified variants to the risk of myopia and hyperopia. Finally, we

207 integrated expression data and bioinformatics on the identified genes to gain insight into the possible

208 mechanisms underlying the genetic associations.

209

210

211

RESULTS

Susceptibility loci for refractive error

We performed a GWAS meta-analysis on adult refractive error using summary statistics from 37 studies from the Consortium for Refractive Error and Myopia (CREAM) and two cohorts from the personal genomics company 23andMe (Supplementary Figure 1). Analyses were based on ~11 million genetic variants (SNPs, insertions and deletions) genotyped or imputed to 1000 Genomes Project Phase I reference panel (version 3, March 2012 release[29]) that passed extensive quality control (Supplementary Figures 2-5, Supplementary Table 1a) and were represented by at least half of the entire study population and by > 13 cohorts from CREAM and both cohorts from 23andMe.

Meta-analyses were conducted in three stages. *Stage 1* focused on CREAM and included a fixed effects inverse variance-weighted meta-analysis on 44,192 individuals of European descent (CREAM-EUR) and 11,935 participants of East or South Asian ancestry (CREAM-ASN) using untransformed spherical equivalent (SphE) as the dependent variable representing refractive error (Supplementary Table 1b). 1,063 variants clustering in 24 loci (Supplementary Excel File 1) were genome-wide significant (P 5 x 10^{-8}). *Stage 2* consisted of a fixed effects inversed variance-weighted meta-analysis of the two European 23andMe cohorts ($N_{23andMe\_V2}$= 12,128; $N_{23andMe\_V3}$ = 92,165) using age of diagnosis of myopia (AODM) as outcome[27]. A total of 5,205 genome-wide significant variants clustered in 112 loci (Supplementary Excel File 1). All 25 loci identified at Stage 1 replicated in Stage 2 (*pBonferroni* 2.00 x 10^{-3}). Vice versa, 29 (26%) of the loci identified at Stage 2 replicated in Stage 1 (*pBonferroni* 4.46 x 10^{-4}), an expected proportion given the lower statistical power in CREAM. *Stage 3* was the joint meta-analysis of Stage 1 and Stage 2. As CREAM and 23andMe applied different phenotype measures, we used signed Z-scores as the mean per-allele effect size and assigned equal weights to CREAM and 23andMe. We identified 7,967 genome-wide significant genetic variants clustering in 140 loci (Figure 1a; Supplementary Figure 5-6,

Supplementary Excel File 1 – 3, Supplementary PDF File 1 and 2) of which 104 were novel. All 37 loci

that were found previously by CREAM and 23andMe using genotype data imputed to the HapMap II

reference panel were replicated (*pBonferroni* 1.85 x 10^{-3}), and 36 of the 37 were genome-wide significant

(Supplementary Table 2)[26,27]. We applied genomic control at each stage and checked for population

stratification using LD score regression[30] (Supplementary Table 3). At Stage 1 and 2, population

stratification was unlikely as inflation factors (λGC) were < 1.1 (Supplementary Figure 7), and LD score

regression intercepts (LDSC$_{intercept}$) ranged from 0.892 to 1.023 (Supplementary Figure 8). At Stage 3, we

observed an inflation of the median test statistic (λGC=1.129; Supplementary Figure 6), probably due to

true polygenicity rather than bias (i.e. population stratification or cryptic relatedness)[31]. The mixed

ancestry did not allow for calculation of LDSC$_{intercept}$.

To detect the presence of multiple independent signals at the discovered loci, a stepwise

conditional analysis was performed with GCTA-COJO[32] on meta-analysis summary statistics from all

European cohorts (*N*=148,485*)* using the Rotterdam Study I-III (RS I-III) as a reference panel for LD

structure ($N_{RSI-III}$ = 10,775). This analysis yielded 27 additional independent variants, resulting in a total of

167 loci (Supplementary Excel File 1). The lead variants at the newly-discovered loci were mainly of

lower minor allele frequencies (MAFs) than those reported in previous refractive error GWAS studies

with lower samples, reflecting the increased statistical power of the current analysis (Figure 1b).

We advanced these loci for replication analysis in a GWAS of refractive error carried out by the

UK Eye & Vision (UKEV) Consortium in 95,505 participants of European ancestry from the UK

BioBank. [33]16 variants were not present in UKEV, and were represented by a surrogate variant in high

LD (r^2 >0.8 LD; Supplementary Excel File 1). Six out of the 167 variants were not considered for

replication analysis: one variant (rs188159083) was not present on the array nor a surrogate was available

in UKEV and five variants showed evidence of departure from HWE (HWE exact test P < 3.0 x 10^{-4},

where 3.0 x 10^{-4} =0.05/167). One of these five variants (rs3138141, *RDH5*) was identified previously and

therefore still considered as a refractive error risk variant[26,27]. The remaining 161 genetic variants were

tested for replication. 86% (138/161) of the candidate variants replicated significantly: 104 (65%)

261   replicated surpassing genome wide significance; 34 replicated surpassing Bonferroni correction (P < 3.0 x

262   10^{-4}; 21.1%); and, 12 showed nominal evidence for replication (0.05 < P < 3.0 x 10^{-4}; 7.5%). Of the total,

263   only 11 (7%) did not replicate (Table 1 and Supplementary Excel File 1).

264   As CREAM and 23andMe employed different phenotypic outcomes, we evaluated consistency of

265   genotypic effects by comparing marker-wise additive genetic effect sizes (in units diopters per copy of the

266   risk allele) for SphE from CREAM-EUR against those (in units log(HR) per copy of the risk allele) for

267   AODM from 23andMe. All variants strongly associated with either outcome ($P < 0.001$), were

268   concordant in direction-of-effect, and had highly correlated effect sizes (Figure 2 a,b; Supplementary

269   Figure 9). For these variants, a 10% decrease in the log(HR) for AODM, indicating an earlier age-at-

270   myopia onset, was associated with a decrease of 0.15 diopters in SphE. A quantitative analysis for all

271   common SNPs (MAF > 0.01; HapMap3) using LD score regression yielded a genetic correlation of 0.93

272   (95% CI 0.86 to 0.99; $P = 2.1$ x 10^{-159}), confirming that effect sizes for both phenotypic outcomes were

273   closely related.

274

Gene annotation of susceptibility loci

276   We annotated all genetic variants with wANNOVAR using the University of California Santa Cruz

277   (UCSC) Known Gene database[34,35]. The identified 139 genetic loci were annotated to 208 genetic

278   structures (i.e. genes and known transcribed RNA genes, Table 1, Supplementary Excel file 1, Online

279   Methods). The physical positions of the lead genetic variants relative to protein-coding genes are shown

280   in Figure 1c. 86% of the identified variants were either intragenic or less than 50 kb from the 5'or 3' end

281   of the transcription start site. We found seven exonic variants (Supplementary Table 4) of which two had

282   MAF ≤ 0.05: rs5442 (*GNB3*) and rs17400325 (*PDE11A*). The index SNP in the *GNB3* locus with MAF

283   0.05 in Europeans is a highly conserved missense variant (G272S) predicted to be damaging by

284   PolyPhen-2[36] and SIFT[37]. *PDE11A* is presumed to play a role in tumorigenesis, brain function, and

285   inflammation[38]. The index SNP in the *PDE11A* locus with MAF 0.03 in Europeans is also a highly

286　conserved missense variant (Y727C); this variant was predicted to be damaging by PolyPhen[36], SIFT[39]

287　and align GVGD[40,41].The other exonic variants, rs1064583 (*COL10A1*), rs807037 (*KAZALD1*), rs1550094

288　(*PRSS56*), rs35337422 (*RD3L*) and rs6420484 (*TSPAN10*), were not predicted to be damaging.

289　　　　The most significant variant (Stage 3; rs12193446, $P = 4.21 \times 10^{-84}$) resides on chromosome 6

290　within a non-coding RNA, *BC035400*, in an intron of the *LAMA2* gene. This locus had been identified

291　previously, but our current fine mapping redefined the most associated variant. The function and potential

292　downstream target sites for BC035400 are currently unknown. The previously most strongly associated

293　variant, rs524952 on chromosome 15 near *GJD2*, was the second most significant variant ($P = 2.28 \times 10^{-65}$).

294　

Post-GWAS analyses identify 22 additional novel candidate loci

295

296　We performed two gene-based tests, fastBAT[42] and EUGENE[43], and applied a functional enrichment

297　approach using fgwas[44] (Online Methods). Fgwas incorporates functional annotation (eg. DNase I

298　hypersensitive sites in various tissues and 3'UTR regions) to reweight data from GWAS, and uses a

299　Bayesian model to calculate a posterior probability of association. With fastBAT, we identified 13 genes

300　at P value $< 2.0 \times 10^{-6}$, one of which (*CHD7*) had been identified previously[26,27]. Using EUGENE, we

301　found 7 genes at P value $< 2.0 \times 10^{-6}$ after incorporation of blood eQTLs. With fgwas, we identified 6

302　loci, which could be annotated to 9 genes, at posterior probability >0.9. Two genes (*HMGN4* and *TLX1*)

303　showed significant associations in two or more approaches. Taken together, these post-GWAS approaches

304　resulted in a total of 22 additional candidate loci for refractive error, annotated to 25 genes

305　(Supplementary Table 5). This increases the overall number of significant genetic associations to 161

306　candidate loci.

307

308

Polygenic risk scores

We calculated polygenic risk scores (PGRS)[45] per individual at various *P* value thresholds (Online Methods) for Rotterdam Study I-III (RS I-III; *N*=10,792) after recalculating *P* values and Z-scores of variants from Stage 3 excluding RS I-III. We found the highest fraction of phenotypic variance (7.8%) explained with 7,307 variants at *P* value threshold 0.005 (Supplementary Table 6). A PGRS based on these variants distinguished well between individuals with hyperopia and myopia at the lower and higher deciles (Figure 3); those in the highest decile had a 40-fold increased risk of myopia. When the PGRS was stratified for the median age (< 63 or > 63+ yrs), we found a significant difference in the variance explained (<63 yrs 8.9%; 63+ yrs 7.4%; *P* value 0.0038). The variance explained by PGRS was not significantly different between males and females (8.3% vs 7.5%; *P* value 0.13). The predictive value (area under the receiver operating characteristic curve, AUC) of the PGRS for myopia versus hyperopia adjusted for age and gender was 0.77 (95% CI = 0.75–0.79), a 10% increase compared to previous estimations[46].

Trans-ethnic comparison of genotypic effects

To explore potential ancestry differences in the identified refractive error loci, we calculated the heritability explained by common genetic variants (SNP-h^2) for Europeans and Asians using LD score regression[47]. SNP-h^2 was 0.214 (95% CI 0.185 to 0.243) and 0.172 (95% CI 0.154 to 0.190) in the European samples (CREAM-EUR and 23andMe), while it was only 0.053 (95% CI -0.025 to 0.131) in the Asian sample (CREAM-EAS). Next, we estimated the genetic correlation between Europeans and Asians by comparing variant effect size for common variants using the novel statistical program Popcorn[48]. Popcorn takes summary GWAS statistics from two populations and LD information from ancestry-matched reference panels, and computes genetic correlations by implementing a weighted likelihood function that accounts for the inflation of Z scores due to LD (Online Methods). Two genetic correlation metrics were calculated (Table 2); first, a genetic effect correlation (ρge) that quantifies the

12

333   correlation in SNP effect sizes between Europeans and Asians without taking into account ancestry-

334   related differences in allele frequency; and second, a genetic impact correlation (ρgi) that estimates the

335   correlation in variance-normalized SNP effect sizes between the two ancestry groups. Estimates of the

336   genetic effect correlation ρge were high between Europeans and Asians, but significantly different from 1

337   (0.79 and 0.80 respectively at P <1.9 x 10^{-6}; Table 2), indicating a clear genetic overlap but a difference in

338   per allele effect size. Estimates of the genetic impact correlation ρgi were similarly high (> 0.8), but not

339   significantly different from 1 for the correlation between CREAM-EUR and CREAM-ASN (P = 0.065),

340   indicating that the genetic impact of these alleles may still be similar.

341

342   **In silico pathway analysis**

343   We used an array of bioinformatics tools to investigate potential functions and pathways of the associated

344   genes. We first employed DEPICT[49] to perform a gene set enrichment analysis, a tissue type enrichment,

345   and a gene prioritization analysis, on all variants with P value < 1.00 x 10^{-5} from Stage 3. The gene set

346   enrichment analysis resulted in 66 reconstituted gene sets, of which 55 (83%) were eye-related. To reduce

347   redundancies between pathways, we clustered the significant pathways into 13 meta gene sets (false

348   discovery rate (FDR) < 5% and a P value < 0.05) (Supplementary Methods 1, Figure 4, Supplementary

349   Excel File 4). The most significant gene set was the 'abnormal photoreceptor inner segment morphology'

350   (MP:0003730; P value = 1.79 x 10^{-7}). The eye-related meta gene sets consisted of the 'thin retinal outer

351   nuclear layer' (MP:0008515; 27 (55%) gene sets), 'detection of light stimulus' (GO:0009583; 13 (24%)

352   gene sets), 'nonmotile primary cilium' (GO:0031513; 4 (6%) gene sets), and 'abnormal anterior eye

353   segment morphology' (MP:0005193; 4 (6%) gene sets). The first three meta gene sets had a Pearson's

354   correlation > 0.6. Interestingly, *RGR*, *RP1L1*, *RORB* and *GNB3* were present in all of these meta gene

355   sets. Retina was the most significant tissue of expression according to the tissue enrichment analysis (P

356   value = 1.11 x 10^{-4}, FDR <0.01). From the gene prioritization according to DEPICT, 7 genes were

357 highlighted as the most likely causal genes at P value $< 7.62 \times 10^{-6}$ and FDR <0.05: *ANO2, RP1L1,*

358 *GNB3, EDN2, RORB* and *CABP4.*

359 Next, we performed a canonical pathway analysis on all genes annotated to the variants of Stage

360 3 using Ingenuity Pathway Analysis (IPA; http://www.ingenuity.com/index.html). All genes were run

361 against the IPA database incorporating functional biological evidence on genomic and proteomic

362 expression based on regulation or binding studies. IPA identified "Glutamate Receptor Signaling" with

363 central player *Nf-kB* gene as the most significant pathway after correction for multiple testing (ratio of the

364 number of molecules 8.8% and Fisher's Exact test P value $= 1.56 \times 10^{-4}$; Figure 5, Supplementary Figure

365 10).

366

From disease associated loci to biological mechanisms

368 We adapted the scoring scheme designed by Fritsche et al.[50] to highlight genes for which there is

369 biological plausibility for a role in eye growth. We used 10 equally rated categories (maximum score 10;

370 genes with score ≥ 5 Table 3; all genes Supplementary Excel File 5; Online Methods, Supplementary

371 Methods 1): internal replication of index genetic variants in the individual cohort GWAS (CREAM-ASN,

372 CREAM-EUR and 23andMe), evidence for an eQTL effect in at least four tissue or cell types, annotation

373 to altered genomic function, ocular phenotype in humans and in mice, expression in human adult and fetal

374 ocular tissue, the presence of genes in the gene set enrichment, the presence of genes in the prioritization

375 analysis of DEPICT, and the presence of genes in the top 5 canonical pathway analysis of IPA. Sixty-five

376 index variants replicated in two or more individual cohorts; we found evidence for seven genetic variants

377 with eQTL effects in multiple tissue types; nine exonic variants, of which seven predicted protein-

378 alterations (Supplementary Table 4); 27 RNA genes, six located in the 3' or 5'UTR (Supplementary Table

379 7, Supplementary Figure 11), 84 genes resulting in an ocular phenotype in humans (Supplementary Excel

380 File 6) and 28 in mice (Supplementary Excel File 7); 169/212 (79%) genes expressed in human ocular

381 tissue (Supplementary Methods 1, Supplementary Excel File 8); 42 genes identified by DEPICT at P

14

382   value $< 5.4 \times 10^{-4}$ and FDR < 0.05 and 45 genes contributed to the most significant canonical pathways of

383   IPA. Notably, 48 of the associated genes encode known drug targets (Supplementary Excel File 9).

384   The gene with the highest biological plausibility score (score = 8) was *GNB3*, a highly conserved

385   gene encoding a guanine nucleotide-binding protein expressed in rod and cone photoreceptors and ON-

386   bipolar cells[51]. *GNB3* participates in signal transduction through G-protein coupled receptors and

387   enhances the temporal accuracy of phototransduction and ON-center signaling in the retina[51]. As

388   described above, the index SNP harbors a missense variant associated with refractive errors. Non-

389   synonymous mutations within *GNB3* are known to cause syndromic congenital stationary night

390   blindness[52] in humans, progressive retinopathy and globe enlargement in chickens[51], and abnormal

391   development of the photoreceptor-bipolar synapse in knock-out mice[53,54].

392   Other genes highly ranked (score = 7) include *CYP26A1, GRIA4, RDH5, RORB* and *RGR*, all

393   previously associated with refractive error, and one newly identified gene: *EFEMP1*. *EFEMP1* encodes a

394   member of the fibulin family of extracellular matrix glycoproteins, and is found pan-ocularly including in

395   the inner nuclear layer and Bruch's membrane. Mutations in this gene lead to specific macular

396   dystrophies[55], while variants have also been shown to co-segregate with primary open-angle glaucoma[56]

397   and associate with optic disc cup area[57].

398   Several other genes from our analysis are noteworthy for their function. *CABP4*, a calcium

399   binding protein expressed in cone and rod photoreceptor cells, mediates Ca^{2+} influx and glutamate release

400   in the photoreceptor-bipolar synapse[58]. Mutations in this gene have been described in congenital cone-rod

401   synaptic disorder[59], a retinal dystrophy associated with nystagmus, photophobia, and, remarkably, high

402   hyperopia. *KCNMA1* encodes pore forming alpha subunits of Ca^{2+}-activated K^+ (BK) channels. These

403   channels regulate synaptic transmission exclusively in the rod pathway[60]. A striking function of the other

404   previously identified genes is retinoic acid signaling and metabolism[61-63]. *ANO2* is a Ca^{2+}-activated Cl^-

405   channel recently reported to regulate retinal pigment epithelial (RPE) cell volume in a light-dependent

406   manner[64]. *EDN2* is a potent vasoconstrictor that binds to two G-protein-coupled receptors, *EDNRA*, which

407   resides on bipolar dendrites, and *EDNRB*, which is present on Mueller and horizontal cells. Both receptors

408   are also present on choroidal vessels[65], implying that the choroid as well as retinal cells are target sites for

409   this gene. *RP1L1* is expressed in cone and rod photoreceptors where it is involved in the maintenance of

410   microtubules in the connecting cilium[66]. Mutations in this gene cause dominant macular dystrophy and

411   retinitis pigmentosa[67]. We replicated two genes known to cause myopia in family studies. *FBN1* harbors

412   mutations causing with Marfan (OMIM #154700) and Weil Marchesani (OMIM #608328) syndrome;

413   *PTPRR* was one of the candidates in the MYP3 locus, which was found by linkage in families with high

414   myopia[68].

415   The location of rs7449443 (*P* value 3.58 x 10^{-8}) is notable as it resides in between *DRD1* and

416   *FLJ16171*. *DRD1* encodes dopamine receptor 1 and is known to modulate dopamine receptor 2-mediated

417   events[69,70]. The dopamine pathway has been implicated in myopia pathogenesis in many studies[69,71]. SNPs

418   in and near other genes involved in the dopamine pathway (dopamine receptors, synthesis, degradation,

419   and transporters)[72-74] did not reveal genome-wide significant associations (Supplementary Methods 2,

420   Supplementary Table 8; Supplementary Figure 12).

421   There were twenty-eight genetic variants in or near DNA structures transcribing RNA genes (non

422   coding RNA, linc RNAs, tRNAs, snoRNas, rRNAs). Notably, five were in the transcription region and

423   thirteen were in the vicinity (>0 kb and ≤ 50 kb) of start or end of the RNA gene transcription region.

424   They received low scores, since many have no reported function or disease association to date (Table 3,

425   Supplementary Excel File 10, Supplementary Figure 11, Supplementary Table 7). Our ranking of genes

426   based on functional information existing in the public domain does not necessarily represent the true

427   order of importance for refractive error pathogenesis. The observation that genes with strong statistical

428   association were distributed over all scores supports this concept. Nevertheless, this list may help to select

429   genes for subsequent functional studies.

430   Finally, integration of all aforementioned data with findings from literature allowed us to annotate

431   a large number of genes to ocular cell types (Figure 6). Remarkably, all cell types of the retina harbored

432   refractive error genes, as well as the RPE, vascular endothelium, and extracellular matrix.

433

Genetic pleiotropy

We performed a GWAS catalogue look up using FUMA to investigate overlap of genes with other common traits (Supplementary Figure 13)[75]. Refractive error and hyperopia were replicated significantly after correcting for multiple testing (adjusted P value 1.44 x 10^{-52} and 9.34 x 10^{-9}, respectively). We found significant overlap with 74 other traits, of which height (adjusted P value 1.11 x 10^{-10}), obesity (adjusted P value 1.38 x 10^{-10}), and BMI (adjusted P value 4.05 x 10^{-7}) were most important. Ocular traits significantly associated were glaucoma (optic cup area, intraocular pressure; adjusted P values 2.69 x 10^{-5} and 3.01 x 10^{-5}, respectively) and age-related macular degeneration (adjusted P value 1.27 x 10^{-3}).

DISCUSSION

Myopia may become the leading cause of world blindness in the near future, a grim outlook for which current counteractions are still insufficient[11,76]. To improve understanding of the genetic landscape and biology of the refractive error trait, we conducted a large GWAS meta-analysis in 160,420 participants of mixed ancestry and replicated in 95,505 participants. This led to the identification of 139 independent susceptibility loci by single variant analysis and 22 additional loci through post-GWAS methods, a four-fold increase in refractive error genes. The majority of annotated genes were found to be expressed in the human posterior segment of the eye. Using in silico analysis, we identified significant biological pathways, of which retinal cell physiology, light processing, and specifically glutamate receptor signaling were the most prominent mechanisms. Our integrated bio-informatic approach highlighted known ocular functionality for many genes.

To ensure robustness of our genetic associations, we included studies of various designs and populations, sought replication in an independent cohort of significant sample size, and stringently accounted for population stratification by performing genomic control at all stages of the meta-analysis[77].

458 With this approach, we internally replicated all loci from CREAM in 23andMe, and replicated a

459 considerable proportion of the 23andMe loci in CREAM. We combined studies with outcomes based on

460 actual refractive error measurements as well as on self-reported age-of-myopia-onset, and found the

461 direction-of-effect of the associated variants, as well as their effect size, to be remarkably consistent.

462 Combining two different outcome measures may appear unconventional, but age of onset and refractive

463 error have been shown to be very tightly correlated[11,28]. Each year of earlier onset leads to a higher degree

464 of myopia[78,79]. Moreover, the high genetic correlation of common SNPs between the two phenotypes

465 underscores their similarity. Most compelling evidence was provided by replication of 86% of the

466 discovered variants in the independent UKEV cohort which also used conventional refractive error

467 measurements. This robustness indicates that both phenotypic outcomes (SphE and AODM) can be used

468 to capture a shared source of genetic variation. In addition, we found trans-ethnic replication of

469 significant loci, and a high per-allele correlation of genetic effects of common variants in the Europeans

470 and Asians. Our findings support a largely shared genetic predisposition to refractive error and myopia in

471 the two ethnicities, although ancestry-specific allelic effects may exist. The low heritability estimate in

472 Asians may, in part, be explained by the low representation of this ethnicity in our study sample.

473 Alternatively, it may imply that environmental factors explain a greater proportion of the phenotypic risk

474 and recent rise in myopia prevalence in this ancestry group[80].

475   Limitations of our study were the possibility of false negative findings due to genomic control,

476 and underrepresentation of studies with Asian ancestry. Heterogeneity of observed effect estimates was

477 large for several associated variants, but not unexpected given the large number of collaborating studies

478 with varying methodology.

479   Although neurotransmission was already a suggested pathway in our previous studies[26,27], our

480 current pathway analyses provide more in depth insights into the retinal circuitry driving refractive error.

481 DEPICT identified 'thin retinal outer nuclear layer', 'detection of light stimulus', and 'nonmotile primary

482 cilium' as the most important meta-gene sets. These are the main characteristics of photoreceptors, which

483 are located in the outer retina and contain cilia. These photosensitive cells drive the phototransduction

484 cascade in response to light, which in turn induces visual information processing. IPA pointed towards

485 glutamate receptor signaling as the most significant pathway. Glutamate is released by photoreceptors and

486 determines conductance of retinal signaling to the ON and OFF bipolar cells[81]. Our functional gene look

487 ups provide evidence that rod as well as cone bipolar cells play a role (rod: *CLU*; cone: *GNB3*). Taken

488 together, these findings strongly suggest that light response and light processing in the retina are initiating

489 factors leading to refractive error.

490 The genetic association with light dependent pathways may also link to the well-established

491 protective effect of outdoor exposure on myopia. We found suggestive evidence for a genetic association

492 with *DRD1,* the dopamine receptor D1 gene. The dopaminergic pathway has been studied extensively in

493 animal models for its role in controlling eye growth in response to light[69,71,82-91]. *DRD1* was found to be a

494 mediator in this process, as bright light increased DRD1 activity in the bipolar ON-pathway, and

495 diminished form deprivation myopia in mice. Blockage of DRD1 reversed this inhibitory effect[92]. We did

496 not find evidence for direct involvement of other genes in the dopamine pathway, but *GNB3* may be an

497 indirect modifier as it is a downstream signaling molecule of dopamine and has been shown to influence

498 availability of the dopamine transporter DAT[93]. Although a promising target for therapy, further evidence

499 of *DRD1* in human myopiagenesis is warranted.

500 Novel pathways elicited by the newly identified genes are anterior segment morphology

501 *(TCF7L2, VIPR2, MAF)* and angiogenesis (*FLT1*). In addition, the high number of variants residing near

502 small RNA genes suggests that post-transcriptional regulation is an important mechanism, as these RNAs

503 are known to play a distinct and central regulatory role in cells[94]. These findings will serve as leads for

504 future studies performing detailed mapping of cellular networks, and functional studies into genes

505 implicated in ocular phenotypes, harboring protein-altering variants, and proven drug targets.

506 Our evaluation of shared genetics between refractive error and other disease-relevant phenotypes

507 highlighted overlap with anthropometric traits such as height, obesity, and BMI. This could give valuable

508 additional clues as to the phenotypic outcomes of perturbations of some of the networks identified.

509   Our genetic observations add credence to the current notion that refractive errors are caused by a

510   retina-to-sclera signaling cascade that induces scleral remodeling in response to light stimuli. The concept

511   of this cascade originates from various animal models showing that form deprivation, retinal defocus and

512   contrast, ambient light, and wavelength can influence eye growth in young animals[95-97]. Cell-specific

513   moieties in this putative signaling cascade in humans were largely unknown, although animal models

514   implicated GABA, dopamine, all-trans-retinoic acid and TGF-β[69,91,98,99]. Our study provides a large

515   number of new molecular candidates for this cascade, and clearly shows that a wide range of neuronal cell

516   types in the retina, the RPE, the vascular endothelium, as well as components of the extracellular matrix

517   are implicated (Figure 6). The many interprotein relationships (Figure 4) exemplify the complexity of eye

518   growth, and provide a challenge to develop strategies to prevent pathological eye elongation.

519   In conclusion, by using a cross-ancestry design in the largest study population on common

520   refractive errors to date, we uncovered numerous novel loci and pathways involved in eye growth. Our

521   multi-disciplinary approach incorporating GWAS data with *in silico* analyses and expression experiments

522   provides an example for the design of future genetic studies for complex traits. Additional genetic insights

523   into refractive errors will be gained by increasing sample size, greater genotyping depth, family studies

524   for identifying rare alleles of large effect, and by evaluating population extremes. Our list of plausible

525   genes and pathways provide a plethora of data for future studies focusing on gene-environment

526   interaction, and on translation of GWAS findings into starting points for therapy.

527

528

ONLINE METHODS

Ethics Statement

All human research was approved by the relevant institutional review boards and conducted according to the Declaration of Helsinki. All CREAM participants provided written informed consent; all 23andMe applicants provided informed consent online, and answered surveys according to 23andMe's human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, an AAHRPP-accredited institutional review board.

Study data

The study populations were participants of the Consortium for Refractive Error and Myopia (CREAM) comprising of 41,793 individuals with European ancestry from 26 cohorts (CREAM-EUR) and 11,935 individuals with Asian ancestry from 8 studies (CREAM-ASN); and customers of the 23andMe genetic testing company who gave informed consent for inclusion in research studies consisting of 104,293 individuals (2 cohorts of individuals with European ancestry, $N = 12,128$ and $N= 92,165$, respectively). All participants included in this analysis from CREAM and 23andMe were aged 25 years or older. Participants with conditions that could alter refraction, such as cataract surgery, laser refractive procedures, retinal detachment surgery, keratoconus as well as ocular or systemic syndromes were excluded from the analyses. Recruitment and ascertainment strategies varied per study (Supplementary Table 1a,b, and Supplementary Methods 3). Refractive error represented by measurements of refraction and analyzed as spherical equivalent (SphE =spherical refractive error + 1/2 cylinder refractive error) was the outcome variable for CREAM; myopic refractive error represented by self-reported age of diagnosis of myopia (AODM) for 23andMe[27].

Genotype calling and imputation

551   Samples were genotyped on different platforms and study specific quality control measures of the

553   genotyped variants were implemented before association analysis (Supplementary Table 1a). Genotypes

554   were imputed using the appropriate ancestry-matched reference panel for all cohorts from the 1000

555   Genomes Project (Phase I version 3, March 2012 release) with either minimac[100] or IMPUTE[101,102]. The

556   metrics for pre-imputation quality control varied amongst studies, but genotype call rate thresholds were

557   set at high level (≥ 0.95 for both CREAM and 23andMe). These metrics were similar to our previous

558   GWAS analyses[26,27]; details per cohort can be found in Supplementary Table 1a.

GWAS per study

560   For each CREAM cohort, a single marker analysis for the SphE (in diopters) phenotype was carried out

561   using linear regression adjusting for age, sex and up to the first five principal components. All non-

562   family-based cohorts removed one of each pair of relatives (after detection using either GCTA or

563   IBS/IBD analysis). In family-based cohorts, a score test-based association was used to adjust for within-

564   family relatedness[103,104]. For the 23andMe participants, Cox proportional hazards analysis testing AODM

565   as the dependent variable were performed as previously described[27], with P values calculated using a

566   likelihood ratio test for the single marker genotype term. We used an additive SNP allelic effect model for

567   all analyses.

568

Centralized quality control per study

570   After individual GWAS, all studies underwent a second round of quality control (QC). Quantile-quantile,

571   effect allele frequency, $P - Z$ test, standard error – sample size, and genomic control inflation factor plots

572   were generated for each individual cohort using EasyQC[105] (Supplementary Figure 2.1, 2.2, 2.3) . All

573   analytical issues discovered during this QC step were resolved per individual cohort.

574

GWAS meta-analyses

575 The GWAS meta-analyses were performed in three stages (Supplementary Figure 1). In Stage 1,

576

577 European (CREAM-EUR) and Asian (CREAM-ASN) participants from the CREAM cohort were meta-

578 analysed separately. Subsequently, all CREAM cohorts (CREAM-ALL) were meta-analysed. Variants

579 with MAF < 1% or imputation quality score < 0.3 (info metric of IMPUTE) or Rsq < 0.3 (minimac) were

580 excluded. A fixed effects inverse variance-weighted meta-analysis was performed using METAL[106]. In all

581 stages, each genetic variant had to be represented by at least half of the entire study population and at

582 least represented by 13 cohorts in CREAM and one cohort in 23andMe. For SNPs with high

583 heterogeneity (at $P < 0.05$), we also performed a random effects meta-analysis using METASOFT[50]. In

584 Stage 2, a meta-analysis of the two 23andMe cohorts was performed, using similar filtering but a lower

585 MAF threshold ($< 0.5\%$). In Stage 3, CREAM-ALL and 23andMe samples were combined using a fixed

586 effects meta-analysis based on P values and direction of effect.

587 In Stage 3, we chose a different weighting scheme due to the differences in effect size scaling; 23andMe

588 used a less accurate phenotype variable (AODM); i.e. the *effective* sample size of the 23andMe was

589 approximately equivalent to the *effective* sample size of CREAM-ALL (Figure 2b), thus weighting by

590 $(1/\sqrt{n_{effective}})$ yielded a final weighting ratio of 1:1[107,108]. Genome-wide statistical significance was defined

591 at $P < 5.0 \times 10^{-8}$ [109]. All three meta-analysis stages were performed under genomic control. Study specific

592 and meta-analysis lambda (λ) estimates are shown in Supplementary Figure 7; to check for confounding

593 biases (e.g. cryptic relatedness and population stratification), LD score intercepts from LD score

594 regressions per ancestry were constructed (Supplementary Figure 8)[30]. To check the robustness of signals,

595 we performed a conventional random effects models using METASOFT, fixed effects models weighted

596 on sample size and on weights estimated from standard error per allele tested using METAL

597 (Supplementary Excel Files 1 and Supplementary Excel Files 2).

598 Manhattan (modified version of package 'qqman'), regional, box, and forest plots were made using R

599 version 3.2.3 and LocusZoom[110]. An overview of the Hardy Weinberg P values of all index variants per

cohort can be found in Supplementary Excel File 3. The comparison between refractive error and age-of-onset was performed using the LDSC program[30].

Population stratification and heritability calculations

Each study assessed the degree of genetic admixture and stratification in their study participants through the use of principal components. Homogeneity of participants was assured by removal of all individuals whose ancestry did not match the prevailing ancestral group. We used genomic inflation factors to control for admixture and stratification, and performed genomic-controlled meta-analysis to account for the effects of any residual heterogeneity. To further distinguish between inflation from a true polygenic signal and population stratification, we examined the relationship between test statistics and linkage disequilibrium (LD) with LDSC. CREAM-EUR, CREAM-ASN and 23andMe were evaluated separately; variants not present in HapMap3 and MAF < 1% were excluded. SNP heritability estimates were calculated using LDSC for the same set of genetic variants.

Locus definition and annotation

All study effect size estimates were oriented to the positive strand of the NCBI Build 37 reference sequence of the human genome. The index variant of a locus was defined as the variant with the lowest *P* value in a region spanning a 100 kb window of the most outer genome wide significant variant of that same region. We annotated all index variants using the web-based version of ANNOVAR[111]based on UCSC Known Gene Database[35]. For variants within the coding sequence or 5' or 3' untranslated regions of a gene, that gene was assigned to the index variant (note that this led to more than 1 gene being assigned to variants located within the transcription units of multiple, overlapping genes). For variants in intergenic regions, the nearest 5' gene and the nearest 3' gene were assigned to the variant. Index variants were annotated to functional RNA elements when described as such in the UCSC Known Gene Database.

623 We used conservation (PhyloP[112]) and prediction tools (SIFT[39], Mutation Taster[113], align GVGD[40,41],

624 PolyPhen-2[36]) to predict the pathogenicity of protein-altering exonic variants.

625

626 Conditional signal analysis

627

628 We performed conditional analysis to identify additional independent signals nearby the index variant at

629 each locus, using GCTA-COJO[32]. We transformed the Z-scores of the summary statistics to beta's using

630 the following formula: Standard Error $= \sqrt{1/2 * N * MAF(1 - MAF)}$. We performed the GCTA-COJO

631 analysis[32], utilizing summary-level statistics from the meta-analysis on all cohorts. Linkage

632 disequilibrium (LD) between variants was estimated from the Rotterdam Study I-III.

633

634 Replication in UK Biobank

635 The UK Biobank Eye & Vision (UKEV) Consortium performed a GWAS of refractive error in 95,505

636 participants of European ancestry aged 37-73 year with no history of eye disorders[33]. Refractive error was

637 measured using an autorefractor; SphE was calculated per eye and averaged between the two eyes. To

638 account for relatedness a mixed model analysis with BOLT-LMM was used[114], including age, gender,

639 genotyping array, and the first 10 principal components as covariates. Analysis was restricted to markers

640 present on the HRC reference panel[115]. We performed lookups for all independent genetic variants

641 identified in our Stage 3 meta-analysis and conditional analysis. For variants not present in UKEV, we

642 performed lookups for a surrogate variant in high LD (r^2 >0.8). When more than one potential surrogate

643 variant was available, the variant in strongest LD with the index variant was selected.

644

645 Post-GWAS analyses

646 We performed two gene-based tests to identify additional significant genes not found in the single variant

647 analysis. First, we applied the gene-based test implemented in fastBAT[42] to the per-variant summary

648 statistics of the meta-analysis of all European cohorts (23andMe and CREAM-EUR). We used the default

649   parameters (all variants in or within 50kb of a gene) and focused on variants with a gene-based P value

650      $<2 \times 10^{-6}$ (Bonferroni correction based on 25,000 genes) and the per-variant P value $>5 \times 10^{-8}$. Secondly,

651   we applied another gene-based test in EUGENE[43] which only includes variants which are eQTLs (GTex,

652   blood[116]). EUGENE tests an hypothesis predicated on eQTLs as key drivers of the association signal.

653   eQTLs within 50kb of a gene were included in the test. Genes with EUGENE P value $<2 \times 10^{-6}$ (and not

654   found in the single variant analysis) were considered to be significant. Finally, we used functional

655   annotation information from genome-wide significant loci to reweigh results using fgwas (version

656   0.3.64[44]). This approach is able to identify risk loci that otherwise might not reach the genome-wide

657   significance threshold in standard GWAS. Details about this approach can be found in Supplementary

658   Methods 4.

659

Refractive errors and myopia risk prediction

661   To assess the risk of the entire range of refractive errors, we computed polygenic risk scores (PGRS) for

662   the population-based Rotterdam Studies (RS) I, RS-II and RS-III using the P values and Z scores from a

663   meta-analysis on CREAM-ALL and 23andMe, excluding the RS I-III cohorts. Only variants with high

664   imputation quality (IMPUTE info score > 0.5 or minimac Rsq > 0.8) and MAF $> 1\%$ were considered. P

665   value-based clumping was performed with PLINK[117], using an r^2 threshold of 0.2 and a physical distance

666   threshold of 500 kb, excluding the MHC region. This resulted in a total of 243,938 variants. For each

667   individual in RS-I, RS-II and RS-III ($N = 10,792$), PGRS were calculated using the --score command in

668   PLINK across strata of P value thresholds: 5.0×10^{-8}, 5.0×10^{-7}, 5.0×10^{-6}, 5.0×10^{-5}, 5.0×10^{-4}, 0.005,

669   0.01, 0.05, 0.1, 0.5, 0.8 and 1.0. The proportion of variance explained by each PGRS model was

670   calculated as the difference in the R^2 between two regression models; one where SphE was regressed on

671   age, sex, the first five principal components, and the other also including the PGRS as an additional

672 covariate. Subsequently, AUCs were calculated for myopia (SphE $\leq$ -3 SD) versus hyperopia (SphE $\geq$ +3

673 SD).

674

Genetic correlation between ancestries

676 We used Popcorn[48] to investigate ancestry-related differences in the genetic architecture of refractive

677 error and myopia. Pairwise analyses were carried out using the GWAS summary statistics from 23andMe

678 ($N = 104,292$), CREAM-EUR ($N = 44,192$) and CREAM-EAS ($N = 9,826$) meta-analyses. Only SNPs

679 with MAF $\geq$ 5% were included, resulting in a final set of 3,625,602 SNPs for analyses involving

680 23andMe and 3,642,928 SNPs for the CREAM-EUR versus CREAM-EAS analysis. Reference panels

681 were constructed using genotype data from 503 European and 504 East Asian individuals sequenced as

682 part of the 1000 Genomes Project (release 2013-05-02 downloaded from: ftp.1000genomes.ebi.ac.uk).

683 The reference panel VCF files were filtered using PLINK[117] to remove indels, strand-ambiguous variants,

684 variants without an "rs" id prefix, and variants located in the MHC region on chromosome 6

685 (chr6:25,000,000-33,500,000; Build 37).

686

Analysis between phenotypes

688 To evaluate consistency of genotypic effects across studies that employed different phenotype definitions,

689 we compared effect sizes from GWAS studies of either SphE or AODM in Europeans, i.e. CREAM-EUR

690 ($N = 44,192$) or 23andMe ($N = 104,293$) respectively. Marker-wise additive genetic effect sizes (in units

691 diopters per copy of the risk allele) for SphE were compared against those (in units log(HR) per copy of

692 the risk allele) for AODM. Data was visualised using R. Genetic correlation between the two phenotypes

693 SphE and AODM was calculated using LD score regression. This analysis included all common SNPs

694 (MAF > 0.01) present in HapMap3.

695

Evidence for functional involvement

697   In order to rank genes according to biological plausibility, we scored annotated genes based on our own

698   findings and published reports for a potential functional role in refractive error. Points were assigned for

699   each gene on the basis of 10 categories (details on the methodology per category are provided in

700   Supplementary Methods 3): internal replication of index genetic variants in the individual cohort GWAS

701   analyses through Bonferroni corrections (CREAM-ASN, CREAM-EUR and 23andMe; *pBonferroni* 1.10

702   x 10^{-4}), evidence for eQTL using the FUMA[32] and extensive look-ups in GtEx, evidence of expression in

703   the eye in developmental and adult ocular tissues, presence of an eye phenotype in knock-out mice (MGI

704   and IMPC database), presence of an eye phenotype in humans (OMIM[118] (http://omim.org),

705   DisGeNET[119]), location in a functional region of a gene (wANNOVAR), presence of the gene in a

706   significant enriched functional pathway with false discovery rate < 0.05 (DEPICT[49]), presence of the gene

707   in the gene priority analysis of DEPICT with false discovery rate < 0.05 and the presence of the gene in

708   the canonical pathway analysis of Ingenuity Pathway Analysis (IPA;

709   http://www.ingenuity.com/index.html). Furthermore, we performed a systematic search for each gene to

710   assess its potential as a drug target (SuperTarget[120], STITCH[121], DrugBank[122], PharmaGkb[123]). All

711   information derived from this study and literature were used to annotate genes to retinal cell types.

712

Genetic pleiotropy

714   To investigate overlap of genes with other common traits, we performed a look-up in the GWAS catalog

715   using FUMA. Multiple testing correction (i.e. Benjamini-Hochberg) was performed. Traits were

716   significantly associated when adjusted P value ≤ 0.05 and the number of genes that overlap with the

717   GWAS catalog gene sets was ≥ 2.

718

719   **Data availability statement**

720   The summary statistics of the Stage 3 meta-analysis are included in supplementary information files of

721   this published article. In order to protect the privacy of the participants in our cohorts, further summary

722   statistics of Stage 1 (CREAM) and Stage 2 (23andMe) will be available upon request. Please contact

723   c.c.w.klaver@erasmusmc.nl (CREAM) and/or apply.research@23andMe.com (23andMe) for more

724   information and to access the data.

725

726   **Acknowledgments**

730

731

References

750   **References**

751
752   1.   Pan, C.W., Ramamurthy, D. & Saw, S.M. Worldwide prevalence and risk factors for myopia.
753   *Ophthalmic Physiol Opt* **32**, 3-16 (2012).
754   2.   Morgan, I.G. What Public Policies Should Be Developed to Deal with the Epidemic of Myopia?
755   *Optom Vis Sci* **93**, 1058-60 (2016).
756   3.   Morgan, I. & Rose, K. How genetic is school myopia? *Prog Retin Eye Res* **24**, 1-38 (2005).
757   4.   Morgan, I.G., Ohno-Matsui, K. & Saw, S.M. Myopia. *Lancet* **379**, 1739-48 (2012).
758   5.   Williams, K.M. *et al.* Increasing Prevalence of Myopia in Europe and the Impact of Education.
759   *Ophthalmology* **122**, 1489-97 (2015).
760   6.   Williams, K.M. *et al.* Prevalence of refractive error in Europe: the European Eye Epidemiology
761   (E(3)) Consortium. *Eur J Epidemiol* **30**, 305-15 (2015).
762   7.   Vongphanit, J., Mitchell, P. & Wang, J.J. Prevalence and progression of myopic retinopathy in an
763   older population. *Ophthalmology* **109**, 704-11 (2002).
764   8.   Seet, B. *et al.* Myopia in Singapore: taking a public health approach. *Br J Ophthalmol* **85**, 521-6
765   (2001).
766   9.   Smith, T.S., Frick, K.D., Holden, B.A., Fricke, T.R. & Naidoo, K.S. Potential lost productivity
767   resulting from the global burden of uncorrected refractive error. *Bull World Health Organ* **87**,
768   431-7 (2009).
769   10.   Verhoeven, V.J. *et al.* Visual consequences of refractive errors in the general population.
770   *Ophthalmology* **122**, 101-9 (2015).
771   11.   Tideman, J.W. *et al.* Association of Axial Length With Risk of Uncorrectable Visual Impairment
772   for Europeans With Myopia. *JAMA Ophthalmol* **134**, 1355-1363 (2016).
773   12.   Flitcroft, D.I. The complex interactions of retinal, optical and environmental factors in myopia
774   aetiology. *Prog Retin Eye Res* **31**, 622-60 (2012).
775   13.   Nakanishi, H. *et al.* A genome-wide association analysis identified a novel susceptible locus for
776   pathological myopia at 11q24.1. *PLoS Genet* **5**, e1000660 (2009).
777   14.   Lam, C.Y. *et al.* A genome-wide scan maps a novel high myopia locus to 5p15. *Invest*
778   *Ophthalmol Vis Sci* **49**, 3768-78 (2008).
779   15.   Stambolian, D. *et al.* Meta-analysis of genome-wide association studies in five cohorts reveals
780   common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive
781   error. *Hum Mol Genet* **22**, 2754-64 (2013).
782   16.   Fan, Q. *et al.* Genetic variants on chromosome 1q41 influence ocular axial length and high
783   myopia. *PLoS Genet* **8**, e1002753 (2012).
784   17.   Fan, Q. *et al.* Meta-analysis of gene-environment-wide association scans accounting for education
785   level identifies additional loci for refractive error. *Nat Commun* **7**, 11008 (2016).
786   18.   Cheng, C.Y. *et al.* Nine Loci for Ocular Axial Length Identified through Genome-wide
787   Association Studies, Including Shared Loci with Refractive Error. *Am J Hum Genet* **93**, 264-77
788   (2013).
789   19.   Shi, Y. *et al.* Exome sequencing identifies ZNF644 mutations in high myopia. *PLoS Genet* **7**,
790   e1002084 (2011).
791   20.   Shi, Y. *et al.* Genetic variants at 13q12.12 are associated with high myopia in the Han Chinese
792   population. *Am J Hum Genet* **88**, 805-13 (2011).
793   21.   Li, Y.J. *et al.* Genome-wide association studies reveal genetic variants in CTNND2 for high
794   myopia in Singapore Chinese. *Ophthalmology* **118**, 368-75 (2011).
795   22.   Li, Z. *et al.* A genome-wide association study reveals association between common variants in an
796   intergenic region of 4q25 and high-grade myopia in the Chinese Han population. *Hum Mol Genet*
797   **20**, 2861-8 (2011).
798   23.   Liu, J. & Zhang, H.X. Polymorphism in the 11q24.1 genomic region is associated with myopia: a
799   comprehensive genetic study in Chinese and Japanese populations. *Mol Vis* **20**, 352-8 (2014).

800 24. Tran-Viet, K.N. *et al.* Mutations in SCO2 are associated with autosomal-dominant high-grade
801   myopia. *Am J Hum Genet* **92**, 820-6 (2013).
802 25. Aldahmesh, M.A. *et al.* Mutations in LRPAP1 are associated with severe myopia in humans. *Am
803   J Hum Genet* **93**, 313-20 (2013).
804 26. Verhoeven, V.J. *et al.* Genome-wide meta-analyses of multiancestry cohorts identify multiple
805   new susceptibility loci for refractive error and myopia. *Nat Genet* (2013).
806 27. Kiefer, A.K. *et al.* Genome-wide analysis points to roles for extracellular matrix remodeling, the
807   visual cycle, and neuronal development in myopia. *PLoS Genet* **9**, e1003299 (2013).
808 28. Wojciechowski, R. & Hysi, P.G. Focusing in on the complex genetics of myopia. *PLoS Genet* **9**,
809   e1003442 (2013).
810 29. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74
811   (2015).
812 30. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in
813   genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
814 31. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* **19**, 807-
815   12 (2011).
816 32. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. FUMA: Functional mapping and
817   annotation of genetic associations. *bioRxiv* (2017).
818 33. Plotnikov, D., Guggenheim, J. & Is a large eye size a risk factor for myopia? A Mendelian
819   randomization study. *bioRxiv* (2017).
820 34. UCSC Genome Browser.
821 35. Hsu, F. *et al.* The UCSC Known Genes. *Bioinformatics* **22**, 1036-46 (2006).
822 36. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat
823   Methods* **7**, 248-9 (2010).
824 37. Ng, P.C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic
825   Acids Res* **31**, 3812-4 (2003).
826 38. Kelly, M.P. Does phosphodiesterase 11A (PDE11A) hold promise as a future therapeutic target?
827   *Curr Pharm Des* **21**, 389-416 (2015).
828 39. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on
829   protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073-81 (2009).
830 40. Mathe, E. *et al.* Computational approaches for predicting the biological effect of p53 missense
831   mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res* **34**, 1317-
832   25 (2006).
833 41. Tavtigian, S.V. *et al.* Comprehensive statistical study of 452 BRCA1 missense substitutions with
834   classification of eight recurrent substitutions as neutral. *J Med Genet* **43**, 295-305 (2006).
835 42. Bakshi, A. *et al.* Fast set-based association analysis using summary data from GWAS identifies
836   novel gene loci for human complex traits. *Sci Rep* **6**, 32894 (2016).
837 43. Ferreira, M.A. *et al.* Gene-based analysis of regulatory variants identifies 4 putative novel asthma
838   risk genes related to nucleotide synthesis and signaling. *J Allergy Clin Immunol* (2016).
839 44. Pickrell, J.K. Joint analysis of functional genomic data and genome-wide association studies of
840   18 human traits. *Am J Hum Genet* **94**, 559-73 (2014).
841 45. International Schizophrenia, C. *et al.* Common polygenic variation contributes to risk of
842   schizophrenia and bipolar disorder. *Nature* **460**, 748-52 (2009).
843 46. Verhoeven, V.J. *et al.* Large scale international replication and meta-analysis study confirms
844   association of the 15q14 locus with myopia. The CREAM consortium. *Hum Genet* **131**, 1467-80
845   (2012).
846 47. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide
847   association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
848 48. Brown, B.C., Asian Genetic Epidemiology Network Type 2 Diabetes, C., Ye, C.J., Price, A.L. &
849   Zaitlen, N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet*
850   **99**, 76-88 (2016).

851 49. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted
852   gene functions. *Nat Commun* **6**, 5890 (2015).
853 50. Fritsche, L.G. *et al.* A large genome-wide association study of age-related macular degeneration
854   highlights contributions of rare and common variants. *Nat Genet* **48**, 134-43 (2016).
855 51. Ritchey, E.R. *et al.* Vision-guided ocular growth in a mutant chicken model with diminished
856   visual acuity. *Exp Eye Res* **102**, 59-69 (2012).
857 52. Vincent, A. *et al.* Biallelic Mutations in GNB3 Cause a Unique Form of Autosomal-Recessive
858   Congenital Stationary Night Blindness. *Am J Hum Genet* **98**, 1011-9 (2016).
859 53. Blake, J.A. *et al.* Mouse Genome Database (MGD)-2017: community knowledge resource for the
860   laboratory mouse. *Nucleic Acids Res* **45**, D723-D729 (2017).
861 54. Nikonov, S.S. *et al.* Cones respond to light in the absence of transducin beta subunit. *J Neurosci*
862   **33**, 5182-94 (2013).
863 55. Stone, E.M. *et al.* A single EFEMP1 mutation associated with both Malattia Leventinese and
864   Doyne honeycomb retinal dystrophy. *Nat Genet* **22**, 199-202 (1999).
865 56. Mackay, D.S., Bennett, T.M. & Shiels, A. Exome Sequencing Identifies a Missense Variant in
866   EFEMP1 Co-Segregating in a Family with Autosomal Dominant Primary Open-Angle Glaucoma.
867   *PLoS One* **10**, e0132529 (2015).
868 57. Springelkamp, H. *et al.* ARHGEF12 influences the risk of glaucoma by increasing intraocular
869   pressure. *Hum Mol Genet* **24**, 2689-99 (2015).
870 58. Haeseleer, F. *et al.* Essential role of Ca2+-binding protein 4, a Cav1.4 channel regulator, in
871   photoreceptor synaptic function. *Nat Neurosci* **7**, 1079-87 (2004).
872 59. Littink, K.W. *et al.* A novel homozygous nonsense mutation in CABP4 causes congenital cone-
873   rod synaptic disorder. *Invest Ophthalmol Vis Sci* **50**, 2344-50 (2009).
874 60. Grimes, W.N., Li, W., Chavez, A.E. & Diamond, J.S. BK channels modulate pre- and
875   postsynaptic signaling at reciprocal synapses in retina. *Nat Neurosci* **12**, 585-92 (2009).
876 61. Parker, R.O. & Crouch, R.K. Retinol dehydrogenases (RDHs) in the visual cycle. *Exp Eye Res*
877   **91**, 788-92 (2010).
878 62. Radu, R.A. *et al.* Retinal pigment epithelium-retinal G protein receptor-opsin mediates light-
879   dependent translocation of all-trans-retinyl esters for synthesis of visual chromophore in retinal
880   pigment epithelial cells. *J Biol Chem* **283**, 19730-8 (2008).
881 63. Luo, T., Sakai, Y., Wagner, E. & Drager, U.C. Retinoids, eye development, and maturation of
882   visual function. *J Neurobiol* **66**, 677-86 (2006).
883 64. Keckeis, S., Reichhart, N., Roubeix, C. & Strauss, O. Anoctamin2 (TMEM16B) forms the Ca2+-
884   activated Cl- channel in the retinal pigment epithelium. *Exp Eye Res* **154**, 139-150 (2016).
885 65. Prasanna, G., Narayan, S., Krishnamoorthy, R.R. & Yorio, T. Eyeing endothelins: a cellular
886   perspective. *Mol Cell Biochem* **253**, 71-88 (2003).
887 66. Yamashita, T. *et al.* Essential and synergistic roles of RP1 and RP1L1 in rod photoreceptor
888   axoneme and retinitis pigmentosa. *J Neurosci* **29**, 9748-60 (2009).
889 67. Davidson, A.E. *et al.* RP1L1 variants are associated with a spectrum of inherited retinal diseases
890   including retinitis pigmentosa and occult macular dystrophy. *Hum Mutat* **34**, 506-14 (2013).
891 68. Hawthorne, F. *et al.* Association mapping of the high-grade myopia MYP3 locus reveals novel
892   candidates UHRF1BP1L, PTPRR, and PPFIA2. *Invest Ophthalmol Vis Sci* **54**, 2076-86 (2013).
893 69. Feldkaemper, M. & Schaeffel, F. An updated view on the role of dopamine in myopia. *Exp Eye*
894   *Res* **114**, 106-19 (2013).
895 70. Paul, M.L., Graybiel, A.M., David, J.C. & Robertson, H.A. D1-like and D2-like dopamine
896   receptors synergistically activate rotation and c-fos expression in the dopamine-depleted striatum
897   in a rat model of Parkinson's disease. *J Neurosci* **12**, 3729-42 (1992).
898 71. Stone, R.A., Lin, T., Laties, A.M. & Iuvone, P.M. Retinal dopamine and form-deprivation
899   myopia. *Proc Natl Acad Sci U S A* **86**, 704-6 (1989).

900   72.   Gardner, M., Bertranpetit, J. & Comas, D. Worldwide genetic variation in dopamine and
901   serotonin pathway genes: implications for association studies. *Am J Med Genet B Neuropsychiatr*
902   *Genet* **147B**, 1070-5 (2008).
903   73.   D'Souza, U.M. & Craig, I.W. Functional polymorphisms in dopamine and serotonin pathway
904   genes. *Hum Mutat* **27**, 1-13 (2006).
905   74.   Beaulieu, J.M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of dopamine
906   receptors. *Pharmacol Rev* **63**, 182-217 (2011).
907   75.   MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies
908   (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901 (2017).
909   76.   Holden, B.A. *et al.* Global Prevalence of Myopia and High Myopia and Temporal Trends from
910   2000 through 2050. *Ophthalmology* **123**, 1036-42 (2016).
911   77.   Cardon, L.R. & Palmer, L.J. Population stratification and spurious allelic association. *Lancet* **361**,
912   598-604 (2003).
913   78.   Chua, S.Y. *et al.* Age of onset of myopia predicts risk of high myopia in later childhood in
914   myopic Singapore children. *Ophthalmic Physiol Opt* **36**, 388-94 (2016).
915   79.   Williams, K.M. *et al.* Age of myopia onset in a British population-based twin cohort. *Ophthalmic*
916   *Physiol Opt* **33**, 339-45 (2013).
917   80.   Dolgin, E. The myopia boom. *Nature* **519**, 276-8 (2015).
918   81.   Connaughton, V. Glutamate and Glutamate Receptors in the Vertebrate Retina. in *Webvision: The*
919   *Organization of the Retina and Visual System* (eds. Kolb, H., Fernandez, E. & Nelson, R.) (Salt
920   Lake City (UT), 1995).
921   82.   Hung, G.K., Mahadas, K. & Mohammad, F. Eye growth and myopia development: Unifying
922   theory and Matlab model. *Comput Biol Med* **70**, 106-18 (2016).
923   83.   Norton, T.T. What Do Animal Studies Tell Us about the Mechanism of Myopia-Protection by
924   Light? *Optom Vis Sci* **93**, 1049-51 (2016).
925   84.   Weiss, S. & Schaeffel, F. Diurnal growth rhythms in the chicken eye: relation to myopia
926   development and retinal dopamine levels. *J Comp Physiol A* **172**, 263-70 (1993).
927   85.   Stone, R.A., Lin, T., Iuvone, P.M. & Laties, A.M. Postnatal control of ocular growth:
928   dopaminergic mechanisms. *Ciba Found Symp* **155**, 45-57; discussion 57-62 (1990).
929   86.   Morgan, I.G. The biological basis of myopic refractive error. *Clin Exp Optom* **86**, 276-88 (2003).
930   87.   Li, X.X., Schaeffel, F., Kohler, K. & Zrenner, E. Dose-dependent effects of 6-hydroxy dopamine
931   on deprivation myopia, electroretinograms, and dopaminergic amacrine cells in chickens. *Vis*
932   *Neurosci* **9**, 483-92 (1992).
933   88.   Iuvone, P.M., Tigges, M., Stone, R.A., Lambert, S. & Laties, A.M. Effects of apomorphine, a
934   dopamine receptor agonist, on ocular refraction and axial elongation in a primate model of
935   myopia. *Invest Ophthalmol Vis Sci* **32**, 1674-7 (1991).
936   89.   Ashby, R., McCarthy, C.S., Maleszka, R., Megaw, P. & Morgan, I.G. A muscarinic cholinergic
937   antagonist and a dopamine agonist rapidly increase ZENK mRNA expression in the form-
938   deprived chicken retina. *Exp Eye Res* **85**, 15-22 (2007).
939   90.   Ashby, R. Animal Studies and the Mechanism of Myopia-Protection by Light? *Optom Vis Sci* **93**,
940   1052-4 (2016).
941   91.   Rymer, J. & Wildsoet, C.F. The role of the retinal pigment epithelium in eye growth regulation
942   and myopia: a review. *Vis Neurosci* **22**, 251-61 (2005).
943   92.   Chen, S. *et al.* Bright Light Suppresses Form-Deprivation Myopia Development With Activation
944   of Dopamine D1 Receptor Signaling in the ON Pathway in Retina. *Invest Ophthalmol Vis Sci* **58**,
945   2306-2316 (2017).
946   93.   Chen, P.S. *et al.* Effects of C825T polymorphism of the GNB3 gene on availability of dopamine
947   transporter in healthy volunteers--a SPECT study. *Neuroimage* **56**, 1526-30 (2011).
948   94.   Scott, M.S. & Ono, M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs.
949   *Biochimie* **93**, 1987-92 (2011).

950   95.   McFadden, S.A. Understanding and Treating Myopia: What More We Need to Know and Future
951   Research Priorities. *Optom Vis Sci* **93**, 1061-3 (2016).

952   96.   Smith, E.L., 3rd, Hung, L.F. & Arumugam, B. Visual regulation of refractive development:
953   insights from animal studies. *Eye (Lond)* **28**, 180-8 (2014).

954   97.   Zhang, Y. & Wildsoet, C.F. RPE and Choroid Mechanisms Underlying Ocular Growth and
955   Myopia. *Prog Mol Biol Transl Sci* **134**, 221-40 (2015).

956   98.   Harper, A.R. & Summers, J.A. The dynamic sclera: extracellular matrix remodeling in normal
957   ocular growth and myopia development. *Exp Eye Res* **133**, 100-11 (2015).

958   99.   Summers, J.A. The choroid as a sclera growth regulator. *Exp Eye Res* (2013).

959   100.   Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate
960   genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-
961   9 (2012).

962   101.   Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for
963   genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).

964   102.   van Leeuwen, E.M. *et al.* Population-specific genotype imputations using minimac or IMPUTE2.
965   *Nat Protoc* **10**, 1285-96 (2015).

966   103.   Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide
967   association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).

968   104.   Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans.
969   *Am J Hum Genet* **81**, 913-26 (2007).

970   105.   Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat*
971   *Protoc* **9**, 1192-212 (2014).

972   106.   Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide
973   association scans. *Bioinformatics* **26**, 2190-1 (2010).

974   107.   Zaykin, D.V. Optimally weighted Z-test is a powerful method for combining probabilities in
975   meta-analysis. *J Evol Biol* **24**, 1836-41 (2011).

976   108.   Whitlock, M.C. Combining probability from independent tests: the weighted Z-method is
977   superior to Fisher's approach. *J Evol Biol* **18**, 1368-73 (2005).

978   109.   Dudbridge, F. & Gusnanto, A. Estimation of significance thresholds for genomewide association
979   scans. *Genet Epidemiol* **32**, 227-34 (2008).

980   110.   Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results.
981   *Bioinformatics* **26**, 2336-7 (2010).

982   111.   Yang, H. & Wang, K. Genomic variant annotation and prioritization with ANNOVAR and
983   wANNOVAR. *Nat Protoc* **10**, 1556-66 (2015).

984   112.   Cooper, G.M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence.
985   *Genome Res* **15**, 901-13 (2005).

986   113.   Schwarz, J.M., Rodelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-
987   causing potential of sequence alterations. *Nat Methods* **7**, 575-6 (2010).

988   114.   Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large
989   cohorts. *Nat Genet* **47**, 284-90 (2015).

990   115.   McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*
991   **48**, 1279-83 (2016).

992   116.   Consortium, G.T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis:
993   multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).

994   117.   Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets.
995   *Gigascience* **4**, 7 (2015).

996   118.   Amberger, J., Bocchini, C.A., Scott, A.F. & Hamosh, A. McKusick's Online Mendelian
997   Inheritance in Man (OMIM). *Nucleic Acids Res* **37**, D793-6 (2009).

998   119.   Bauer-Mehren, A., Rautschka, M., Sanz, F. & Furlong, L.I. DisGeNET: a Cytoscape plugin to
999   visualize, integrate, search and analyze gene-disease networks. *Bioinformatics* **26**, 2924-6 (2010).

1000 120. Gunther, S. *et al.* SuperTarget and Matador: resources for exploring drug-target relationships.
1001   *Nucleic Acids Res* **36**, D919-22 (2008).
1002 121. Kuhn, M. *et al.* STITCH 4: integration of protein-chemical interactions with user data. *Nucleic*
1003   *Acids Res* **42**, D401-7 (2014).
1004 122. Wishart, D.S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and
1005   exploration. *Nucleic Acids Res* **34**, D668-72 (2006).
1006 123. Whirl-Carrillo, M. *et al.* Pharmacogenomics knowledge for personalized medicine. *Clin*
1007   *Pharmacol Ther* **92**, 414-7 (2012).
1008
1009

1012   **Figure 1. GWAS meta-analysis identifies 140 loci for refractive error (Stage 3)**

1013   A

1023   B   C
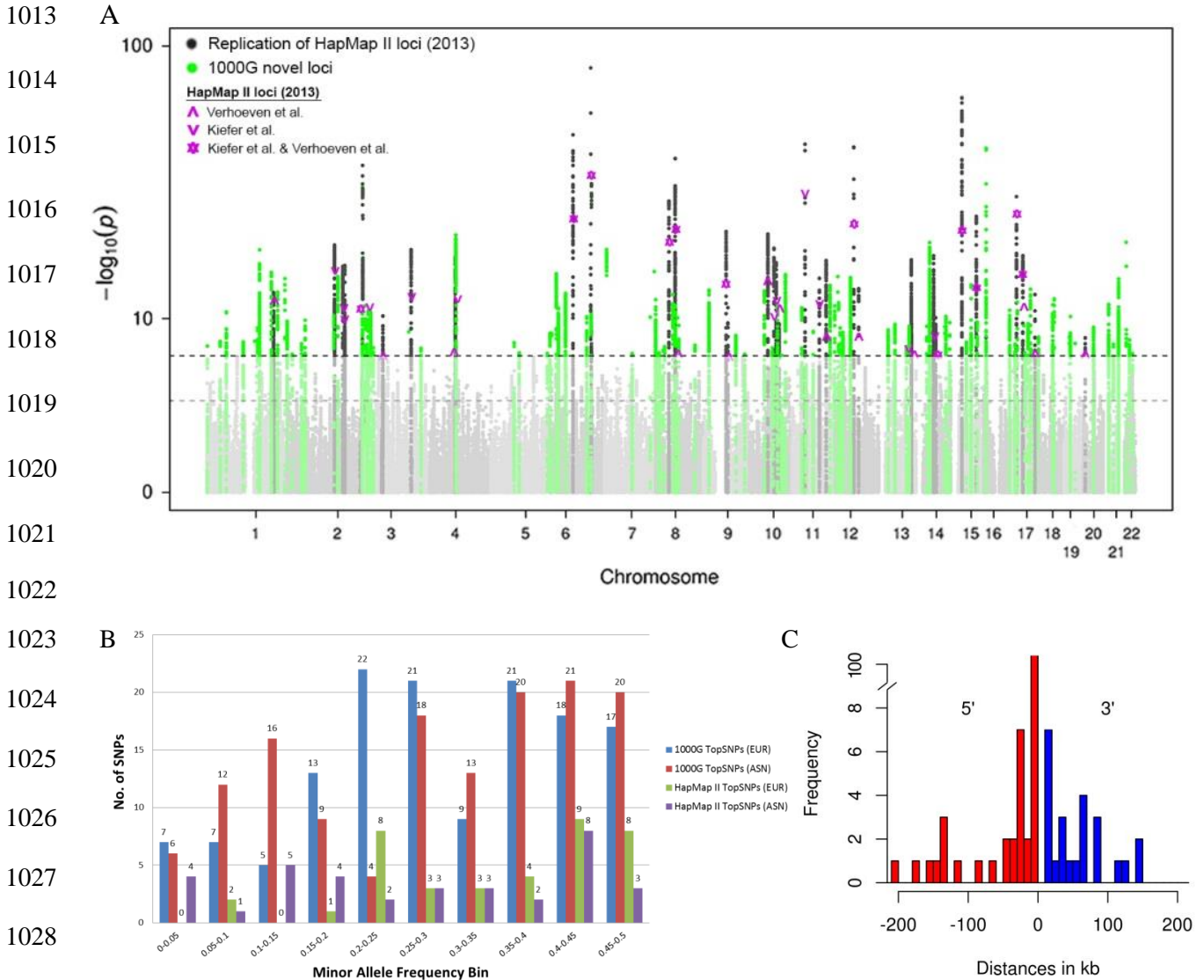
1030   **(a)** We conducted a meta-analysis of genome-wide single-variant analyses for >10 million variants in

1031   160,420 participants of CREAM and 23andMe (Stage 3). Shown is the Manhattan plot depicting *P* values

1032   for association, highlighting new ($P < 5 \times 10^{-8}$ for the first time; green) and known (dark grey) refractive

1033   error loci previously found using HapMap II imputations from Kiefer et al.[27] and Verhoeven et al.[26]

1034   (**Table 1**). The horizontal lines indicate suggestive significance ($P = 1 \times 10^{-5}$) or genome-wide

1035   significance ($P = 5 \times 10^{-8}$). **(b)** We compared the minor allele frequencies of the 140 discovered index

1036   variants based on 1000G (blue: Europeans; red: Asians) to the minor allele frequencies of the previously

1037   found genetic variants based on HapMap II (green: Europeans; purple: Asians). Observed are an increase

1038   in genetic variants found across all minor allele frequency bins increase, including the lower minor allele

1039   frequency bins. **(c)** We annotated the 167 loci to genes using wANNOVAR. Shown are the distances

1040   between index variants from the nearest gene and its gene on the 5' and/or 3' site. The majority of index

1041   variants (84%) were at a distance of less than 50 kb up- or downstream from the annotated gene.

1042
1043

Table 1. Results of the meta-analysis of CREAM and 23andMe for the previously-identified loci and a subset of the newly-identified loci, and replication in UK Biobank

Table 1a Replication of the HapMap II index variants for refractive error per locus in the Stage 3 meta-analysis

| SNP	Chr.	Position	Nearest Loci And Gene(s)	Effect Allele	Other Allele	EAF EUR	EAF ASN	Z-score	Direction	P value	HetISq	HetPVal	Sample Size (N)	HapMap II Discovery (2013)	Category (I = both GWS in Stage 1 and 2, 2=one of two cohorts (CREAM or 23andMe) GWS, 3= both not GWS in Stage 1 or 2)	P value Replication UKBB
rs10500355	16	7459347	RBFOX1	A	T	0.354	0.1334	-13.73	--	6.49E-43	9.1	2.93E-07	160139	Kiefer et al. & Verhoeven et al.	I	2.50E-48
rs11145465	9	71766593	TJP2	A	C	0.212	NA	-9.55	--	1.35E-21	46.3	0.1722	153174	Kiefer et al. & Verhoeven et al.	I	1.00E-10
rs11178469	12	71275137	PTPRR	T	C	0.752	0.6384	-7.40	--	1.33E-13	0	0.6989	160139	Verhoeven et al.	II (CREAM)	2.60E-04
rs11602008	11	40149305	LRRC4C	A	T	0.822	0.7488	13.98	++	2.12E-44	22.5	1.56E-10	157505	Kiefer et al.	II (23andMe)	2.90E-47
rs12193446	6	129820038	BC035400, LAMA2	A	G	0.906	NA	-19.43	--	4.21E-84	16.8	5.72E-15	150269	Kiefer et al. & Verhoeven et al.	I	4.60E-106
rs1550094	2	233385396	CHRNG,PRSS56	A	G	0.701	0.705	12.74	++	3.64E-37	26.3	0.002705	159422	Kiefer et al. & Verhoeven et al., Kiefer et al.	I	4.10E-59
rs1649068	10	60304864	BICC1	A	C	0.475	0.5044	-9.44	--	3.77E-21	0	0.7118	160144	Verhoeven et al.	I	7.50E-11
rs17382981	10	94953258	CYP26A1,MYOF	T	C	0.417	0.1901	-6.31	--	2.72E-10	67.9	0.07737	155332	Verhoeven et al.	II (CREAM)	4.10E-07
rs17428076	2	172851936	HAT1,METAP1D	C	G	0.768	0.8542	-8.18	--	2.77E-16	0	0.002854	160151	Kiefer et al.	II (23andMe)	7.50E-08
rs1858001	1	207488004	C4BPA,CD55	C	G	0.676	0.4151	7.28	++	3.45E-13	59.6	0.02007	160149	Verhoeven et al.	II (CREAM)	6.70E-20
rs1954761	11	105596885	GRIA4	T	C	0.371	0.3772	-8.40	--	4.57E-17	0	0.911	160122	Verhoeven et al.	II (CREAM)	1.20E-16
rs2155413	11	84634790	DLG2	A	C	0.482	0.6549	-7.76	--	8.85E-15	0	0.0002987	159504	Kiefer et al.	II (23andMe)	1.10E-17
rs235770	20	6761765	BMP2	T	C	0.372	0.3875	-5.93	--	3.11E-09	0	0.5474	157521	Verhoeven et al.	II (23andMe)	4.80E-11
rs2573081	2	178828507	PDE11A	C	G	0.524	0.5378	8.21	++	2.18E-16	47.6	0.1672	160126	Kiefer et al.	II (23andMe)	1.60E-29
rs2753462	14	60850703	JB175233,C14orf39	C	G	0.296	0.5679	-6.49	--	8.37E-11	73.9	0.05032	157352	Verhoeven et al.	II (CREAM)	2.00E-15
rs2855530	14	54421917	BMP4	C	G	0.507	0.4736	-8.58	--	9.87E-18	41.7	0.1904	160092	Kiefer et al.	I	4.80E-22
rs2908972	17	11407259	SHISA6	A	T	0.415	0.4879	-11.13	--	9.46E-29	23	0.2544	160123	Kiefer et al. & Verhoeven et al.	I	6.10E-29
rs3138141	12	56115778	BLOC1S1-RDH5,RDH5	A	C	0.214	0.1472	13.80	++	2.46E-43	3.2	5.05E-07	157531	Kiefer et al. & Verhoeven et al.	I	2.30E-56
rs4687586	3	53837971	CACNA1D	C	G	0.691	NA	-6.55	--	5.86E-11	0	0.6046	150217	Verhoeven et al.	III	1.60E-08
rs4793501	17	68718734	KCNJ2,BC039327	T	C	0.575	0.444	-7.21	--	5.53E-13	0	0.5917	160150	Verhoeven et al.	II (CREAM)	3.70E-12
rs524952	15	35005886	GOLGA8B,GJD2	A	T	0.475	0.5077	-17.08	--	2.28E-65	67.2	0.01544	160150	Kiefer et al. & Verhoeven et al.	I	1.60E-103
rs56075542	2	146882415	BC040861,PABPC1P2	T	G	0.552	0.4726	-8.99	--	2.39E-19	13.9	0.001284	159478	Kiefer et al.	II (23andMe)	1.30E-18
rs62070229	17	31227593	MYO1D,TMEM98	A	G	0.807	0.8747	8.58	++	9.64E-18	0	0.4158	156570	Verhoeven et al.	I	1.30E-18
rs6495367	15	79375347	RASGRF1	A	G	0.408	0.3988	-10.20	--	1.95E-24	0	0.667	160144	Kiefer et al. & Verhoeven et al.	I	7.20E-37
rs7042950	9	77149837	RORB	A	G	0.732	0.3924	6.80	++	1.07E-11	0	0.9122	160153	Verhoeven et al.	III	2.90E-18
rs72621438	8	60178580	SNORA51,CA8	C	G	0.642	0.6089	-13.14	--	2.03E-39	38.4	0.00559	160128	Kiefer et al. & Verhoeven et al.	I	1.80E-49
rs745480	10	85986554	LRIT2,LRIT1	C	G	0.511	0.4182	8.31	++	9.26E-17	67.3	0.0805	159504	Kiefer et al.	II (23andMe)	8.20E-18
rs7624084	3	141093285	ZBTB38	T	C	0.568	0.6332	-8.81	--	1.24E-18	18.5	0.01802	160151	Kiefer et al.	II (23andMe)	6.50E-17
rs7662551	4	80537638	LOC100506035,PCAT4	A	G	0.723	0.5577	8.53	++	1.47E-17	19.4	0.2653	160147	Verhoeven et al.	I	6.00E-12
rs7692381	4	81903049	C4orf22,BMP3	A	G	0.763	0.6308	9.40	++	5.55E-21	0	0.01253	160134	Kiefer et al.	I	7.50E-13
rs7744813	6	73643289	KCNQ5	A	C	0.591	0.6017	-14.56	--	5.43E-48	35	0.001132	160091	Kiefer et al. & Verhoeven et al.	I	1.00E-75
rs7829127	8	40726394	ZMAT4	A	G	0.792	0.8974	-10.91	--	1.02E-27	15.9	0.0002774	160132	Kiefer et al. & Verhoeven et al.	II (23andMe)	3.10E-22
rs7895108	10	79061458	KCNMA1	T	G	0.351	0.1182	-8.87	--	7.56E-19	32.8	0.02115	160140	Kiefer et al.	II (23andMe)	1.10E-16
rs79266634	16	7309047	RBFOX1	C	G	0.093	0.1151	-5.93	--	3.00E-09	0	0.5614	156268	Kiefer et al. & Verhoeven et al.	III	1.50E-08
rs837323	13	101175664	PCCA	T	C	0.512	0.7625	6.32	++	2.65E-10	35.6	0.2129	160142	Verhoeven et al.	II (23andMe)	5.30E-16
rs9517964	13	100717833	ZIC2,PCCA	T	C	0.589	0.786	8.42	++	3.68E-17	0	0.01962	160121	Kiefer et al.	II (23andMe)	3.40E-20

Table 1b Subset of the new loci harboring the smallest p-values for refractive error in the Stage 3 meta-analysis

SNP	Chr.	Position	Nearest Loci And Gene(s)	Effect Allele	Other Allele	EAF EUR	EAF ASN	Z-score	Direction	P value	HetISq	HetPVal	Sample Size (N)	Category (I = both GWS in Stage 1 and 2, 2=one of two cohorts (CREAM or 23andMe) GWS, 3= not GWS in Stage 1 or 2)	P value Replication UKBB
rs36024104	14	42294993	LRFN5	A	G	0.823	NA	9.09	++	9.86E-20	15.9	0.01414	152585	II (23andMe)	2.20E-12
rs7456039	7	6901710	CCZ1B,LOC100131257	C	G	0.183	NA	8.82	++	1.18E-18	42.1	3.79E-12	121337	II (23andMe)	6.50E-01
rs1556867	1	164213686	5S_rRNA,PBX1	T	C	0.264	0.494	-8.81	--	1.29E-18	71.1	0.06266	160155	II (23andMe)	4.20E-17
rs12667032	7	154406581	DPP6	A	G	0.152	0.317	7.99	++	1.31E-15	82.3	1.02E-11	130790	II (23andMe)	2.10E-01
rs2225986	1	200311910	LINC00862	A	T	0.381	0.169	-7.96	--	1.68E-15	40.2	0.196	160152	II (23andMe)	7.50E-17
rs1207782	6	22059967	LINC00340	T	C	0.577	0.265	-7.92	--	2.47E-15	0	0.8946	160149	I	4.90E-13
rs72826094	10	114801488	TCF7L2	A	T	0.799	0.838	7.88	++	3.20E-15	64.5	0.09323	156825	II (23andMe)	4.90E-02
rs297593	2	157363743	GPD2	T	C	0.286	0.257	-7.82	--	5.45E-15	0	0.5285	159461	II (23andMe)	7.80E-11
rs5442	12	6954864	GNB3	A	G	0.068	NA	-7.82	--	5.48E-15	8.8	0.03693	146217	II (23andMe)	1.20E-33
rs10880855	12	46144855	ARID2	T	C	0.507	0.464	-7.78	--	7.35E-15	0	0.9683	160144	I	4.80E-08
rs12405776	1	242431557	PLD5	T	C	0.220	0.521	7.75	++	9.52E-15	64.9	3.56E-10	153784	II (23andMe)	1.50E-01
rs2150458	21	47377296	PCBP3,COL6A1	A	G	0.455	0.641	7.74	++	1.04E-14	55.7	0.1329	160151	II (23andMe)	1.80E-13
rs12898755	15	63574641	APH1B	A	G	0.245	0.456	7.53	++	4.98E-14	7.9	0.2974	159506	II (23andMe)	1.40E-16
rs7122817	11	117657679	DSCAML1	A	G	0.507	0.662	7.51	++	5.73E-14	73.8	0.05077	160147	II (23andMe)	1.10E-16
rs10511652	9	18362865	SH3GL2,ADAMTSL1	A	G	0.416	0.445	7.36	++	1.91E-13	44.8	0.1782	160149	II (23andMe)	3.50E-18
rs11101263	10	49414181	FRMPD2	T	C	0.258	0.105	-7.33	--	2.33E-13	0	0.3477	160155	II (23andMe)	2.20E-13
rs11118367	1	219790221	LYPLAL1	T	C	0.482	0.630	-7.29	--	3.16E-13	0	0.8576	160141	III	1.20E-13
rs9395623	6	50757699	TFAP2D,TFAP2B	A	T	0.315	0.381	7.25	++	4.16E-13	0	0.9579	160151	III	2.20E-10
rs284816	8	53362145	ST18,FAM150A	A	G	0.163	0.198	-7.21	--	5.52E-13	0	0.9242	160140	III	1.60E-08
rs12965607	18	47391025	MYO5B	T	G	0.857	0.923	7.07	++	1.52E-12	20.8	0.01674	157604	II (23andMe)	8.10E-16
rs7747	4	80827062	ANTXR2	T	C	0.202	0.093	7.03	++	2.05E-12	5.4	0.01267	150327	II (23andMe)	7.70E-16
rs12451582	17	54734643	NOG,C17orf67	A	G	0.369	0.308	7.02	++	2.22E-12	0	0.5925	160155	II (23andMe)	8.80E-18
rs80253120	17	14138507	CDRT15	T	C	0.626	0.723	6.97	++	3.25E-12	58.6	0.12	156054	II (23andMe)	7.20E-11
22:23069851:I	22	23069851	DKFZp667J0810,abParts	ATG	A	0.084	0.1582	6.95	-+	3.56E-12	98.5	4.80E-16	120481	II (23andMe)	9.30E-01
rs7968679	12	9313304	PZP	A	G	0.700	0.894	6.95	++	3.65E-12	0	0.01951	160076	II (23andMe)	4.20E-10
rs11202736	10	90142203	RNLS	A	T	0.717	0.762	-6.92	--	4.53E-12	0	0.4007	160150	II (23andMe)	9.40E-07
rs11088317	21	16574122	NRIP1,USP25	T	C	0.287	0.299	-6.90	--	5.38E-12	72.5	0.05657	160116	II (23andMe)	6.50E-06
rs10853531	18	42824449	SLC14A2	A	G	0.200	0.182	6.88	++	5.89E-12	0	0.6755	160104	III	2.60E-10
rs72655575	8	60556509	SNORA51,CA8	A	C	0.201	0.124	6.87	++	6.54E-12	0	0.8811	156566	I	7.10E-07
rs12998513	2	242879499	CXXC11,AK097934	A	G	0.880	0.676	-6.86	+-	7.15E-12	65.2	4.51E-14	117611	II (23andMe)	7.80E-01
rs1790165	11	131928971	NTM	A	C	0.411	0.283	6.85	++	7.17E-12	0	0.003708	160131	II (23andMe)	1.80E-10
rs511217	11	30029948	METTL15,KCNA4	A	T	0.738	0.729	-6.79	--	1.10E-11	0	0.3626	160143	II (23andMe)	1.40E-17
rs1150687	6	28162469	ZNF192P1,TRNA_Ser	T	C	0.619	0.504	6.78	++	1.17E-11	56.2	0.131	159448	II (23andMe)	3.10E-10
rs56055503	16	80532694	MAF,DYNLRB2	A	G	0.751	0.539	-6.72	--	1.83E-11	0	0.8407	160145	II (23andMe)	8.00E-06
rs9681162	3	8194734	AK124857,LMCD1-AS1	T	C	0.680	0.437	6.70	++	2.10E-11	63	0.1002	160152	II (23andMe)	6.30E-13
rs11589487	1	61342229	AK097193,BC030753	A	G	0.445	0.089	6.67	++	2.64E-11	34.6	0.2163	160143	II (23andMe)	2.20E-10

We identified 140 loci for refractive error with genome-wide significance (P < 5 × 10−8) on the basis the meta-analyses of the genome-wide single-variant linear regressions performed in 160,420 participants of mixed ancestries (CREAM-ASN, CREAM-EUR and 23andMe). Shown are the replication of the previously found loci from HapMap II and a subset of the new loci harboring the smallest p-values. For each locus, represented by an index variant (the variant with smallest p-value in that locus), Effect Allele, Other Allele, effect allele frequencies per ancestry (EAF AZN and EAF EUR), effect size (Z-score), direction of the effect (Direction), the P value, heterogeneity I square (HetISq), heterogeneity P value (HetPval), Sample Size (N), Category and P value of the replication in UK Biobank are shown (Full table: Supplementary Excel Table 1). Chr., chromosome; EAF, effect allele frequency; ASN, Asian; EUR, European; GWS, genome wide significant; UKBB, United Kingdom Biobank.

1049
1050
1051
1052

1053

1054 **Figure 2. Correlation of statistical significance and effect size of SNPs based on spherical equivalent**

1055 **(SphE) in diopters and age of diagnosis of myopia (AODM) in years.**
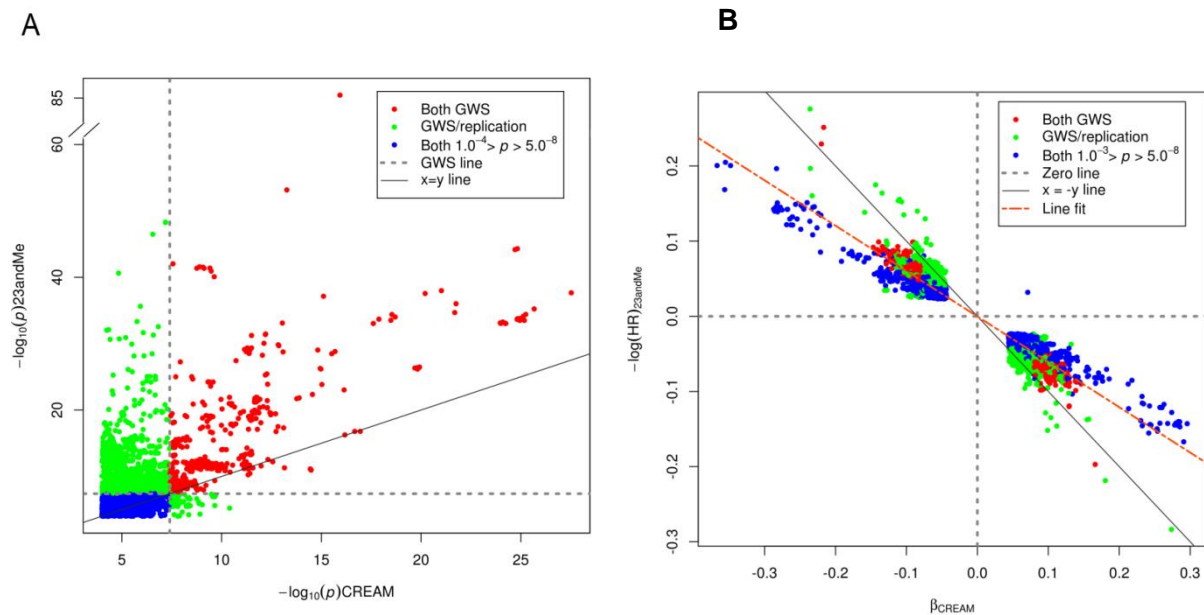
1056

1057

1058 **(a)** *P* value per genetic variant comparison between CREAM meta-analysis (Stage 1) and 23andMe

1059 (Stage 2) meta-analysis. Shown is the overlap (red) and the difference (green) in *P* value signals per

1060 cohort for genetic variants. Green genetic variants are only genome wide significant in either CREAM or

1061 23andMe. Blue: genetic variants with P value between 5.0×10^{-8} and 1.0×10^{-3} in both CREAM and

1062 23andMe. **(b)** Comparison of effects (SphE and logHR of AODM in years) between CREAM and

1063 23andMe. Same color code was applied as in (a). The effects were concordant in their direction of effect

1064 on refractive error. The regression slope is -0.15 diopters per logHR of AODM in years.

1065 **Figure 3. Risk of refractive error per decile of polygenic risk score (Rotterdam Study I-III, N =**

1066 **10,792)**
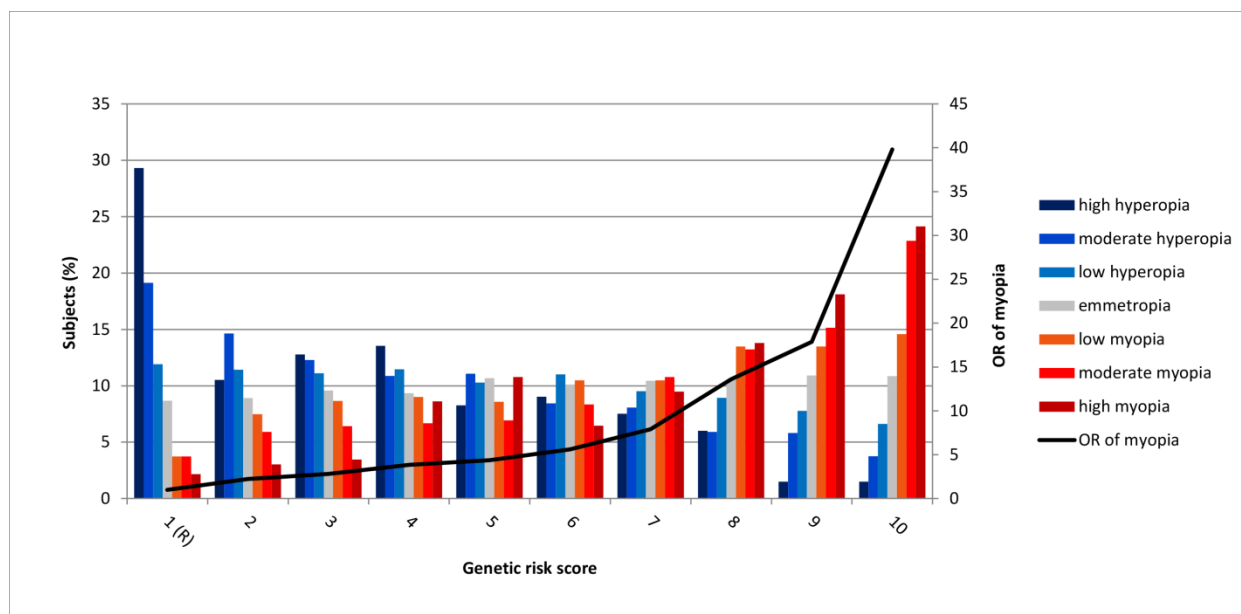
1067

1068 Distribution of refractive error in subjects from Rotterdam Study I–III ($N = 10,792$) as a function

1069 of the optimal polygenic risk score (including 7,303 variants at P value ≤ 0.005 explaining 7.8% of the

1070 variance of SphE; Supplementary Table 9). Mean OR of myopia (black line) was calculated per polygenic

1071 risk score category using the lowest category as a reference. High myopia (SphE ≤-6 D), moderate

1072 myopia (SphE >-6 D & ≤ −3 D), low myopia (SphE > −3 D & <-1.5 D), emmetropia (SphE ≥ −1.5 D and

1073 ≤ 1.5 D), low hyperopia (SphE > 1.5 D & < 3 D), moderate hyperopia (SphE ≥ 3 D & 6 D), high

1074 hyperopia (SphE ≥ 6 D).

1075

1076 **Table 1. Genetic correlation for refractive error between Europeans and East Asians**

Sample 1	Sample 2	Genetic effect correlation (*pge*) [a]	Genetic impact correlation (*pgi*) [a]
EUR CREAM	EAS CREAM	0.804 (se=0.041) $P = 1.83 \times 10^{-6}$	0.888 (se=0.061) $P = 0.065$
EUR 23andMe	EAS CREAM	0.788 (se=0.041) $P = 2.48 \times 10^{-7}$	0.865 (se=0.054) $P = 0.014$
Abbreviations: EUR, European; EAS, East Asian.			
[a] P-value relates to a test of the null hypothesis that *pge*=1 or *pgi*=1.			

1077

1078 We calculated the genetic correlation of effect (pge) and impact (pgi) using Popcorn to compare the

1079 genetic associations between Europeans (CREAM-EUR, N= 44,192; 23andMe, N=104,292) and East

1080 Asians (CREAM-ASN, N= 9,826). Reference panels for Popcorn were constructed using genotype data

1081 for 503 EUR and 504 EAS individuals sequenced as part of the 1000 Genomes Project. SNPs used had a

1082 MAF of at least 5% in both populations, resulting in a final set of 3,625,602SNPs for the analyses using

1083 the 23andMe GWAS sample and 3,642,928 SNPs for those using the CREAM-EUR sample. These

1084 findings support a largely common genetic predisposition to refractive error and myopia in Europeans and

1085 Asians, although ancestry-specific risk alleles may exist.

Figure 4. Visualization of the DEPICT gene-set enrichment analysis based on loci associated with

1087 **refractive error and the correlation between the (meta)gene sets**
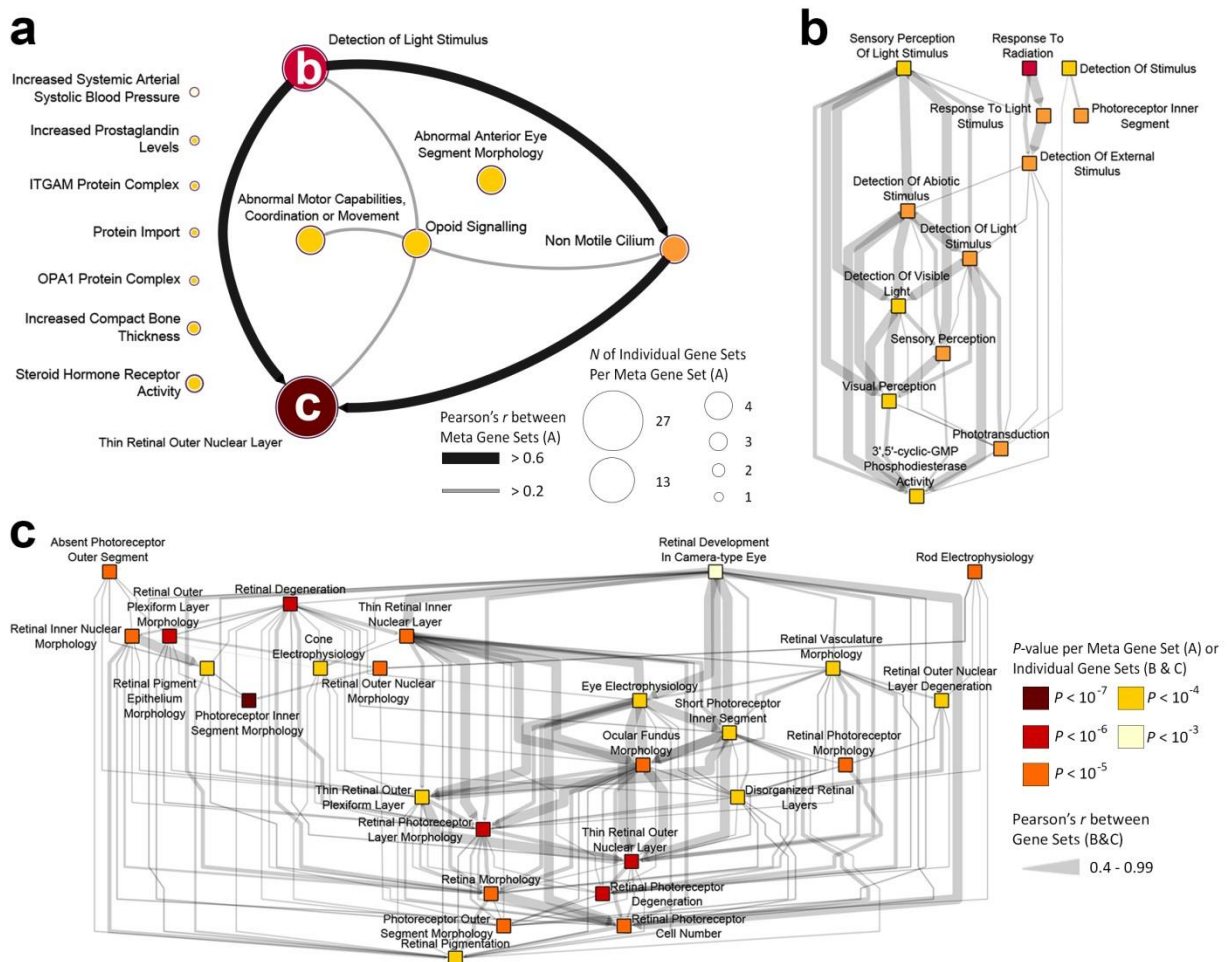
1088

1089 **(a)** Shown are the 66 gene sets clustered into thirteen meta gene sets based on the gene set enrichment

1090 analysis of DEPICT (*P* value < 1 × 10−5 in the GWAS, FDR < 0.05). **(b)** Visualization of the

1091 interconnectivity between gene sets (n=13) of the meta gene set 'Detection of Light Stimulus'

1092 (GO:0009583). **(c)** Visualization of the interconnectivity between gene sets (n=27) of the largest meta

1093 gene set 'Thin Retinal Outer Nuclear Layer' (MP:0008515). In all panels, (meta)gene sets are represented

1094 by nodes colored according to statistical significance, and similarities between them are indicated by

1095 edges scaled according to their correlation; r ≥ 0.2 are shown in panel **(a)** and r ≥ 0.4 are shown in panel

1096 **(b,c).**

Table 2. Genes ranked according to biological and statistical evidence

Locus	Gene Priority Score (Σ)	Internal Replication (≥2 cohorts)	ANNOTATION*: Exonic: Protein Altering	Exonic: Non-Protein Altering	5' or 3' UTR	RNA (nc, sno, linc, other)	EXPRESSION: eQTL	Expression in human adult ocular tissue	Expression in human developing ocular tissue (fetal - 24 weeks)	BIOLOGY: Ocular phenotype in mice	Ocular phenotype in human	PATHWAYS: DEPICT gene-set enrichment	DEPICT gene prioritization	IPA Canonical Pathways	Known Drug Target
Score	Σ	1	1	1	1	1	1	1	1	1	1	1	1	1	
GNB3	8	■	■				■	■	■	■	■				
RDH5	7	■			■		■	■	■	■	■				
CYP26A1	7	■						■	■	■	■	■		■	■
EFEMP1	7	■						■	■	■	■	■		■	
GRIA4	7	■						■	■	■	■	■		■	■
RGR	7	■						■	■	■	■	■	■		
RORB	7	■						■	■	■	■	■		■	■
RCBTB1	6	■					■	■	■		■	■			
MAF	6	■						■	■	■	■			■	
ZEB2	6	■						■	■	■	■			■	
KCNMA1	6	■						■	■	■		■		■	
TJP2	6	■			■			■	■			■		■	
ST18	6	■						■	■	■	■			■	
FBN1	6	■						■	■	■	■			■	
KCNJ2	6	■						■	■	■	■			■	
GJD2	6	■							■	■	■	■		■	
CABP4	6							■	■	■	■		■	■	
TCF7L2	6	■						■	■	■	■			■	
PRSS56	6	■	■					■	■	■	■				

Genes were ranked (orange) based on 10 equal categories which can be divided in four categories: internal replication of genetic variant in more than two cohorts (purple; CREAM-EUR, CREAM-ASN and/or 23andMe), annotation (light blue; genetic variant harboring an exonic protein altering variant or non-protein altering variant, genetic variant residing in a 5' or 3' UTR region of a gene or transcribing an RNA structure), expression (yellow; eQTL, expression in adult human ocular tissue, expression in developing ocular tissue), biology (dark yellow; ocular phenotype in mice, ocular phenotype in humans), pathways (green; DEPICT gene-set enrichment, DEPICT gene prioritization analysis and canonical pathway analysis of IPA). We assessed genes harboring drug targets (salmon red), but did not assign a scoring point to this category.
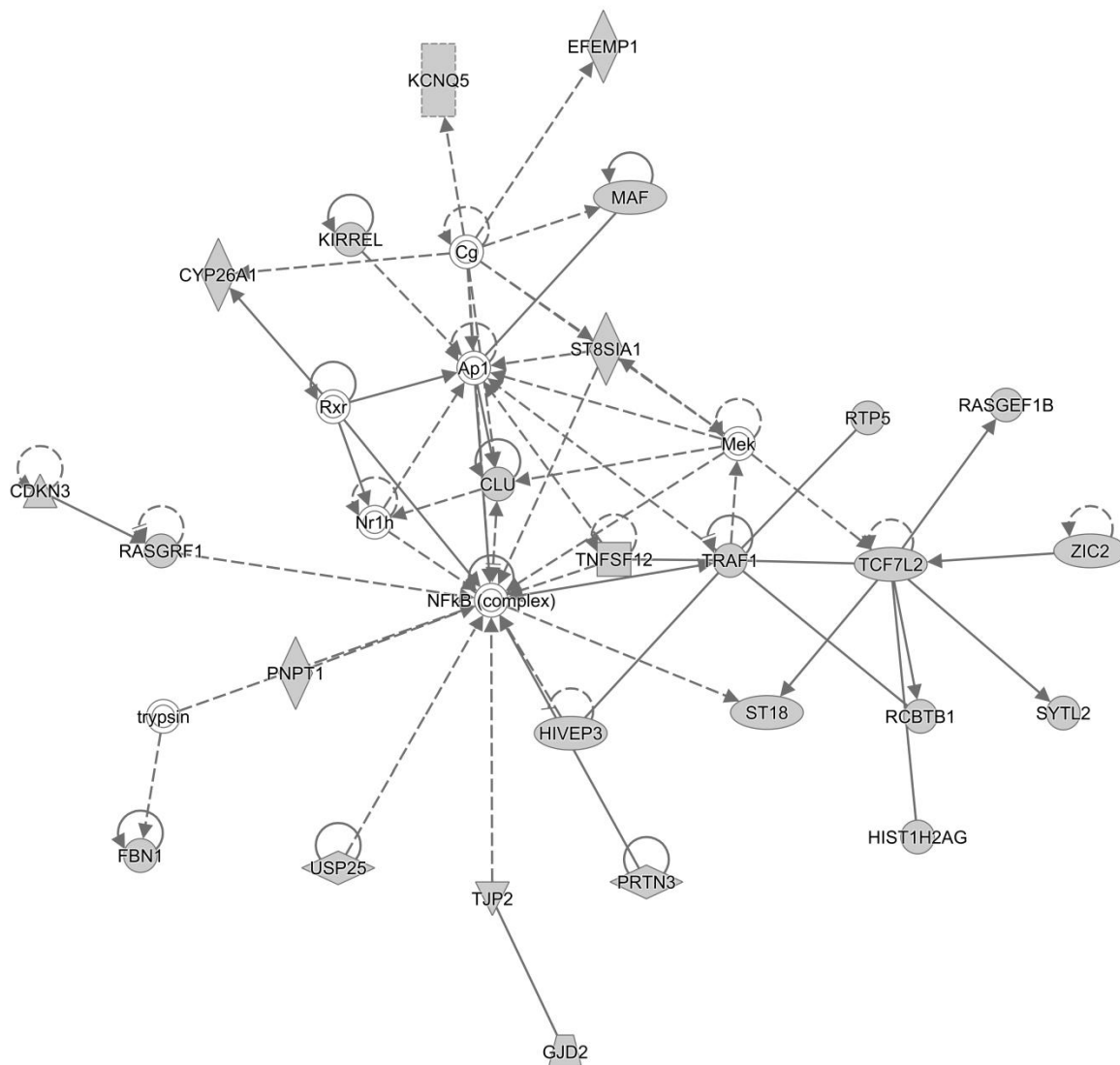
*Only one point can be assigned in the category 'ANNOTATION', even though it has four columns (i.e. a genetic variant is located in only 1 of these four categories).

1111   **Figure 5. Top molecular network identified by Ingenuity Pathway Analysis (IPA)**


1112

1113

1114   The genes annotated to the top hits identified at Stage 3 were mapped to networks and pathways present

1115   in the IPA database. The most significant network identified by Fisher's exact test was 'Glutamate

1116   Receptor Signaling' (P value= 1.56 x 10^{-4}). Genes within the network indicated in grey are genes

1117   associated with refractive error. Other significant pathways are depicted in Supplementary Figure 10.

1118 **Figure 6. Schematic representation of the human eye, retinal cell types, and functional sites of associated genes**
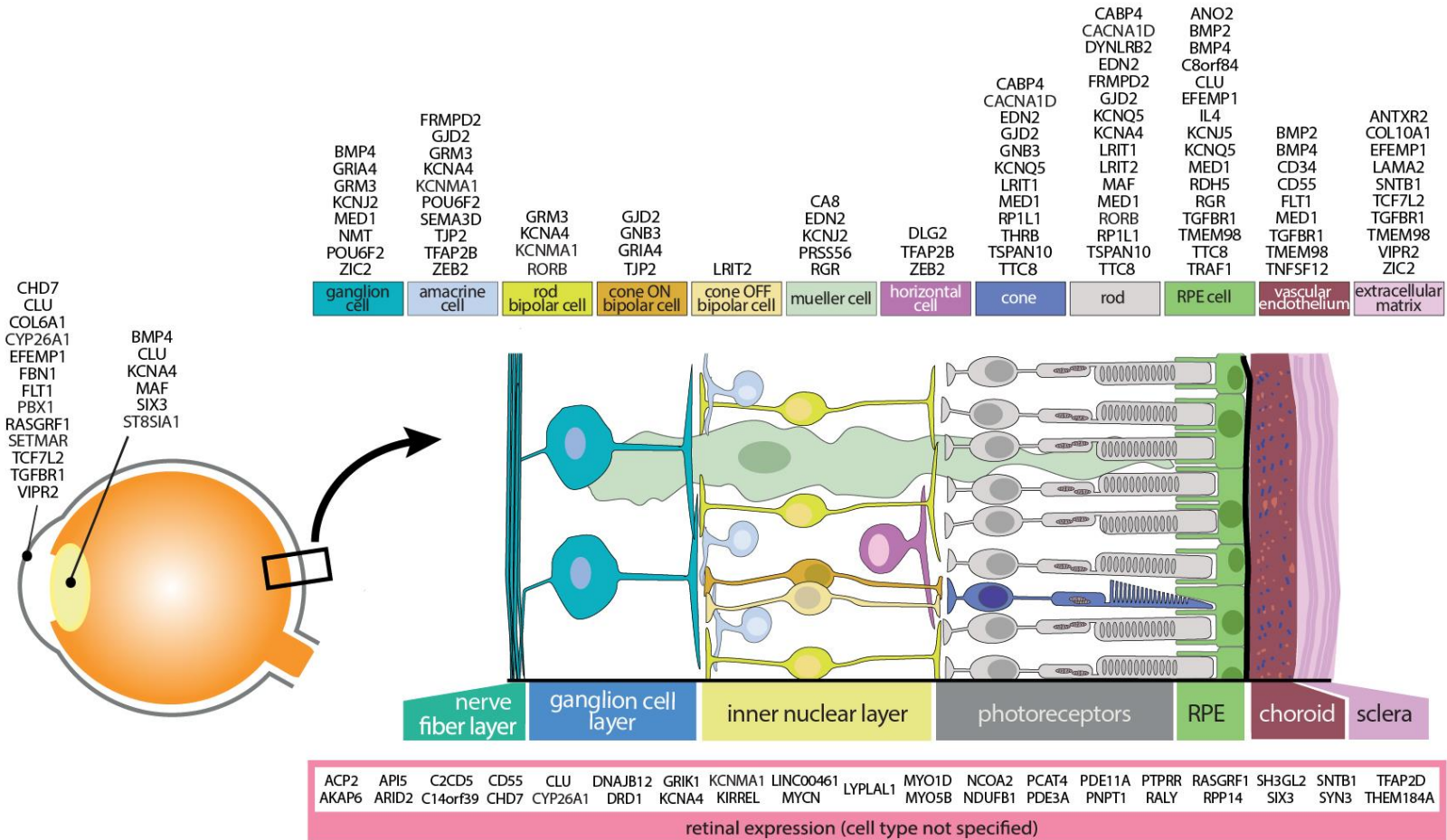
1119

1120 We assessed gene expression sites and/or functional target cells in the eye for all genes using our expression data and literature and data present in

1121 the public domain. The genes appear to be distributed across virtually all cell types in the neurosensory retina, in the RPE, vascular endothelium

1122 and extracellular matrix; i.e., the route of the myopic retina-to-sclera signalling cascade.

1

The CREAM Consortium

Tin Aung[1,2], Joan E. Bailey-Wilson[3], Paul Nigel Baird[4], Amutha Barathi Veluchamy[1,5], Ginevra Biino[6], Kathryn P. Burdon[7], Harry Campbell[8], Li Jia Chen[9], Peng Chen[2], Wei Chen[10], Ching-Yu Cheng[1,2,5], Emily Chew[11], Jamie E. Craig[7], Phillippa Cumberland[12,13], Margaret M. Deangelis[14], Cécile Delcourt[15,16], Xiaohu Ding[17], Angela Döring[18], Cornelia M. van Duijn[19], David M. Evans[20,21], Qiao Fan[22], Lindsay Farrer[14], Sheng Feng[23], Brian Fleck[24], Rhys D. Fogarty[7], Jeremy R. Fondran[25], Maurizio Fossarello[26], Paul J. Foster[27,28], Puya Gharahkhani[29], Christian Gieger[18], Adriana I. Iglesias Gonzalez[19,30],, Jeremy A. Guggenheim[31], Xiaobo Guo[17,32], Toomas Haller[33], Christopher J. Hammond[34,35], Caroline Hayward[36], Mingguang He[4,17], Alex W. Hewitt[4,37], René Höhn[38], S. Mohsen Hosseini[39], Laura D. Howe[20,40], Pirro G Hysi[35], Robert P. Igo Jr[25], Sudha K. Iyengar[11,25,41], Sarayut Janmahasatian[25], Vishal Jhanji[9], Jost B. Jonas[42-44], Mika Kähönen[45], Jaakko Kaprio[46], John P. Kemp[20], Kay-Tee Khaw[47], Anthony P. Khawaja[47], Chiea-Chuen Khor[2,48-50], Caroline C. W. Klaver[19,30], Barbara E. Klein[51], Ronald Klein[51], Eva Krapohl[52], Jean-François Korobelnik[15,16], Jonathan H. Lass[25,41], Elisabeth M. van Leeuwen[19,30], Terho Lehtimäki[53], Yi Lu[29], Robert N. Luben[47], Stuart MacGregor[29], David A. Mackey[4,37], Kari-Matti Mäkelä[53], Nicholas G. Martin[54], George McMahon[20], Akira Meguro[55], Thomas Meitinger[56], Andres Metspalu[33], Evelin Mihailov[33], Paul Mitchell[57], Masahiro Miyake[58], Nobuhisa Mizuki[55], Margaux Morrison[14], Vinay Nangia[59],Songhomitra Panda-Jonas[59], Chi Pui Pang[9], Olavi Pärssinen[60,61], Andrew D. Paterson[39], Norbert Pfeiffer[38], Mario Pirastu[62], Robert Plomin[52], Ozren Polasek[8,63], Jugnoo S. Rahi[12,13,64], Olli Raitakari[65,66], Taina Rantanen[60,67], Janina S. Ried[18], Igor Rudan[8], Seang-Mei Saw[1,48], Maria Schache[4], Ilkka Seppälä[53], George Davey Smith[20], Dwight Stambolian[68], Beate St Pourcain[20,69], Claire L. Simpson[3,70], E-Shyong Tai[5,48,71], Pancy O. Tam[9], Milly S. Tedja[19,30], Yik-Ying Teo[48,72], J. Willem L. Tideman[19,30], Nicholas J. Timpson[20], Simona Vaccargiu[62], Zoran Vatavuk[73], Virginie J.M. Verhoeven[19,30], Veronique Vitart[36], Jie Jin Wang[4,57], Nick J. Wareham[74], Juho Wedenoja[46,75], Cathy Williams[40], Katie M Williams[34,35], James F. Wilson[8], Robert Wojciechowski[3,76,77], Ya Xing Wang[42,43], Tien-Yin Wong[1,2], Alan F. Wright[36], Jing Xie[4],

1148   Liang Xu[42], Kenji Yamashiro[58], Maurice K.H. Yap[78], Seyhan Yazar[37], Shea Ping Yip[79], Nagahisa

1149   Yoshimura[58], Alvin L. Young[9], Terri L. Young[5,51], Jing Hua Zhao[74], Wanting Zhao[1], Xiangtian Zhou[10]

1150
1151
1152   1.   Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore.
1153   2.   Department of Ophthalmology, National University Health Systems, National University of
1154   Singapore, Singapore.
1155   3.   Computational and Statistical Genomics Branch, National Human Genome Research Institute,
1156   National Institutes of Health, Baltimore, MD, USA.
1157   4.   Centre for Eye Research Australia (CERA), Ophthalmology, Department of Surgery, University
1158   of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia.
1159   5.   Duke-NUS Medical School, Singapore, Singapore.
1160   6.   Institute of Molecular Genetics, National Research Council, Pavia, Italy.
1161   7.   Department of Ophthalmology, Flinders University, Adelaide, Australia.
1162   8.   Centre for Global Health Research, The Usher Institute for Population Health Sciences and
1163   Informatics, University of Edinburgh, Scotland, UK.
1164   9.   Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong Hong
1165   Kong Eye Hospital, Kowloon, Hong Kong.
1166   10.   Department of Radiology, Southwest Hospital, The Third Military Medical University,
1167   Chongqing, China.
1168   11.   Department of Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
1169   12.   Institute of Child Health, University College London, London, UK.
1170   13.   Ulverscroft Vision Research Group, University College London, London, UK.
1171   14.   Department of Ophthalmology and Visual Sciences, John Moran Eye Center, University of Utah,
1172   Salt Lake City, Utah, USA.
1173   15.   Université de Bordeaux, Bordeaux, France.
1174   16.   INSERM (Institut National de la Santé Et de la Recherche Médicale), ISPED (Institut de Santé
1175   Publique d'Épidémiologie et de Développement), Centre INSERM U897-Epidemiologie-
1176   Biostatistique, Bordeaux, France.
1177   17.   State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University,
1178   Guangzhou, China.
1179   18.   Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for
1180   Environmental Health, Neuherberg, Germany.
1181   19.   Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.
1182   20.   MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, UK.
1183   21.   University of Queensland Diamantina Institute, Translational Research Institute, Brisbane,
1184   Queensland, Australia.
1185   22.   Centre for Quantitative Medicine, Duke-NUS Medical School, Singapore, Singapore.
1186   23.   Department of Pediatric Ophthalmology, Duke Eye Center For Human Genetics, Durham, North
1187   Carolina, USA.
1188   24.   Princess Alexandra Eye Pavilion, Edinburgh, UK.
1189   25.   Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland,
1190   Ohio, USA.
1191   26.   University Hospital 'San Giovanni di Dio', Cagliari, Italy.
1192   27.   Division of Genetics and Epidemiology, UCL Institute of Ophthalmology, London, UK.
1193   28.   NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL
1194   Institute of Ophthalmology, London, UK.
1195   29.   Statistical Genetics Laboratory, QIMR Berghofer Medical Research Institute, Herston, Brisbane,
1196   Queensland, Australia.

1197 30. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands.
1198 31. School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
1199 32. Department of Statistical Science, School of Mathematics & Computational Science, Sun Yat-
1200 Sen University, Guangzhou.
1201 33. Estonian Genome Center, University of Tartu, Tartu, Estonia.
1202 34. Department of Ophthalmology, King's College London, St Thomas' Hospital campus, London,
1203 UK.
1204 35. Department of Twin Research and Genetic Epidemiology, King's College London School of
1205 Medicine, London, UK.
1206 36. Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine,
1207 University of Edinburgh, Edinburgh, UK.
1208 37. Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western
1209 Australia, Perth, Australia.
1210 38. Department of Ophthalmology, University Medical Center Mainz, Mainz, Germany.
1211 39. Genetics and Genome Biology Program, Hospital for Sick Children and University of Toronto,
1212 Toronto, Ontario, Canada.
1213 40. School of Social and Community Medicine, University of Bristol, Bristol, UK.
1214 41. Department of Ophthalmology and Visual Sciences, Case Western Reserve University and
1215 University Hospitals Eye Institute, Cleveland, Ohio, USA.
1216 42. Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University,
1217 Beijing, China.
1218 43. Beijing Ophthalmology and Visual Science Key Lab, Beijing, China.
1219 44. Department of Ophthalmology, Medical Faculty Mannheim, Ruprecht-Karls-University
1220 Heidelberg, Mannheim, Germany.
1221 45. Department of Clinical Physiology, Tampere University Hospital and School of Medicine,
1222 University of Tampere, Tampere, Finland.
1223 46. Department of Public Health, University of Helsinki, Helsinki, Finland.
1224 47. Department of Public Health and Primary Care, Institute of Public Health, University of
1225 Cambridge School of Clinical Medicine, Cambridge, UK.
1226 48. Saw Swee Hock School of Public Health, National University Health Systems, National
1227 University of Singapore, Singapore, Singapore.
1228 49. Division of Human Genetics, Genome Institute of Singapore, Singapore, Singapore.
1229 50. Department of Ophthalmology and Visual Sciences, University of Utah, Moran Eye Center, Salt
1230 Lake City, USA.
1231 51. Ophthalmology and Visual Sciences, University of Wisconsin-Madison, Madison, WI, USA.
1232 52. MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology
1233 & Neuroscience, King's College London, London, UK.
1234 53. Department of Clinical Chemistry, Fimlab laboratories and School of Medicine, University of
1235 Tampere, Tampere, Finland.
1236 54. Genetic Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Herston,
1237 Brisbane, Queensland, Australia.
1238 55. Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama,
1239 Kanagawa, Japan.
1240 56. Institute of Human Genetics, Technical University Munich, Munich, Germany.
1241 57. Centre for Vision Research, Department of Ophthalmology and Westmead Institute for Medical
1242 Research, University of Sydney, Sydney, Australia.
1243 58. Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of
1244 Medicine, Kyoto, Japan.
1245 59. Suraj Eye Institute, Nagpur, Maharashtra, India.
1246 60. Department of Health Sciences and Gerontology Research Center, University of Jyväskylä,
1247 Jyväskylä, Finland.

1248   61.   Department of Ophthalmology, Central Hospital of Central Finland, Jyväskylä, Finland.
1249   62.   Institute of Genetic and Biomedic Research, National Research Council, Cagliari, Italy.
1250   63.   Faculty of Medicine, University of Split, Split, Croatia.
1251   64.   Institute of Ophthalmology, Moorfields Eye Hospital, London, UK.
1252   65.   Research Centre of Applied and Preventive Medicine, University of Turku, Turku, Finland.
1253   66.   Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku,
1254   Finland.
1255   67.   Gerontology Research Center, University of Jyväskylä, Jyväskylä, Finland.
1256   68.   Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
1257   69.   Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands.
1258   70.   Deptartment of Genetics, Genomics and Informatics, University of Tennessee Health Science
1259   Center, Memphis, TN, USA.
1260   71.   Department of Medicine, National University of Singapore, Singapore, Singapore.
1261   72.   Department of Statistics and Applied Probability, National University of Singapore, Singapore,
1262   Singapore.
1263   73.   Department of Ophthalmology, Sisters of Mercy University Hospital, Zagreb, Croatia.
1264   74.   MRC Epidemiology Unit, Institute of Metabolic Sciences, University of Cambridge, Cambridge,
1265   UK.
1266   75.   Department of Ophthalmology, Helsinki University Central Hospital, Helsinki, Finland.
1267   76.   Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore,
1268   Maryland, USA.
1269   77.   Wilmer Eye Institute, Johns Hopkins School of Medicine, Baltimore, MD, USA.
1270   78.   Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University,
1271   Hong Kong, Hong Kong.
1272   79.   Department of Health Technology and Informatics, The Hong Kong Polytechnic University,
1273   Hong Kong, Hong Kong.
1274