Human Molecular Genetics, 2010, Vol. 19, No. 13 **2716–2724**
doi:10.1093/hmg/ddq144
Advance Access published on April 15, 2010

Genome-wide association identifies *ATOH7* as a major gene determining human optic disc size

Stuart Macgregor[1,†], Alex W. Hewitt[2,†], Pirro G. Hysi[3], Jonathan B. Ruddle[2], Sarah E. Medland[1], Anjali K. Henders[1], Scott D. Gordon[1], Toby Andrew[3], Brian McEvoy[1], Paul G. Sanfilippo[2], Francis Carbonaro[3], Vikas Tah[3], Yi Ju Li[4,5], Sonya L. Bennett[2], Jamie E. Craig[6], Grant W. Montgomery[1], Khanh-Nhat Tran-Viet[4], Nadean L. Brown[7,8], Timothy D. Spector[3], Nicholas G. Martin[1], Terri L. Young[4], Christopher J. Hammond[3] and David A. Mackey[2,9,10,*]

[1]Genetics and Population Health, Queensland Institute of Medical Research, Brisbane, Australia, [2]Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, University of Melbourne, Melbourne, Australia, [3]Department of Twin Research and Genetic Epidemiology, King's College London School of Medicine, St Thomas' Hospital, London, UK, [4]Center for Human Genetics and [5]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, USA, [6]Department of Ophthalmology, Flinders Medical Centre, Flinders University, Adelaide, Australia, [7]Division of Developmental Biology, Cincinnati Children's Research Foundation and [8]Department of Pediatrics, University of Cincinnati School of Medicine, Cincinnati, OH, USA, [9]Lions Eye Institute, Centre for Ophthalmology and Visual Science, University of Western Australia, Perth, Australia and [10]Discipline of Medicine, University of Tasmania, Hobart, Australia

Received December 22, 2009; Revised March 31, 2010; Accepted April 9, 2010

Optic nerve assessment is important for many blinding diseases, with cup-to-disc ratio (CDR) assessments commonly used in both diagnosis and progression monitoring of glaucoma patients. Optic disc, cup, rim area and CDR measurements all show substantial variation between human populations and high heritability estimates within populations. To identify loci underlying these quantitative traits, we performed a genome-wide association study in two Australian twin cohorts and identified rs3858145, $P = 6.2 \times 10^{-10}$, near the *ATOH7* gene as associated with the mean disc area. *ATOH7* is known from studies in model organisms to play a key role in retinal ganglion cell formation. The association with rs3858145 was replicated in a cohort of UK twins, with a meta-analysis of the combined data yielding $P = 3.4 \times 10^{-10}$. Imputation further increased the evidence for association for several SNPs in and around *ATOH7* ($P = 1.3 \times 10^{-10}$ to 4.3×10^{-11}, top SNP rs1900004). The meta-analysis also provided suggestive evidence for association for the cup area at rs690037, $P = 1.5 \times 10^{-7}$, in the gene *RFTN1*. Direct sequencing of *ATOH7* in 12 patients with optic nerve hypoplasia, one of the leading causes of blindness in children, revealed two novel non-synonymous mutations (Arg65Gly, Ala47Thr) which were not found in 90 unrelated controls (combined Fisher's exact $P = 0.0136$). Furthermore, the Arg65Gly variant was found to have very low frequency (0.00066) in an additional set of 672 controls.

INTRODUCTION

Since the 1850s, the evaluation of the optic nerve head has been recognized as critical in the assessment of many blinding diseases, particularly glaucoma. The optic nerve is the major afferent input to the brain and is composed of retinal ganglion cell (RGC) axons and their supporting tissue. Clinically, the

*To whom correspondence should be addressed at: Lions Eye Institute, Centre for Ophthalmology and Visual Science, University of Western Australia, 2 Verdun Street, Nedlands, WA 6009, Australia. Tel: +61 893810777; Fax: +61 893810700; Email: d.mackey@utas.edu.au
†These authors contributed equally to this work.

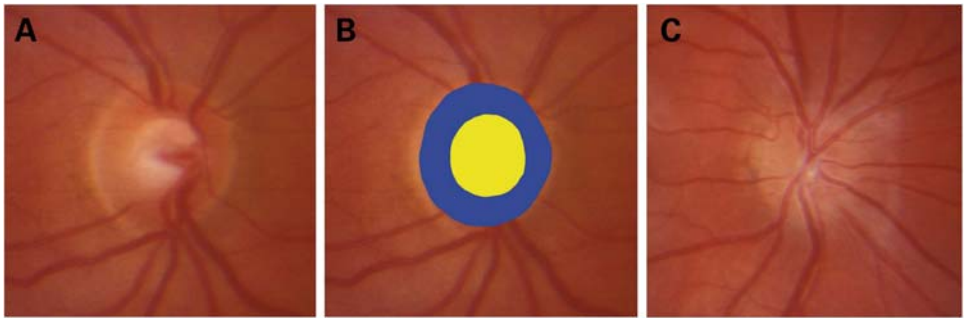

Figure 1. Colour photograph of the normal optic nerve head (**A**). In (**B**), the optic cup and neuroretinal rim areas are coloured yellow and blue, respectively. The optic disc area encompasses the sum of the optic cup and neuroretinal rim areas. The appearance of optic nerve hypoplasia is displayed in (**C**).

optic nerve can be directly assessed by observing the optic cup and neuroretinal rim areas (Fig. 1), as the RGCs exit the eye in a crude retinotopic pattern (1,2). Diseases of the optic nerve generally result in RGC axonal loss, which manifests clinically as optic atrophy or optic neuropathy with corresponding reduced visual acuity, altered colour perception and visual field defects (3).

There is considerable variation in the size of the optic disc both within populations and between ethnic groups, with individuals of African descent generally having larger disc diameters compared with people of Asian or European descent (4). Interestingly, the number of optic nerve fibres is correlated with optic disc size (5), and many components of the optic nerve have been found to have a high heritability (6–10). Small optic nerves are predisposed to non-arteritic anterior ischaemic optic neuropathy (11), the development of optic disc drusen (12) and visual loss from Leber's hereditary optic neuropathy (LHON) (4,13,14). To date, several genes have known association with uncommon Mendelian forms of optic atrophy or optic neuropathy such as that seen in LHON; autosomal dominant optic atrophy; Wolfram syndrome; a small proportion of open-angle glaucoma cases as well as other rarer neurodegenerative diseases (3). Clinically, change in cup-to-disc ratio (CDR) over time is commonly used to follow disease progression in glaucoma. Other diseases, such as optic nerve hypoplasia, result in congenitally small optic discs. Herein, we present a genome-wide association to identify genes associated with the endophenotype of optic disc and neuroretinal rim area, with the aim of uncovering potential candidates for optic nerve diseases and thereby allowing an improved understanding of human optic nerve development.

RESULTS

Data summaries are provided in Table 1. Differences in means and standard deviations between countries are largely attributable to differing measurement protocols. All four traits are inter-correlated. The cup area has a much higher coefficient of variation than the rim area. Thus, the majority of variation in CDR is driven by the observed variation in the cup area. CDR, the trait commonly used clinically, is highly correlated with the cup area (Pearson correlation 0.89). Since the disc area is the sum of the cup and rim areas, both the disc rim and disc cup correlations are also high (>0.6). Given these high correlations, the reported *P*-values are not corrected for the testing of multiple traits.

Genotyping quality appeared excellent with an average discordance rate of 49 per million across the seven Australian individuals genotyped at both the Center for Inherited Disease Research (CIDR, The Johns Hopkins University, Baltimore, MD, USA) and deCODE Genetics (Reykjavik, Iceland). In the Australian data, after the removal of ancestry outliers, there was no evidence for population stratification. Supplementary Material Figures S1–S4 show that the Q–Q plots for each trait approximate what would be expected in a homogeneous population. The only strong deviation observed was for the disc area, a trait for which we have found clear evidence for genome-wide significant associations. The genomic inflation factor (λ) in the Australian cohort was 1.01 for all four optic nerve traits here. The λ values in the UK data were 0.99, 1.04, 0.98 and 1.02 for the disc area, cup area, rim area and CDR, respectively. Results are presented uncorrected for λ.

The results for association testing across the genome in the Australian cohort are provided in Figure 2. There is a clear genome-wide significant signal at rs3858145 for the disc area ($P = 6.2 \times 10^{-10}$), <20 kb from the nearest gene, *atonal homolog 7* (*ATOH7*) on chromosome 10q21. SNP rs3858145 was also associated with the cup area ($P = 1.4 \times 10^{-5}$), CDR ($P = 1.1 \times 10^{-3}$) and the rim area ($P = 2.0 \times 10^{-3}$). Imputing SNPs in HapMap slightly increased the evidence for association in the *ATOH7* region, with the most associated ($P = 2.0 \times 10^{-10}$) SNP for the disc area being rs10762201, at 69 710 117 bp. No other region revealed genome-wide significant signals ($P < 5 \times 10^{-8}$) in the Australian cohort alone, and no SNPs reached genome-wide significance in the UK cohort alone.

To confirm the *ATOH7* locus and identify further SNPs of smaller effect, we performed a meta-analysis of the Australian and UK cohorts (Table 2). The *ATOH7* region replicated in the UK cohort, with $P = 1.3 \times 10^{-2}$ for rs3858145 for the disc area. After imputation, the meta-analysis of both cohorts indicated maximum evidence for association at SNP rs1900004 for the disc area ($P = 4.2 \times 10^{-11}$). The effect size estimates (standard errors) at rs1900004 (C allele) were 0.298 (0.049), $P = 1.7 \times 10^{-9}$ and 0.179 (0.06), $P = 2.9 \times 10^{-3}$ in the Australian and UK data, respectively. All SNPs in the *ATOH7* region with minor allele frequency (MAF) >0.1 in HapMap Centre d'Etude du Polymorphisme Humain (CEU) samples could be successfully imputed in both the Australian and UK cohorts;

Table 1. Descriptive statistics for optic disc areas in the two cohorts

	Australia	UK
Number of subjects	1368	848
Number of families	666	531
Age (years)	20 ± 10	56 ± 13
Gender (%female)	55	90
Optic disc area (mm^2)	2.06 ± 0.39	2.58 ± 0.65
Optic cup area (mm^2)	0.45 ± 0.29	0.31 ± 0.23
Neuroretinal rim area (mm^2)	1.62 ± 0.31	2.27 ± 0.57
CDR	0.45 ± 0.13	0.32 ± 0.10

Data are presented as the mean ± standard deviation. Full details of the Australian and UK cohorts are given in Mackey *et al*. (34) and Healey *et al*. (10).

meta-analysis results for both imputed and genotyped SNPs are displayed in Figure 3. There is little recombination across the region from the 5′ (centromeric) end of *ATOH7* at 69.660 mb to the 5′ (centromeric) end of the *phenazine biosynthesis-like protein domain-containing gene* (*PBLD*) at 69.712 mb. SNP s1900004 is <5 kb from the telomeric end of *ATOH7*, whereas several other SNPs, from rs7916697 ($P = 1.3 \times 10^{-10}$) at 69661859 to rs10762201 ($P = 6.9 \times 10^{-11}$) at 69710117, have *P*-values around 10^{-10}. SNP rs7916697 lies within *ATOH7*. Two additional SNPs in the region typed in the HapMap CEU samples, rs34888891 and rs7896916, have MAF in the range 0.05–0.1 but could not be imputed in the Australian or UK data sets. All of the most associated SNPs in the region had MAFs close to 0.25. We examined publicly available data on the effects of SNPs in the *ATOH7* region on gene expression in transformed lymphocytes (15,16). None of the SNPs in or around *ATOH7*, that were typed by Dixon *et al*. (15), were significantly associated (all $P > 0.001$) with *ATOH7* expression. Similarly, rs3858145 had no significant influence on *ATOH7* expression in the data described by Stranger *et al*. (16).

Atoh7/*Ath5* embryonic expression and function have been studied in multiple vertebrate model organisms (reviewed in 17). In the prenatal mouse retina, *Atoh7*/*Math5* is activated just prior to the initiation of RGC histogenesis (18) and expressed by exiting/postmitotic retinal progenitor cells (19,20). As displayed in Figure 4, these cells coexpress *Atoh7*, and the transcription factor Pax6, which was recently shown to bind directly to a highly conserved 5′ retinal enhancer (21–23).

Sequence results in our optic nerve hypoplasia cohort revealed distinct cases with two SNPs in the 5′-UTR (untranslated gene region) (rs7916697, rs61854782), one novel coding synonymous SNP (g.487C>T; Gly22Gly) and two novel coding non-synonymous SNPs: a single nucleotide substitution with an amino acid change of arginine to glycine (g.614A>G; Arg65Gly) within helix 1 of the encoded ATOH7 protein, and a single base pair substitution resulting in an amino acid sequence change of an alanine to a threonine (g.560G>A; Ala47Thr) located N terminal to the basic domain. Neither of these non-synonymous coding SNPs were identified in 90 unrelated control individuals (combined Fisher's exact test $P = 0.0136$).

The predicted effects of the novel coding non-synonymous variants (g.560G>A; Ala47Thr and g.614A>G; Arg65Gly) were examined using SIFT analysis (http://blocks.fhcrc.org/sift/SIFT.html) and PolyPhen prediction (http://genetics.bwh.harvard.edu/pph). Both variants were predicted to be damaging based on SIFT analysis, although PolyPhen analysis determined Ala47Thr to be benign. The Ala47Thr variant is highly conserved across eutherian mammals (Supplementary Material, Fig. S5) in which a hydrophobic-aliphatic alanine is converted to a neutral threonine that consists of polar side chains. The Arg65Gly variant is highly conserved across eutherian mammals (Supplementary Material, Fig. S5) and is predicted to alter the hydrophobicity of the amino acid from a hydrophilic arginine to a neutral glycine.

Based on these results, we decided to pursue determination of variant frequency of the Arg65Gly variant only. As mentioned earlier, none of the 90 bidirectionally sequenced controls had the variant. A TaqMan assay of the SNP was designed and an additional 672 genomic DNA samples from a Duke myopia cohort with ophthalmic examination information were genotyped. In this cohort, eight DNA samples had the SNP by TaqMan genotyping, but only one was confirmed with additional bidirectional sequencing of these specific DNA samples. Therefore, out of 672 Duke DNA samples with full eye examinations and no optic nerve hypoplasia noted clinically, 1 sample had the Arg to Gly SNP variant by TaqMan genotyping and confirmed with bidirectional sequencing. In the total set of 762 individuals typed at this SNP, the allele frequency was estimated as $1/(2 \times 762) = 0.00066$. Comparing the frequency of this SNP in 12 hypoplasia cases ($1/24 = 0.042$) compared with 762 controls yielded a Fisher's exact test $P = 0.035$.

An SNP in the *raftlin lipid raft linker 1* gene (*RFTN1*) on chromosome 3p24, rs690037, was associated with the cup area in both cohorts, with the meta-analysis providing suggestive evidence for association ($P = 1.5 \times 10^{-7}$, Table 2 and Supplementary Material, Table S3). Imputation for this region of chromosome 3 did not improve the evidence for association, with the genotyped SNP rs690037 remaining the most associated SNP after imputation.

In the meta-analysis for the disc area, there was also suggestive evidence ($P = 3.1 \times 10^{-7}$) for association at rs1192415 near *heat shock protein 90 kDa beta (Grp94) member 3* (*HSP90B3P*), a pseudogene on chromosome 1p22 (Table 2 and Supplementary Material, Table S2). Similarly, there were signals with $P < 5 \times 10^{-7}$ for the disc area and SNPs near the genes *low density lipoprotein-related protein 1B* (*LRP1B*, chromosome 2q21) and *zinc finger protein 157* (*ZNF157*, chromosome Xp11). The meta-analysis results for disc, cup, rim and CDR for the top genotyped SNPs are shown in Supplementary Material, Tables S2–S5.

DISCUSSION

We have identified and replicated a genome-wide significant association with the *ATOH7* locus for the optic disc area, with the same region also affecting the cup area, rim area and CDR. *ATOH7* is a small, single exon gene, which has a key role in ocular embryogenesis and is a vertebrate semi-orthologue of *Drosophila atonal*, which acts during peripheral or central nervous system development (24,25). Targeted deletion of *Atoh7* causes ≥95% loss of RGCs and absence of

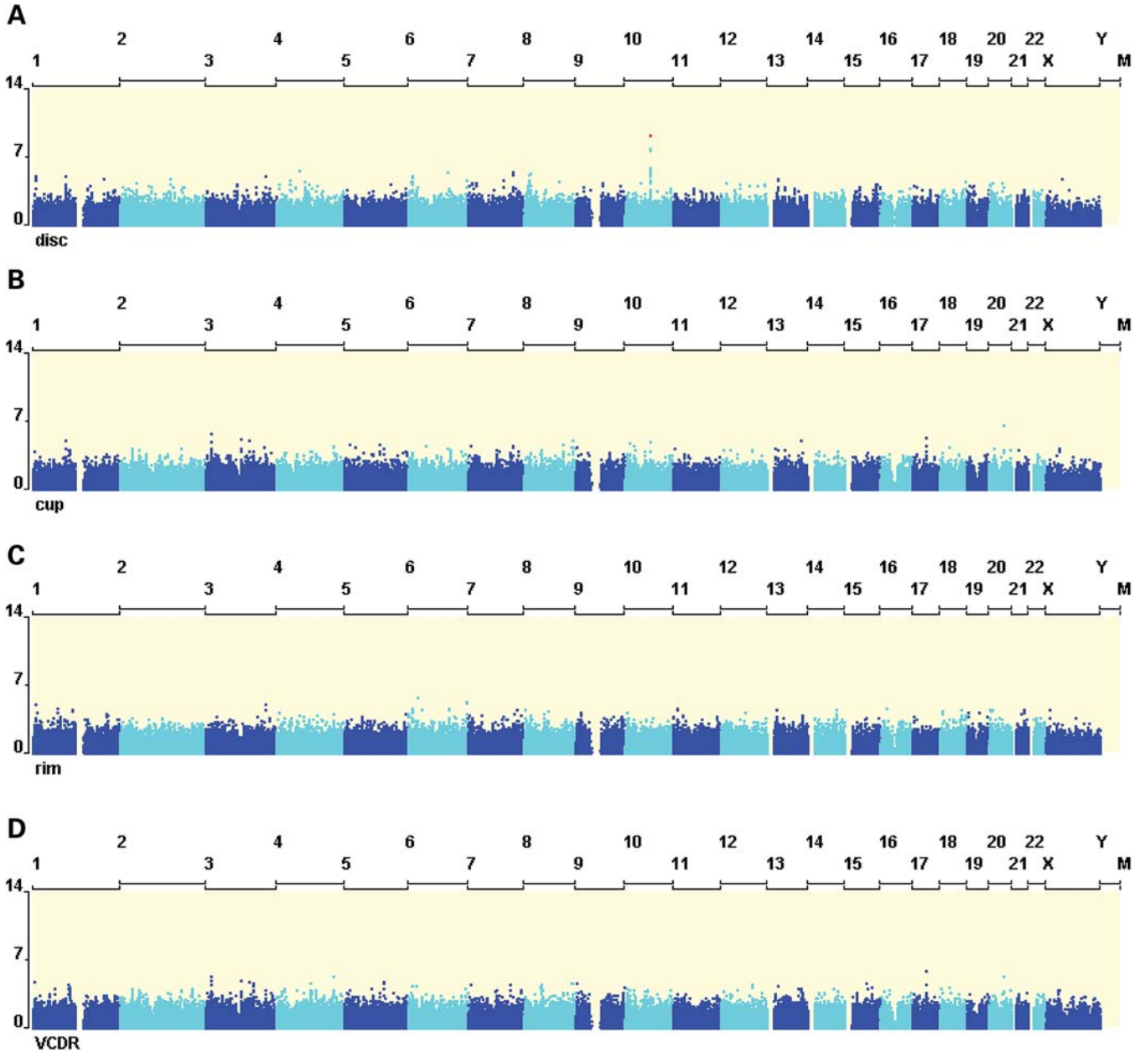

Figure 2. Genome-wide association of optic disc endophenotypes. Results are presented for the optic disc area (**A**), optic cup area (**B**), neuroretinal rim area (**C**) and the vertical optic CDR (**D**) for the discovery Australian twin cohort. Variants in red indicate genome-wide significance ($P = 5 \times 10^{-8}$).

optic nerves, chiasm and tracts in homozygous mutant mice (26,27). *Atoh7* occupies a key node in the genetic hierarchy of RGC development (17), and was recently shown to be directly activated by Pax6 in the embryonic retina (21,23). This regulatory relationship has clinical implications, as human *PAX6* is well associated with dysplastic or hypoplastic optic nerves (28).

The results from our meta-analysis show a broad region of association which includes the *ATOH7* gene itself. Two HapMap SNPs appear within *ATOH7* (rs6480320, which is monomorphic in the CEU population; and rs7916697, which appears on the Illumina 610 arrays but failed quality control in the Australian samples). SNP rs7916697 was successfully genotyped in the UK samples and was the most significant SNP in the region in the UK samples ($P = 2.7 \times 10^{-3}$). SNP rs7916697 could be imputed with good confidence ($r^2 = 0.93$) in the Australian samples, giving a meta-analysis *P*-value of 1.3×10^{-10}. There are also four SNPs in *ATOH7* listed in dbSNP that are derived from sequencing projects of small numbers of individuals (e.g. 1000 genomes pilot data). Hence, the population frequency estimates for these SNPs are not accurate. None of these dbSNP-listed polymorphisms within the transcribed region of *ATOH7* are located in the protein-coding portion of the gene. Further sequencing of *ATOH7* would possibly uncover further variation, although the extensive linkage disequilibrium (LD) in CEU and

Table 2. Directly genotyped variants identified on meta-analysis associated with optic nerve endophenotypes ($P < 5 \times 10^{-7}$)

Trait	Chr	Coordinate (build 36)	Marker	Allele	Nearest gene	Australia Effect	SE	*P*-value	UK Effect	SE	*P*-value	Meta-analysis *P*-value
Disc	10	69 689 377	rs12571093	G	*PBLD*	0.313	0.055	1.6E − 08	0.242	0.082	3.1E − 03	1.9E − 10
Disc	10	69 681 844	rs3858145	A	*ATOH7*	0.295	0.048	6.2E − 10	0.138	0.055	1.3E − 02	3.4E − 10
Disc	10	69 701 073	rs2241970	A	*PBLD*	0.311	0.055	2.0E − 08	0.121	0.066	6.7E − 02	3.4E − 08
Disc	10	69 682 295	rs17231602	G	*ATOH7*	0.286	0.059	1.4E − 06	0.233	0.087	7.6E − 03	3.5E − 08
Disc	10	69 733 976	rs7067601	T	*PBLD*	0.278	0.059	2.2E − 06	0.221	0.089	1.2E − 02	1.2E − 07
Disc	10	69 734 515	rs10733843	G	*PBLD*	0.276	0.059	2.6E − 06	0.215	0.087	1.3E − 02	1.5E − 07
Disc	10	69 733 542	rs4517412	G	PBLD	0.276	0.059	2.6E − 06	0.215	0.087	1.3E − 02	1.5E − 07
Cup	3	16 370 672	rs690037	C	*RFTN1*	0.208	0.043	1.6E − 06	0.137	0.061	2.5E − 02	1.5E − 07
Disc	10	69 637 453	rs6480314	A	*MYPN*	0.244	0.055	9.1E − 06	0.221	0.08	5.6E − 03	1.8E − 07
Cup	10	69 681 844	rs3858145	A	*ATOH7*	0.209	0.048	1.4E − 05	0.158	0.054	3.5E − 03	2.0E − 07
Disc	1	91 849 685	rs1192415	A	*HSP90B3P*	−0.245	0.055	8.6E − 06	−0.198	0.077	1.0E − 02	3.1E − 07
Disc	2	141 078 528	rs491391	A	*LRP1B*	−0.407	0.095	1.8E − 05	−0.364	0.131	5.5E − 03	3.4E − 07
Disc	X	47 059 220	rs6611365	G	*ZNF157*	0.178	0.042	2.0E − 05	0.173	0.063	5.9E − 03	4.4E − 07
Disc	1	91 841 555	rs1192404	A	*HSP90B3P*	−0.23	0.058	7.5E − 05	−0.256	0.082	1.9E − 03	4.6E − 07
Disc	10	69 738 803	rs10762217	T	*PBLD*	0.263	0.058	5.2E − 06	0.194	0.085	2.2E − 02	4.9E − 07

Japanese/Han Chinese samples suggests that statistically distinguishing between any of these variants would be difficult. It was also particularly interesting that no coding variant was identified on the direct sequencing of 90 people who had no signs of optic nerve hypoplasia.

For the cup area, we identified relatively strong evidence for association with rs690037 in the *RFTN1* gene on chromosome 3p24 in the Australian cohort ($P = 1.6 \times 10^{-6}$, variance explained: 2.1%), with a replication in the UK cohort ($P = 0.025$, variance explained 1.0%). *RFTN1* is a large (198 kb) gene thought to be important in the formation or maintenance of membrane lipid rafts. Interestingly, other lipid-raft molecules have been implicated in eye abnormal ocular development in *Drosophila* and maintenance of the mammalian central nervous system (29,30).

Despite the difficulties in obtaining unbiased estimates of effect size, it appears that, in contrast to traits such as stature (where effect sizes of 0.1% of the variance are the norm) (31), at least some of the genetic basis of the optic disc measures studied lies in the range of moderately large effect sizes. The most associated genotyped SNP (rs3858145) identified near the *ATOH7* locus in the Australian cohort explains 3.5% of the phenotypic variance in the optic disc area, though this is likely to be an overestimate of the true effect (32). The effect size in the UK replication cohort is more likely to represent an unbiased estimate of the true effect size. The proportion of variance explained by the SNP rs3858145 in the UK cohort is 0.8%, with the SNP rs7916697 within *ATOH7* explaining 1.7% of the variance in the UK cohort. For the association of rs690037 with the cup area, the effect size in the replication cohort was 1.0%.

In summary, we have demonstrated that the *ATOH7* locus, a gene known to be important in ocular development in model organisms, has an important role in determining optic disc size in humans (25). Sequencing data support the association of this gene with human optic nerve hypoplasia. Although the two non-synonymous amino acid substitutions are unlikely to affect ATOH7 protein binding to DNA or dimerization with other bHLH proteins, the changes could theoretically affect bHLH protein dimer interactions with the transcriptional machinery (33). Our results are consistent with sequence variants in or near *ATOH7* determining the optic disc area (and likely the cup area as well as CDR), although fine mapping of this locus in Caucasian or Asian samples will be difficult due to extensive LD. Relative to other quantitative traits measured in humans such as stature, optic disc measures appear to have a less complex genetic architecture more amenable to quantitative trait locus mapping.

MATERIALS AND METHODS

Subject recruitment

Twins in our discovery cohort were recruited in Australia from the Twins Eye Study in Tasmania (TEST) as well as the Brisbane Adolescent Twin Study (BATS) (34). A population-based twin cohort from the UK was used for replication (10). Participants ranged in age from 5 to 90 years. The majority of Australian twins were adolescents or young adults, whereas the majority of the UK twins were aged >50 years. All participants underwent a comprehensive ophthalmic examination (34), with the main aim of identifying underlying genetic and environmental factors associated with ocular biometry parameters as endophenotypes for eye diseases.

The UK data functioned as a replication cohort in this study. The UK data are also being used as a replication cohort for a genome-wide association study based on discovery cohorts from the Netherlands (Ramdas *et al*., submitted). The Australian and Dutch cohorts both independently implicated *ATOH7* as playing a role in optic disc phenotypes and both utilize the UK data to replicate their findings. Future studies seeking to combine the results from this study and the Ramdas *et al*. study should take into account the overlap in data used.

Simultaneous stereoscopic 15° optic disc photographs were taken with a Nidek 3-Dx fundus camera (Nidek, Gamagori, Japan) in both cohorts (early TwinsUK subjects were imaged with non-simultaneous 30° stereo-photographs with a Kowa camera, Kowa-Europe, Dusseldorf, Germany). All images were developed on colour 35 mm slides (Ektachrome,

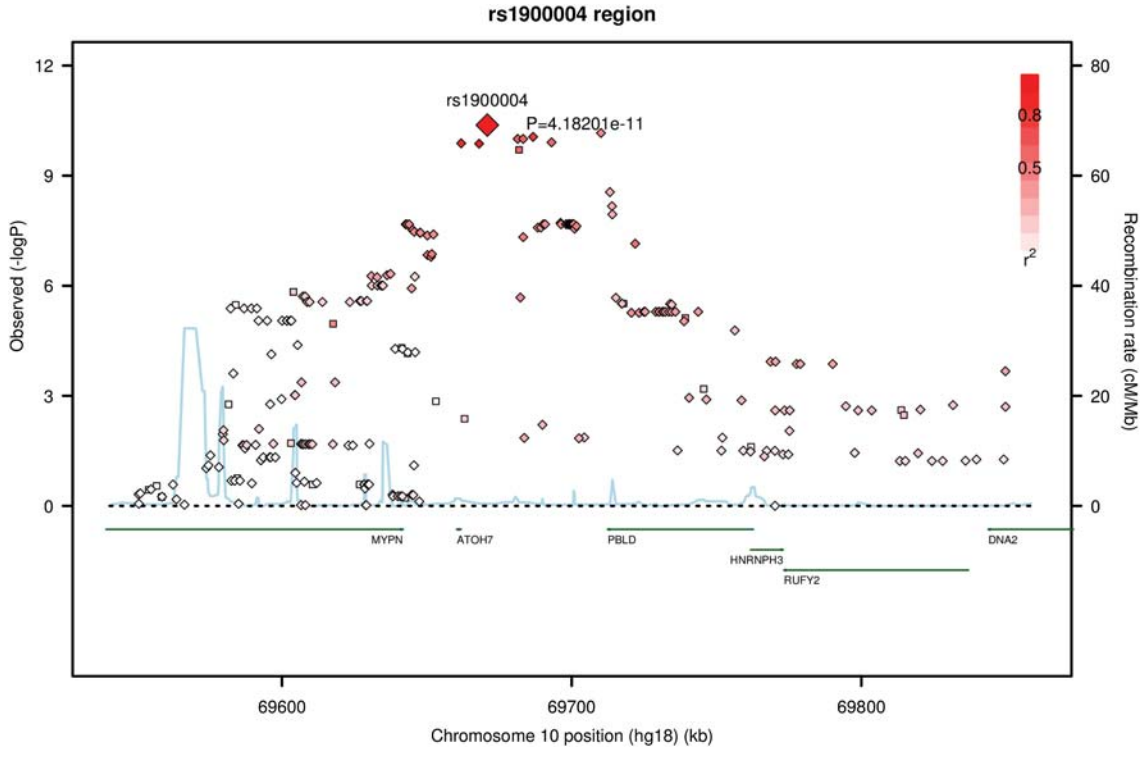

Figure 3. Significant association of the optic nerve endophenotype disc area at the chromosome 10q21 locus. Results from the meta-analysis of both the Australian and UK twin cohorts are presented. Imputed SNPs are indicated by diamonds and genotyped SNPs by squares. The degree of linkage disequilibrium between rs1900004 and other SNPs is indicated by red shading. The recombination rate is displayed in light blue line, with scale on the right-hand axis.

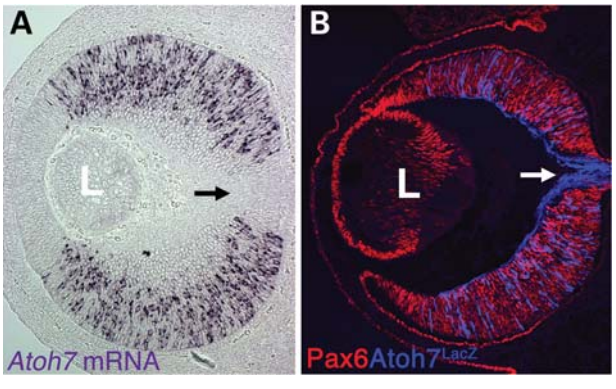

Figure 4. Mouse *Atoh7* is expressed by retinal progenitor cells that give rise to RGCs. (**A**) *Atoh7* mRNA (visualized in purple) is expressed by the majority of E15.5 retinal progenitor cells, but not within differentiated RGCs, or their axons at the optic nerve head (arrow). (**B**) In E13.5 $Atoh7^{tmGla}$ eyes, β-gal protein (in blue) reports retinal progenitor cell expression of *Atoh7*. Slow turnover of β-gal protein also shows subsequent differentiation into RGCs, which rapidly form into an optic nerve at the nerve head (arrow). All β-gal+ cells coexpress the transcription factor Pax6 (in red). Magnification in both panels is ×200. L, lens.

Eastman Kodak, Rochester, NY, USA) and digitized (Nikon CoolScan IV ED slide scanner, Nikon Corp., Tokyo, Japan). At both centres, images were analysed stereoscopically with custom planimetric software (StereoDx, using a Z-screen; StereoGraphics Corp., Beverly Hills, CA, USA) (35). As detailed previously, the inner margin of the optic disc and the neuroretinal rim were delineated at the depth of the scleral plane, and images were modified for magnification using refraction and keratometry data (35). Intra- and inter-grader correlations were high for both optic disc and the optic cup area measurements (intra-class correlation coefficients ranged from 0.75 to 0.96). In total, four related traits were considered for analysis: the optic disc area (total area of two concentric circles), optic cup area (area of the inner of two concentric circles), derived neuroretinal rim area (disc area minus cup area) and the vertical optic CDR.

Genotyping and data quality control

DNA was extracted from peripheral blood or saliva samples. All Australian samples used for this analysis were genotyped on the Illumina HumanHap 610W Quad arrays (Illumina Inc., San Diego, CA, USA). Most participants recruited as part of the BATS were genotyped as part of a larger project contracted to deCODE Genetics. A small number of BATS individuals (50) and the participants of the TEST were genotyped at the CIDR. The majority of UK twin samples were also genotyped on Hap610W arrays at the CIDR, with the remainder genotyped using the Illumina HumHap 300k Duo arrays at the Wellcome Trust Sanger Institute.

For the Australian data, the exclusion criteria for SNPs were: MAF $\leq$1%; Hardy–Weinberg equilibrium (HWE) (in all individuals) $P < 10^{-6}$; an SNP call rate $\leq$95%; or an Illumina Beadstudio GenCall Score $<$0.7; resulting in data on 543 862 SNPs. For the UK data, similar parameters were

used: MAF ≤4%, HWE $P < 10^{-4}$ and SNP call rate ≤95%. Seven Australian individuals were genotyped at both the CIDR and deCODE to verify genotyping quality.

Investigation of participant ancestry and population stratification

To identify ancestry outliers, we used principal component analysis (PCA) of similarly genotyped data from 16 global populations sourced from HapMap Phase 3 (HM3; 11 populations) and a previously published study of Northern European genetic diversity (GEUT; 5 populations) (36). PCA was conducted using autosomal SNPs that were genotyped in common between the Australian, HM3 and GEUT cohorts, and also had low missing rates (<2.5%) in all populations. The EIGENSOFT package was used to conduct the PCA (37). Only those individuals in the 16 global populations ($n = 2317$) were included in the initial PCA used to generate the top 10 eigenvectors or principal components (PCs). The Australian data were then projected onto this 'genetic space' or background. PC1 largely reflects the difference between Africans and non-Africans, whereas PC2 separates East Asians from others. Subsequent PCs show dramatically lower eigenvalues and were not considered further. Most Australian individuals cluster with Europeans as expected. However, several show evidence of African or Asian ancestry. To remove outliers, we used the non-Australian European populations to calculate mean reference PC1 and PC2 scores. Any Australian individual more than 6 standard deviations from this mean was deemed to be an ancestry outlier and removed from further analysis. Any full siblings from offspring of excluded individuals were also excluded. In total, 2% of the original samples was excluded as outliers. Ancestry for the UK data was determined initially through self-reporting and was verified using PCA comparing it with the three reference populations from HapMap Phase 2 in a manner similar to the Australian data. To further investigate for effects of population stratification in the cleaned data set, we generated Q–Q plots in our discovery cohort for each of the four analysed traits.

Genomic imputation

Imputation for the Australian twins was performed using MACH2 with HapMap data obtained from people of northern and western European ancestry collected by the CEU (HapMap release 22) (38). As outlined earlier, the Australian data were part of a larger subset of samples typed on two different Illumina arrays (Hap610W and CNV370). Thus, imputation was based on a set of SNPs common to all Australian subsamples ($n = 274\,604$). Imputation was run in two stages. First, data from a set of representative Australian sample individuals were compared with the phased haplotype data from the HapMap CEU I + II data (release 22, build 36) to generate recombination and error maps. For the second stage, data were imputed for all individuals using the phased HapMap data as the reference panel and the recombination as well as error files generated in the first stage, for customization. After imputation, data for the 2.74 million HapMap SNPs were available for study. Locus-specific imputation was undertaken with reference to HapMap release 22 CEU, using IMPUTE version 2 for the UK sample (39).

Association analysis

Association analysis was performed using Merlin to account for the relatedness of individuals, based on the score test (–fastAssoc option) (40). For the Australian data, age and sex were fitted as covariates. In the (predominantly female) UK data, only age was fitted as a covariate. Traits were transformed to a standard normal distribution to ensure robustness to outliers.

Meta-analysis

The Australian and UK samples were meta-analysed based on the allelic effect estimated for each SNP from Merlin. The meta-analysis *P*-value was generated by comparing:

$$\frac{(\text{b\_AUS} \times \text{w\_AUS} + \text{b\_UK} \times \text{w\_UK})^2}{\text{w\_AUS} + \text{w\_UK}}$$

with a chi-squared 1 distribution, where b_AUS and b_UK are the allelic (additive) effects and w_AUS and w_UK are the inverse variance of allelic effects. Given that some of the trait standard deviations differed by sample site, reflecting the different photographic methods, traits were scaled so that the additive effects were on the scale of a standard normal distribution rather than the original scale of the trait.

Atoh7 expression in the mouse prenatal retina

Mouse embryos were staged, collected and processed for retinal sections as described previously (18). Section *in situ* hybridization with an *Atoh7* digoxygenin-labelled cRNA probe was performed as in Wallace and Raff (41). Anti-Pax6 (Covance Research Products, Princeton, NJ, USA) and anti-βgal co-labelling of $Atoh7^{tmGla?+}$ embryonic eye sections was performed as described previously (19,21). Digital imaging was performed with a Zeiss Axioplan microscope, colour camera and software, or a Zeiss LS M510 confocal microscope and software. Adobe Photoshop, version 7.0, was used for image formatting.

ATOH7 sequence screening

Genomic DNA samples of 12 individuals with optic nerve hypoplasia were obtained and screened. All patients had no definitive signs of septo-optic dysplasia. Primers for polymerase chain reaction (PCR) and sequencing were designed to cover coding and UTR, including intron–exon boundaries and promoter region 1000 bp 5′ upstream, using the ExonPrimer (Helmholtz Center Munich) and Primer3 programs. Primers were selected to produce amplification product sizes not to exceed 900 bp for optimal sequence output and reading. All amplicons overlapped at a minimum of 50 bp of sequence. Bidirectional sequencing (forward and reverse primers) was used. Supplementary Material, Table S1, displays the optimized primer sequences used.

PCR was performed and amplicons were visualized by agarose gel electrophoresis by standard procedures. PCR amplicon purification was conducted using the Quickstep™ 2 SOPE™ Resin (Edge BioSystems, Gaithersburg, MD, USA). BigDye™ Terminator 3.1 was used to perform sequencing reactions, and ABI3730 robotics was used to process the DNA fragments (Applied Biosystems Inc., Foster City, CA, USA). Base pair calls were made using the Sequencher 4.9™ software (Gene Codes, Ann Arbor, MI, USA). Generated sequences were aligned to a known reference genomic sequence (UCSC Genome Browser; Assembly hg18) and compared for sequence variation. The coding regions of ATOH7 were screened by direct sequencing in 90 unrelated Caucasian control individuals. The test of association was computed using Fisher's exact test, comparing 2 out of 24 case chromosomes with a non-synonymous mutation (8.3% carry a mutation) with 0 out of 180 control chromosomes (0% carry a mutation).

Cross-species homology of the non-synonymous variants was compared using the Ensembl database (Ensembl gene transcript, ENST00000373673).

WEB RESOURCES

EIGENSOFT: http://genepath.med.harvard.edu/~reich/Software.htm
HapMap: http://www.hapmap.org/
IMPUTE: https://mathgen.stats.ox.ac.uk/impute/impute.html
MACH2: http://www.sph.umich.edu/csg/abecasis/MACH/index.html
Merlin: http://www.sph.umich.edu/csg/abecasis/merlin/
Primer3: http://ihg2.helmholtz-muenchen.de/ihg/ExonPrimer.html
UCSC Genome Browser: http://genome.ucsc.edu
1000 Genomes: http://www.1000genomes.org
SIFT: (http://blocks.fhcrc.org/sift/SIFT.html)
PolyPhen prediction: (http://genetics.bwh.harvard.edu/pph).
Ensembl: http://www.ensembl.org/

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank our laboratory staff led by Gabriela Surdulescu, the staff from the genotyping facilities at the Wellcome Trust Sanger Institute for sample preparation, quality control and genotyping led by Leena Peltonen and Panos Deloukas; Le Centre National de Génotypage, France, led by Mark Lathrop, for genotyping; Duke University, NC, USA, led by David Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki, led by Aarno Palotie. Genotyping was also performed by the CIDR as part of an NEI/NIH project grant, and we are grateful to Dr Camilla Day and staff there.

Conflict of Interest statement. None declared.

FUNDING

The Australian Twin Registry is supported by an Australian National Health and Medical Research Council (NHMRC) Enabling Grant (2004–2009). We also thank the following organizations for their financial support: Clifford Craig Medical Research Trust, Ophthalmic Research Institute of Australia (ORIA), Glaucoma Australia, American Health Assistance Foundation (AHAF), Peggy and Leslie Cranbourne Foundation, Foundation for Children, NHMRC project grant (2005–2007), Jack Brockhoff Foundation, NEI Project Grant (2007–2010). Genotyping for part of the Australian sample was funded by an NHMRC Medical Genomics Grant. Genotyping for the remainder was performed by the National Institutes of Health (NIH) Center for Inherited Research (CIDR) as part of an NIH/National Eye Institute (NEI) grant 1RO1EY018246, and we are grateful to Dr Camilla Day and staff. Australian sample imputation analyses were carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003). We thank Dale R. Nyholt, Margaret J. Wright, Megan J. Campbell and Anthony Caracella for their assistance in processing the Australian genotyping data. D.A.M. is a recipient of the Pfizer Australia Senior Research Fellowship. S.M. is supported by an Australian NHMRC Career Development Award. We are also grateful to Jane MacKinnon, Shayne Brown, Lisa Kearns, Sandra Staffieri, Olivia Bigault, Colleen Wilkinson, Julie Barbour and James Elder, who assisted with clinical examinations. The authors of King's College London acknowledge funding from 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). The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. T.D.S. is an NIHR senior investigator and C.J.H. an NIHR senior research fellow. The project also received support from a Biotechnology and Biological Sciences Research Council (BBSRC) project grant (G20234). The authors acknowledge the funding and support of the NIH/NEI grant 1RO1EY018246, an NIH/Center for Inherited Diseases Research genotyping project grant (PI: T.L.Y.) and NIH/NEI grant 1R01EY013612 (PI: N.L.B.). T.L.Y. acknowledges funding from Research to Prevent Blindness, Inc. We would also like to thank all the research team who phenotyped the subjects, including S.H. Melissa Liew, MD, and Christine Smoliner. We are grateful to the volunteer twins who made available their time to research. Funding to pay the Open Access publication charges for this article was provided by The Wellcome Trust.

REFERENCES

1. Fitzgibbon, T. and Taylor, S.F. (1996) Retinotopy of the human retinal nerve fibre layer and optic nerve head. *J. Comp. Neurol.*, **375**, 238–251.
2. Jonas, J.B., Budde, W.M. and Panda-Jonas, S. (1999) Ophthalmoscopic evaluation of the optic nerve head. *Surv. Ophthalmol.*, **43**, 293–320.

3. O'Neill, E.C., Mackey, D.A., Connell, P.P., Hewitt, A.W., Danesh-Meyer, H.V. and Crowston, J.G. (2009) The optic nerve head in hereditary optic neuropathies. *Nat. Rev. Neurol.*, **5**, 277–287.
4. Hoffmann, E.M., Zangwill, L.M., Crowston, J.G. and Weinreb, R.N. (2007) Optic disk size and glaucoma. *Surv. Ophthalmol.*, **52**, 32–49.
5. Jonas, J.B., Schmidt, A.M., Muller-Bergh, J.A., Schlotzer-Schrehardt, U.M. and Naumann, G.O. (1992) Human optic nerve fiber count and optic disc size. *Invest. Ophthalmol. Vis. Sci.*, **33**, 2012–2018.
6. Klein, B.E., Klein, R. and Lee, K.E. (2004) Heritability of risk factors for primary open-angle glaucoma: the Beaver Dam Eye Study. *Invest. Ophthalmol. Vis. Sci.*, **45**, 59–62.
7. Chang, T.C., Congdon, N.G., Wojciechowski, R., Munoz, B., Gilbert, D., Chen, P., Friedman, D.S. and West, S.K. (2005) Determinants and heritability of intraocular pressure and cup-to-disc ratio in a defined older population. *Ophthalmology*, **112**, 1186–1191.
8. Hewitt, A.W., Poulsen, J.P., Alward, W.L., Bennett, S.L., Budde, W.M., Cooper, R.L., Craig, J.E., Fingert, J.H., Foster, P.J., Garway-Heath, D.F. *et al.* (2007) Heritable features of the optic disc: a novel twin method for determining genetic significance. *Invest. Ophthalmol. Vis. Sci.*, **48**, 2469–2475.
9. He, M., Liu, B., Huang, W., Zhang, J., Yin, Q., Zheng, Y., Wang, D. and Ge, J. (2008) Heritability of optic disc and cup measured by the Heidelberg retinal tomography in Chinese: the Guangzhou twin eye study. *Invest. Ophthalmol. Vis. Sci.*, **49**, 1350–1355.
10. Healey, P., Carbonaro, F., Taylor, B., Spector, T.D., Mitchell, P. and Hammond, C.J. (2008) The heritability of optic disc parameters: a classic twin study. *Invest. Ophthalmol. Vis. Sci.*, **49**, 77–80.
11. Beck, R.W., Servais, G.E. and Hayreh, S.S. (1987) Anterior ischemic optic neuropathy. IX. Cup-to-disc ratio and its role in pathogenesis. *Ophthalmology*, **94**, 1503–1508.
12. Lam, B.L., Morais, C.G. Jr and Pasol, J. (2008) Drusen of the optic disc. *Curr. Neurol. Neurosci. Rep.*, **8**, 404–408.
13. Ramos, C.V., Bellusci, C., Savini, G., Carbonelli, M., Berezovsky, A., Tamaki, C., Cinoto, R., Sacai, P.Y., Moraes-Filho, M.N., Miura, H.M. *et al.* (2009) Association of optic disc size with development and prognosis of Leber's hereditary optic neuropathy. *Invest. Ophthalmol. Vis. Sci.*, **50**, 1666–1674.
14. Roberts, M.D., Liang, Y., Sigal, I.A., Grimm, J., Reynaud, J., Bellezza, A., Burgoyne, C.F. and Downs, J.C. (2009) Correlation between local stress and strain and lamina cribrosa connective tissue volume fraction in normal monkey eyes. *Invest. Ophthalmol. Vis. Sci.*, **51**, 295–307.
15. Dixon, A.L., Liang, L., Moffatt, M.F., Chen, W., Heath, S., Wong, K.C., Taylor, J., Burnett, E., Gut, I., Farrall, M. *et al.* (2007) A genome-wide association study of global gene expression. *Nat. Genet.*, **39**, 1202–1207.
16. Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C. *et al.* (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*, **315**, 848–853.
17. Mu, X. and Klein, W.H. (2004) A gene regulatory hierarchy for retinal ganglion cell specification and differentiation. *Semin. Cell Dev. Biol.*, **15**, 115–123.
18. Brown, N.L., Kanekar, S., Vetter, M.L., Tucker, P.K., Gemza, D.L. and Glaser, T. (1998) Math5 encodes a murine basic helix–loop–helix transcription factor expressed during early stages of retinal neurogenesis. *Development*, **125**, 4821–4833.
19. Le, T.T., Wroblewski, E., Patel, S., Riesenberg, A.N. and Brown, N.L. (2006) Math5 is required for both early retinal neuron differentiation and cell cycle progression. *Dev. Biol.*, **295**, 764–778.
20. Trimarchi, J.M., Stadler, M.B. and Cepko, C.L. (2008) Individual retinal progenitor cells display extensive heterogeneity of gene expression. *PLoS One*, **3**, e1588.
21. Riesenberg, A.N., Le, T.T., Willardsen, M.I., Blackburn, D.C., Vetter, M.L. and Brown, N.L. (2009) Pax6 regulation of Math5 during mouse retinal neurogenesis. *Genesis*, **47**, 175–187.
22. Hutcheson, D.A., Hanson, M.I., Moore, K.B., Le, T.T., Brown, N.L. and Vetter, M.L. (2005) bHLH-dependent and -independent modes of Ath5 gene regulation during retinal development. *Development*, **132**, 829–839.
23. Willardsen, M.I., Suli, A., Pan, Y., Marsh-Armstrong, N., Chien, C.B., El Hodiri, H., Brown, N.L., Moore, K.B. and Vetter, M.L. (2009) Temporal regulation of Ath5 gene expression during eye development. *Dev. Biol.*, **326**, 471–481.
24. Jarman, A.P., Grell, E.H., Ackerman, L., Jan, L.Y. and Jan, Y.N. (1994) Atonal is the proneural gene for *Drosophila* photoreceptors. *Nature*, **369**, 398–400.
25. Brown, N.L., Dagenais, S.L., Chen, C.M. and Glaser, T. (2002) Molecular characterization and mapping of ATOH7, a human atonal homolog with a predicted role in retinal ganglion cell development. *Mamm. Genome*, **13**, 95–101.
26. Wee, R., Castrucci, A.M., Provencio, I., Gan, L. and Van Gelder, R.N. (2002) Loss of photic entrainment and altered free-running circadian rhythms in math5$^{-/-}$ mice. *J. Neurosci.*, **22**, 10427–10433.
27. Brzezinski, J.A., Brown, N.L., Tanikawa, A., Bush, R.A., Sieving, P.A., Vitaterna, M.H., Takahashi, J.S. and Glaser, T. (2005) Loss of circadian photoentrainment and abnormal retinal electrophysiology in Math5 mutant mice. *Invest. Ophthalmol. Vis. Sci.*, **46**, 2540–2551.
28. Azuma, N., Yamaguchi, Y., Handa, H., Tadokoro, K., Asaka, A., Kawase, E. and Yamada, M. (2003) Mutations of the PAX6 gene detected in patients with a variety of optic-nerve malformations. *Am. J. Hum. Genet.*, **72**, 1565–1570.
29. Hoehne, M., de Couet, H.G., Stuermer, C.A. and Fischbach, K.F. (2005) Loss- and gain-of-function analysis of the lipid raft proteins Reggie/Flotillin in *Drosophila*: they are posttranslationally regulated, and misexpression interferes with wing and eye development. *Mol. Cell Neurosci.*, **30**, 326–338.
30. Schaeren-Wiemers, N., Bonnet, A., Erb, M., Erne, B., Bartsch, U., Kern, F., Mantei, N., Sherman, D. and Suter, U. (2004) The raft-associated protein MAL is required for maintenance of proper axon–glia interactions in the central nervous system. *J. Cell Biol.*, **166**, 731–742.
31. Soranzo, N., Rivadeneira, F., Chinappen-Horsley, U., Malkina, I., Richards, J.B., Hammond, N., Stolk, L., Nica, A., Inouye, M., Hofman, A. *et al.* (2009) Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. *PLoS Genet.*, **5**, e1000445.
32. Beavis, W.D. (1994) The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the Corn and Sorghum Industry Research Conference, American Seed Trade Association, Washington, DC, pp.250–266.
33. Atchley, W.R., Wollenberg, K.R., Fitch, W.M., Terhalle, W. and Dress, A.W. (2000) Correlations among amino acid sites in bHLH protein domains: an information theoretic analysis. *Mol. Biol. Evol.*, **17**, 164–178.
34. Mackey, D.A., MacKinnon, J.R., Brown, S.A., Kearns, L.S., Ruddle, J.B., Sanfilippo, P.G., Sun, C., Hammond, C.J., Young, T.L., Martin, N.G. *et al.* (2009) Twins Eye Study in Tasmania (TEST): rationale and methodology to recruit and examine twins. *Twin. Res. Hum. Genet.*, **12**, 441–454.
35. Morgan, J.E., Sheen, N.J., North, R.V., Goyal, R., Morgan, S., Ansari, E. and Wild, J.M. (2005) Discrimination of glaucomatous optic neuropathy by digital stereoscopic analysis. *Ophthalmology*, **112**, 855–862.
36. McEvoy, B.P., Montgomery, G.W., McRae, A.F., Ripatti, S., Perola, M., Spector, T.D., Cherkas, L., Ahmadi, K.R., Boomsma, D., Willemsen, G. *et al.* (2009) Geographical structure and differential natural selection among North European populations. *Genome Res.*, **19**, 804–814.
37. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
38. International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.
39. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.*, **5**, e1000529.
40. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
41. Wallace, V.A. and Raff, M.C. (1999) A role for Sonic hedgehog in axon-to-astrocyte signalling in the rodent optic nerve. *Development*, **126**, 2901–2909.