nature genetics

Common variants near *ABCA1*, *AFAP1* and *GMDS* confer risk of primary open-angle glaucoma

Puya Gharahkhani[1,24], Kathryn P Burdon[2,3,24], Rhys Fogarty[2], Shiwani Sharma[2], Alex W Hewitt[4], Sarah Martin[2], Matthew H Law[1], Katie Cremin[5], Jessica N Cooke Bailey[6,7], Stephanie J Loomis[8], Louis R Pasquale[8,9], Jonathan L Haines[6,7], Michael A Hauser[10,11], Ananth C Viswanathan[12], Peter McGuffin[13], Fotis Topouzis[14], Paul J Foster[12], Stuart L Graham[15], Robert J Casson[16], Mark Chehade[16], Andrew J White[17], Tiger Zhou[2], Emmanuelle Souzeau[2], John Landers[2], Jude T Fitzgerald[2], Sonja Klebe[18], Jonathan B Ruddle[4], Ivan Goldberg[19], Paul R Healey[17], Wellcome Trust Case Control Consortium 2[20], NEIGHBORHOOD Consortium[20], Richard A Mills[2], Jie Jin Wang[17], Grant W Montgomery[1], Nicholas G Martin[1], Graham Radford-Smith[1,21], David C Whiteman[1], Matthew A Brown[5], Janey L Wiggs[8], David A Mackey[3,22], Paul Mitchell[17], Stuart MacGregor[1,25] & Jamie E Craig[2,23,25]

Primary open-angle glaucoma (POAG) is a major cause of irreversible blindness worldwide. We performed a genome-wide association study in an Australian discovery cohort comprising 1,155 cases with advanced POAG and 1,992 controls. We investigated the association of the top SNPs from the discovery stage in two Australian replication cohorts (932 cases and 6,862 controls total) and two US replication cohorts (2,616 cases and 2,634 controls total). Meta-analysis of all cohorts identified three loci newly associated with development of POAG. These loci are located upstream of *ABCA1* (rs2472493[G], odds ratio (OR) = 1.31, $P = 2.1 \times 10^{-19}$), within *AFAP1* (rs4619890[G], OR = 1.20, $P = 7.0 \times 10^{-10}$) and within *GMDS* (rs11969985[G], OR = 1.31, $P = 7.7 \times 10^{-10}$). Using RT-PCR and immunolabeling, we show that these genes are expressed within human retina, optic nerve and trabecular meshwork and that ABCA1 and AFAP1 are also expressed in retinal ganglion cells.

POAG, the most common subtype of glaucoma, is characterized by a progressive loss of peripheral vision, but cases may remain undiagnosed until central vision is affected[1,2]. The etiology and pathogenesis of POAG are poorly understood. Linkage studies, candidate gene studies and genome-wide association studies (GWAS) have identified several loci reproducibly associated with development of POAG[3–7]. Our previous GWAS of advanced POAG identified two loci at *TMCO1* and *CDKN2B-AS1* (ref. 6), with studies of non-advanced POAG also having implicated *CAV1* (ref. 5), *SIX6* and a region at 8q22 (ref. 7). Here we used a three-stage GWAS to identify additional genetic loci associated with POAG in participants of European descent.

The stage 1 discovery cohort comprised 1,155 cases with advanced glaucoma from the Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG) and 1,992 controls genotyped on Illumina Omni1M or OmniExpress arrays (**Supplementary Note** and **Supplementary Table 1**). We combined and cleaned the genotype data from cases and controls and used 569,249 SNPs as the base of imputation against the 1000 Genomes phase 1 European-ethnicity data set. We successfully imputed 7,594,768 SNPs with minor allele frequency (MAF) >0.01 and imputation quality score >0.8.

[1]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [2]Department of Ophthalmology, Flinders University, Adelaide, South Australia, Australia. [3]Menzies Research Institute Tasmania, University of Tasmania, Hobart, Tasmania, Australia. [4]Centre for Eye Research Australia (CERA), University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia. [5]University of Queensland Diamantina Institute, Brisbane, Queensland, Australia. [6]Center for Human Genetics Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA. [7]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, USA. [8]Department of Ophthalmology, Harvard Medical School and Massachusetts Eye and Ear Infirmary, Boston, Massachusetts, USA. [9]Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [10]Department of Ophthalmology, Duke University Medical Center, Durham, North Carolina, USA. [11]Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA. [12]National Institute for Health Research (NIHR) Biomedical Research Centre, Moorfields Eye Hospital National Health Service (NHS) Foundation Trust and University College London (UCL) Institute of Ophthalmology, London, UK. [13]Medical Research Council (MRC) Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College, De Crespigny Park, London, UK. [14]Department of Ophthalmology, School of Medicine, Aristotle University of Thessaloniki, AHEPA Hospital, Thessaloniki, Greece. [15]Ophthalmology and Vision Science, Macquarie University, Sydney, New South Wales, Australia. [16]South Australian Institute of Ophthalmology, University of Adelaide, Adelaide, South Australia, Australia. [17]Centre for Vision Research, Westmead Millennium Institute, University of Sydney, Westmead, New South Wales, Australia. [18]Department of Anatomical Pathology, Flinders University, Flinders Medical Centre, Adelaide, South Australia, Australia. [19]Department of Ophthalmology, University of Sydney, Sydney Eye Hospital, Sydney, New South Wales, Australia. [20]Full lists of members and affiliations are provided in the **Supplementary Note**. [21]School of Medicine, University of Queensland, Herston Campus, Brisbane, Queensland, Australia. [22]Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western Australia, Perth, Western Australia, Australia. [23]South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia. [24]These authors contributed equally to this work. [25]These authors jointly directed this work. Correspondence should be addressed to J.E.C. (jamie.craig@flinders.edu.au) or P.G. (puya.gharahkhani@qimrberghofer.edu.au).

Received 9 April; accepted 4 August; published online 31 August 2014; doi:10.1038/ng.3079

We performed association analysis using an additive model adjusted for sex and six principal components. We corrected the *P* values from the association analysis for the estimated genomic inflation factor, λ, of 1.06 (the quantile-quantile plot is shown in **Supplementary Fig. 1**).

Table 1 Association results for the top SNPs in previously unreported regions with $P < 1 \times 10^{-7}$ in the discovery cohort

Chr.	SNP	Position[a]	Gene	Risk allele	*P*[b]	OR	s.e.m.	Frequency[c]
9	rs2472493	107695848	*ABCA1*[d]	G	2.0×10^{-10}	1.43	0.05	0.51/0.43
11	rs11827818	120198728	*ARHGEF12*[d]	G	9.2×10^{-9}	1.52	0.07	0.20/0.14
6	rs114096562	1984385	*GMDS*	A	7.0×10^{-8}	1.55	0.08	0.88/0.83

[a]Position in build 37. [b]*P* corrected for the genomic inflation factor ($\lambda = 1.06$). [c]Allele frequency in cases/controls. [d]The corresponding SNP is not in the indicated gene; instead, characterized genes located near these SNPs are shown. Chr., chromosome.

The stage 1 association results across the genome are shown in **Supplementary Figure 2**, and the association results for all SNPs with $P < 1 \times 10^{-7}$ are shown in **Supplementary Table 2**. Two previously unreported regions reached genome-wide significance ($P < 5 \times 10^{-8}$) in the stage 1 discovery cohort, with a further previously unreported region showing association at close to genome-wide significance (**Table 1**). The top newly associated SNPs were rs2472493[G] upstream of *ABCA1* (encoding ATP-binding cassette, subfamily A, member 1) on chromosome 9 (OR = 1.43, $P = 2.0 \times 10^{-10}$), rs11827818[G] close to *ARHGEF12* (encoding Rho guanine nucleotide exchange factor 12) (OR = 1.52, $P = 9.2 \times 10^{-9}$) on chromosome 11 and rs114096562[A] in *GMDS* (encoding GDP-mannose 4,6-dehydratase) (OR = 1.55, $P = 7.0 \times 10^{-8}$) on chromosome 6. The regional association results for these three SNPs are shown in **Figure 1**. We also performed the analysis after removing controls affected by other diseases (**Supplementary Note**) and found that the effect sizes were similar (**Supplementary Table 3**).

We then investigated the associations of top SNPs in the discovery cohort in a stage 2 set comprising two Australian replication data sets (the ANZRAG and Blue Mountains Eye Study (BMES) data sets, totaling 932 cases and 6,862 controls; **Supplementary Note** and **Supplementary Table 1**). All replication cohort participants were of European descent. To make maximum valid use of our cohorts, for replication we focused on SNPs directly genotyped on the Illumina Human610 and Human670 arrays; we used proxy genotyped SNPs

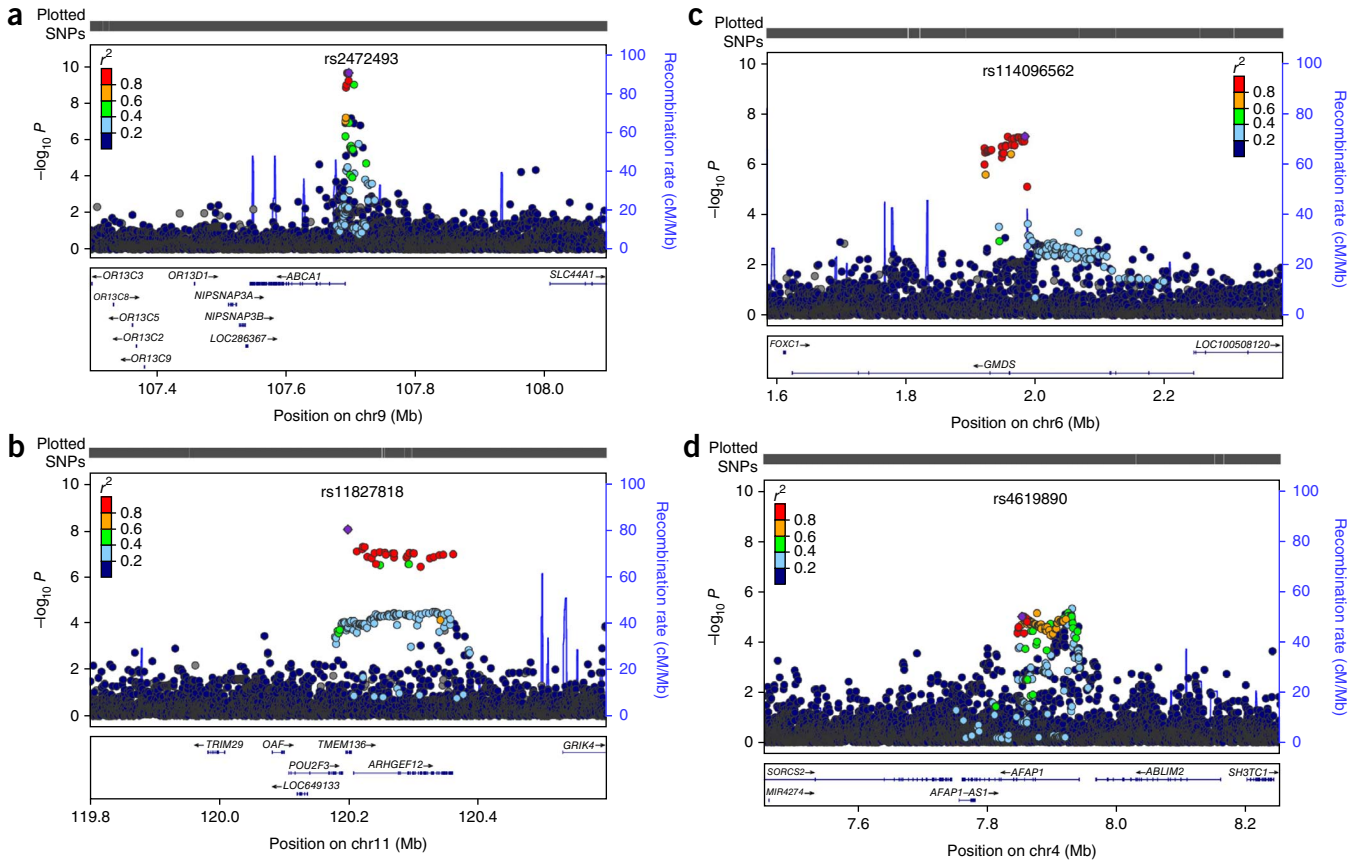

Figure 1 Association results for the regions reaching genome-wide significance. These plots show the regional association (using logistic regression with sex and the first six principal components fitted as covariates) and recombination rates for the top SNPs in the discovery data set (1,155 cases with advanced POAG and 1,992 controls). In each plot, the solid diamond indicates the top-ranked SNP in the region based on two-sided *P* values. The colored box at the right or left corner of each plot indicates the pairwise correlation (r^2) between the top SNP and the other SNPs in the region. The blue spikes show the estimated recombination rates. The box underneath each plot shows the gene annotations in the region. Each plot was created using LocusZoom (http://csg.sph.umich.edu/locuszoom/) for the top-ranked SNP in each region with a 400-kb region surrounding it. (**a**) The top-ranked SNP for this plot is rs2472493 on chromosome 9 upstream of *ABCA1* with $P = 2.0 \times 10^{-10}$. (**b**) The top-ranked SNP for this plot is rs11827818 on chromosome 11 near *ARHGEF12* with $P = 9.2 \times 10^{-9}$. (**c**) The top-ranked SNP for this plot is rs114096562 on chromosome 6 in *GMDS* with $P = 7.0 \times 10^{-8}$. (**d**) This plot is centered on rs4619890 on chromosome 4 in *AFAP1* with $P = 9.7 \times 10^{-6}$. This SNP clearly reached genome-wide significance ($P = 7.0 \times 10^{-10}$) in the meta-analysis of the results between the discovery and replication cohorts.

Table 2 Association and meta-analysis of the discovery and replication cohorts for the top-ranked loci

Chr.	SNP	Position[a]	A1[b]	A2	Gene	ANZRAG (discovery)		ANZRAG (replication)		BMES		NEIGHBOR		MEEI		Meta-analysis			
						OR	*P*[c]	OR	*P*	OR	*P*[c]	OR	*P*	OR	*P*	OR	*P*	Het $I^{2,d}$	Het *P*[e]
3	rs2710323	52815905	T	C	*ITIH1*	1.25	9.16×10^{-5}	1.14	0.005	1.44	0.01	1.06	0.25	0.87	0.31	1.14	4.53×10^{-6}	64.6	0.02
4	rs4619890	7853160	G	A	*AFAP1*	1.26	9.76×10^{-6}	1.20	0.0004	1.07	0.62	1.14	0.008	1.13	0.38	1.20	7.03×10^{-10}	0	0.57
4	rs4478172	7902003	C	A	*AFAP1*	1.29	2.73×10^{-5}	1.15	0.02	1.21	0.26	1.16	0.005	1.11	0.47	1.19	2.19×10^{-8}	0	0.64
6	rs11969985	1922907	G	A	*GMDS*	1.53	3.18×10^{-7}	1.23	0.009	0.92	0.71	1.28	0.001	1.28	0.24	1.31	7.70×10^{-10}	46.4	0.11
6	rs2761233	1949101	T	C	*GMDS*	1.53	3.35×10^{-7}	1.19	0.02	0.94	0.79	1.28	0.001	1.28	0.23	1.29	2.17×10^{-9}	48.5	0.10
9	rs2472493	107695848	G	A	*ABCA1*[f]	1.43	2.08×10^{-10}	1.26	4.84×10^{-6}	1.44	0.01	1.26	7.05×10^{-6}	0.99	0.89	1.31	2.16×10^{-19}	53.6	0.07
11	rs2276035	120346360	A	G	*ARHGEF12*	1.47	1.13×10^{-7}	1.08	0.29	1.05	0.81	1.09	0.15	NA[g]	NA[g]	1.18	7.83×10^{-6}	77.1	0.004

Association results for three loci that reached genome-wide significance in the discovery cohort, as well as other top-ranked loci showing replication. Proxy SNPs are presented where imputed data were not available for the replication cohorts.
[a]Position in build 37. [b]Effect allele in all the cohorts. [c]*P* value corrected for the genomic inflation factor. [d]I^2 statistic measuring heterogeneity (Het) on a scale of 0% to 100%. [e]*P* values for the heterogeneity test. [f]The corresponding SNP is not in the indicated gene; instead, a characterized gene located near the SNP is shown. [g]rs2276035 SNP was not available (NA) in the MEEI cohort. Chr., chromosome.

where imputed data were not available for the replication cohorts (Online Methods).

Examining all autosomal SNPs with $P < 1 \times 10^{-4}$ in stage 1 (24 SNPs with the best *P* values were used as the lead SNPs; **Supplementary Table 4**), four regions showed nominal evidence ($P < 0.05$ for seven SNPs in or near *ABCA1*, *GMDS*, *ITIH1* and *AFAP1*) for replication in the ANZRAG replication samples (**Supplementary Table 4**). When we combined stages 1 and 2, SNPs near *ABCA1* and in *AFAP1* exceeded genome-wide significance ($P < 5 \times 10^{-8}$ for rs2472493 and rs4619890, respectively) with consistent effect sizes and directions of effects among the cohorts (**Table 2** and **Supplementary Table 4**).

In the stage 3 replication, we examined the newly identified top SNPs from stage 2 in data available from two additional replication cohorts (**Supplementary Note** and **Supplementary Table 1**): the National Eye Institute Glaucoma Human Genetics Collaboration (NEIGHBOR) and the Massachusetts Eye and Ear Infirmary (MEEI) (totaling 2,616 cases and 2,634 controls). We also performed a meta-analysis of the results for these SNPs between all cohorts (the discovery stage and all four replication cohorts) using the effect sizes and their standard errors. In the meta-analysis results, SNPs in or near *ABCA1*, *AFAP1* and *GMDS* clearly reached genome-wide significance ($P < 5 \times 10^{-8}$) (**Table 2**).

The top SNP within *ARHGEF12* (rs2276035) did not reach the significance level ($P < 5 \times 10^{-8}$) in our standard meta-analysis (**Table 2**), primarily because of heterogeneity between stage 1 and stages 2 and 3. This heterogeneity could be explained by the difference in glaucoma status in these cohorts, the 'winner's curse' effect that leads to inflated OR estimates in GWAS or chance. The top SNP within *ITIH1* (rs2710323) was not genome-wide significant in our meta-analysis (**Table 2**).

At each of the newly discovered loci, the effect size was larger in the discovery cohort than in the replication cohorts (**Table 2**). The discovery cohort comprises cases with advanced POAG only, whereas the replication cohorts contained cases with POAG representing a range of disease severity. One cannot directly infer, however, that the true effect size is largest in advanced POAG. A winner's curse effect in the ANZRAG discovery cohort would inflate the OR estimates. Furthermore, there may have been greater diagnostic certainty in the cases with advanced POAG. To investigate further whether the newly discovered loci conferred higher risk in advanced compared to non-advanced POAG, we performed a subanalysis on the ANZRAG replication cohort. We found no consistent difference between the ORs for the cases with non-advanced ($n = 605$) and advanced ($n = 220$) POAG separately (**Supplementary Table 5**). This subanalysis, together with the significant results in the replication cohorts taken alone, suggest that the newly discovered loci in this study are associated with POAG in general (and not advanced POAG only), indicating the generalizability of our findings.

Intraocular pressure (IOP) was not a criterion in the definition of POAG in this study because patients with POAG may have either normal or elevated IOP[8]. Thus, the new loci identified in this study are associated with POAG in general regardless of IOP levels. However, we had peak IOP measures available for 1,039 of the 1,155 cases in the ANZRAG discovery cohort. Of these cases, 330 (31.8%) had normal-tension glaucoma (NTG) (IOP ≤ 21 mm Hg) and 709 (68.2%) had high-tension glaucoma (HTG) (IOP >21 mm Hg). We investigated the association of the new loci identified in this study in 330 cases with NTG and 709 cases with HTG compared to 1,992 population controls in the discovery cohort (**Supplementary Table 6**). The direction and magnitude of effects of the risk alleles were similar for NTG, HTG and all POAG (**Table 2** and **Supplementary Table 6**). However, the

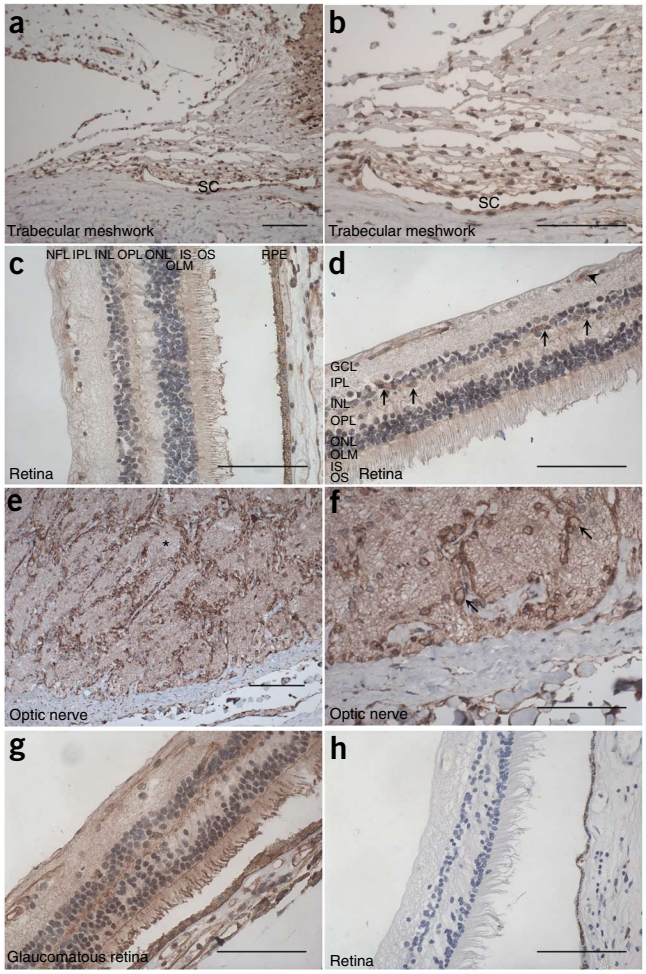

Figure 2 Distribution of the ABCA1 protein in human ocular tissues. (**a**–**f**) Sections of a normal human eye were immunolabeled with the ABCA1-specific antibody (brown) and counterstained with haematoxylin to visualize nuclei (blue). Positive immunolabeling was detected in the trabecular meshwork (**a**,**b**), throughout the retina (**c**,**d**) and in the optic nerve (**e**,**f**). In the retina (**c**), comparatively pronounced ABCA1 immunolabeling was observed at the tips of photoreceptors and in the outer limiting membrane (OLM), outer plexiform layer (OPL) and nerve fiber layer (NFL). (**d**) Labeling was also pronounced in some cells in the inner nuclear layer (INL, arrows), in retinal ganglion cells in the ganglion cell layer (GCL, arrowhead) and in the retinal blood vessel wall (not shown). In the optic nerve (**e**,**f**), the protein was distributed in the nerve fiber bundles (**e**, asterisk) and at the cell boundary of astrocytes in the glial columns (**f**, arrows). (**g**,**h**) In sections of a glaucomatous eye (data not shown), including in the retina (**g**), similar distribution of the protein to that in the normal eye was observed. The experiment was repeated once for reproducibility. (**h**) Section hybridized with the secondary detection reagent alone as a negative control. SC, Schlemm's canal; RPE, retinal pigment epithelium; OS, outer segment; IS, inner segment; ONL, outer nuclear layer; IPL, inner plexiform layer. Scale bars, 100 μm.

analysis for NTG and HTG was less powerful compared to that for POAG because of the smaller sample sizes of the subgroups.

None of our newly identified POAG loci overlapped with the previously published loci associated with POAG subphenotypes, including IOP and vertical cup-disk ratio[9–11]. We also investigated the association of the new loci identified in this study with peak measured IOP in 1,039 cases with POAG with available data in the ANZRAG discovery cohort. The new loci were not associated with peak IOP in the ANZRAG discovery cohort (**Supplementary Table 7**), although the *ABCA1* SNP showed a trend toward significance ($P = 0.0675$, two-sided test). The *ABCA1* glaucoma risk–increasing allele acts in the expected direction on IOP (the allele increases IOP), resulting in a P value of 0.034 in a one-sided test. Larger sample sizes and further meta-analyses of multiple studies will unambiguously determine whether the new loci identified in this study are associated with subphenotypes such as IOP.

We also investigated previously reported GWAS hits identified in other studies[5–7] in the meta-analysis of results between our discovery and replication cohorts (**Supplementary Table 8**). The *TMCO1*, *CDKN2B-AS1* and *SIX6* loci were clearly genome-wide significant ($P < 5 \times 10^{-8}$), whereas *CAV1-CAV2* and the locus on chromosome 8 were associated with POAG but not at a genome-wide significance level. SNP rs11669977 at *NTF4* was not associated with POAG in our meta-analysis.

We used ENCODE project data[12] and the Genevar database[13] (expression quantitative trait locus (eQTL) database) to predict the possible functional effects of the top SNPs identified in this study. The top SNP rs2472493 located upstream of *ABCA1* is an eQTL in lymphoblastoid cell lines (Genevar database) and may alter the sequence of motifs for proteins such as FOXJ2 and SIX5 (HaploReg v2)[14]. One of the SNPs in high linkage disequilibrium (LD, $r^2 > 0.8$) with the top SNP near *ABCA1* (rs2472494) alters the regulatory motif for binding of PAX6 (HaploReg v2). *PAX6* is an established master control gene in eye development[15]. A SNP (rs28495790) in high LD ($r^2 > 0.8$) with the best SNP in *AFAP1* (rs4619890) is likely to affect the binding of proteins (score 2b in RegulomeDB)[16] such as CTCF and RAD21 in a variety of cell lines, including WERI-Rb-1 (retinoblastoma). rs28495790 alters the sequence of regulatory motifs for binding of several proteins, including PAX6 (HaploReg v2). This may suggest a regulatory role for this SNP in gene expression in a pathway similar to that of rs2472494 near *ABCA1*. In *GMDS*, rs3046543 (in high LD, $r^2 = 0.8$, with the top imputed SNP rs114096562) alters the sequence of the regulatory motif for binding of SIX6; *SIX6* variants confer glaucoma risk[7]. SNPs close to *SIX6* also clearly reached genome-wide significance in our meta-analysis (top SNP rs10483727[T], OR = 1.32, $P = 1.56 \times 10^{-17}$). These data suggest that the top SNPs identified in this study may have important regulatory roles.

ABCA1 is a membrane-bound receptor involved in phospholipid and cholesterol efflux from cells. In monkey retinas, *ABCA1* is expressed in retinal ganglion cells[17], the cells that undergo apoptosis in glaucoma. We analyzed the expression of *ABCA1* mRNA in human ocular tissues by RT-PCR and found that the iris, ciliary body, retina, optic nerve head, optic nerve and trabecular meshwork cell lines derived from normal and glaucomatous eyes expressed the main transcript that encodes the full-length protein (**Supplementary Fig. 3a**). We also detected an alternative transcript in the ocular tissues (**Supplementary Fig. 3a**) that had unknown function[18,19]. Immunolabeling of sections of normal human eye with ABCA1-specific antibody (**Supplementary Fig. 4**) showed a distribution of the protein in the trabecular meshwork, all layers of the retina (including retinal ganglion cells) and the optic nerve (**Fig. 2**). We observed similar ABCA1 labeling in a glaucomatous eye, including in the layers of the retina (**Fig. 2g**). *ABCA1* has been reported to regulate neuroinflammation and neurodegeneration through coordinated activity in various cell types in mouse brains[20], and it may be involved in glaucoma through a similar function in the retina.

AFAP1 encodes a protein that binds to actin filaments and allows their crosslinking[21,22]. Actin cytoskeleton-modulating signals have been shown to be involved in the regulation of aqueous outflow and

intraocular pressure[23–25], which are important parts of glaucoma pathogenesis. *AFAP1* encodes two isoforms, the neuronal cell–specific A isoform and the ubiquitously expressed B isoform. By RT-PCR, we detected expression of both the A and B isoforms in human retina (**Supplementary Fig. 3c**) and expression of the B isoform in other ocular tissues, including the iris, ciliary body, lens, optic nerve and optic nerve head, as well as in cultured trabecular meshwork cells (**Supplementary Fig. 3b**). Consistent with the mRNA expression data, we observed AFAP1-positive immunolabeling in the trabecular meshwork, retina (including retinal ganglion cells) and optic nerve of normal human eye (**Supplementary Fig. 5**) using AFAP1-specific antibody (**Supplementary Fig. 6**). We observed similar AFAP1 labeling in a glaucomatous eye, including in the retina (**Supplementary Fig. 5g,h**). These data indicate that the function of AFAP1 in the trabecular meshwork and retina may be relevant in the pathogenesis of glaucoma.

GMDS encodes a protein that is required for the first step in *de novo* synthesis of fucose[26]. Fucose is required for diverse biological functions, including growth factor receptor signaling[27]. Several studies have suggested effects of growth factors on the development of glaucoma[23,28–32]. *GMDS* expresses two variant transcripts, 1 and 2. We detected expression of the variant 1 transcript in human ocular tissues and cultured trabecular meshwork cells by RT-PCR (**Supplementary Fig. 3d**), which indicates ubiquitous expression of the gene in the eye.

In this study we identified three new risk loci for POAG and highlighted related candidate genes and pathways that might be involved in developing POAG. These new loci, in addition to the previously known risk loci, will improve risk profiling for glaucoma, with better opportunities for the management of high-risk individuals. Currently, many cases of glaucoma remain undiagnosed until severe visual loss occurs; early detection and treatment can slow disease progression and prevent blindness[33]. Further dissection of these new POAG risk loci will likely lead to insights into the etiology of this common, irreversible cause of blindness.

URLs. PLINK, http://pngu.mgh.harvard.edu/~purcell/plink; R Project, http://www.r-project.org; LocusZoom, http://csg.sph.umich.edu/locuszoom; Ensembl, http://www.ensembl.org/index.html; NCBI, http://www.ncbi.nlm.nih.gov; UCSC Genome Bioinformatics, http://genome.ucsc.edu; GeneCards, http://www.genecards.org; UniprotKB, http://www.uniprot.org; ENCODE project, http://www.genome.gov/10005107; RegulomeDB, http://regulome.stanford.edu; HaploReg v2, http://www.broadinstitute.org/mammals/haploreg/haploreg.php; Genevar, http://www.sanger.ac.uk/resources/software/genevar.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG): support for recruitment of ANZRAG was provided by the Royal Australian and New Zealand College of Ophthalmology (RANZCO) Eye Foundation. Genotyping was funded by the National Health and Medical Research Council (NHMRC) of Australia (#535074 and #1023911). This work was also supported by funding from NHMRC #1031362 awarded to J.E.C., NHMRC #1037838 awarded to A.W.H., NHMRC #1048037 awarded to S.L.G., NHMRC #1009844 awarded to R.J.C. and I.G., NHMRC #1031920 and an Alcon Research Institute grant awarded to D.A.M., an Allergan Unrestricted grant awarded to A.J.W., the BrightFocus Foundation and a Ramaciotti Establishment Grant. The authors acknowledge the support of B. Usher-Ridge in patient recruitment and data collection and P. Danoy and J. Hadler for genotyping.

Controls for the ANZRAG discovery cohort were drawn from the Australian Cancer Study, the Study of Digestive Health and a study of inflammatory bowel diseases. The Australian Cancer Study was supported by the Queensland Cancer Fund and the NHMRC of Australia (program number 199600, awarded to D.C.W., A.C. Green, N.K. Hayward, P.G. Parsons, D.M. Purdie and P.M. Webb, and program number 552429, awarded to D.C.W.). The Study of Digestive Health was supported by grant number 5 R01 CA 001833 from the US National Cancer Institute (awarded to D.C.W.).

The Barrett's and Esophageal Adenocarcinoma Genetic Susceptibility Study (BEAGESS) sponsored the genotyping of cases with esophageal cancer and Barrett's esophagus, which were used as unscreened controls in the ANZRAG discovery cohort. BEAGESS was funded by grant R01 CA136725 from the US National Cancer Institute.

Genotyping for part of the Australian twin control sample included in the ANZRAG replication cohort was funded by an NHMRC Medical Genomics Grant. Genotyping for the remainder of twin controls was performed by the US National Institutes of Health (NIH) Center for Inherited Research (CIDR) as part of NIH/National Eye Institute (NEI) grant 1RO1EY018246, and we are grateful to C. Day and staff. We acknowledge with appreciation all women who participated in the QIMR endometriosis study. We thank Endometriosis Associations for supporting study recruitment. We thank S. Nicolaides and the Queensland Medical Laboratory for *pro bono* collection and delivery of blood samples and other pathology services for assistance with blood collection. The QIMR twin and endometriosis studies were supported by grants from the NHMRC of Australia (241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496739, 552485 and 552498), the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne) and donations from N. and S. Hawkins. We thank M.J. Wright, M.J. Campbell, A. Caracella, S. Gordon, D.R. Nyholt, A.K. Henders, B. Haddon, D. Smyth, H. Beeby, O. Zheng and B. Chapman for their input into project management, databases, sample processing and genotyping. We are grateful to the many research assistants and interviewers for assistance with the studies contributing to the QIMR twin collection.

Blue Mountains Eye Study (BMES): BMES was supported by the NHMRC, Canberra Australia (974159, 211069, 457349, 512423, 475604 and 529912), the Centre for Clinical Research Excellence in Translational Clinical Research in Eye Diseases, NHMRC Senior Research Fellowships (358702 and 632909 to J.J.W.) and the Wellcome Trust, UK, as part of Wellcome Trust Case Control Consortium 2 (A.C.V., P. McGuffin, P. Mitchell, F.T. and P.J.F.) for genotyping costs of the entire BMES population (085475B08Z, 08547508Z, 076113). P.J.F. is also supported by Medical Research Council (MRC) G0401527, Research Into Ageing (Ref 262) and NIHR (UK) Biomedical Research Centre at Moorfields Eye Hospital and University College London Institute of Ophthalmology (BRC2_009) funds.

The BMES acknowledges E. Rochtchina from the Centre for Vision Research, Department of Ophthalmology and Westmead Millennium Institute University of Sydney, J. Attia, R. Scott and E.G. Holliday from the University of Newcastle, J. Xie and P.N. Baird from the Centre for Eye Research Australia, University of Melbourne, M.T. Inouye, Medical Systems Biology, Department of Pathology and Department of Microbiology and Immunology, University of Melbourne and X. Sim, National University of Singapore.

NEI Glaucoma Human Genetics Collaboration (NEIGHBOR): genotyping services for the NEIGHBOR study were provided by the CIDR and were supported by the NEI through grant HG005259-01 (J.L.W.). Additionally, CIDR is funded through a federal contract from the NIH to The Johns Hopkins University, contract number HHSN268200782096C. Collecting and processing samples for the NEIGHBOR data set was supported by the NEI through American Recovery and Reinvestment Act (ARRA) grants 3R01EY015872-05S1 (J.L.W.) and 3R01EY019126-02S1 (M.A.H.). Genotype imputation and meta-analysis were supported by EY022305 (J.L.W.). Funding for the collection of cases and controls was provided by the following NIH grants: EY015543 (R.R. Allingham); EY006827 (D. Gaasterland); HL73042, HL073389, EY13315, EY023646 (M.A.H.); CA87969, CA49449, CA55075 (J.H. Kang); EY009149 (P.R. Lichter); HG004608 (C. McCarty); EY008208 (F.A. Medeiros); EY015473 (L.R.P.); EY012118 (M. Pericak-Vance); EY015682 (A. Realini); EY011671, EY09580 (J.E. Richards); EY013178 (J.S. Schuman); RR015574, EY015872, EY010886, EY009847, EY014104 (J.L.W.); EY011008, EY144428, EY144448 and EY18660 (K. Zhang). J.L.W. and L.R.P. are also supported by the Harvard Glaucoma Center for Excellence and the Margolis Fund. Y. Liu is supported by the Glaucoma Research Foundation, the American Health Assistance Foundation and the Glaucoma Foundation. J.L.W., L.R.P., D.C. Musch and J.E. Richards are supported by Research to Prevent Blindness. J.N.C.B. is supported by NIH T32 EY007157 (CWRU) and T32 EY21453-2 (VUMC).

MEEI case-control sample: genotyping for the Massachusetts Eye and Ear Infirmary (MEEI) case-control sample was performed at the Broad Institute of MIT and Harvard with funding support from the NIH GEI (Gene Environment Initiative) (U01HG04424 and U01HG004728). The GENEVA Coordinating Center (U01HG004446) assisted with genotype cleaning. Imputation was supported by NIH EY022305. Collection of cases and controls was supported by NIH EY015872.

Support for molecular analysis of the associated genes was provided by the Ophthalmic Research Institute of Australia. The authors acknowledge the support of M. Philpott in collection of cadaveric human eye tissues and N. Mabarrack for initial optimization of the antibody to AFAP.

S. MacGregor is supported by Australian Research Council (ARC) and NHMRC Fellowships. G.W.M., M.A.B., K.P.B., D.C.W. and J.E.C. are supported by Australian NHMRC Fellowships. D.C.W. was funded by the ARC and G.R.-S. was funded by NHMRC during the period of this study. The authors acknowledge C. Abbot (Alcon Research Ltd.) for providing normal and glaucomatous trabecular meshwork cell lines, NTM-5 and GTM-3, respectively, as a kind gift.

AUTHOR CONTRIBUTIONS

K.P.B., S. MacGregor and J.E.C. were involved in designing the study. A.W.H., K.C., L.R.P., M.A.H., A.C.V., P. McGuffin, F.T., P.J.F., J.J.W., G.W.M., N.G.M., G.R.-S., D.C.W., M.A.B., J.L.W., D.A.M., P. Mitchell and J.E.C. were involved in participant recruitment, sample collection or genotyping. Analysis was performed by P.G., R.F., K.P.B., S.S., M.H.L., J.N.C.B., S.J.L., L.R.P., J.L.H., J.L.W. and S. MacGregor. Designing and conducting the laboratory experiments were performed by K.P.B., S.S., S. Martin and R.F. Clinician assessments were performed by S.L.G., R.J.C., M.C., A.J.W., T.Z., E.S., J.L., J.T.F., S.K., J.B.R., I.G., P.R.H., R.A.M., D.A.M. and J.E.C. The initial draft was written by P.G., K.P.B., S.S. and S. MacGregor.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Quigley, H.A. & Broman, A.T. The number of people with glaucoma worldwide in 2010 and 2020. *Br. J. Ophthalmol.* **90**, 262–267 (2006).
2. Casson, R.J., Chidlow, G., Wood, J.P., Crowston, J.G. & Goldberg, I. Definition of glaucoma: clinical and experimental concepts. *Clin. Experiment. Ophthalmol.* **40**, 341–349 (2012).
3. Stone, E.M. *et al.* Identification of a gene that causes primary open angle glaucoma. *Science* **275**, 668–670 (1997).
4. Pasutto, F. *et al.* Heterozygous *NTF4* mutations impairing neurotrophin-4 signaling in patients with primary open-angle glaucoma. *Am. J. Hum. Genet.* **85**, 447–456 (2009).
5. Thorleifsson, G. *et al.* Common variants near *CAV1* and *CAV2* are associated with primary open-angle glaucoma. *Nat. Genet.* **42**, 906–909 (2010).
6. Burdon, K.P. *et al.* Genome-wide association study identifies susceptibility loci for open angle glaucoma at *TMCO1* and *CDKN2B–AS1*. *Nat. Genet.* **43**, 574–578 (2011).
7. Wiggs, J.L. *et al.* Common variants at 9p21 and 8q22 are associated with increased susceptibility to optic nerve degeneration in glaucoma. *PLoS Genet.* **8**, e1002654 (2012).
8. Quigley, H.A. Open-angle glaucoma. *N. Engl. J. Med.* **328**, 1097–1106 (1993).
9. Ozel, A.B. *et al.* Genome-wide association study and meta-analysis of intraocular pressure. *Hum. Genet.* **133**, 41–57 (2014).
10. van Koolwijk, L.M. *et al.* Common genetic determinants of intraocular pressure and primary open-angle glaucoma. *PLoS Genet.* **8**, e1002611 (2012).
11. Ramdas, W.D. *et al.* A genome-wide association study of optic disc parameters. *PLoS Genet.* **6**, e1000978 (2010).
12. ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **9**, e1001046 (2011).
13. Yang, T.P. *et al.* Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* **26**, 2474–2476 (2010).
14. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40**, D930–D934 (2012).
15. Hanson, I.M. *et al. PAX6* mutations in aniridia. *Hum. Mol. Genet.* **2**, 915–920 (1993).
16. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).
17. Tserentsoodol, N. *et al.* Intraretinal lipid transport is dependent on high density lipoprotein-like particles and class B scavenger receptors. *Mol. Vis.* **12**, 1319–1333 (2006).
18. Bellincampi, L. *et al.* Identification of an alternative transcript of *ABCA1* gene in different human cell types. *Biochem. Biophys. Res. Commun.* **283**, 590–597 (2001).
19. Singaraja, R.R. *et al.* Alternate transcripts expressed in response to diet reflect tissue-specific regulation of ABCA1. *J. Lipid Res.* **46**, 2061–2071 (2005).
20. Karasinska, J.M. *et al.* ABCA1 influences neuroinflammation and neuronal death. *Neurobiol. Dis.* **54**, 445–455 (2013).
21. Qian, Y. *et al.* PC phosphorylation increases the ability of AFAP-110 to cross-link actin filaments. *Mol. Biol. Cell* **13**, 2311–2322 (2002).
22. Qian, Y., Baisden, J.M., Zot, H.G., Van Winkle, W.B. & Flynn, D.C. The carboxy terminus of AFAP-110 modulates direct interactions with actin filaments and regulates its ability to alter actin filament integrity and induce lamellipodia formation. *Exp. Cell Res.* **255**, 102–113 (2000).
23. Junglas, B. *et al.* Connective tissue growth factor causes glaucoma by modifying the actin cytoskeleton of the trabecular meshwork. *Am. J. Pathol.* **180**, 2386–2403 (2012).
24. Kwon, H.S., Lee, H.S., Ji, Y., Rubin, J.S. & Tomarev, S.I. Myocilin is a modulator of Wnt signaling. *Mol. Cell. Biol.* **29**, 2139–2154 (2009).
25. Inoue, T. & Tanihara, H. Rho-associated kinase inhibitors: a novel glaucoma therapy. *Prog. Retin. Eye Res.* **37**, 1–12 (2013).
26. Becker, D.J. & Lowe, J.B. Fucose: biosynthesis and biological function in mammals. *Glycobiology* **13**, 41R–53R (2003).
27. Miyoshi, E., Moriwaki, K. & Nakagawa, T. Biological function of fucosylation in cancer biology. *J. Biochem.* **143**, 725–729 (2008).
28. Inatani, M. *et al.* Transforming growth factor-β2 levels in aqueous humor of glaucomatous eyes. *Graefes Arch. Clin. Exp. Ophthalmol.* **239**, 109–113 (2001).
29. Picht, G., Welge-Luessen, U., Grehn, F. & Lutjen-Drecoll, E. Transforming growth factor β2 levels in the aqueous humor in different types of glaucoma and the relation to filtering bleb development. *Graefes Arch. Clin. Exp. Ophthalmol.* **239**, 199–207 (2001).
30. Ozcan, A.A., Ozdemir, N. & Canataroglu, A. The aqueous levels of TGF-β2 in patients with glaucoma. *Int. Ophthalmol.* **25**, 19–22 (2004).
31. Pena, J.D., Taylor, A.W., Ricard, C.S., Vidal, I. & Hernandez, M.R. Transforming growth factor β isoforms in human optic nerve heads. *Br. J. Ophthalmol.* **83**, 209–218 (1999).
32. Fleenor, D.L. *et al.* TGFβ2-induced changes in human trabecular meshwork: implications for intraocular pressure. *Invest. Ophthalmol. Vis. Sci.* **47**, 226–234 (2006).
33. Heijl, A., Leske, M.C., Bengtsson, B., Hyman, L. & Hussein, M. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch. Ophthalmol.* **120**, 1268–1279 (2002).

ONLINE METHODS

Study design. In total, 1,155 cases with glaucoma and 1,992 controls, genotyped on Illumina Omni1M or OmniExpress arrays and imputed to the 1000 Genomes phase 1 Europeans panel, were used as the discovery cohort in this study to perform a GWAS for POAG (stage 1). The association results for the top SNPs from the discovery cohort were replicated in stage 2 and 3 replication cohorts. The cohort details, genotyping platforms for each cohort and diagnostic criteria are listed in the **Supplementary Note**. In addition, we performed a meta-analysis for the top SNPs in the discovery and replication cohorts. In this methods section, we describe the methods used for imputation and statistical analyses in the discovery cohort. Methods used for each replication cohort are included in the **Supplementary Note**.

Quality control. The quality-control (QC) methods for the discovery cohort were performed in PLINK[34] by removing individuals with more than 3% missing genotypes, SNPs with call rate <97%, MAF < 0.01 and Hardy-Weinberg equilibrium $P < 0.0001$ in controls and $P < 5 \times 10^{-10}$ in cases. We used the same QC protocol before merging the cases and controls in our discovery cohort to avoid mismatches between the merged data sets. After merging, the genotypes for 569,249 SNPs common to the arrays were taken forward for analysis. Identity by descent was computed in PLINK based on autosomal markers, with one of each pair of individuals with relatedness >0.2 removed. Principal components were computed for all participants and reference samples of known northern European ancestry (1000G British, CEU and Finland participants) using the smartpca package from EIGENSOFT software[35,36]. Participants with principal component 1 (PC1) or PC2 values >6 s.d. from the known northern European ancestry group were excluded.

Imputation. Imputation was conducted using IMPUTE2 (ref. 37) in 1-Mb sections, with the 1000 Genomes phase 1 Europeans (March 2012 release) used as the reference panel[38]. Genotyped SNPs that were strand ambiguous (for example, A/T or C/G) were dropped from the input genotype panel before imputation; given that these are deliberately under-represented on Illumina arrays, this step has limited effects on the ability to impute data but gives greater confidence in the imputation's quality. Imputation was performed with the recommended settings for IMPUTE2, including a 250-kb buffer flanking the imputation sections and an effective size of the sampled population of 20,000 (ref. 37). Reference panel SNPs with MAF < 0.001 in Europeans were not imputed. SNPs with imputation quality score (INFO) >0.8 and MAF > 0.01 were carried forward for analysis.

Statistical analyses. Association testing on the imputed data was performed in SNPTEST[39,40] using an additive model (-frequentist 1) and full dosage scores (-method expected) with sex and the first six principal components fitted as covariates. The genomic inflation factor λ was calculated to investigate the presence of population stratification and inflation. The *P* values were corrected for the genomic inflation factor λ. Quantile-quantile and Manhattan plots were created in R[41]. Regional association plots for the regions reaching genome-wide significance were created using LocusZoom[42].

To investigate whether any hits identified in the discovery cohort were driven by a subset of controls affected by the other diseases (i.e., esophageal cancer, Barrett's esophagus or inflammatory bowel diseases), we also performed a genome-wide association analysis after removing the controls that were affected by these other diseases (the structure of controls in the discovery cohort is provided in the **Supplementary Note**). This analysis included 1,155 cases with glaucoma and 1,147 controls.

Associations of the top autosomal SNPs in the discovery cohort ($P < 1 \times 10^{-4}$) (stage 1) were investigated in the replication cohorts (stages 2 and 3) (the structure of the replication cohorts, QC protocols and statistical analyses for each cohort are provided in the **Supplementary Note**). Stage 2 included two Australian replication data sets (total of 932 cases and 6,862 controls), and stage 3 included two US cohorts (total of 2,616 cases and 2,634 controls). For replication in stage 2, 24 SNPs with the best *P* values in the discovery cohort were used as the lead SNPs for the autosomal regions with $P < 1 \times 10^{-4}$ (**Supplementary Table 4**). The SNPs that were nominally replicated in stage 2 ($P < 0.05$) were taken forward for replication in stage 3. To make maximum valid use of our cohorts, for replication we focused on SNPs directly genotyped on the Illumina Human610 and Human670 arrays. Because some of the cases in stage 2 were genotyped on a non–genome wide platform (Sequenom), we could not accurately evaluate the imputed SNPs from stage 1. Hence, the most-associated SNP upstream of *ABCA1* (rs2472493) and SNPs in high LD with the most-associated SNP near *ARHGEF12* (rs11217878 and rs2276035, $r^2 = 1$ and $r^2 = 0.94$, respectively, with rs11827818) were used in the replication studies. Similarly, SNPs in high LD with the most-associated SNP in *GMDS* (rs2761233 and rs11969985, $r^2 = 0.93$ and $r^2 = 0.87$, respectively, with rs114096562) were used for replication studies.

Fixed-effects meta-analysis for the top SNPs was performed between the discovery and replication cohorts in METAL[43] using the effect sizes and their standard errors for the risk alleles. The presence of heterogeneity between cohorts for the effect sizes of risk alleles was investigated using the I^2 statistic as implemented in METAL.

Identifying candidate genes. Candidate genes in the regions of association were selected on the basis of the location and function of the genes, the pathways that the genes are involved in, the tissue location of expression of the gene and whether similar phenotypes have been reported to be caused by mutations in these genes. This information was found in Ensembl[44], NCBI, UCSC genome Bioinformatics[45], Genecards[46] and UniprotKB[47], as well as in available published data. To predict functional effects of the top POAG-associated SNPs identified in this study, we used the ENCODE project data[12] and the associated databases RegulomeDB[16] and HaploReg v2 (ref. 14). We used the Genevar database[48] to investigate eQTLs within genetic regions of interest.

Expression analysis of genes at associated loci in ocular tissues and cells. Ocular tissues from post-mortem human eyes were obtained through the Eye Bank of South Australia according to guidelines of the Southern Adelaide Clinical Human Research Ethics Committee. Normal and glaucomatous trabecular meshwork cell lines, NTM-5 and GTM-3, respectively, were a kind gift from C. Abbot (Alcon Research Ltd.). Both the cell lines tested negative for mycoplasma contamination. Total RNA was extracted using the RNeasy Mini Kit (Qiagen Pty Ltd.). First-strand cDNA was synthesized using Superscript III reverse transcriptase (Invitrogen, Life Technologies Australia Pty Ltd.) and random hexamers. PCR was performed using Hot Star Taq Plus polymerase (Qiagen) and gene-specific primers (**Supplementary Table 9**). PCR was performed at the conditions specified in **Supplementary Table 9**. The enzyme was activated at 95 °C for 5 min, denaturation was at 95 °C for 30 s and elongation was at 72 °C. Additional elongation at 72 °C for 5 min was allowed after completion of the amplification cycles. The specificity of each amplified product was confirmed by sequencing.

Immunohistochemical labeling. Eye tissue was fixed in neutral buffered formalin and embedded in paraffin. For immunolabeling, 4-µm sections were blocked with 5% normal goat serum and incubated with the mouse anti-ABCA1 (1:2,000, Ab66217, Sapphire Biosciences) or anti-AFAP (1:1,000, 610200, BD Transduction Laboratories) primary antibody at 4 °C overnight. Primary antibody binding was detected with the Novolink Polymer detection kit (Leica Microsystems) and Chromogen substrate coloration (Dako). Sections were counterstained with haematoxylin and mounted in dePeX (Merck KGaA). Light microscopy was performed on an Olympus BX50 brightfield upright microscope with a Q-Imaging color charge-coupled device (CCD) camera attached; images were taken using the QCapture software (Q-Imaging Corporate).

Western blotting. For western blotting, proteins from NTM-5 and GTM-3 human trabecular meshwork cells established from a normal individual and an individual with glaucoma, respectively, were extracted in RIPA buffer, analyzed by SDS-PAGE using the mini-PROTEAN TGX gel and transferred onto a polyvinylidene difluoride membrane (Bio-Rad Laboratories Pty. Ltd.). Western blotting was performed using the mouse anti-ABCA1 (1:500, Ab66217, Sapphire Biosciences) or anti-AFAP (1:250, 610200, BD Transduction Laboratories) primary antibody followed by hybridization with the hydrogen peroxide–conjugated goat anti-mouse IgG secondary antibody (1:1,000, 115-035-003, Jackson ImmunoResearch Laboratories Inc.). ABCA1 antibody binding was detected using the Pierce SuperSignal West Pico (Jackson ImmunoResearch

Laboratories Inc.), and AFAP1 antibody binding was detected using ECL Prime (GE Healthcare Australia and New Zealand) chemiluminescence reagents.

34. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
35. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
36. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
37. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
38. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
39. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
40. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
41. R Core Team. *R: A Language and Environment for Statistical Computing* (Vienna, 2013).
42. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
43. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
44. Flicek, P. *et al.* Ensembl 2014. *Nucleic Acids Res.* **42**, D749–D755 (2014).
45. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
46. Safran, M. *et al.* GeneCards Version 3: the human gene integrator. *Database* (*Oxford*) **2010**, baq020 (2010).
47. UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42**, D191–D198 (2014).
48. Yang, T.P. *et al.* Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* **26**, 2474–2476 (2010).

doi:10.1038/ng.3079