Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma

Stuart MacGregor[1], Jue Sheng Ong[1], Jiyuan An[1], Xikun Han[1], Tiger Zhou[2], Owen Siggs[2], Matthew H Law[1], Emmanuelle Souzeau[2], Shiwani Sharma,[2] David Lynn[3], Jonathan Beesley[1], Bronwyn Sheldrick[2], Richard A Mills,[2] John Landers[2], Jonathan B Ruddle[4], Stuart L Graham[5], Paul R Healey,[6] Andrew JR White[6], Robert J Casson[7], Stephen Best,[8] John R Grigg[9], Ivan Goldberg[9], Joseph E Powell[10], David C Whiteman[1], Graham Radford Smith[1], Nicholas G Martin[1], Grant W Montgomery[10], Kathryn P Burdon[11], David A Mackey[11,12]*, Puya Gharahkhani[1]*, Jamie E Craig[2]*, Alex W Hewitt[4,11]*

1. QIMR Berghofer Medical Research Institute, Brisbane, Australia.
2. Department of Ophthalmology, Flinders University, Flinders Medical Centre, Bedford Park, Australia.
3. South Australian Health and Medical Research Institute, Adelaide, Australia.
4. Centre for Eye Research Australia, University of Melbourne, Australia.
5. Faculty of Medicine and Health Sciences, Macquarie University, Australia.
6. Centre for Vision Research, Westmead Institute for Medical Research, University of Sydney, Australia.
7. South Australian Institute of Ophthalmology, University of Adelaide, Adelaide, South Australia, Australia.
8. Eye Department, Greenlane Clinical Centre, Auckland District Health Board, Auckland, New Zealand.
9. Discipline of Ophthalmology, University of Sydney, Sydney Eye Hospital, Sydney, Australia.
10. Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia.
11. Menzies Institute for Medical Research, University of Tasmania, Australia.
12. Centre for Ophthalmology and Visual Science, University of Western Australia, Australia.

* Contributed equally
Corresponding author: Stuart MacGregor, stuart.macgregor@qimrberghofer.edu.au

Intraocular pressure (IOP) is currently the sole modifiable risk factor for primary open-angle glaucoma (POAG), one of the leading causes of blindness worldwide.[1] **Both IOP and POAG are highly heritable.**[2] **We report a combined analysis of participants from the UK Biobank (N=103,914) and previously published data from the International Glaucoma Genetic Consortium (N=29,578)**[3,4] **that identified 101 statistically independent genome-wide significant SNPs for IOP, 85 of which had not been previously reported.**[4–12] **We examined these SNPs in 11,018 glaucoma cases and 126,069 controls, with 53 showing evidence for association. Gene-based tests implicated an additional 22 independent genes for IOP. We derived an allele score based on the IOP loci, and loci influencing optic nerve head morphology. In 1,734 people with advanced glaucoma and 2,938 controls, participants in the top decile of the allele score were at increased risk (OR=5.6; 95% CI:4.1-7.6) of glaucoma relative to the bottom decile.**

Over the past decade genome-wide association studies (GWAS) have implicated 14 independent loci involved in the pathogenesis of POAG,[4–12] and an additional eight loci have been associated with primary angle-closure glaucoma (PACG).[13,14] The classification of POAG and PACG is based on the anatomical configuration of the iridocorneal angle, where outflow of aqueous humor occurs through the trabecular meshwork. Regardless of the glaucoma classification, it is well established that elevated IOP can lead to irreversible optic nerve degeneration and corresponding visual field deficits.[1] Currently all drugs used to treat glaucoma lower IOP by either increasing aqueous outflow (through the trabecular meshwork or uveoscleral tracts), or decreasing aqueous production.[1] Understanding which genes influence IOP may open new avenues for glaucoma treatment. We report results from a large GWAS for IOP and glaucoma, and explore the genetic relationship between the endophenotype and the disease.

To identify SNPs influencing IOP, we undertook a meta-analysis of IOP GWAS from the publicly available UK Biobank (UKBB; see URLs) and previously published data from the International Glaucoma Genetic Consortium (IGGC; see URLs; **Supplementary Fig. 1**).[4] To determine which of the peak SNPs were statistically independent and thus potentially informative in allelic risk profiling, we used the program GCTA-COJO to perform conditional analysis on the summary meta-analysis (see URLs and methods section for full description).[15] A total of 106 independent SNPs (uncorrelated with other peak SNPs) surpassed the genome-wide significance threshold ($P < 5 \times 10^{-8}$, **Fig. 1; Supplementary Table 1 & 2, Supplementary Fig. 2**). For downstream

analysis, we removed five peak SNPs influencing IOP measurement through corneal biomechanics. The removed SNPs were rs66724425 in *ADAMTS6,* previously shown to be associated with central corneal thickness,[16] and SNPs rs1570204, rs78658973, rs12492846 and rs2797560, which were more strongly associated (*i.e.* lower P-value) with corneal hysteresis (a measure of viscous damping in the cornea that influences IOP measurement), than with IOP (**Supplementary Table 2**). Among the remaining 101 SNPs, we found strong concordance (Pearson's correlation coefficient = 0.85; P < 0.001) in the effect sizes between IGGC and UKBB (**Fig. 2a**). Of the 101 associated SNPs, 85 had not been previously associated with IOP, whilst 16 had been previously associated with either IOP or glaucoma at the genome-wide significant level (marked in blue in **Fig. 1**).[4–12] The only previously identified IOP locus that we did not replicate at the genome-wide significant level was *ADAMTS8* (peak SNP rs56009602, P = 6.2×10^{-6}).

Similar to other complex traits, it is likely that additional SNPs beyond the 101 described above, are also associated with IOP, but do not reach genome-wide significance.[17] To estimate the overall contribution of all common variants (*i.e.* SNP MAF > 0.01) to IOP, we applied LD Score regression,[18] which yielded a SNP heritability estimate of 0.16 (standard error, SE = 0.01). We then considered the distribution of association P-values across the genome. Since there was genomic inflation (genomic control lambda = 1.26, **Supplementary Fig. 3**), we computed the LD Score regression intercept to assess whether this genomic inflation was attributable to many variants of small effect (polygenes) or due to the effect of issues such as population structure. The LD Score regression intercept was 1.06 (SE = 0.01), indicating that the majority of the inflation was due to polygenes.

We then performed a GWAS meta-analysis for glaucoma by combining data from UKBB glaucoma cases and controls (selected to be independent of those in our IOP GWAS; 7947 cases, 119318 controls) with 3,071 cases from the Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG) and 6,750 historic controls (see the methods section for full description; **Supplementary Fig. 1**). Our genome-wide analysis of glaucoma found 24 genome-wide significant loci (**Table 1, Supplementary Fig. 4, Supplementary Fig. 5**). Similar to IOP, there was genomic inflation due to the effect of polygenes (**Supplementary Fig. 6**), but the intercept of the univariate LD score regression obtained from the meta-analysed data was close to 1 (0.95, SE = 0.01), suggesting that our results were not biased by population substructure or cryptic relatedness.

Of the 24 genome-wide significant loci for glaucoma, two (rs944801 within *CDKN2B-AS1* and rs2093210 within the *SIX6* locus; orange dots on **Fig. 2b**) are known to be associated with vertical cup-disc ratio (VCDR), an important optic nerve head parameter which is often used to define or diagnose glaucoma.[19] An additional locus (rs61861119 near *MYOF*) was found to have no association with IOP but did have a suggestive level of evidence for association with VCDR (P= 1.6×10^{-5}; pink dot on **Fig. 2b**).[4] The remaining 21 glaucoma loci are likely to influence disease development wholly or partly via IOP, with all showing at least P < 0.01 (15 were genome-wide significant) for IOP (**Fig. 2b** and **Table 1**). Seven of the 21 also showed association with VCDR at P < 0.01 (**Table 1**).

The relationship between IOP and glaucoma beyond the 24 SNPs which were genome-wide significantly associated with glaucoma was also examined. At the individual SNP level, of the 101 independent genome-wide significant IOP SNPs, 53 were significantly associated with glaucoma after Bonferroni correction (P < 0.05/101 = 0.000495). The Pearson's correlation coefficient between IOP effect size and the glaucoma log odds ratio was 0.93 (P < 0.001; **Fig. 2b**). Using bivariate LD score regression, we estimated the genome-wide genetic correlation between IOP and glaucoma to be 0.71 (SE = 0.04).[20]

We also undertook a series of gene-based and pathway-based analyses for IOP and glaucoma. An additional 22 independent genes associated with IOP were identified through FastBAT gene-based tests.[21] Of these 22 genes, four were associated with glaucoma following Bonferroni correction for 22 genes (P < 0.00227), with an additional seven achieving P < 0.05 (**Supplementary Table 3**). In MAGMA pathway analysis,[22] 11 Gene Ontology (GO) annotations were significantly associated with IOP, including extracellular matrix, collagen and vascular development. Among the 11 pathways highlighted by the IOP analysis, 9 showed at least P < 0.05 in pathway analysis in the glaucoma samples, with the strongest GO annotation result for vascular development (P = 0.0015; **Supplementary Table 4**). Seven pathways were significant in our DEPICT analysis of IOP.[23] As displayed in **Supplementary Table 5**, the most significant IOP pathways were positive regulation of locomotion, cell motility and cell migration. These pathways were also significant in glaucoma (P = 0.0021 to 0.0025).

Next, we tested whether the IOP loci could be used to predict POAG in the ANZRAG cohort. Allele scores were derived based on the 101 genome-wide significant primary IOP SNPs identified in

this study (see the Methods section for inclusion criteria), as well as two loci with established associations with optic nerve head morphology (*CDKN2B-AS1* and *SIX6*). These were tested in an independent dataset comprising 1,734 Australians of European ancestry with advanced POAG and 2,938 controls. Relative to a base model without the allelic scores, the scores were strongly associated with POAG status ($P < 2 \times 10^{-16}$, Nagelkerke $R^2 = 7.7\%$, AUC = 0.65 [95% CI: 0.63 - 0.66]). Fitting only the IOP and only the VCDR SNPs in the allele score reduced the Nagelkerke R^2 to 5.4% and 2.7%, respectively. Individuals in the top 5%, 10%, and 20% of the allele scores were at significantly ($P < 0.0001$) increased risk of POAG relative to the bottom 5%, 10%, and 20%, respectively (OR = 7.8, 5.6, and 4.2, respectively).

We sought to characterize the expression profile of genes at the novel IOP loci that were also associated with glaucoma (**Supplementary Fig. 7**) across a range of human ocular tissues (corneal epithelium, corneal stroma, corneal endothelium, trabecular meshwork, ciliary body pigmented epithelium, neurosensory retina, optic nerve head and the optic nerve). Expression of newly-associated genes was more highly enriched ($P = 6.1 \times 10^{-59}$, Wilcoxon rank sum test for novel genes versus all other genes) in the trabecular meshwork compared to other ocular tissues. We then computed the ranks of the novel genes amongst all genes for each tissue and found that four of the other seven tissues (ciliary body pigmented epithelium, corneal stroma, optic nerve head and the optic nerve) were not significantly different, in terms of enrichment, compared to trabecular meshwork ($P > 0.05$ for each pairwise comparison, Wilcoxon rank sum test, the similar tissues are the 5 leftmost columns in **Supplementary Fig. 7**). For the other three tissue types (neurosensory retina, corneal epithelium, corneal endothelium, clustered as the 3 rightmost columns in **Supplementary Fig. 7**), the degree of enrichment was less than that seen in trabecular meshwork ($P < 0.05$ for each pairwise comparison, Wilcoxon rank sum test). Finally, using FANTOM5 Cap Analysis of Gene Expression data, we found evidence of correlation between enhancers harboring associated SNPs and the promoters of nine genes, including *PTPN1*, *BCLAF1* and *GAS7* in stromal and eye tissues (**Supplementary Table 6**), which is noteworthy given that hypoplasia of the anterior iris stroma is the most common iris defect associated with developmental glaucoma,[24] and that these genes may act in a similar, albeit subclinical, manner.

Many of the loci we identified are associated with other eye conditions. Loss-of-function variants in *LTBP2* have been found to cause primary congenital glaucoma (PCG)[25]; we now report that common variants at this locus influence IOP in the general population. Similarly, rare loss-of-

function variants in *TEK* have been associated with PCG,[26] and we identified common IOP-influencing variants in genes encoding the two known TEK ligands (*ANGPT1; ANGPT2*), as well as a third related protein (*ANGPTL2*).

Anterior segment dysgenesis, iris abnormalities, nanophthalmos, and microcornea are known causes of secondary glaucoma.[24] Interestingly, four genes influencing the variation of IOP in the general population have been associated with anterior segment dysgenesis or other abnormalities of the iris, lens or cornea: *FOXC1* with ocular anterior segment dysgenesis; *TRAF3IP1* with iris furrows;[27] *MFRP* with nanophthalmos;[28] and *ADAMTS18* with microcornea, myopic chorioretinal atrophy and telecanthus.[29] Loss-of-function variants in *LMX1B* lead to nail-patella syndrome; common variants at this locus are now definitively associated with both POAG and IOP.[30,31] Interestingly three loci (*PLEKHA7; FERMT2; GLIS3*) have been previously associated with PACG,[13,14] and we have now implicated these regions with IOP, with two of them (*PLEKHA7; FERMT2)* also showing association with POAG (**Supplementary Table 1**). It is acknowledged that UKBB participants were not subjected to detailed clinical examination of their ocular anterior segment, hence some associations with IOP or POAG could be at least in part related to undiagnosed narrow drainage angles or subtle variations of ocular development.

Although the Australian glaucoma samples used were confirmed POAG cases,[32] a limitation of the UKBB glaucoma cases was that only a small subset had documented disease subtype. Nevertheless, the proportion of non-POAG glaucoma cases in UKBB would be expected to be small.[33] Applanation-based methods for IOP measurement are influenced by corneal biomechanical properties, such as corneal thickness and hysteresis.[34] A strength of our work is the large sample size for standardised IOP measurement, with corneal compensation data available for approximately three-quarters of the dataset (corneal compensated IOP data was available for UKBB samples but not for IGGC samples). SNPs more strongly associated with corneal hysteresis than with IOP were excluded and this allowed us to identify a set of SNPs that have greater relevance to glaucoma development, rather than spuriously influencing IOP measurement.

In conclusion, we leveraged large sample sets from the UKBB and the IGGC to dramatically expand the number of genomic regions associated with IOP. We identified 101 statistically independent SNPs for IOP and found that 53 of them were associated with glaucoma. This work highlights the high genetic correlation between IOP and glaucoma. A number of previously

implicated (extracellular matrix and collagen), and novel (vascular development and cell migration) pathways were associated with both IOP and glaucoma. Finally, an allele score based on the IOP loci and loci influencing optic nerve head morphology was able to enhance risk stratification.

URLs:

BOLT-LMM: https://data.broadinstitute.org/alkesgroup/BOLT-LMM/

DEPICT: https://data.broadinstitute.org/mpg/depict/index.html

Drug Gene Interaction Database: http://dgidb.genome.wustl.edu/

EdgeR bioconductor package: https://bioconductor.org/packages/release/bioc/html/edgeR.html

FANTOM5 data: http://enhancer.binf.ku.dk/

GCTA software: http://cnsgenomics.com/software/gcta/

Haplotype Reference Consortium: http://www.haplotype-reference-consortium.org/

International Glaucoma Genetic Consortium dataset: https://goo.gl/73qHqk

HTseq-count v0.6.0 software: https://pypi.python.org/pypi/HTSeq

LOCUSZOOM: http://locuszoom.sph.umich.edu/

LD-hub database: http://ldsc.broadinstitute.org/

MAGMA: https://ctg.cncr.nl/software/magma

METAL software: http://csg.sph.umich.edu/abecasis/Metal/

PLINK software: http://www.cog-genomics.org/plink2

TopHat v2.1.1 software: https://ccb.jhu.edu/software/tophat/index.shtml

UK Biobank: http://www.ukbiobank.ac.uk/

Reporting Summary

Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data Availability

The data that support the findings of this study are available from the corresponding author upon request. International Glaucoma Genetic Consortium results are available from

https://goo.gl/73qHqk. UK Biobank data are available through the UK Biobank Access Management System (see URLs).

Acknowledgments:

This work was conducted using the UK Biobank Resource (application number 25331) and publicly available data from the International Glaucoma Genetics Consortium. This work was supported by grants from the National Health and Medical Research Council (NHMRC) of Australia (#1107098 (JEC); 1116360 (DAM), 1116495 (JEC), 1023911 (DAM)), the Ophthalmic Research Institute of Australia, the BrightFocus Foundation. SM is supported by an Australian Research Council Future Fellowship. KPB, JEC and AWH are supported by NHMRC Fellowships. We thank Scott Wood and John Pearson from QIMR Berghofer for IT support.

Author Contribution:

SM, AWH, JEC, PG, DAM designed the study and obtained funding. SM, JSO, JA, XH, TZ, MHL, SS, JEP, DL, JB analyzed the data. SM, TZ, OS, ES, SS, BS, RAM, JL, JBR, SLG, PRH, AJRW, RJC, SB, JRG, IG, DCW, GRS, NGM, GWM, KPB, DAM, JEC, AWH contributed to data collection and contributed to genotyping. SM, JSO, DAM, PG, AWH wrote the first draft of the paper. All authors contributed to the final version of the paper.

Competing Interests:

The authors declare no competing financial interests.

References:

1. Weinreb, R. N. *et al.* Primary open-angle glaucoma. *Nat Rev Dis Primers* **2,** 16067 (2016).
2. Sanfilippo, P. G., Hewitt, A. W., Hammond, C. J. & Mackey, D. A. The heritability of ocular traits. *Surv. Ophthalmol.* **55,** 561–583 (2010).
3. Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. (2017). doi:10.1101/166298
4. Springelkamp, H. *et al.* New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. *Hum. Mol. Genet.* **26,** 438–453 (2017).
5. van Koolwijk, L. M. E. *et al.* Common genetic determinants of intraocular pressure and primary open-angle glaucoma. *PLoS Genet.* **8,** e1002611 (2012).
6. Springelkamp, H. *et al.* ARHGEF12 influences the risk of glaucoma by increasing intraocular pressure. *Hum. Mol. Genet.* **24,** 2689–2699 (2015).
7. Bailey, J. N. C. *et al.* Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. *Nat. Genet.* **48,** 189–194 (2016).
8. Gharahkhani, P. *et al.* Common variants near ABCA1, AFAP1 and GMDS confer risk of primary open-angle glaucoma. *Nat. Genet.* **46,** 1120–1125 (2014).
9. Hysi, P. G. *et al.* Genome-wide analysis of multi-ancestry cohorts identifies new loci influencing intraocular pressure and susceptibility to glaucoma. *Nat. Genet.* **46,** 1126–1130 (2014).
10. Chen, Y. *et al.* Common variants near ABCA1 and in PMM2 are associated with primary open-angle glaucoma. *Nat. Genet.* **46,** 1115–1119 (2014).
11. 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).

12. Thorleifsson, G. *et al.* Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma. *Nat. Genet.* **42,** 906–909 (2010).
13. Vithana, E. N. *et al.* Genome-wide association analyses identify three new susceptibility loci for primary angle closure glaucoma. *Nat. Genet.* **44,** 1142–1146 (2012).
14. Khor, C. C. *et al.* Genome-wide association study identifies five new susceptibility loci for primary angle closure glaucoma. *Nat. Genet.* **48,** 556–562 (2016).
15. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44,** 369–75, S1–3 (2012).
16. Lu, Y. *et al.* Genome-wide association analyses identify multiple loci associated with central corneal thickness and keratoconus. *Nat. Genet.* **45,** 155–163 (2013).
17. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42,** 565–569 (2010).
18. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).
19. Ramdas, W. D. *et al.* A genome-wide association study of optic disc parameters. *PLoS Genet.* **6,** e1000978 (2010).
20. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47,** 1236–1241 (2015).
21. Bakshi, A. *et al.* Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. *Sci. Rep.* **6,** 32894 (2016).
22. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11,** e1004219 (2015).
23. Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6,** 5890 (2015).

24. Reis, L. M. & Semina, E. V. Genetics of anterior segment dysgenesis disorders. *Curr. Opin. Ophthalmol.* **22,** 314–324 (2011).
25. Ali, M. *et al.* Null mutations in LTBP2 cause primary congenital glaucoma. *Am. J. Hum. Genet.* **84,** 664–671 (2009).
26. Souma, T. *et al.* Angiopoietin receptor TEK mutations underlie primary congenital glaucoma with variable expressivity. *J. Clin. Invest.* **126,** 2575–2587 (2016).
27. Larsson, M. *et al.* GWAS findings for human iris patterns: associations with variants in genes that influence normal neuronal pattern development. *Am. J. Hum. Genet.* **89,** 334–343 (2011).
28. Sundin, O. H. *et al.* Extreme hyperopia is the result of null mutations in MFRP, which encodes a Frizzled-related protein. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 9553–9558 (2005).
29. Khan, A. O. Microcornea with myopic chorioretinal atrophy, telecanthus and posteriorly-rotated ears: a distinct clinical syndrome. *Ophthalmic Genet.* **33,** 196–199 (2012).
30. Vollrath, D. *et al.* Loss-of-function mutations in the LIM-homeodomain gene, LMX1B, in nail-patella syndrome. *Hum. Mol. Genet.* **7,** 1091–1098 (1998).
31. Sweeney, E., Fryer, A., Mountford, R., Green, A. & McIntosh, I. Nail patella syndrome: a review of the phenotype aided by developmental biology. *J. Med. Genet.* **40,** 153–162 (2003).
32. Souzeau, E. *et al.* Australian and New Zealand Registry of Advanced Glaucoma: methodology and recruitment. *Clin. Experiment. Ophthalmol.* **40,** 569–575 (2012).
33. Quigley, H. A. & Broman, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *Br. J. Ophthalmol.* **90,** 262–267 (2006).
34. Ehlers, N., Bramsen, T. & Sperling, S. Applanation tonometry and central corneal thickness. *Acta Ophthalmol.* **53,** 34–43 (1975).

Figure 1:

Manhattan plot displaying associations with intraocular pressure (IOP) in people of Northern European descent. The dashed line represents the threshold for genome-wide significance ($P < 5 \times 10^{-8}$). Loci highlighted in blue are established regions known to be associated with primary open-angle glaucoma (POAG). The top SNP and its effect allele (SNP-EA) at each genome-wide significant locus is displayed with the corresponding effect allele frequency (EAF); odds ratio (OR); and p-value (P) for association in glaucoma cases (full details are contained in **Supplementary Table 1**). The results are in black for SNPs with $P > 0.05$ with glaucoma, red text for SNPs with $0.05 < P < 0.05/101$ (not significant following correction for multiple comparisons) and bold red text for SNPs with $P < 0.05/101$ (significant following correction for multiple comparisons). *These loci were either reported central corneal thickness loci (*ADAMTS6*) or more strongly associated with corneal hysteresis and were removed from subsequent analysis (**Supplementary Table 2**).

Figure 2:

Regression coefficients (β in mmHg) or effect size for the top associated SNPs at each locus associated with intraocular pressure (IOP) at the genome-wide significant level. 95% confidence intervals are displayed in gray. (**a**) Comparison of regression coefficients in the UK Biobank (y-axis) and the International Glaucoma Genetic Consortium dataset (x-axis; Pearson's correlation coefficient = 0.85). The solid line indicates the line of best fit. (**b**) Concordance between regression coefficients for IOP in 133,492 people of Northern European descent (x-axis) and the direct effect size (log odds ratio) in 11,018 glaucoma cases versus 126,069 controls (y-axis; Pearson's correlation coefficient = 0.93). The solid line indicates the line of best fit through the 101 IOP SNPs. The 101 IOP SNPs are shown as black dots. SNPs identified in the GWAS of glaucoma are superimposed in red/pink/orange; those in red show $P < 0.05$ with IOP, those in pink show $P < 0.05$ with VCDR but not IOP and the SNPs in orange are at *CDKN2B-AS1* and *SIX6*, which are known to act independently of IOP.

Table 1:

Genome-wide significant loci identified in meta-analysis of glaucoma (UKBB + ANZRAG), with their corresponding GWAS statistics from meta-analysis of intraocular pressure (UKBB + IGGC) or vertical cup-disc ratio (IGGC). The results are presented from the smallest to the largest P value for glaucoma. Bold indicates previously unreported risk loci for primary open-angle glaucoma.

Chr	Position	SNP	EA	NEA	OR POAG	95% CIs POAG	P POAG	Effect IOP	SE IOP	P IOP	P VCDR^	Nearest gene
1	165736880	rs7518099	T	C	0.73	0.70-0.76	2.35×10^{-52}	-0.33	0.02	3.96×10^{-67}	0.058	*LOC100147773, TMCO1*
9	22051670	rs944801	C	G	1.22	1.17-1.27	8.00×10^{-36}	0.02	0.01	0.232	3.85×10^{-32}	*CDKN2B-AS1*
9	107695848	rs2472493*	A	G	0.84	0.80-0.87	4.30×10^{-30}	-0.19	0.01	3.62×10^{-50}	4.85×10^{-07}	*LOC105376196, ABCA1*
14	60957279	rs2093210	T	C	0.86	0.83-0.90	6.29×10^{-22}	-0.009	0.01	0.483	1.22×10^{-09}	*C14orf39*, *SIX6*
17	10031183	rs9913911	A	G	1.16	1.12-1.21	2.13×10^{-21}	0.21	0.01	1.59×10^{-57}	5.62×10^{-06}	*GAS7*
4	7891545	rs28795989	A	G	1.15	1.11-1.20	1.90×10^{-20}	0.15	0.01	2.94×10^{-32}	0.019	*AFAP1*
9	129378026	rs945686	C	G	0.86	0.83-0.90	2.58×10^{-17}	-0.14	0.01	4.25×10^{-22}	0.016	*LMX1B*
6	1548369	rs2745572	A	G	1.13	1.08-1.17	1.35×10^{-13}	0.13	0.01	2.66×10^{-22}	5.41×10^{-06}	*LOC102723944, GMDS*
3	**85095766**	**rs9284802**	**A**	**G**	**0.90**	**0.86-0.93**	**1.56×10^{-12}**	**-0.05**	**0.01**	**4.74×10^{-05}**	**0.665**	***CADM2***
11	120248493	rs58073046	A	G	0.85	0.82-0.89	1.99×10^{-12}	-0.20	0.02	1.03×10^{-22}	0.189	*ARHGEF12*
7	**11679113**	**rs12699251**	**A**	**G**	**0.90**	**0.86-0.93**	**4.16×10^{-12}**	**-0.05**	**0.01**	**9.98×10^{-05}**	**0.100**	***THSD7A***
8	**108278616**	**rs10505100**	**A**	**C**	**0.84**	**0.81-0.88**	**4.86×10^{-12}**	**-0.21**	**0.02**	**1.45×10^{-27}**	**0.043**	***ANGPT1***
7	116153025	rs2024211	A	C	0.90	0.86-0.93	9.48×10^{-12}	-0.22	0.01	2.90×10^{-55}	0.004	*CAV1, CAV2*
3	186131600	rs9853115	A	T	0.90	0.87-0.94	4.35×10^{-11}	-0.18	0.01	2.84×10^{-43}	0.026	*DGKG*, *LOC107986164, TBCCD1*
5	**14851094**	**rs61394862**	**T**	**C**	**0.90**	**0.86-0.93**	**4.13×10^{-10}**	**-0.09**	**0.01**	**8.42×10^{-11}**	**0.781**	***ANKH***
6	**170454915**	**rs2935057**	**A**	**G**	**1.15**	**1.11-1.20**	**8.02×10^{-10}**	**0.11**	**0.02**	**1.30×10^{-08}**	**0.250**	***LOC101929614, LOC105378153***
6	**637465**	**rs2073006**	**T**	**C**	**1.14**	**1.10-1.18**	**1.20×10^{-09}**	**0.11**	**0.02**	**2.29×10^{-09}**	**1.81×10^{-05}**	***EXOC2***
10	94942423	rs61861119	A	G	0.91	0.88-0.95	2.56×10^{-09}	0.01	0.01	0.313	1.56×10^{-05}	*MYOF, XRCC6P1*
22	19854006	rs8141433	A	G	1.15	1.11-1.20	3.04×10^{-09}	0.08	0.02	2.85×10^{-06}	0.235	*TXNRD2*
10	**60338753**	**rs4141671**	**T**	**C**	**0.91**	**0.88-0.95**	**8.67×10^{-09}**	**-0.05**	**0.01**	**0.0004**	**0.0001**	***BICC1***
3	**169252883**	**rs73174345**	**T**	**G**	**0.84**	**0.80-0.89**	**1.53×10^{-08}**	**-0.07**	**0.03**	**0.008**	**0.732**	***MECOM***

7	**117603820**	**rs1013278**	**C**	**G**	**1.09**	**1.05-1.14**	**2.99×10^{-08}**	**0.08**	**0.01**	**3.32×10^{-10}**	**0.003**	***CTTNBP2, CFTR***
11	**128380742**	**rs7924522**	**A**	**C**	**1.09**	**1.05-1.14**	**3.33×10^{-08}**	**0.11**	**0.01**	**3.99×10^{-15}**	**0.090**	***ETS1***
3	**150059342**	**rs11710139**	**A**	**G**	**0.90**	**0.87-0.94**	**5.00×10^{-08}**	**-0.06**	**0.01**	**3.89×10^{-05}**	**0.463**	***LOC107986141, LOC107986142***

Abbreviations: Chr, Chromosome; EA, effect allele; NEA, non-effect allele; SE, Standard error; IOP, intraocular pressure; VCDR, vertical cup-disc ratio.
^P value obtained from the VCDR GWAS in IGGC.
*This SNP was not present in the quality control passed IOP data, and hence, the corresponding statistics for IOP is reported for rs2472496 (effect allele A, non-effect allele G), a SNP in high LD (r^2 = 0.967) with rs2472493.

ONLINE METHODS:

Analysis of UK Biobank (UKBB) Data:

For a complete description of the UKKB genotype curation, please see the report by Bycroft and colleagues.[3] All participants provided informed written consent, the study was approved by the National Research Ethics Service Committee North West – Haydock, and all study procedures were performed in accordance with the World Medical Association Declaration of Helsinki ethical principles for medical research. In brief, approximately 488,000 participants were genotyped on custom-designed Affymetrix UK BiLEVE Axiom or UK Biobank Axiom arrays (Affymetrix Santa Clara, USA), which produced a combined total of 805,426 markers in the released data. Following standard quality control (QC) the dataset was phased and ~96M genotypes were imputed using Haplotype Reference Consortium (HRC; see URLs) and UK10K haplotype resources.[3,35,36] Due to the UKBB's reported QC issues with non-HRC SNPs, we retained only the ~40M HRC SNPs for analysis.

Among the 487,409 individuals who passed initial genotyping QC, 409,694 participants had white-British ancestry, based on self-reported ethnicity and genetic principal components. To maximise our effective sample size, we also included UKBB participants if their self-reported ancestry was not white-British (this includes a substantial number of individuals reporting their ancestry as "Irish" or "any other white background") but their first two genetic principal components fell within the region of those that are classified white-British in the N = 409,694 set in Bycroft et al. (see **Supplementary Fig. 8**). Using these criteria, we identified 438,870 individuals for this study who are genetically similar to those of white-British ancestry.

Individuals were selected for analysis to ensure independence of the IOP and glaucoma arms of the study. Selection was based on the following (**Supplementary Fig. 1**): 1). glaucoma cases were selected, 2). individuals participating in the ocular examination (approximately a quarter of the UKBB cohort) were selected (with glaucoma cases and their relatives [$\hat{\pi}$> 0.2)] omitted) and 3). individuals who self-reported having no eye disease were selected (controls were screened to be unrelated [$\hat{\pi}$> 0.2]) for use as controls with the glaucoma cases. Among the 438,870 with suitable genetic data, we extracted 7,947 individuals with glaucoma; cases were those who either 1) had an ICD-10 diagnosis ("Primary Open Angle Glaucoma", "Other Glaucoma", "Glaucoma, unspecified", 2) responded "Glaucoma" to "Has a doctor told you that you have any of the following problems with your eyes?", 3) responded "Glaucoma" to the question "In the touch screen you

selected that you have been told by a doctor that you have other serious illnesses or disabilities, could you now tell me what they are? (non-cancer illness)". Although this glaucoma definition is broad, ~80% of "glaucoma" cases amongst white British individuals are likely to meet diagnostic criteria for POAG.[33] The number of individuals with ICD-10 POAG was over five times less, limiting the power of the study. A subset (127,468) of UKBB participants took part in the ocular examination, which included IOP measurements using the Ocular Response Analyzer non-contact tonometer.[37] Our primary IOP analysis was based on corneal-compensated IOP (IOPcc) measurements because these are expected to be less affected by corneal factors than Goldmann-correlated IOP measures. The mean IOPcc for each participant was calculated, with measurements < 5 or > 60 mmHg set to missing. Mean corneal hysteresis and mean non-corneal-compensated (Goldmann-correlated) IOP were also derived and tested at loci of interest from the IOPcc analysis. 103,914 individuals with ocular examinations had both phenotype and genotype data available. Finally, controls for the glaucoma cases were selected based on a reply of "None" to "Has a doctor told you that you have any of the following problems with your eyes?" and no ocular examination.

Genotyping and Analysis of the Australian & New Zealand Registry of Advanced Glaucoma (ANZRAG) Cohort:

The clinical recruitment and characterisation of the ANZRAG cohort has been described previously.[32] In this analysis a total of 3,071 POAG cases and 6,750 historic controls of European descent were used. Case and control samples were genotyped on Illumina Omni1M, OmniExpress or HumanCoreExome arrays (Illumina, San Diego, USA).[8,11] This dataset involves three phases of POAG data collection, and hence, QC, imputation, and association analysis was conducted separately for each phase before combining the results in a meta-analysis. The first phase was previously published and comprises 1,155 advanced POAG cases and 1,992 historic controls genotyped on Illumina Omni1M or OmniExpress arrays.[11] In this phase the historic controls were obtained from 225 oesophageal cancer cases, 317 Barrett's oesophagus cases and their 552 controls, as well as 303 inflammatory bowel diseases cases and their corresponding 595 control cohort. The second phase includes a further 579 advanced POAG cases genotyped on Illumina HumanCoreExome array and 946 controls selected from parents of twins previously genotyped on the same array.[8] The third phase comprises 1,337 POAG cases genotyped on Illumina HumanCoreExome array and 3,812 controls selected from a study of endometriosis previously genotyped on the same array.[38] There is strong female bias in the control set in phase three, but not in phases one and two (our allele score prediction work below uses only phases

one and two). Human research ethics approval was obtained from the relevant committees of the Southern Adelaide Clinical Human Research Ethics Committee/Flinders University, the University of Tasmania, QIMR Berghofer Institute of Medical Research and the Royal Victorian Eye and Ear Hospital. Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki.

As described previously, QC was performed using PLINK (see URLs).[39] Individuals with more than 3% missing genotypes, and SNPs with call rate less than 97%, minor allele frequency (MAF) < 0.01, and Hardy-Weinberg equilibrium (HWE) $P < 0.0001$ in controls or $P < 5 \times 10^{-10}$ in cases were removed from the analysis. Identity by descent was determined based on autosomal markers in PLINK,[39] and only one of each pair of individuals with relatedness ($\hat{\pi}$) of > 0.2 was used in the analysis. PLINK was used to compute principal components for all participants and reference samples of known northern European ancestry (1000G British, CEU, Finland participants). Participants with PC1 or PC2 values > 6 standard deviations from the mean of known Northern European ancestry group were excluded. All statistical tests throughout the manuscript were two-sided.

Phasing of the genotyped SNPs was conducted using ShapeIT,[40] and imputation was performed using Minimac3 through the Michigan Imputation Server,[41] with the HRC as the reference panel.[35] SNPs with imputation quality ($\hat{r}^2$) > 0.3 and MAF > 0.01 were used for analysis.

Association testing: IOP IGGC

We obtained publicly available GWAS summary statistics from an IGGC study on IOP.[4] 29,578 individuals had 1000G imputed GWAS data available, with IOP corrected for age and sex.[4] The vast majority of IGGC sites used Goldmann-corrected IOP; these IOP measures do not account for corneal differences between individuals and in large samples an “IOP” analysis may identify loci that are primarily driven by corneal parameters.[4]

Association testing: IOP UKBB

Association analysis was performed using a linear mixed model framework to account for cryptic relatedness and population stratification in the UKBB samples using BOLT-LMM version 2.3 (see URLs).[42] We used a sparse set of 360,087 genotyped SNPs across the autosomes to estimate the Bayesian Gaussian mixture prior to characterising the random-effects genetic component. The infinitesimal model in BOLT-LMM was used to generate GWAS p-values. The IGGC and

UKBB IOP results were combined using meta-analysis, implemented in METAL (2011-03-25 release; see URLs).[43]

To identify statistically independent genome-wide significant SNPs, an initial list of SNPs with meta-analysis p-values < 5×10^{-8} was pruned into discrete regions by LD clumping in PLINK v1.9 (r^2 threshold for clumping 0.1, physical distance threshold for clumping 2 megabases). This initial list of SNPs was then further explored for additional independent signals by conditioning the meta-analysis summary data using GCTA version 1.26 (see URLs). To calculate LD, a reference panel was constructed from 5,000 individuals randomly selected from the UKBB white British ancestry individuals. Imputed SNPs with a rsq > 0.3 and MAF > 0.001 were converted to best guess genotypes, and then cleaned for 3% missingness and HWE < 1×10^{-6}. Initially a given peak SNP was used to condition all SNPS within 2 megabases (--cojo-cond option). Where there were multiple SNP within 2 megabases of each other, they were analysed together using boundaries at least ±2 megabases from the furthest apart SNP. Following this, a SNP was deemed independent if its initial single SNP p-value was < 5×10^{-8} and remained <5×10^{-8} following conditioning. Newly identified SNPs were iteratively added to the regional conditioning until no more SNPs had a p-value < 5×10^{-8}. As a final check the joint effect (--cojo-joint) of all putatively genome-wide significant SNPs was estimated, with any SNPs which then had joint p-values > 5×10^{-8} discarded.

Association testing: UK Biobank glaucoma case-control analysis

We assessed associations between SNPs and glaucoma status adjusted for sex and the first six principal components, under an additive genetic model using the dosage scores obtained from imputation. Association analysis was performed using PLINK version 2.0. [39] Identity by descent was determined based on autosomal markers in PLINK version 1.90b, and only one of each pair of individuals with $\hat{\pi} > 0.2$ was used in the analysis. Figure 1 was produced by using Adobe Illustrator to juxtaposition a manhattan plot in R with a table produced in R.

We used mean χ^2 and the univariate LD score regression approach to investigate presence of model or structural bias in the glaucoma GWAS data.[18] An LD score intercept close to 1 in a univariate analysis indicates that there is no model misspecification and that other sources of bias such as population stratification and cryptic relatedness are not adversely affecting results.

Exclusion of loci based on association with corneal parameters

All loci that were genome-wide significantly associated with IOP were tested for association with corneal hysteresis (hyst - a measure of viscous damping in the cornea that may affect the measurement of IOP). SNPs with a larger effect on hyst than on IOP are unlikely to be truly associated with IOP and hence we filtered SNPs with $P_{hyst} < P_{IOP}$ (SNPs with effects on various aspects of eye morphology of large effect, such as *TMCO1*, influenced both IOP and hyst and so we did not filter simply on P_{hyst}). Similarly, putative IOP loci were examined for previous association with central corneal thickness.[16]

LD-score regression:

We applied univariate LD-score regression (see URLs),[18] to estimate the SNP-heritability of IOP and bivariate LD-score regression[20] to estimate the genetic correlation between IOP and glaucoma.

Gene-based tests:

Gene-based tests were conducted using the fast and flexible set-Based Association Test (fastBAT), a gene-based approach that calculates the association p-values for a set of SNPs (within ± 50 kb of a gene for this study) using GWAS summary data while accounting for LD between SNPs.[21] Only loci distinct from those found in the per-SNP tests (>1 megabase away) were tested. fastBAT was applied to the IOP meta-analysis results, with a significance threshold of 2×10^{-6} (0.05/24,654 genes tested). Genes exceeding this threshold were then tested for association with glaucoma (ANZRAG+UKBB) using fastBAT.

Pathway-based tests

Pathway-based tests were conducted on the IOP meta-analysis results using MAGMA and DEPICT (see URLs).[23,44] We opted to use both approaches because they use different pathway databases as well as a different method for annotating SNPs to genes. In MAGMA, Z-scores from a gene-based step were combined based on 5,917 pre-specified Gene Ontology gene sets. DEPICT is an integrative tool that, for each gene, uses the likelihood of membership in each gene set based on the co-regulation of gene expression data, and tests whether any of the 14,462 preconstituted gene sets are significantly enriched for genes in the trait-associated loci. SNPs exceeding $P < 5 \times 10^{-8}$ were used to define trait-associated loci in a pathway analysis in DEPICT. Pathways exceeding $P < 0.05/5917$ (MAGMA) or $P<0.05/14463$ (DEPICT) were then tested using the same approach in glaucoma (ANZRAG+UKBB).

Allele Scores:

We used the allele score approach to investigate whether the genome-wide significant IOP loci identified in this study, as well as the two previously known VCDR loci with established association with POAG (rs2157719 within the *CDKN2B-AS1* locus and rs8015152 within the *SIX6* locus), can significantly predict risk of glaucoma. We used only statistically independent SNPs to create the profile scores and excluded the known published central corneal thickness loci as well as corneal hysteresis SNPs whose P values in this study were lower than the IOP P values. This was to rule out those SNPs that may not truly affect IOP but have been detected as IOP loci through their effects on corneal hysteresis. The SNPs passing the above criteria were used to score individuals in a target cohort, a subset of the ANZRAG data with advanced POAG (1,734 cases and 2,938 controls). Our ANZRAG dataset was non-overlapping with the cohort used to identify the IOP SNPs (and their estimated effect sizes). The score for each individual in ANZRAG was calculated by summing the number of risk alleles weighted by their effect sizes obtained from the IOP and VCDR analyses. As IOP and VCDR are measured on different scales, we benchmarked their relative weights (*in terms of their effect on glaucoma*) using the well-established large effect associations with IOP (*TMCO1* rs10918274 - estimated to increase IOP by 0.33 units and in a POAG meta-analysis, to increase risk 1.39 fold) and with VCDR (*CDKN2B-AS1* rs2157719- estimated to increase VCDR by 0.13 and POAG 1.44 fold).[4] Based on these benchmarks, each 1 unit increase in IOP leads to a 0.998 log(OR) increase (log(1.39)/0.33) in POAG risk. Similarly, each 1 unit increase in VCDR leads to a 28.049 log(OR) increase (log(1.44)/0.13) in POAG risk. Hence, prior to combining the IOP and VCDR allele scores for analysis, we multiplied the VCDR risk score by 28.049/0.998 to place it on an equivalent scale to IOP.

To estimate the contribution of the profile scores with the POAG status in the ANZRAG target cohort, we first performed a logistic regression with sex and the first four principal components used as covariates (base model). We then added the profile scores into the logistic model and computed the increase in the Nagelkerke's pseudo R^2 from the logistic regression over and above the base model (Nagelkerke's pseudo R^2 is a measure of the goodness of fit in the prediction model, analogous to phenotypic variance explained in a linear regression). We also compared the POAG risk for the top versus bottom 5%, 10% (decile), and 20% of the profile score distribution.

Drug Pathway:

The Drug Gene Interaction Database (DGIdb 3.0 release; see URLs) was used to identify compounds that act on genes at each locus and could be repurposed in the treatment of glaucoma.[45]

Gene Target Prediction

FANTOM5 data representing enhancer-promoter Cap Analysis of Gene Expression (CAGE) expression correlation from all cell types were downloaded and processed (see URLs).[44] Enhancers active in eye and stromal tissues were tested for overlap with SNPs correlated with lead SNPs ($r^2 > 0.8$ in 1000Genomes EUR populations). Genes for which CAGE promoter expression signals were correlated with enhancers were selected as potential target genes.

Ocular Expression Analysis:

The gene expression profiles of all genes within IOP-associated loci were examined in relevant ocular tissues. Data were available from a total of 16 donor eyes from 16 individuals. RNA was extracted from 48 samples of distinct ocular tissue (corneal epithelium, corneal stroma, corneal endothelium, trabecular meshwork, ciliary body pigmented epithelium, neurosensory retina, optic nerve head and the optic nerve) and sequenced using Illumina NextSeq 500 (Catalog# FC-404-2005, San Diego, USA) with Bioo Scientific NEXTflex rapid directional mRNA-seq Kit (Catalog# 5138-10, Austin, Texas, USA). We obtained an average of 56 million 75 bp paired-end reads per sample. Following QC and trimming these were mapped to the reference human genome (hg19) using TopHat v2.1.1 and HTseq-count v0.6.0 (see URLs).[46,47] Normalized counts per million (CPM) data were calculated using trimmed mean of M-values (TMM) normalization method using edgeR v.3.10.2 (see URLs).[48] Transcripts from a total of 21,962 RefSeq protein-coding genes were captured and mapped. We had 94.5% of the reads mapped to the human genome after QC filtering. The mean TMM value across all available samples for each gene in each tissue was calculated and to test whether there was enrichment for genes at the novel loci associated with glaucoma in each tissue we used a Wilcoxon rank sum test for novel genes versus all other genes. We then computed the ranks of the novel genes amongst all genes for each tissue and compared each tissue in turn to the tissue showing most enrichment (Wilcoxon rank sum test).

Methods-only References:

35. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48,** 1279–1283 (2016).

36. UK10K Consortium *et al.* The UK10K project identifies rare variants in health and disease. *Nature* **526,** 82–90 (2015).
37. Chan, M. P. Y. *et al.* Associations with Intraocular Pressure in a Large Cohort: Results from the UK Biobank. *Ophthalmology* **123,** 771–782 (2016).
38. Nyholt, D. R. *et al.* Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nat. Genet.* **44,** 1355–1359 (2012).
39. 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).
40. Delaneau, O., Marchini, J. & Zagury, J.-F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9,** 179–181 (2011).
41. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48,** 1284–1287 (2016).
42. Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47,** 284–290 (2015).
43. Willer, C. J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* **40,** 161–169 (2008).
44. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507,** 455–461 (2014).
45. Griffith, M. *et al.* DGIdb: mining the druggable genome. *Nat. Methods* **10,** 1209–1210 (2013).
46. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).
47. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
48. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40,** 4288–

4297 (2012).