A genome-wide association study identifies a susceptibility locus for refractive errors and myopia at 15q14

Abbas M. Solouki[1,2#], Virginie J.M. Verhoeven[1,2#], Cornelia M. van Duijn[2], Annemieke J.M.H. Verkerk[3], M. Kamran Ikram[1,2,4], Pirro G. Hysi[5], Dominiek D.G. Despriet[1,2], Leonieke M. van Koolwijk[1,6], Lintje Ho[1,2], Wishal D. Ramdas[1,2], Monika Czudowska[1,2], Robert W.A.M. Kuijpers[1], Najaf Amin[2], Maksim Struchalin[2], Yurii S. Aulchenko[2], Gabriel van Rij[1], Frans C.C. Riemslag[7], Terri L. Young[8], David A. Mackey[9], Timothy D. Spector[5], Theo G.M.F. Gorgels[10], Jacqueline J.M. Willemse-Assink[11], Aaron Isaacs[2], Rogier Kramer[12], Sigrid M.A. Swagemakers[3,13,14], Arthur A.B. Bergen[10,15,16], Andy A.L.J. van Oosterhout[2,17], Ben A. Oostra[17], Fernando Rivadeneira[2,18], André G. Uitterlinden[2,18], Albert Hofman[2], Paulus T.V.M. de Jong[2,10], Christopher J. Hammond[5], Johannes R.Vingerling[1,2], Caroline C.W. Klaver[1,2]

[1]Department of Ophthalmology, Erasmus Medical Center, Rotterdam, the Netherlands
[2]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands
[3]Department of Bioinformatics, Erasmus Medical Center, Rotterdam, the Netherlands
[4]Department of Neurology, Erasmus Medical Center, Rotterdam, the Netherlands
[5]Department of Twin Research and Genetic Epidemiology, King's College London School of Medicine, London, United Kingdom
[6]The Rotterdam Eye Hospital, Rotterdam, the Netherlands
[7]Bartiméus, Institute for the Visually Impaired, Zeist, the Netherlands
[8]Center for Human Genetics, Duke University, Durham, North Carolina, United States of America
[9]Lions Eye Institute, University of Western Australia, Centre for Ophthalmology and Visual Science, Perth, Australia
[10]Department of Clinical and Molecular Ophthalmogenetics, Netherlands Institute of Neurosciences (NIN), An Institute of the Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, the Netherlands
[11]Department of Ophthalmology, Amphia Hospital, Breda, the Netherlands
[12]Department of Ophthalmology, Fransiscus Hospital, Roosendaal, the Netherlands
[13]Department of Genetics, Erasmus Medical Center, Rotterdam, the Netherlands
[14]The Cancer Genomics Center, the Netherlands
[15]Department of Clinical Genetics, Academic Medical Center, Amsterdam, the Netherlands
[16]Department of Ophthalmology, Academic Medical Center, Amsterdam, the Netherlands
[17]Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, the Netherlands
[18]Department of Internal Medicine and Clinical Chemistry, Erasmus Medical Center, Rotterdam, the Netherlands

[#]These authors contributed equally to this work.

Correspondence should be addressed to: Dr. C.C.W. Klaver (C.C.W.K.), Dept. of Ophthalmology, Erasmus Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, Telephone: (+31) 6-51934491; Fax: (+31) 10-7033692; E-mail: c.c.w.klaver@erasmusmc.nl

Refractive errors are the most common ocular disorders worldwide, and may lead to blindness. Although this trait is highly heritable, identification of susceptibility genes has been challenging. We conducted a genome-wide association study testing single nucleotide polymorphisms for association with refractive error in 5,328 unrelated individuals of a Dutch population-based study, and replicated findings in four independent cohorts (10,280 persons). We identified a significant association at chromosome 15q14 with $P=2.21x10^{-14}$ for rs634990. The odds ratio of myopia versus hyperopia for the minor allele (MAF 0.47) was 1.41 (95% CI 1.16-1.70) for heterozygous, and 1.83 (95% CI 1.42-2.36) for homozygous subjects. The associated region lies in the vicinity of genes which are expressed in the retina, *GJD2* and *ACTC1*, and appears to harbor regulatory elements which may influence transcription of these genes. Our data suggest that common variants at 15q14 influence susceptibility for refractive errors in the general population.

Refractive errors are by far the most common cause of visual impairment in humans[1-5]. They result from aberrant coordinated effects of the ocular biometric components, most notably of axial length. Elongation of the eye axis leads to myopia (nearsightedness), while a shortened axis causes hyperopia (farsightedness). Refractive errors often cause alterations in the anatomical structure of the eye, increasing the risk of complications[6]. Myopia may lead to ocular morbidity such as glaucoma and retinal detachment, and high myopia in particular can cause posterior staphyloma and macular degeneration[7-11]. Treatment options for myopia are limited; it is the fifth most common cause of impaired vision, and the seventh most common cause of legal blindness worldwide[10,11].

The etiology of refractive errors and myopia is complex and largely unknown. The current notion is that eye growth is triggered by a visually evoked signaling cascade, which

begins in the retina, traverses the choroid, and subsequently mediates scleral remodeling. Established risk factors are education, reading, outdoor exposure, and familial predisposition[11-14]. Familial aggregation studies quantified a strong genetic basis; the estimated sibling recurrence risk (λ_s) varied between 1.5-3.0 for low myopia- and between 4.9-19.8 for high myopia, and heritability estimates (h^2) ranged from 0.60-0.90[15]. Segregation analyses suggested the involvement of multiple genes rather than a single major gene effect[11,13,15]. In an attempt to identify causal genes, previous mapping studies mainly focussed on highly myopic probands with multiple affected relatives, and thereby identified at least 20 putative genetic loci[11]. Replication of these results has been limited, and proposed loci were shown to have little to no effect in unselected populations. Genome-wide mapping has not been conducted in refractive error studies of the general population. Hence, the genetic basis of common refractive errors and myopia remains to be elucidated.

We performed a genome-wide association study (GWAS) in the population-based Rotterdam Study (RS-I, *n*=5,328), and investigated refractive error as a quantitative trait. Study design and baseline characteristics are provided in the **Online Methods** and **Supplementary Table 1**. The mean spherical equivalent in this older population of European descent was +0.86 (standard deviation (SD) 2.45) dioptres. Refractive errors occurred in 52% (*n*=2790) of the participants, ranging from -19 to +10 diopters (D).

We genotyped the entire sample using the Illumina HumanHap 550k and 610Q arrays (**Online Methods)**. Genotypes for more than 2.5 million autosomal single nucleotide polymorphisms (SNPs) were imputed with reference to the HapMap Phase II CEU build 36. Comparison of the observed and expected distributions (Q-Q plot, **Supplementary Figure 1**) showed modest inflation of the test statistics (λ_{GC}=1.054 for RS-1). Using an additive model, we identified a novel genome-wide significant (P=1.76x10^{-8}) locus on chromosome 15q14 (**Table 1, Figure 1**). Subsequently, we investigated 31 SNPs spread across four loci on

chromosome 15q14, 14q24, 1q41, and 10p12.3 reaching $P<10^{-6}$ (**Supplementary Table 2**) for further investigation in four independent replication cohorts, i.e., RS-II (*n*=2008; λ_{GC}=1.012), RS-III (*n*=1970; λ_{GC}=1.012), Erasmus Rucphen Family Study (ERF, *n*=2032; λ_{GC}=1.037) from the Netherlands; and a twin study from the United Kingdom (TwinsUK; *n*=4270; λ_{GC}=1.04. The designs of RS-II and RS-III were population-based; those of ERF and TwinsUK family-based. Cohorts were not selected on a disease phenotype. All studies consisted predominantly of individuals of European ancestry, and all used similar protocols to evaluate refractive error (**Online Methods**, **Supplementary Table 2**).

At validation, meta-analysis confirmed a significant association between refractive errors and locus 15q14 **(Table 1).** Frequencies of the risk alleles at this region were similar across the studies. The *P*-values were nominally significant for the 14 top SNPs in RS-II, RS-III, and TwinsUK, and the direction of the effect (regression coefficient beta) of the minor alleles was consistent. The strongest signal in the meta-analysis was observed for rs634990 ($P=2.21\text{x}10^{-14}$; **Table 1**), and this SNP accounted for 0.5% of the variance in spherical equivalent.

To determine the impact of this locus on the risk of clinically relevant outcomes, we compared subjects with myopia to those with hyperopia in a logistic regression analysis. We found strong evidence that the C allele of rs634990 carried a higher risk of myopia (**Figure 2**). The odds ratio (OR) of mild or severe myopia versus mild or severe hyperopia was 1.41 (95% Confidence Interval (CI) 1.16-1.70) for heterozygous individuals, and 1.83 (95% CI 1.42-2.36) for homozygous persons.

The 15q14 region of highly significant SNPs (**Figure 3**) lies in an intergenic region in the vicinity of the genes *GJD2* (39 kb from rs634990 at 3' end), *ACTC1* (74 kb at 3'end), and *GOLGA8B* (180 kb at 5'end). We investigated a potential function for these genes in eye growth development by examining gene expression levels in the retina of postmortem human

eyes (**Supplementary Table 3**), and observed a moderate to high expression for *GJD2* and *ACTC1*, and a much lower expression for *GOLGA8B*. *GOLGA8B* (Golgi autoantigen golgin-67) encodes a 67 kDa protein, belongs to a family of Golgi auto-antigens, and is localized at the cytoplasmic surface of the Golgi complex[16]. A specific function of this gene in the retina has not been reported. *ACTC1* (cardiac muscle alpha actin 1) encodes a 42 kDa smooth muscle actin. The functional role of *ACTC1* in the eye is currently unclear, but actins which are similar, such as α-SMA, have been shown to be increased in developing myopic eyes[17]. α-SMA influences the number of contractile myofibroblasts in the sclera, and contributes to extracellular matrix remodelling. As these are key factors occurring in eye enlargement, it is intriguing to know whether *ACTC1* has these characteristics as well.

The functional properties of *GJD2* make this gene an interesting candidate to explain our findings. *GJD2* (gap junction protein delta 2) encodes the 36 kDa connexin36 (CX36), which is a neuron-specific protein belonging to a multi-gene family of integral membrane proteins[18]. CX36 forms gap junction channels between adjacent membranes of neuronal cells, is present in photoreceptors, amacrine, and bipolar cells, and plays a critical role in the transmission process of the retinal electric circuitry by enabling intercellular transport of small molecules and ions[18-21]. Further exploration of *GJD2* using Ingenuity analysis (**Online Methods**, **Supplementary Figure 2)** alluded to a role in eye growth regulation as well as lens fiber maturation in knock-out animals[22,23]. To identify possible causal variants in this gene, we performed direct sequencing of all exons and intron-exon boundaries of *GJD2* in 47 subjects with either high myopia, high hypermetropia, or emmetropia. We found neither new mutations nor frequency differences of variants between groups (**Supplementary Table 4),** and conclude that linkage disequilibrium with common functional variants in *GJD2* is unlikely to explain the observed association.

The next step was to assess whether the intergenic region itself can have functional consequences. We evaluated the expression of SNPs of our associated region in lymphoblastoid cell lines. At least two of our most associated SNPs significantly altered expression, providing evidence that elements of our locus are transcribed and may alter cell function (**Supplementary Table 5**). Subsequently, we searched for regulatory elements[24,25] in the entire 53 kb locus of highly significantly correlated SNPs using UCSC Genome Browser, and found the predicted presence of seven DNase I hypersensitive sites, six enhancers based on experimentally validated H3 chromatin signatures in Hela and K562 cells[24,25], 20 peaks of sequence conservation in alignments of multiple species of placental mammals, and one insulator site (**Supplementary Figure 3)**[25]**.** Enhancers are known to facilitate transcription of distal genes, and its range of activity is confined by insulators[25]. Remarkably, the greatest peak of our association coincided with an insulator site. Precedents of genomic alterations of insulators causing hereditary disease have been reported[26,27]. We speculate that variants or mutations in regulatory elements at 15q14 may lead to illegitimate transcription of genes in the area, e.g., of *ACTC1* and *GJD2.*

In GWA studies, sources of heterogeneity may cause spurious findings. To address this issue and minimize potential biases, we applied genomic control to the cohort-level test statistics in the population cohorts, and correction using the identity by descent structure for the family-based cohorts. Three studies significantly replicated our initial findings. The fourth study, ERF, showed the same direction of association, albeit non-significant, and revealed similar risks of myopia for carriers of the risk allele (**Figure 2).** Thus, the observed effects of the genetic variants at 15q14 are relatively homogeneous among the 5 studies, enhancing credibility of the findings.

In the same issue of this journal, Hysi et al. report the results of a GWAS for refractive errors in the TwinsUK study[28]. The authors find genome-wide significance (best combined

P=1.85x10^{-9} for rs939658 and P=2.07x10^{-9} for rs8027411) for a locus on chromosome 15q25, explaining 0.81% of the variance in spherical equivalent. The locus includes the promoter of the *RASGRF1* gene. This gene is known to be functionally involved in eye development[29], and, similar to *GJD2*, is involved in synaptic transmission of photoreceptor responses[30]. TwinsUK and RS-I are two of the largest existing refractive error cohorts with GWAS data. Our studies identified different genome-wide significant tophits in terms of *P*-values, and we both estimated the variation in refractive error explained by these SNPs to be small. Therefore, it is likely that common variants with a substantial disease risk do not play a role in the pathogenesis of this trait. The findings of our studies suggest that the genetic variance of refractive error is mostly determined by multiple variants with a low to moderate penetrance, resembling traits such as height[31].

Nevertheless, the mutual validation of the direction and beta of the effect of variants at 15q14 and 15q25 suggests that alterations at these genomic loci lead to refractive error and myopia. To unravel the mechanism, next steps should include comprehensive resequencing of the entire associated regions and flanking genes, validation in cohorts of other ethnicities, functional assays, and study of risk modulation by environmental factors. This may help to launch new pathogenic pathways for refractive errors, and may eventually lead to novel strategies to reduce the sight-threatening consequences of myopia.

Acknowledgments: Lameris Ootech Major funding of the work performed in the Netherlands came from the Netherlands Organisation of Scientific Research (NWO); Erasmus Medical Center and Erasmus University, Rotterdam; Netherlands Organization for the Health Research and Development (ZonMw); UitZicht; the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (DG XII); the Municipality of Rotterdam; the Netherlands Genomics Initiative (NGI)/NWO; Center for Medical Systems Biology of NGI; Lijf en Leven; MD Fonds; Prof. Dr. Henkes Stichting; Oogfonds Nederland; Stichting Nederlands Oogheelkundig Onderzoek; Swart van Essen; Bevordering van Volkskracht; Blindenhulp; Landelijke Stichting voor Blinden en Slechtzienden; Rotterdamse Vereniging voor Blindenbelangen; OOG; Algemene Nederlandse Vereniging ter Voorkoming van Blindheid; the Rotterdam Eye Hospital Research Foundation; Laméris Ootech BV; Topcon Europe BV; Heidelberg Engineering. We thank Ada Hooghart, Corina Brussee, Riet Bernaerts-Biskop, and Patricia van Hilten, Pascal Arp, Mila Jhamai, Michael Moorhouse, Jeannette Vergeer, Marijn Verkerk, Sander Bervoets, and Peter van der Spek for help in execution of the study.

TwinsUK acknowledges the Wellcome Trust, the EU MyEuropia Marie Curie Research Training Network, Guide Dogs for the Blind Association, the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F2-2008-201865-GEFOS and (FP7/2007-2013), ENGAGE project grant agreement HEALTH-F4-2007-201413 and the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254); Dept of Health; Biotechnology and Biological Sciences Research Council (G20234); National Institutes of Health National Eye Institute grant RO1EY018246; Center for Inherited Disease Research. TwinsUK thanks Gabriela Surdulescu, Leena Peltonen, Panos Deloukas, Mark Lathrop, David Goldstein, Aarno Palotie, Camilla Day for help in execution of the study and analyses.

Figure 1. Genome-wide signal intensity (Manhattan) plot of discovery cohort Rotterdam Study-I

The statistical significance values across the 22 autosomes of each SNP association with refractive error (measured as spherical equivalent) are plotted as $-\log_{10}$ *P*-values. SNPs with minor allele frequency >=0.01 were included. The blue horizontal line indicates *P*-value of 10^{-5}; the red line *P* value of $5x10^{-8}$.

Figure 2. Forest plot of associations for myopia (SE<= -3D) versus hyperopia (SE >= +3D)

Forest plot of the estimated per-genotype odds ratio for topSNP rs634990 for the 5 studies separately, and for the meta-analysis of all studies. Abbreviations: RS-I, Rotterdam Study I; RS-II, Rotterdam Study II; RS-III, Rotterdam Study III; ERF, Erasmus Rucphen Family Study; TwinsUK, the Twin Cohort recruited in the UK; OR, odds ratio; 95%CI, 95% Confidence Interval.

Figure 3. Regional plot at chromosome 15q14

Log_{10} *P*-values from the discovery cohort Rotterdam Study-I as a function of genomic position (HapMap release 22 build 36). The *P*-value for the top SNP is denoted by the large diamond; *P*-values for other genotyped and imputed SNPs are shown as smaller diamonds. *P*-values for SNPs of unknown type are presented as squares. Superimposed on the plot are gene locations (green) and recombination rates (blue).

Table 1. Genome-wide association and replication for refractive error at locus 15q14

Abbreviations: RS-I, Rotterdam Study I; RS-II, Rotterdam Study II; RS-III, Rotterdam Study III; ERF, Erasmus Rucphen Family Study; TwinsUK, the Twin Cohort recruited in London; SNP, single nucleotide polymorphism; MA, Minor Allele; MAF, Minor Allele Frequency; Beta, effect size on spherical equivalent in diopters; sem, standard error of the mean.

		Discovery cohort:				**Replication**									
		RS-I (*n* = 5328)				**RS-II (*n* = 2008)**		**RS-III (*n* = 1970)**		**ERF (*n* = 2032)**		**TwinsUK (*n* = 4270)**		**Meta-analysis (*n* = 15608)**	
SNP	**Position**	**MA**	**MAF**	**Beta (sem)**	***P***	**Beta (sem)**	***P***	**Beta (sem)**	***P***	**Beta (sem)**	***P***	**Beta (sem)**	***P***	**Beta (sem)**	***P***
rs688220	32786167	A	0.45	-0.27 (0.05)	$1.76x10^{-8}$	-0.28 (0.08)	$3.80x10^{-4}$	-0.22 (0.08)	$9.27x10^{-3}$	-0.03 (0.07)	$6.24x10^{-1}$	-0.15 (0.07)	$2.60x10^{-2}$	-0.20 (0.0009)	$2.79x10^{-11}$
rs580839	32786121	A	0.44	-0.27 (0.05)	$1.89x10^{-8}$	-0.27 (0.08)	$4.96x10^{-4}$	-0.22 (0.08)	$7.95x10^{-3}$	-0.03 (0.07)	$6.34x10^{-1}$	-0.16 (0.07)	$1.92x10^{-2}$	-0.20 (0.0009)	$2.53x10^{-11}$
rs619788	32782398	A	0.44	-0.27 (0.05)	$1.92x10^{-8}$	-0.27 (0.08)	$4.94x10^{-4}$	-0.22 (0.08)	$7.72x10^{-3}$	-0.03 (0.07)	$6.27x10^{-1}$	-0.16 (0.07)	$1.85x10^{-2}$	-0.20 (0.0009)	$2.53x10^{-11}$
rs4924134	32781857	G	0.44	-0.27 (0.05)	$2.04x10^{-8}$	-0.27 (0.08)	$4.76x10^{-4}$	-0.27 (0.08)	$6.58x10^{-3}$	-0.06 (0.07)	$4.10x10^{-1}$	-0.16 (0.07)	$1.85x10^{-2}$	-0.21 (0.0009)	$1.36x10^{-12}$
rs560766	32788234	A	0.44	-0.26 (0.05)	$4.27x10^{-8}$	-0.28 (0.08)	$4.54x10^{-4}$	-0.21 (0.08)	$1.29x10^{-2}$	-0.03 (0.07)	$6.65x10^{-1}$	-0.18 (0.07)	$7.68x10^{-3}$	-0.20 (0.0009)	$2.49x10^{-11}$
rs7176510	32786771	T	0.45	-0.26 (0.05)	$5.16x10^{-8}$	-0.28 (0.08)	$5.10x10^{-4}$	-0.22 (0.08)	$9.62x10^{-3}$	-0.02 (0.07)	$7.51x10^{-1}$	-0.16 (0.07)	$1.76x10^{-2}$	-0.20 (0.0009)	$6.25x10^{-11}$
rs7163001	32777866	A	0.44	-0.26 (0.05)	$5.23x10^{-8}$	-0.28 (0.08)	$4.08x10^{-4}$	-0.23 (0.08)	$5.89x10^{-3}$	-0.07 (0.07)	$3.01x10^{-1}$	-0.16 (0.07)	$1.87x10^{-2}$	-0.21 (0.0009)	$5.61x10^{-12}$
rs11073060	32777143	A	0.44	-0.26 (0.05)	$5.76x10^{-8}$	-0.28 (0.08)	$4.05x10^{-4}$	-0.23 (0.08)	$5.82x10^{-3}$	-0.08 (0.07)	$2.72x10^{-1}$	-0.16 (0.07)	$1.91x10^{-2}$	-0.21 (0.0009)	$3.65x10^{-12}$
rs8032019	32778782	G	0.40	-0.26 (0.05)	$6.09x10^{-8}$	-0.28 (0.08)	$5.57x10^{-4}$	-0.13 (0.09)	$1.30x10^{-1}$	-0.05 (0.07)	$5.12x10^{-1}$	-0.16 (0.07)	$1.96x10^{-2}$	-0.19 (0.0009)	$3.71x10^{-10}$
rs685352	32795627	G	0.44	-0.25 (0.05)	$8.80x10^{-8}$	-0.25 (0.08)	$1.28x10^{-3}$	-0.19 (0.08)	$1.98x10^{-2}$	-0.07 (0.07)	$3.06x10^{-1}$	-0.24 (0.07)	$4.43x10^{-4}$	-0.21 (0.0009)	$4.19x10^{-12}$
rs524952	32793178	A	0.47	-0.25 (0.05)	$1.03x10^{-7}$	-0.30 (0.08)	$2.09x10^{-4}$	-0.19 (0.08)	$2.56x10^{-2}$	-0.06 (0.07)	$4.13x10^{-1}$	-0.32 (0.07)	$4.15x10^{-6}$	-0.23 (0.0009)	$3.18x10^{-14}$
rs634990	32793365	C	0.47	-0.25 (0.05)	$1.03x10^{-7}$	-0.30 (0.08)	$2.15x10^{-4}$	-0.20 (0.08)	$2.03x10^{-2}$	-0.05 (0.07)	$5.11x10^{-1}$	-0.33 (0.07)	$2.93x10^{-6}$	-0.23 (0.0009)	$2.21x10^{-14}$
rs11073059	32776966	A	0.44	-0.25 (0.05)	$1.20x10^{-7}$	-0.28 (0.08)	$3.96x10^{-4}$	-0.23 (0.08)	$5.83x10^{-3}$	-0.08 (0.07)	$2.72x10^{-1}$	-0.16 (0.07)	$1.91x10^{-2}$	-0.20 (0.0009)	$8.45x10^{-12}$
rs11073058	32776918	T	0.44	-0.25 (0.05)	$1.30x10^{-7}$	-0.28 (0.08)	$3.93x10^{-4}$	-0.23 (0.08)	$5.84x10^{-3}$	-0.08 (0.07)	$2.71x10^{-1}$	-0.16 (0.07)	$1.90x10^{-2}$	-0.20 (0.0009)	$8.45x10^{-12}$

Online methods

Participants

Discovery cohort

The Rotterdam Study (**RS-I**) is a prospective population-based cohort study of 7,983 residents aged 55 years and older living in Ommoord, a suburb of Rotterdam, the Netherlands[32]. The baseline examination for the ophthalmic part took place between 1991 and 1993, and included 6,775 persons. Subjects were excluded if they had undergone bilateral cataract surgery, laser refractive procedures, or other intra-ocular procedures which might alter refraction. Complete data on refractive error and genome-wide SNPs were available on 5,328 persons, of whom 99% were of European ancestry.

Replication cohorts

The first three replication studies originated from the Netherlands. The first cohort was **RS-II**, an independent cohort which included 2,157 new participants aged 55+ years living in Ommoord since 2000[32], who had good quality genotyping data. Baseline examinations took place between 2000 and 2002; follow-up examination from 2004 to 2005. The second replication cohort was **RS-III**, a study which included 2,082 new participants aged 45 and older living in Ommoord since 2006, who had good quality genotyping data. Baseline examination took place between 2006 and 2009. The third replication study was the Erasmus Rucphen Family (**ERF**) Study, a family-based study in a genetically isolated population in the southwest of the Netherlands. This study included 2,032 living descendants aged 18 years and older originating from 22 families who had at least six children baptized in the community church between 1880 and 1900, and who had good quality genotyping data. The fourth replication cohort was derived from the United Kingdom (**TwinsUK**). This study is an adult twin registry of over 10,000 healthy volunteer twins based at St Thomas' Hospital in London.

Participants were recruited and phenotyped between 1998 and 2008. A total of 4,270 Caucasian participants had complete data on ocular phenotype and genotype[33].
As in the discovery cohort, participants in the four replication cohorts had been excluded if they had undergone bilateral surgery which inhibited evaluation of the original refractive error.

Measurements of refractive error

All studies used a similar protocol for phenotyping. Participants underwent an ophthalmologic examination which included non-dilated automated measurement of refractive error (RS I – III, ERF: Topcon RM-A2000 autorefractor; TwinsUK cohort: Humphrey-670 (Humphrey Instruments, San Leandro, CA) from 1998 to 2002; and then ARM-10 (Takagi Seiko, Japan), best-corrected visual acuity, and keratometry. Spherical equivalent (SE) was calculated from the standard formula: spherical equivalent = sphere + (cylinder/2). In addition to investigating SE as a quantitative trait, we stratified SE into categories of refractive error to evaluate findings from a clinical viewpoint. Myopia was categorized into low (SE –1.5 to –3 diopters (D)), moderate (SE –3 to –6D), and high (SE –6 D or lower). For hyperopia, these categories were mild (SE +1.5 to +3D), moderate (SE +3 to +6D), and high (SE +6D or higher), respectively. We considered SE -1.5 to +1.5D as emmetropia.

Ethics

All measurements in RS-I-III and ERF were conducted after the Medical Ethics Committee of the Erasmus University had approved the study protocols, and all participants had given a written informed consent in accordance with the Declaration of Helsinki. In the TwinsUK study, all twins gave fully informed consent under a protocol reviewed by the St Thomas' Hospital Local Research Ethics Committee.

Genotyping

Discovery cohort

All persons attending the baseline examination in 1990-1993 consented to genotyping, and had DNA extracted from blood leucocytes. Genotyping of autosomal SNPs was performed in persons with high-quality extracted DNA (n=6,449) using the Illumina Infinium II HumanHap550chip v3.0® array according to the manufacturer's protocols. Samples with low call rate (<97.5%, n=209), with excess autosomal heterozygosity (>0.336, n=21), and with sex-mismatch (n=36) were excluded, as were outliers identified by the identity-by-state (IBS) clustering analysis (>3 standard deviations from population mean, n=102 or IBS probabilities >97%, n=129). The total sample of individuals with good quality genotyping data was 5,974.

Replication cohorts

In RS-II, the majority of the 2,516 DNA samples were genotyped using the HumanHap 550 Duo Arrays; 133 (5%) were genotyped using the Human 610 Quad Arrays (Illumina). In the RS-III cohort, all DNA samples were genotyped using the Illumina Infinium II HumanHap550chip v3.0® array. In ERF, DNA was genotyped on four different platforms (Illumina 6k, Illumina 318K, Illumina 370K and Affymetrix 250K). Genotyping for the TwinsUK cohort took place in stages; in the first stage 1,810 individuals were genotyped using Illumina's HumanHap 300k duo chip, at a later stage 2,578 persons were genotyped using Illumina's HumanHap610 Quad.

Imputation

The set of genotyped input SNPs used for imputation in each study was selected based on highest quality GWA data. The callrate was set at >98% in Rotterdam Study I-III; the minor allele frequency at >0.01; and the Hardy-Weinberg $P > 10^{-6}$. We used the Markov Chain

Haplotyping (MaCH) package version 1.0.15 software (Rotterdam; imputed to plus strand of NCBI build 36, HapMap release #22) for the analyses. For each imputed SNP, a reliability of imputation was estimated (as the ratio of the empirically observed dosage variance to the expected binomial dosage variance: O/E ratio).

Statistical analysis

Discovery cohort

Refractive error measured at baseline as a continuous variable was used as outcome in the analysis. We calculated the mean SE for those with measurements on both eyes, and included the SE of only one eye if data from the other eye were missing. Linear regression models with 1-degree of freedom trend test were used to examine the associations between SNPs and SE, adjusted for age and gender. Using these linear regression models, we calculated regression coefficients with corresponding 95% confidence intervals (CIs). Odds ratios (ORs) of myopia and hyperopia were calculated with logistic regression analysis, adjusting for age and gender. GWAS analyses were performed using GRIMP[34].

We used genomic control to obtain optimal and unbiased results, and applied the inverse variance method of each effect size estimated for both autosomal SNPs that were genotyped and imputed in both cohorts. A *P*-value $<5x10^{-8}$ was considered genome-wide significant.

Replication analyses

The topSNPs with *P*-value $<1 \times 10^{-6}$ from the discovery analysis were examined in the replication cohorts RS-II, RS-III, ERF and TwinsUK cohorts using SPSS version 15.0.0 for Windows (SPSS inc., Chicago, IL, USA; 2006), and R statistical package version 2.8.1 for Linux. A meta-analysis was performed on all 5 studies using Metal for Linux.

GRIMP[34] was used for the analysis of the population-based replication cohorts. To adjust for family relationships, the GenABEL package[35] was used in the ERF study, and Merlin in the TwinsUK Study[36]. SNPs which deviated significantly from Hardy-Weinberg equilibrium ($P<10^{-6}$), or which had minor allele frequency <0.01 were excluded.

Gene expression data in human eye tissue

Human gene expression data were obtained essentially as described[37]. In short, postmortem eye bulbs (RPE: 6 donor eyes, choroid: 3 donor eyes, photoreceptors: 3 donor eyes), provided by the Corneabank Amsterdam, were rapidly frozen using liquid N2. Donors were between 63 and 78 years old and had no known history of eye pathology.

Cryosections were cut from the macula, and histology confirmed a normal histological appearance. RPE, photoreceptor and choroidal cells were isolated from macular sections using a Laser Microdissection System (PALM, Bernried, Germany). Total RNA was isolated and the mRNA component was amplified, labelled, and hybridized to a 44k microarray (Agilent Technologies, Amstelveen, The Netherlands)[38]. At least 3-6 microarrays were performed per tissue. Sample isolation, procedures, and expression microarray analysis were carried out according to obligatory MIAMI guidelines and the relevant expression data are deposited in the GEO database (2010) with accession number GSE20191. As a measure of the level of expression we sorted all the genes represented on the 44k microarray by increasing expression and calculated the corresponding percentiles (**Supplementary Table 3**).

Ingenuity database search

We explored the Ingenuity knowledge database using the keyword ‘eye development’ for all genes involved in ‘function or diseases’. This search provided approximately 100 genes, which formed a new network for eye development. We subsequently added the *GJD2* gene to

the network, and used the Path Explorer tool to search for possible functional relationships between *GDJ2* and these eye development genes in human, mouse, rat, and in vitro models (**Supplementary Figure 2a**). We continued the search using the keyword ‘eye growth’ for all genes involved in ‘function or diseases’, and investigated functional links between molecules using the connect tool and upstream-downstream analysis (**Supplementary Figure 2b**).

References

1. Bourne, R.R., Dineen, B.P., Ali, S.M., Noorul Huq, D.M. & Johnson, G.J. Prevalence of refractive error in Bangladeshi adults: results of the National Blindness and Low Vision Survey of Bangladesh. *Ophthalmology* **111**, 1150-60 (2004).
2. Dandona, R. et al. Population-based assessment of refractive error in India: the Andhra Pradesh eye disease study. *Clin Experiment Ophthalmol* **30**, 84-93 (2002).
3. Kempen, J.H. et al. The prevalence of refractive errors among adults in the United States, Western Europe, and Australia. *Arch Ophthalmol* **122**, 495-505 (2004).
4. Sawada, A. et al. Refractive errors in an elderly Japanese population: the Tajimi study. *Ophthalmology* **115**, 363-370 e3 (2008).
5. Vitale, S., Ellwein, L., Cotch, M.F., Ferris, F.L., 3rd & Sperduto, R. Prevalence of refractive error in the United States, 1999-2004. *Arch Ophthalmol* **126**, 1111-9 (2008).
6. McBrien, N.A. & Gentle, A. Role of the sclera in the development and pathological complications of myopia. *Prog Retin Eye Res* **22**, 307-38 (2003).
7. Saw, S.M. et al. How blinding is pathological myopia? *Br J Ophthalmol* **90**, 525-6 (2006).
8. Curtin, B.J. & Karlin, D.B. Axial length measurements and fundus changes of the myopic eye. *Am J Ophthalmol* **1**, 42-53 (1971).
9. Saw, S.M., Gazzard, G., Shih-Yen, E.C. & Chua, W.H. Myopia and associated pathological complications. *Ophthalmic Physiol Opt* **25**, 381-91 (2005).
10. Tano, Y. et al. Pathologic myopia: where are we now? *Am J Ophthalmol* **134**, 645-60 (2002).
11. Young, T.L. et al. Molecular genetics of human myopia: an update. *Optom Vis Sci* **86**, E8-E22 (2009).
12. Dirani, M. et al. Outdoor activity and myopia in Singapore teenage children. *Br J Ophthalmol* **93**, 997-1000 (2009).
13. McBrien, N.A. et al. Myopia: Recent Advances in Molecular Studies; Prevalence, Progression and Risk Factors; Emmetropization; Therapies; Optical Links; Peripheral Refraction; Sclera and Ocular Growth; Signalling Cascades; and Animal Models. *Optom Vis Sci* (2008).
14. Saw, S.M., Hong, C.Y., Chia, K.S., Stone, R.A. & Tan, D. Nearwork and myopia in young children. *Lancet* **357**, 390 (2001).
15. Young, T.L., Metlapally, R. & Shay, A.E. Complex trait genetics of refractive error. *Arch Ophthalmol* **125**, 38-48 (2007).

16. Eystathioy, T., Jakymiw, A., Fujita, D.J., Fritzler, M.J. & Chan, E.K. Human autoantibodies to a novel Golgi protein golgin-67: high similarity with golgin-95/gm 130 autoantigen. *J Autoimmun* **14**, 179-87 (2000).
17. Jobling, A.I., Gentle, A., Metlapally, R., McGowan, B.J. & McBrien, N.A. Regulation of scleral cell contraction by transforming growth factor-beta and stress: competing roles in myopic eye growth. *J Biol Chem* **284**, 2072-9 (2009).
18. Kihara, A.H. et al. Connexin36, an essential element in the rod pathway, is highly expressed in the essentially rodless retina of Gallus gallus. *J Comp Neurol* **512**, 651-63 (2009).
19. Deans, M.R., Volgyi, B., Goodenough, D.A., Bloomfield, S.A. & Paul, D.L. Connexin36 is essential for transmission of rod-mediated visual signals in the mammalian retina. *Neuron* **36**, 703-12 (2002).
20. Striedinger, K. et al. Loss of connexin36 increases retinal cell vulnerability to secondary cell loss. *Eur J Neurosci* **22**, 605-16 (2005).
21. Guldenagel, M. et al. Visual transmission deficits in mice with targeted disruption of the gap junction gene connexin36. *J Neurosci* **21**, 6036-44 (2001).
22. Rong, P. et al. Disruption of Gja8 (a8 connexin) in mice leads to microphthalmia associated with retardation of lens growth and lens fiber maturation. *Development* **129**, 167-174 (2002).
23. White, W. Targeted Ablation of Connexin50 in Mice Results in Microphthalmia and Zonular Pulverulent Cataracts. *The Journal of Cell Biology* **143**, 815-825 (1998).
24. Heintzman, N.D. et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108-12 (2009).
25. Heintzman, N.D. & Ren, B. Finding distal regulatory elements in the human genome. *Curr Opin Genet Dev* **19**, 541-9 (2009).
26. Delaloy, C. et al. Deletion of WNK1 first intron results in misregulation of both isoforms in renal and extrarenal tissues. *Hypertension* **52**, 1149-54 (2008).
27. Mihaly, J. et al. Dissecting the regulatory landscape of the Abd-B gene of the bithorax complex. *Development* **133**, 2983-93 (2006).
28. Hysi, P.G. & Young, T.L. Predisposition to Myopia and Refractive Error is Conferred by a Locus on Chromosome 15q25 Which Contains the RASGRF1 Gene. *Nat Genet* (2010).
29. Jones, C. & Moses, K. Cell-cycle regulation and cell-type specification in the developing Drosophila compound eye. *Semin Cell Dev Biol* **15**, 75-81 (2004).

30. Fernandez-Medarde, A. et al. RasGRF1 disruption causes retinal photoreception defects and associated transcriptomic alterations. *J Neurochem* **110**, 641-52 (2009).
31. Lettre, G. et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* **40**, 584-91 (2008).
32. Hofman, A. et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* (2009).
33. Spector, T.D. & MacGregor, A.J. The St. Thomas' UK Adult Twin Registry. *Twin Res* **5**, 440-3 (2002).
34. Estrada, K. et al. GRIMP: A web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* (2009).
35. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-6 (2007).
36. Abecasis, G.R., Cherny, S.S., Cookson, W.O. & Cardon, L.R. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* **30**, 97-101 (2002).
37. Booij J.C. et al. Functional annotation of the human retinal pigment epithelium transcriptome. *BMC Genomics* **20**, 10-164 (2009).
38. Van Soest S.S. et al. Comparison of human retinal pigment epithelium gene expression in macula and periphery highlights potential topographic differences in Bruch's membrane. *Mol Vis* **10**, 13:1608-17 (2007).

URL references

Ingenuity	http://www.ingenuity.com
MaCH	http://www.sph.umich.edu/csg/abecasis/MACH
R	http://www.r-project.org
Metal	http://www.sph.umich.edu/csg/abecasis/metal

Author contribution

A.M.S., V.J.M.V., and C.C.W.K. performed analyses and drafted the manuscript; C.M.van D., B.A.O., F.R., A.G.U., A.H., P.T.V.M.d.J., J.R.V., and C.C.W.K. designed the study and obtained funding; D.D.G.D., L.M.v.K., L.H., W.R., M.C., R.K., J.J.M.W.-A., T.G.M.F.G., F.C.C.R., and S.M.A.S. helped in data collection; A.J.M.H.V., M.K.I., N.A.M.S., Y.S.A., A.A.B.B., A.A.L.J.v.O., and A.I. participated in the genetic analyses; P.G.H., T.L.Y., D.A.M., T.D.S., and C.J.H. were responsible for data from the TwinsUK study; M.K.I., R.W.A.M.K., G.v.R., P.G.H., C.J.H., C.M.v.D., A.J.M.H.V., B.A.O., J.R.V., and A.A.B.B. critically reviewed the manuscript.

Competing financial interests

The authors declare no competing financial interests.