OXFORD

Human Molecular Genetics, 2022, Vol. 31, 17, 3012–3019
https://doi.org/10.1093/hmg/ddac048
Advance access publication date: 26 February 2022
Association Studies Article

A genome-wide analysis of 340 318 participants identifies four novel loci associated with the age of first spectacle wear

Karina Patasova[1,2], Anthony P. Khawaja[3], Robert Wojciechowski[4], Omar A. Mahroo[1,2,3,5], Mario Falchi[2], Jugnoo S. Rahi[6,7,8], Chris J. Hammond[1,2], Pirro G. Hysi[1,2,7,*] and the UK Biobank Eye & Vision Consortium

[1]Section of Ophthalmology, School of Life Course Sciences, King's College London, London SE1 7EH, UK
[2]Department of Twin Research and Genetic Epidemiology, School of Life Course Sciences, King's College London, London SE1 7EH, UK
[3]NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and the UCL Institute of Ophthalmology, London WC1E 6BT, UK
[4]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA
[5]Department of Ophthalmology, St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust, London SE1 7EH, UK
[6]Institute of Ophthalmology, University College London, London WC1E 6BT, UK
[7]UCL Great Ormond Street Hospital Institute of Child Health, London WC1N 1EH, UK
[8]Ulverscroft Vision Research Group, University College London, London WC1N 1EH, UK
*To whom correspondence should be addressed at: St Thomas' Hospital, Westminster Bridge Road, London SE1 7EH, UK. Tel: +44 (0)2071888545; Fax: +44 (0)2071886761; Email: pirro.hysi@kcl.ac.uk

Abstract

Refractive errors, particularly myopia, are the most common eye conditions, often leading to serious visual impairment. The age of onset is correlated with the severity of refractive error in adulthood observed in epidemiological and genetic studies and can be used as a proxy in refractive error genetic studies. To further elucidate genetic factors that influence refractive error, we analysed self-reported age of refractive error correction data from the UK Biobank European and perform genome-wide time-to-event analyses on the age of first spectacle wear (AFSW). Genome-wide proportional hazards ratio analyses were conducted in 340 318 European subjects. We subsequently assessed the similarities and differences in the genetic architectures of refractive error correction from different causes. All-cause AFSW was genetically strongly correlated ($r_g = -0.68$) with spherical equivalent (the measured strength of spectacle lens required to correct the refractive error) and was used as a proxy for refractive error. Time-to-event analyses found genome-wide significant associations at 44 independent genomic loci, many of which (*GJD2, LAMA2*, etc.) were previously associated with refractive error. We also identified six novel regions associated with AFSW, the most significant of which was on chromosome 17q ($P = 3.06 \times 10^{-09}$ for rs55882072), replicating in an independent dataset. We found that genes associated with AFSW were significantly enriched for expression in central nervous system tissues and were involved in neurogenesis. This work demonstrates the merits of time-to-event study design in the genetic investigation of refractive error and contributes additional knowledge on its genetic risk factors in the general population.

Introduction

Refractive errors, particularly myopia, are the most common eye conditions, often leading to serious visual impairment (1). The prevalence of myopia has increased over the past decades, reaching the highest rates in East Asia (2), but also in Europe (3) and the United States (4). Refractive errors arise from a mismatch between the cornea's refractive power and the crystalline lens on one side and the eye's axial length on the other. The physiological process that normally balances them, called emmetropization, consists of a gradual elongation of the sagittal diameter of the eye to match the eyes' refractive power (5). Refractive error results when light converges in front of the retina (myopia), behind the retina (hypermetropia) or follows other non-optimal patterns of light convergence. The strength of spectacles or contact lenses to correct refractive errors and focus light on the retina in these adult volunteers is summarized by the spherical equivalent, with a minus number denoting a concave lens for myopia correction or a plus number for a convex lens correcting hyperopia or long-sightedness. Refractive errors are often underdiagnosed, and delays in correcting them can result in productivity loss. They may also lead to complications causing visual impairment and potentially blindness. High myopia is associated with later-life posterior staphylomas, retinal detachment, cataract and other complications (6–8). The likelihood of high and pathological myopia increases proportionally with the gravity of refractive error, which is correlated with the age at which myopia first developed.

Environmental factors, such as educational attainment (9) and time spent outdoors, vastly influence the development and progression of myopia. Their effects depend on lifestyles and cultural trends, but

Received: July 6, 2021. **Revised:** February 14, 2022. **Accepted:** February 24, 2022

they typically affect whole cohorts across countries and societies (3) sharing similar living environments. Within a society at any given time, the environmental exposures are stable and relatively homogenously distributed, and heritable factors explain over half of the spherical equivalent and risk to refractive error (10). Several genetic studies conducted in the general population have identified DNA variations associated with the risk of refractive error (11,12) and age of first lens or spectacle correction for myopia (13). Genes associated with the age at first correction for myopia usually overlap with those associated with spherical equivalent (14), and both predispose to pathological myopia (15). Yet, the timing of the individual genes' effects is not evenly distributed throughout the childhood years or lifetime. Different genes have varying strength of effect and association throughout the years, and among the genetic factors associated with spherical equivalent, some genes predispose to earlier refractive correction than others (16). There is also considerable genetic pleiotropy in the eye and the same genetic factors may be independently associated with several endophenotypes (17) each a potential to alter the age in which correction of refractive errors is needed.

This study aims to explore the genetic factors that contribute to the risk of early onset of refractive error, using as a proxy the self-reported age of first spectacle wear (AFSW) in a sample of 340 318 UK Biobank participants. This study also further explores the genetic relationship between age of refractive correction and mean spherical equivalent.

Results

The final study sample included 340 318 UK Biobank participants of European ancestry who reported the AFSW in the electronic questionnaire; of them, 46% ($N = 156\,388$) were men and 54% ($N = 183\,930$) were women with a mean age of 58 years (±7.5 years). The AFSW followed bimodal distribution with the first mode between 1 and 35 years, peaking at the age of 13, and the second mode between the ages of 36 and 72 years with a peak at the age of 43 (Supplementary Material, Fig. S1). Participants that started wearing glasses/contact lenses before the age of 35 tended to be more myopic, while subjects with AFSW over 35 years were more likely to have hyperopia (Supplementary Material, Fig. S2). For the vast majority of the study participants, the cause of spectacle wearing was not known. For the subset of participants who specified the reason for wearing glasses/contact lenses ($N = 93\,067$), 41% ($N = 37\,762$) reported myopia. Presbyopia (33%; $N = 31\,137$) and hypermetropia (21%; $N = 19\,178$) were the second and the third most commonly self-reported reasons for refractive correction (18).

The large study sample size ($N = 340\,318$) resulted in relatively high genomic inflation factor ($\lambda = 1.23$) in our analyses of time to the first spectacle correction, but the low intercept of the linkage disequilibrium score regression 0.93, and (intercept-1)/(mean (X2) − 1) ratio (−0.19, SE = 0.02), reassuringly indicate a conservative control for potential confounding in our study.

Table 1. Genetic correlations between SPHE GWAS effects and genome-wide survival analyses

	AFSW all	AFSW, myopia only	AFSW, hyperopia only
Spherical equivalent	–0.683	–0.968	0.808
AFSW all		0.889	–0.085
AFSW, myopia only			–0.651

Each value represents the pairwise genetic correlation (r_g) observed between the trait shown in the table headers and rows.

We first assessed the degree of similarity between the genomic architectures of the spherical equivalent, self-reported age of first lens or spectacle correction for myopia, first self-reported correction for hyperopia and self-reported first correction for any reason. Consistent with previous reports, we found a strong genetic correlation between age of first myopia correction and spherical equivalent ($r_g = -0.97$). We also noted that the age of first correction in participants with myopia alone is also strongly genetically correlated with the age of the first correction of any refractive error ($r_g = 0.89$, Table 1) and less so with the age of the first correction among hyperopic subjects ($r_g = -0.65$). Spherical equivalent and all-cause AFSW shared most of their heritability and were significantly correlated ($r_g = -0.68$, $P = 9.6 \times 10^{-171}$). Because of the strong correlations and the expectation of superior statistical power arising from the larger sample sizes phenotyped for AFSW, we focused this work on the analysis of all-cause AFSW.

Our genome-wide association study for time to the first lens or spectacle wear found a significant association with 44 independent genomic regions (Fig. 1), many of which previously reported in relation to refractive errors (12). The statistically strongest association was observed between AFSW and *TSPAN10* gene (rs7405453, HR = 1.03, $P = 1.71 \times 10^{-35}$). The second strongest association was found at another locus previously associated with spherical equivalent (rs4736886, near the *ZMAT4* gene, $P = 3.36 \times 10^{-27}$). Interestingly, both genes that show the most significant associations with AFSW, although known for associations with refractive error, have relatively low effect sizes over the spherical equivalent. Only further down the list of our genome-wide associations with AFSW do we find the genes usually considered as the strongest risk factors to refractive error, such as *GJD2*, *LAMA2* and *PRSS56* ($P = 1.63 \times 10^{-12}$, $P = 6.27 \times 10^{-24}$ and $P = 1.31 \times 10^{-18}$), respectively).

Although the effects of association with spherical equivalent were usually linearly correlated with their effect over the AFSW for the same alleles, there were notable exceptions. For example, the single-nucleotide polymorphism (SNP) alleles at the known *BMP3*, *ZMAT4* and *TSPAN10* loci predispose to much earlier correction compared to the final spherical equivalent status in

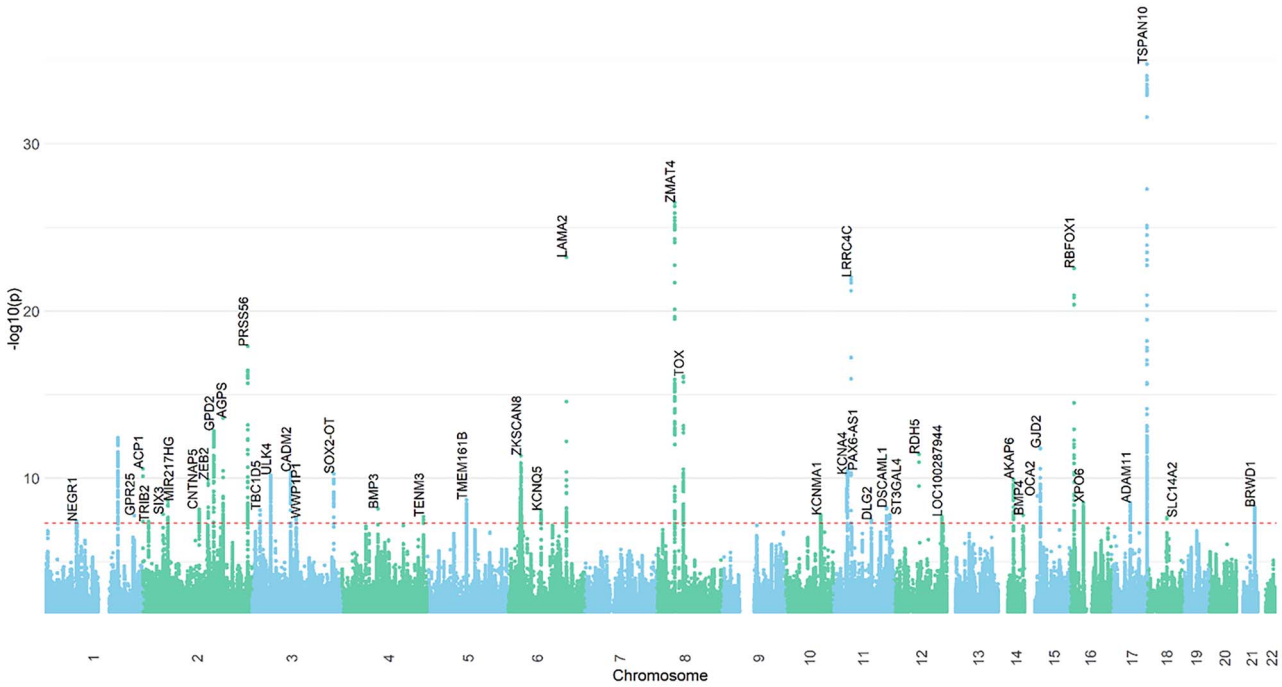

Figure 1. Manhattan plot displaying 44 genome-significant associations with the AFSW in UK Biobank cohort ($N = 340\,318$). The plot shows log10 transformed *P*-values for each marker plotted against the chromosomal location. The red dashed line indicates the genome-wide significance threshold (*P*-value $<5 \times 10^{-08}$). Regions are named with symbols of the transcript-coding genes nearest to the most strongly associated variant in the region.

adulthood than most other loci. Conversely, alleles in the *SOX2-OT* gene seem to confer a low risk towards myopia, but at a much later age than the general regression line across all loci (Fig. 2). Interestingly, there were some examples (e.g. BMP3), where the association with AFSW was different in individuals with myopia only compared to the entire sample that included corrections for all sources of refractive error. This maybe attributable to the particularities of the effects of these genes on the AFSW among myopes. However, most of the effects observed showed similar patterns of relationship between AFSW and spherical equivalent in subgroup analyses, such as time-to-event analyses conducted on a sample of 45 404 UK Biobank participants that excluded causes of refractive error other than myopia (Supplementary Material, Fig. S3).

We observe a genome-wide association with six additional loci that, to our knowledge, were not described in any previous genome-wide associations study (GWAS) for refractive error (12). We discovered new associations with polymorphisms within the genomic sequence of the *NEGR1* gene (rs1204700722, HR = 1.013, $P = 3.72 \times 10^{-08}$), a member of an immunoglobulin superfamily cell adhesion molecule supergroup, implicated in neuronal growth and connectivity (12), where previous studies have identified association with depression and affective disorders (19). Novel significant association was also found at a chromosome 2 intergenic region between the *TRIB2* and *LOC1005064* gene sequences (rs10164589, HR = 1.013, $P = 3.96 \times 10^{-08}$). The *TRIB2* gene is a pseudokinase family member that regulates intracellular cell

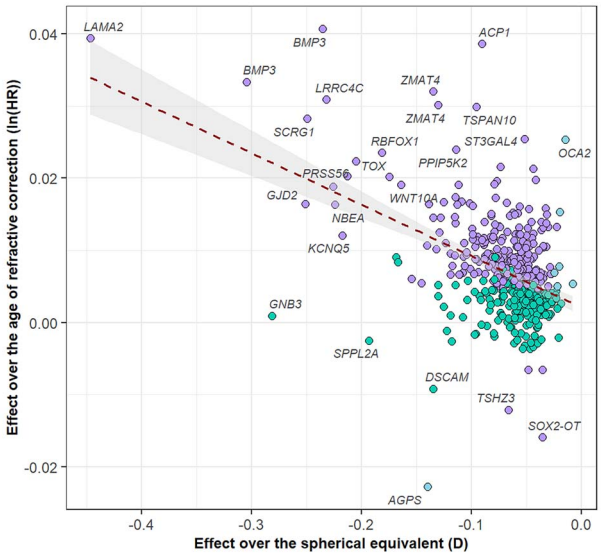

Figure 2. Scatterplot displaying the correlation between the AFSW hazards ratios and spherical equivalent beta coefficients. Hazard ratios shown here as (ln (HR)) represent the multiplicative change in the rate of first spectacle wear per copy of the myopia risk allele calculated in the full sample of 340 318 UK Biobank participants, which was taken as reference. The results are shown for the most strongly associated SNPs in their respective loci. The purple labels depict names of some of the gene loci exhibiting stronger effects over the AFSW, SNPs in blue are associated with spherical equivalent but not AFSW and SNPs in turquoise are associated with AFSW but not spherical equivalent.

signalling through ubiquitination and scaffolding (20). Additionally, we found an association for a locus on chromosome 3 (rs6577621, HR = 1.014, $P = 8.15 \times 10^{-09}$) in a region located between the *TBC1D5* gene, a regulator

Table 2. Replication of six novel loci associated with AFSW

CHR	BP	SNP	Gene	A1	A2	Freq.	Discovery HR	Discovery P-value	Replication HR	Replication P-value
1	72 720 383	rs1194277*	*NEGR1*	G	T	0.69	1.013	3.72×10^{-08}	1.009	0.96
2	**13 042 958**	**rs10164589**	***TRIB2***	**T**	**G**	**0.48**	**1.013**	$\mathbf{3.96 \times 10^{-08}}$	**1.022**	**0.001**
3	**18 192 988**	**rs6577621**	***TBC1D5***	**G**	**A**	**0.45**	**1.014**	$\mathbf{8.15 \times 10^{-09}}$	**1.016**	**0.01**
12	106 927 958	rs7295942	*LOC100287944*	C	T	0.75	1.015	1.96×10^{-08}	1.007	0.38
17	**42 847 438**	**rs55882072**	***ADAM11***	**C**	**G**	**0.72**	**1.015**	$\mathbf{3.06 \times 10^{-09}}$	**1.028**	**0.0002**
21	**40 575 426**	**rs8131965**	***BRWD1***	**G**	**A**	**0.64**	**0.986**	$\mathbf{5.41 \times 10^{-09}}$	**0.974**	**0.00007**

Replication was carried out using the results of a genome-wide time-to-event study on age of first correction for myopia by Kiefer et al. (13). The field 'SNP' includes the polymorphic variants with the strongest associations (Discovery *P*-value) for each region, for which the Chromosome number (CHR) and genomic position (BP) are displayed. A1 lists the alleles at each SNP locus for which the effect sizes (Discovery HR as hazard ratios) and frequencies (Freq.) are reported, and the field 'A2' lists alleles alternative to effect allele. 'Gene' includes the symbol of transcript-coding gene nearest to the most strongly associated variant in the region. The columns 'Replication HR' and 'Replication *P*-value' display hazard ratios and *P*-values for the genetic associations in Kiefer et al. survival analyses. The associations with replication *P*-value below the threshold of multiple testing correction ($P = 0.01$) are shown in bold font. *The rs1194277 SNP, the second-best associated SNP in the AFSW analysis, was used as a replacement for rs1204700722, which was not available in the *23andMe* dataset

of GTPase-activating proteins (21), and the *SATB1* gene, which participates in chromatin remodelling (22). Finally, we found an association for polymorphisms located within the *ADAM11* gene (rs55882072, HR = 1.015, $P = 3.06 \times 10^{-09}$), a metalloprotease that regulates cell and matrix communications (23) and markers within *BRWD1* gene (rs8131965, HR = 0.98, $P = 5.41 \times 10^{-09}$).

Four out of six novel regions replicated in a slightly smaller but independent cohort (13) (Table 2) at a Bonferroni multiple testing correction level (*P*-value $<0.05/6 = 0.01$, Table 2). Specifically, *NEGR1*, *TRIB2*, *TBC1D5*, *LOC100287944*, *ADAM11* risk alleles were associated with earlier age myopia, while *BRWD1* showed significant association with later-age refractive error correction (Table 2). Most SNPs were associated, at various levels of statistical significance, with spherical equivalent in the refracted subgroup of European UK Biobank participants (Supplementary Material, Table S1).

The Cox proportional hazards model assumes that the effects of the tested SNPs have a constant, linear relationship with age. Proportionality of the hazards analyses showed that this assumption held true for many loci, for example *BMP4, TMEM161B, XPO6* (Supplementary Material, Table S2). By contrast, many loci exhibited non-linear effects with age, including *TSPAN10*, *OCA2* loci and interestingly *PAX6*, a gene known to harbour variants that cause microphthalmia and severe eye malformation (24) (Supplementary Material, Table S2), with effects peaking around adolescence. For example, the *LAMA2* variant had a stronger effect over AFSW hazard at an early age, peaking around 16 and a more subdued effect after the age of 40, similar to the effects of other well-known refractive error genes such as *GJD2*, *ZMAT4*, *RDH5* and interestingly *PRSS56*, a gene also known to be associated with eye structural malformations (25) (Supplementary Material, Fig. S4). Among novel loci, *TBC1D5* exerted its influence at an early age, whereas *NEGR1*, *TRIB2* and *ADAM11* were exerted their effect throughout the entire lifespan with the strongest effects over AFSW observed in adolescence (Supplementary Material, Fig. S5).

Our associations with AFSW showed significant enrichment in different body tissues, particularly in the nervous system and retina (Supplementary Material, Tables S3–S5), particularly the brain prefrontal cortex, especially in late infancy. Consistent with a higher than expected expression in cerebral tissues, AFSW genes showed strong genetic correlations with neurological traits such as cognitive ability ($r_g = -0.43$, $P = 2.69 \times 10^{-05}$), neuroticism ($r_g = -0.49$, $P = 0.0039$), insomnia ($r_g = -0.29$, $P = 0.01$) and measures of educational attainment (years of schooling, $r_g = -0.39$, $P = 1.95 \times 10^{-08}$, Supplementary Material, Table S6) and several socio-economically influenced traits.

Gene set enrichment analyses showed that, similar to the findings of other published refractive error GWAS (12), genes associated with AFSW were involved in nervous system development (Supplementary Material, Table S7) and other processes, such as cell signalling and intracellular communications that were other biological processes highlighted in our analyses (Supplementary Material, Table S7). Gene Ontology enrichment analysis results also supported previous conclusions that genes involved in refractive error influence RNA polymerase transcription and gene expression (12).

Discussion

AFSW is a heterogeneous phenotype that is influenced by several different forms of refractive error. Observationally and genetically, this phenotype is strongly correlated with presence and age of developing myopia, the most common form of refractive error in the general population, although other forms of refractive error are also correlated with it. Our study demonstrated that AFSW survival analysis is a powerful statistical method that could be used to augment the existing information available from directly measured refractive error. We found evidence that refractive error and AFSW were strongly correlated and shared most of their heritability and genetic risk loci. Additionally, we have identified

six novel regions associated with the age of refractive correction and replicated four of them. One of the new genes, *TRIB2*, was previously reported for several different ocular traits and disorders. Similar to previous observed genetic associations, polymorphisms within and around the *TRIB2* gene are associated with, among others, optic cup disc area (26) and primary open-angle glaucoma (27), which are consistent with previous observations of genetic pleiotropy between refractive error and optic nerve changes described previously (12). In addition, three other AFSW-associated genes were linked to neurological and neurodevelopmental traits, for which genetic correlation with the refractive error was previously reported: polymorphisms within or near the *TBC1D5/SATB1* genes are associated, among others, with cortical thickness (28), Parkinson's disease (29), schizophrenia (30), general cognitive function (30) and educational attainment (31). Interestingly, the *ADAM11* gene is implicated in familial epilepsy (32,33), while the *BRWD1* gene polymorphisms are associated with general cognitive function (34). Both cognitive ability and educational attainment correlated with the genetic risk of refractive disorders (12). Similarly, another newly associated gene, *NEGR1,* influences neurite outgrowth (35), a process where extracellular cues attach to transmembrane receptors, initiating signalling cascade and reorganizing neuronal structure (36). Neurite outgrowth was essential for functional wiring and building connectivity in the developing brain. *NEGR1* was linked to several neurodevelopmental disorders—intellectual disability, dyslexia (37) and autism (38) due to its function in brain connectivity.

We independently replicated four out of six novel loci associated with AFSW at robust multiple testing correction levels. Alleles of the *TRIB2*, *TBC1D5* and *ADAM11* genes that were associated with myopia were significantly associated with correction at an earlier age, while those at the *BRWD1* locus showed association with myopia correction at older ages. Although *NEGR1* and *LOC100287944* were not significantly associated with the AFSW to correct for myopia in the replication dataset, the estimated effects had the same direction of the effects as in the discovery GWAS, and it is possible that a lack of statistical significance in replication analyses could be due to sample size and power limitations.

The strongest genetic association in our study was identified with a variant located within *TSPAN10*. This gene showed a moderate association with refractive error (12) but was strongly associated with corneal astigmatism (17) as well as with strabismus and amblyopia (39), which manifest early in childhood. Notably, the association between *TSPAN10* strabismus was independent of refractive error (39). Because our study sample was not limited to individuals with myopia and hyperopia, the observed association between early AFSW and *TSPAN10* may have reflected contributions from other ocular disorders such as astigmatism, strabismus or amblyopia.

Our study also found strong associations with markers located near or within *ZMAT4*, *LAMA2* and *GJD2* genes. Similar to previously published results, we found that *LAMA2* and *GJD2* had an early effect that increased with age (16). In particular, these genes were observed to have the strongest effect on myopia in 10- to 25-year-olds but were also expressed during the entire age span of myopia development (16).

The results of this study confirm the strong correlation between AFSW and myopia. These results also demonstrate that AFSW is a complex phenotype that is likely to capture pleiotropic genetic effects that influence phenotypic traits other than myopia. SNP loci associated with AFSW appear to exert their effects at different time, although it is not clear whether the effect size changes over time of these loci are due in part to that pleiotropy or can simply be explained by their effects over myopia.

A potential limitation of our study is that the phenotype used in our study was based on self-reported data and not on clinical evaluations. Although self-reported data are occasionally prone to recall bias that could affect the results, its wider availability compared to directly measured refractive error may lead to statistical power gains. Other potential limitation includes the generalizability of these results. The effect sizes we report were largely consistent in the two large European population cohorts in which they were initially estimated and replicated. However, both cohorts are likely to be enriched for myopic participants. Findings in these cohorts may not be generalizable to other general population cohorts, and particularly, they may not apply to more diverse populations.

Our study identified genome-significant associations with 44 independent loci, most of which were documented in refractive error and myopia GWAS. We demonstrate that the effects of many of these regions strongly correlate with myopic refraction but vary with age, which to date was reported for a handful of spherical equivalent genes. Additionally, we find associations with six novel regions and successfully replicate four of them in an independent cohort. Our results support the role of neural development and signalling in the pathogenesis of myopia. The findings of our study further our knowledge on the genetic basis of refractive disorders and demonstrate the value of large-scale population-based genetic studies.

Materials and Methods

Study population and phenotyping

The UK Biobank cohort is a large population-based longitudinal study that includes 502 682 volunteers from across the United Kingdom, aged between 40 and 69 years at the time of recruitment (40). The study participants were recruited via the UK National Health Service register based on their living proximity to the 22 assessment centres (40). At the baseline assessment, the data on socio-economic, lifestyle and health-related factors were collected via touch-screen questionnaires and face-to-face interviews (40). The

electronic questionnaire contained several eyesight-related inquires, including the questions about the AFSW (The UK Biobank field number: 2217) and reasons for refractive correction (Field number: 6147) (40). About 23% of all UK Biobank participants ($N = 117\,279$) undertook ophthalmic examination (41), including non-cycloplegic autorefraction carried out using Tomey RC 5000 device (Tomey Corp., Nagoya, Japan). For each participant, the spherical equivalent was calculated (SPHE = sphere + ½ cylinder power) (UK Biobank field numbers: 5084–5085; 5086–5087), and subsequently, the average measurement of the two eyes was estimated. The UK Biobank enrollees who had ocular surgery or eye infection 4 weeks before the assessment did not participate in the ophthalmic examination. The spherical equivalent readings of participants who had eye surgery, infection, bilateral eye injury before the assessment or self-reported cataract with mild myopia, as described before (41), were excluded from the analyses. To minimize confounding arising from population genetic structure, we limited the study sample to individuals of European ancestry, as ascertained by using genetic information. Ancestry was defined based on principal component analyses of the participants' genotypes, pre-computed and calculated by the UK Biobank working group.

Genetic data

Genotyping was performed on 488 377 subjects from the UK Biobank cohort as described before (40) using two similar and mutually compatible genotyping platforms (Applied Biosystems UK BiLEVE Axiom Array and the UK Biobank Axiom Array), which although not fully identical, shared approximately 95% of genetic markers. However, our analyses used a subset of Biobank participants, for whom information about the refractive error was available. Specifically, our spherical equivalent analyses were conducted in $N = 102\,117$ subjects, the all-cause age of spectacle wear in $N = 340\,318$ subjects, age of spectacle wear in individuals with myopia in $N = 24\,363$ and in individuals with hypermetropia in $N = 24\,711$ subjects. To avoid bias arising from genetic stratification and admixture, all subjects were of European ancestry.

Phasing and further genomic imputation were conducted as described before (40). Briefly, imputation was carried out using Haplotype Reference Consortium (HRC) data as a primary reference panel, but also merged 1000 Genomes phase 3 and UK10K reference panels. Only markers shared between HRC and 1000 genomes/UK10K datasets were selected for imputation; therefore, a final dataset covered 93 095 623 autosomal SNPs in conjunction with large structural variants indels (40).

Statistical analyses

Descriptive analyses were carried out using *epiDisplay* package in R. We calculated frequencies and percentages and means and standard errors for categorical and continuous variables. For our time-to-event genetic association analyses, we build Cox proportional hazards regression model adjusted for age and sex. Likelihood ratio test was used to compute *P*-values for each SNP in the model. We used two R packages, *gwasurvivr* (42) and *SPACox* (43), to calculate hazard ratios (HR) and their corresponding *P*-values. The genetic variants with *P*-values below the customary genome-wide significance level of 5×10^{-08} were considered statistically significant. The proportionality of the hazards for significant associations was assessed using the *survival* package in R (https://cran.r-project.org/web/packages/survival/). Subgroup sensitivity analyses were conducted in samples that only included participants with available spherical equivalent measurements that were consistent with myopia ($N = 24\,363$).

We sought replication of the novel genetic associations using time-to-event results previously published by Kiefer et al. (13). Replication was considered significant if the association probabilities were below the Bonferroni multiple testing correction level (observed *P*-value multiplied by the number of tests no higher than 0.05). The genomic inflation arising from sample stratification and uncontrolled admixture was tested ld score regression (44).

Data from 45 771 research volunteers recruited among the customer base of the *23andMe* genomics company (Sunnyvale, CA, USA) were used for replication. More detailed information can be found in the original publication (13), but briefly, the phenotypic status was ascertained through an online medical history questionnaire or an eyesight questionnaire. Participants were genotyped and additional SNP genotypes were imputed against the 1000 genomes data and the imputed genotypes from individuals of European ancestry were used for Cox proportional hazards models. Although the analyses conducted in this replication set are in many ways comparable to those in the discovery UK Biobank cohort, there is one difference in the study designs. The 23andMe cohort analysed exclusively individuals who self-reported correction for myopia and not other forms of refractive error.

Genetic correlations between identified loci and other phenotypic traits were assessed using ld-score regression (45) and the summary statistics from GWAS Catalog (46).

The shared functionality of associated genes was further explored through gene set enrichment analyses, as implemented in MAGENTA software (47). The relationship between genotypes and gene expression was modelled using Mendelian Randomization tests implemented in the SMR program (48), using expression data from GTEx release v8 (https://gtexportal.org/home/datasets), the Atlas of the Developing Human Brain (49) (BrainSpan 11) and retinal cis-eQTL data from healthy donors (50).

Data availability

The UK Biobank data are available to all bona fide researchers through a dedicated electronic Access

Management System (https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access). Full GWAS summary statistics for the *23andMe* discovery data set will be made available through *23andMe* to qualified researchers under an agreement with *23andMe* that protects the privacy of the *23andMe* participants. Please visit https://research.23andme.com/collaborate/#dataset-access/ for more information and to apply to access the data.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

The UK Biobank data were accessed as part of the UK Biobank project 17615 and we are grateful to all the participants for their willingness to support research. We would also like to thank the research participants and employees of *23andMe*, Inc. for making this work possible. R.W. and P.G.H. acknowledge funding from the NUS NEI (R21EY027880). K.P. is a grateful recipient of a Fight for Sight PhD studentship (grant number 5037-5038). APK is funded by a UK Research and Innovation Future Leaders Fellowship (MR/T040912/1) and an Alcon Young Investigator Award.

Conflict of Interest statement. APK has consulted for Abbvie, Aerie, Google Health, Novartis, Reichert, Santen and Thea.

References

1. Steinmetz, J.D., Bourne, R.R.A., Briant, P.S., Flaxman, S.R., Taylor, H.R.B., Jonas, J.B., Abdoli, A.A., Abrha, W.A., Abualhasan, A., Abu-Gharbieh, E.G. *et al.* (2021) Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the right to sight: an analysis for the Global Burden of Disease Study. *Lancet Glob. Health*, **9**, e144–e160.
2. Morgan, I.G., French, A.N., Ashby, R.S., Guo, X., Ding, X., He, M. and Rose, K.A. (2018) The epidemics of myopia: aetiology and prevention. *Prog. Retin. Eye Res.*, **62**, 134–149.
3. Williams, K.M., Bertelsen, G., Cumberland, P., Wolfram, C., Verhoeven, V.J.M., Anastasopoulos, E., Buitendijk, G.H.S., Cougnard-Grégoire, A., Creuzot-Garcher, C., Erke, M.G. *et al.* (2015) Increasing prevalence of myopia in Europe and the impact of education. *Ophthalmology*, **122**, 1489–1497.
4. Vitale, S., Sperduto, R.D. and Ferris, F.L., III (2009) Increased prevalence of myopia in the United States between 1971-1972 and 1999-2004. *Arch. Ophthalmol.*, **127**, 1632–1639.
5. Mutti, D.O., Zadnik, K. and Adams, A.J. (1996) Myopia. The nature versus nurture debate goes on. *Invest. Ophthalmol. Vis. Sci.*, **37**, 952–957.
6. Saw, S.-M., Gazzard, G., Shih-Yen, E.C. and Chua, W.-H. (2005) Myopia and associated pathological complications. *Ophthalmic Physiol. Opt.*, **25**, 381–391.
7. Holden, B.A., Wilson, D.A., Jong, M., Sankaridurg, P., Fricke, T.R., Smith, E.L. and Resnikoff, S. (2015) Myopia: a growing global problem with sight-threatening complications. *Community Eye Health*, **28**, 35.
8. Haarman, A.E.G., Enthoven, C.A., Tideman, J.W.L., Tedja, M.S., Verhoeven, V.J.M. and Klaver, C.C.W. (2020) The complications of myopia: a review and meta-analysis. *Invest. Ophthalmol. Vis. Sci.*, **61**, 49.
9. Mountjoy, E., Davies, N.M., Plotnikov, D., Smith, G.D., Rodriguez, S., Williams, C.E., Guggenheim, J.A. and Atan, D. (2018) Education and myopia: assessing the direction of causality by Mendelian randomisation. *BMJ*, **361**, k2022.
10. Sanfilippo, P.G., Hewitt, A.W., Hammond, C.J. and Mackey, D.A. (2010) The heritability of ocular traits. *Surv. Ophthalmol.*, **55**, 561–583.
11. 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. *et al.* (2018) Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nat. Genet.*, **50**, 834–848.
12. Hysi, P.G., Choquet, H., Khawaja, A.P., Wojciechowski, R., Tedja, M.S., Yin, J., Simcoe, M.J., Patasova, K., Mahroo, O.A., Thai, K.K. *et al.* (2020) Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. *Nat. Genet.*, **52**, 401–407.
13. Kiefer, A.K., Tung, J.Y., Do, C.B., Hinds, D.A., Mountain, J.L., Francke, U. and Eriksson, N. (2013) Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet.*, **9**, e1003299.
14. Wojciechowski, R. and Hysi, P.G. (2013) Focusing in on the complex genetics of myopia. *PLoS Genet.*, **9**, e1003442.
15. Wong, Y.-L., Hysi, P., Cheung, G., Tedja, M., Hoang, Q.V., Tompson, S.W.J., Whisenhunt, K.N., Verhoeven, V., Zhao, W., Hess, M. *et al.* (2019) Genetic variants linked to myopic macular degeneration in persons with high myopia: CREAM Consortium. *PLoS One*, **14**, e0220143.
16. Tideman, J.W.L., Fan, Q., Polling, J.R., Guo, X., Yazar, S., Khawaja, A., Höhn, R., Lu, Y., Jaddoe, V.W.V., Yamashiro, K. *et al.* (2016) When do myopia genes have their effect? Comparison of genetic risks between children and adults. *Genet. Epidemiol.*, **40**, 756–766.
17. Shah, R.L., Guggenheim, J.A. and UK Biobank Eye and Vision Consortium (2018) Genome-wide association studies for corneal and refractive astigmatism in UK Biobank demonstrate a shared role for myopia susceptibility loci. *Hum. Genet.*, **137**, 881–896.
18. Cumberland, P.M., Chianca, A., Rahi, J.S. and for the UK Biobank Eye and Vision Consortium (2016) Accuracy and utility of self-report of refractive error. *JAMA Ophthalmol.*, **134**, 794–801.
19. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. and van der Sluis, S. (2018) Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat. Commun.*, **9**, 905.
20. Eyers, P.A., Keeshan, K. and Kannan, N. (2017) Tribbles in the 21st century: the evolving roles of tribbles pseudokinases in biology and disease. *Trends Cell Biol.*, **27**, 284–298.
21. Xie, J., Heim, E.N., Crite, M. and DiMaio, D. (2020) TBC1D5-catalyzed cycling of Rab7 is required for Retromer-mediated human papillomavirus trafficking during virus entry. *Cell Rep.*, **31**, 107750.
22. Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P. and Kohwi-Shigematsu, T. (2002) SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*, **419**, 641–645.
23. Hsia, H.-E., Tüshaus, J., Brummer, T., Zheng, Y., Scilabra, S.D. and Lichtenthaler, S.F. (2019) Functions of 'A disintegrin and metalloproteases (ADAMs)' in the mammalian nervous system. *Cell. Mol. Life Sci.*, **76**, 3055–3081.

24. Hever, A.M., Williamson, K.A. and van Heyningen, V. (2006) Developmental malformations of the eye: the role of PAX6, SOX2 and OTX2. *Clin. Genet.*, **69**, 459–470.
25. Gal, A., Rau, I., El Matri, L., Kreienkamp, H.-J., Fehr, S., Baklouti, K., Chouchane, I., Li, Y., Rehbein, M., Fuchs, J. *et al.* (2011) Autosomal-recessive posterior microphthalmos is caused by mutations in PRSS56, a gene encoding a trypsin-like serine protease. *Am. J. Hum. Genet.*, **88**, 382–390.
26. Springelkamp, H., Mishra, A., Hysi, P.G., Gharahkhani, P., Höhn, R., Khor, C.-C., Cooke Bailey, J.N., Luo, X., Ramdas, W.D., Vithana, E. *et al.* (2015) Meta-analysis of genome-wide association studies identifies novel loci associated with optic disc morphology. *Genet. Epidemiol.*, **39**, 207–216.
27. Springelkamp, H., Iglesias, A.I., Mishra, A., Höhn, R., Wojciechowski, R., Khawaja, A.P., Nag, A., Wang, Y.X., Wang, J.J., Cuellar-Partida, G. *et al.* (2017) New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. *Hum. Mol. Genet.*, **26**, 438–453.
28. van der Meer, D., Frei, O., Kaufmann, T., Shadrin, A.A., Devor, A., Smeland, O.B., Thompson, W.K., Fan, C.C., Holland, D., Westlye, L.T. *et al.* (2020) Understanding the genetic determinants of the brain with MOSTest. *Nat. Commun.*, **11**, 3512.
29. Riessland, M., Kolisnyk, B., Kim, T.W., Cheng, J., Ni, J., Pearson, J.A., Park, E.J., Dam, K., Acehan, D., Ramos-Espiritu, L.S. *et al.* (2019) Loss of SATB1 induces p21-dependent cellular senescence in post-mitotic dopaminergic neurons. *Cell Stem Cell*, **25**, 514–530.e8.
30. Smeland, O.B., Frei, O., Kauppi, K., Hill, W.D., Li, W., Wang, Y., Krull, F., Bettella, F., Eriksen, J.A., Witoelar, A. *et al.* (2017) Identification of genetic loci jointly influencing schizophrenia risk and the cognitive traits of verbal-numerical reasoning, reaction time, and general cognitive function. *JAMA Psychiatry*, **74**, 1065–1075.
31. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R. *et al.* (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.*, **50**, 1112–1121.
32. Owuor, K., Harel, N.Y., Englot, D.J., Hisama, F., Blumenfeld, H. and Strittmatter, S.M. (2009) LGI1-associated epilepsy through altered ADAM23-dependent neuronal morphology. *Mol. Cell. Neurosci.*, **42**, 448–457.
33. Yamagata, A. and Fukai, S. (2020) Insights into the mechanisms of epilepsy from structural biology of LGI1–ADAM22. *Cell. Mol. Life Sci.*, **77**, 267–274.
34. Davies, G., Armstrong, N., Bis, J.C., Bressler, J., Chouraki, V., Giddaluru, S., Hofer, E., Ibrahim-Verbaas, C.A., Kirin, M., Lahti, J. *et al.* (2015) Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N = 53 949). *Mol. Psychiatry*, **20**, 183–192.
35. Pischedda, F. and Piccoli, G. (2015) The IgLON family member Negr1 promotes neuronal arborization acting as soluble factor via FGFR2. *Front. Mol. Neurosci.*, **8**, 89.
36. Raper, J. and Mason, C. (2010) Cellular strategies of axonal pathfinding. *Cold Spring Harb. Perspect. Biol.*, **2**, a001933.
37. Veerappa, A.M., Saldanha, M., Padakannaya, P. and Ramachandra, N.B. (2013) Family-based genome-wide copy number scan identifies five new genes of dyslexia involved in dendritic spinal plasticity. *J. Hum. Genet.*, **58**, 539–547.
38. Neale, B.M., Kou, Y., Liu, L., Maayan, A., Samocha, K.E., Sabo, A., Lin, C.-F., Stevens, C., Wang, L.-S., Makarov, V. *et al.* (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*, **485**, 242–245.
39. Plotnikov, D., Shah, R.L., Rodrigues, J.N., Cumberland, P.M., Rahi, J.S., Hysi, P.G., Atan, D., Williams, C., Guggenheim, J.A. and UK Biobank Eye and Vision Consortium (2019) A commonly occurring genetic variant within the NPLOC4–TSPAN10–PDE6G gene cluster is associated with the risk of strabismus. *Hum. Genet.*, **138**, 723–737.
40. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J. *et al.* (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature*, **562**, 203–209.
41. Cumberland, P.M., Bao, Y., Hysi, P.G., Foster, P.J., Hammond, C.J., Rahi, J.S. and Consortium, U.B.E.& V (2015) Frequency and distribution of refractive error in adult life: methodology and findings of the UK Biobank Study. *PLoS One*, **10**, e0139780.
42. Rizvi, A.A., Karaesmen, E., Morgan, M., Preus, L., Wang, J., Sovic, M., Hahn, T. and Sucheston-Campbell, L.E. (2019) gwasurvivr: an R package for genome-wide survival analysis. *Bioinforma. Oxf. Engl.*, **35**, 1968–1970.
43. Bi, W., Fritsche, L.G., Mukherjee, B., Kim, S. and Lee, S. (2020) A fast and accurate method for genome-wide time-to-event data analysis and its application to UK Biobank. *Am. J. Hum. Genet.*, **107**, 222–233.
44. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L. and Neale, B.M. (2015) LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, **47**, 291–295.
45. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., Duncan, L., Perry, J.R.B., Patterson, N., Robinson, E.B. *et al.* (2015) An atlas of genetic correlations across human diseases and traits. *Nat. Genet.*, **47**, 1236–1241.
46. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E. *et al.* (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.*, **47**, D1005–D1012.
47. Segrè, A.V., Consortium, D., Investigators, M., Groop, L., Mootha, V.K., Daly, M.J. and Altshuler, D. (2010) Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.*, **6**, e1001058.
48. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M. *et al.* (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.*, **48**, 481–487.
49. Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K. *et al.* (2014) Transcriptional landscape of the prenatal human brain. *Nature*, **508**, 199–206.
50. Strunz, T., Kiel, C., Grassmann, F., Ratnapriya, R., Kwicklis, M., Karlstetter, M., Fauser, S., Arend, N., Swaroop, A., Langmann, T. *et al.* (2020) A mega-analysis of expression quantitative trait loci in retinal tissue. *PLoS Genet.*, **16**, e1008934.