

# Retinal ganglion cell-specific genetic regulation in primary open angle glaucoma

Maciej S. Daniszewski<sup>1,2,3,#</sup>, Anne Senabouth<sup>4#</sup>, Helena H. Liang<sup>2,3</sup>, Xikun Han<sup>5</sup>, Grace E. Lidgerwood<sup>1,2,3</sup>, Damián Hernández<sup>1,2,3</sup>, Priyadharshini Sivakumaran<sup>3</sup>, Jordan E. Clarke<sup>3</sup>, Shiang Y. Lim<sup>2,6</sup>, Jarmon G. Lees<sup>6,7</sup>, Louise Rooney<sup>1,2,3</sup>, Lerna Gulluyan<sup>1,2,3</sup>, Emmanuelle Souzeau<sup>8</sup>, Stuart L. Graham<sup>9</sup>, Chia-Ling Chan<sup>4</sup>, Uyen Nguyen<sup>4</sup>, Nona Farbehi<sup>4</sup>, Vikkitharan Gnanasambandapillai<sup>4</sup>, Rachael A. McCloy<sup>4</sup>, Linda Clarke<sup>3</sup>, Lisa Kearns<sup>3</sup>, David A Mackey<sup>10,11</sup>, Jamie E. Craig<sup>8</sup>, Stuart MacGregor<sup>5</sup>, Joseph E. Powell<sup>4,12^</sup>, Alice Pébay<sup>1,2,3^</sup>, Alex W. Hewitt<sup>2,3,11^</sup>

<sup>1</sup> Department of Anatomy and Neuroscience, The University of Melbourne, Parkville, VIC 3010, Australia

<sup>2</sup> Department of Surgery, The University of Melbourne, Parkville, VIC 3010, Australia

<sup>3</sup> Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, VIC 3002, Australia

<sup>4</sup> Garvan Weizmann Centre for Cellular Genomics, Garvan Institute of Medical Research, The Kinghorn Cancer Centre, Darlinghurst, NSW 2010, Australia

<sup>5</sup> QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

<sup>6</sup> O'Brien Institute Department of St Vincent's Institute of Medical research, Melbourne, Fitzroy, VIC 3065, Australia

<sup>7</sup> Department of Medicine, St Vincent's Hospital, The University of Melbourne, Parkville, VIC 3010, Australia

<sup>8</sup> Department of Ophthalmology, Flinders University, Flinders Medical Centre, Bedford Park, SA 5042, Australia

<sup>9</sup> Faculty of Medicine and Health Sciences, Macquarie University, NSW 2109, Australia.

<sup>10</sup> Lions Eye Institute, Centre for Vision Sciences, University of Western Australia, WA 6009, Australia

<sup>11</sup> School of Medicine, Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS 7005, Australia

<sup>12</sup> UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, NSW 2052, Australia

\*Correspondence: [j.powell@garvan.org.au](mailto:j.powell@garvan.org.au); [apecbay@unimelb.edu.au](mailto:apecbay@unimelb.edu.au); [hewitt.alex@gmail.com](mailto:hewitt.alex@gmail.com)

# Equal first authors

^ Equal senior authors

# **ABSTRACT**

To assess the transcriptomic profile of disease-specific cell populations, fibroblasts from patients with primary open-angle glaucoma (POAG) were reprogrammed into induced pluripotent stem cells (iPSCs) before being differentiated into retinal organoids and compared to those from healthy individuals. We performed single-cell RNA-sequencing of a total of 330,569 cells and identified cluster-specific molecular signatures. Comparing the gene expression profile between cases and controls, we identified novel genetic associations for this blinding disease. Expression quantitative trait mapping identified a total of 2,235 significant loci across all cell types, 58 of which are specific to the retinal ganglion cell subpopulations, which ultimately degenerate in POAG. Transcriptome-wide association analysis identified genes at loci previously associated with POAG, and analysis, conditional on disease status, implicated 54 statistically significant retinal ganglion cell-specific expression quantitative trait loci. This work highlights the power of large-scale iPSC studies to uncover context-specific profiles for a genetically complex disease.

# **KEYWORDS**

human induced pluripotent stem cells, retinal organoids, retinal ganglion cells, single-cell RNA sequencing, glaucoma, transcriptomics, eQTL

# INTRODUCTION

Glaucoma is the leading cause of irreversible blindness worldwide and experts predict it will affect approximately 80 million people by 2040 <sup>1</sup>. The most common subtype -- primary open-angle glaucoma (POAG) -- is characterized by an open iridotrabeular meshwork angle and progressive degeneration of retinal ganglion cells (RGCs), which culminates in loss of visual field <sup>2</sup>. Therapeutic options are currently limited; all are directed at lowering intraocular pressure, which has been shown to slow but not fully prevent or reverse visual loss <sup>3</sup>.

POAG has one of the highest heritabilities of all common and complex diseases <sup>4,5</sup>, and much work has focussed on dissecting its genetic architecture. The genetic aetiology of POAG is varied: rare genetic variants, e.g. myocilin (*MYOC*)<sup>6</sup> and optineurin (*OPTN*)<sup>7</sup>, cause disease with high penetrance, while common variants have smaller effect sizes. Genome-Wide Association Analyses (GWAS) have identified over 100 independent loci that carry a common risk allele associated with an increased risk of open-angle glaucoma <sup>8</sup>. Unlike rare variants that largely contribute to disease by changing protein coding, common variants predominantly act via changes in gene regulation <sup>9</sup>. Mapping expression quantitative trait loci (eQTL) is one of the powerful approaches used to provide functional evidence of the mechanisms of the common genetic variants, allowing the allelic effect of a variant on disease risk to be linked to changes in gene expression. To avoid spurious associations, and to best understand the cellular effects of changes in gene expression, eQTL mapping needs to be conducted for cells that are pathophysiologically relevant to the disease. In the case of glaucoma, this includes trabecular meshwork cells and RGCs.

The molecular profiling of RGCs in normal and diseased tissue would improve our understanding of the mechanisms that bear a disease risk or contribute to the glaucoma development. Unfortunately, it is impossible to obtain or molecularly profile RGCs from living donors in a non-invasive manner. To overcome this, somatic cells can be reprogrammed into patient-specific induced pluripotent stem cells (iPSCs) <sup>10,11</sup>, which can then be differentiated into RGCs <sup>12,13</sup>. Over the years, multiple protocols have been

developed to generate RGCs *in vitro*<sup>14</sup>. Human retinal organoids show a stratified cellular organisation closely resembling the developing human neural retina<sup>15–20</sup>, and thus it is now possible to generate organoid-derived RGCs for downstream applications, including disease modelling<sup>13,21,22</sup> and cell transplantation<sup>23</sup>. These constructs can also be subjected to single-cell RNA-sequencing (scRNA-seq) to distinguish cell types based on their unique transcriptional signature and examine rare populations that would be missed using bulk RNA-seq<sup>24–26</sup>. Here, we used scRNA-seq to characterize the transcriptomic profile of the organoid-derived retinal cells, in particular RGCs, generated from a large cohort of the patient-derived iPSCs. We identified a number of cell type and disease-specific eQTLs. Using an additive linear model, a total of 54,786 eQTLs were found to be associated with 21,512 SNPs. By combining our data with recent GWAS results in a transcriptome-wide association study (TWAS), seven genes were identified to be significantly associated with glaucoma development.

## RESULTS

### ***Large-scale generation of patient iPSCs, differentiation into retinal organoids and scRNA-seq.***

We recruited a large cohort of individuals, which included healthy (n=57) and patients with advanced POAG (n=55), who were matched by sex. The mean  $\pm$  SD age at recruitment for controls was : 70.1  $\pm$  8.8 years, and the mean  $\pm$  SD age at diagnosis for case subjects was 59.7  $\pm$  10.3 years. Participants underwent skin biopsy and their cultured fibroblasts were reprogrammed to iPSCs using episomal vectors as we previously described<sup>27</sup>. iPSC lines were differentiated to neural retina for 28 days in adherent cultures, after which retinal organoids were then excised, cultured in suspension for 7 days and plated onto Matrigel for an additional 2-week period to allow neuronal outgrowth from RGCs, and harvested for scRNA-seq (**Figure 1A**). Differentiation was performed in batches, with cells harvested and multiplexed for scRNA-Seq with a targeted capture of 16,000 cells from up to eight cell lines per pool.

Twenty-three lines did not differentiate to retinal organoids and were discarded. A total number of 330,569 cells from 160 individual cell lines were captured and sequenced at a mean read depth of 41,020 per cell (**Table S1**). Following cell-specific quality control, 21,346 poor-quality cells were identified and discarded. Genotyping data were generated for all individuals and after quality control and imputation, yielded 7,691,208 autosomal SNPs from 162 individuals at a minor allele frequency (MAF) above 0.01. 514,087 exon-only SNPs were used in conjunction with scRNA-seq data to successfully assign 276,831 cells to 128 donors (**Table S2**). 32,382 cells were designated doublets or could not be traced to a donor, and were removed from the dataset. Following genotype, virtual karyotyping and scRNA-seq quality control assessment, a total of 24 batches containing 247,520 cells derived from 110 donor iPSC lines were used for subsequent analyses.

### ***Identification and characterisation of 23 subpopulations from 253,107 cells***

Clustering identified 23 subpopulations distributed evenly across cell lines from all donors (**Figure 1B, S1 and S2**). We compared the percentage of cell types between patients with POAG and healthy controls, and observed no statistically significant differences between the groups ( $p\text{-value} \geq 0.174$ , **Figure S2**). Differential expression markers were used to classify the subpopulations to different retinal cell classes based on canonical markers<sup>28–31</sup> (**Figure 1C, 1D, Table 1, Table S3**). Retinal progenitor cells (RPCs) represented 77.4% of all cells and localised across 16 subpopulations (**Table 1**). RPCs expressed *PAX6* and *SOX2* transcription factors that are key regulators of neuronal fate<sup>32,33</sup>, *LHX2*, required for maintenance of open chromatin during retinogenesis<sup>34</sup> and gliogenesis<sup>35</sup>, and the RPC markers *VSX2* and *ASCL1* (**Figure 1D, Figure S1**). Cell cycle genes were not evenly distributed within progenitor subpopulations. The G2/M phase marker *MKi67* was predominantly expressed by cells in RPC1, RPC2 and RPC5. The S phase marker *PCNA* was distributed more broadly;

however, the majority of *PCNA*-positive cells were identified in RPC clusters 1, 2, 4 and 5 (**Figure 1D**).

RGCs were the second most predominant cell group, representing 17.0% of all cells across the cohort. Based on previous work, RGCs were classified by the expression of the following genes: *POU4F2*, *ISL1*, *RBPMS*, *SNCG*, *GAP43*, *NEFL/M*, *ELAVL4*, *EOMES* and *DCX*<sup>36–38</sup>. Three distinct RGC subpopulations (RGC1-3), arising from one subpopulation of RPCs (RPC9), were identified (**Figure 1B-D**, **Table 1**). Pseudotime analysis confirmed the lineage development of RGC1-3 cell types from a single progenitor population (**Figure S3**), with *POU4F2* and *ISL1* showing increasing expression as cells differentiated further from a progenitor state (**Figure 1D**). The expression of both genes are required for RGC specification and differentiation<sup>39–41</sup>. *POU4F2* expression was generally higher in the RGC1 and RGC2 subpopulations than in RGC3, while *ISL1* expression was higher in RGC1 and RGC3 compared to RGC2. The low levels of *ATOH7* expression in RGC1 and RGC3, in conjunction with the fact that cells in these subpopulations expressed markers of differentiated RGCs such as *SNCG*, *RBPMS*, *GAP43* and *NEFM* (**Figure 1D**), suggests that these subpopulations represent more mature RGCs compared to those from RGC2. We also identified cells expressing markers for photoreceptors/bipolar cells (2.6%) and interneurons (1.7%, **Table 1**). Retinal Pigmented Epithelial (RPE) cells localized in one subpopulation and constituted 1.3% of all cells (**Table 1**). No Müller cells were identified. The various subpopulations are fully described in the **Supplemental Results**. These data are consistent with Sridhar and colleagues<sup>30</sup>, who also found that RPCs and RGCs are the predominant populations of cells within early retinal organoids.

### ***The genetic control of gene expression is highly cell type-specific***

To explore cell type-specific genetic control of gene expression, we tested for *cis*-eQTL for each cell population independently. We identified a total of 2,235 eQTL across all cell types, which surpassed a study-wide significance threshold of FDR < 0.05 and where the proportion of gene expression in subpopulation-donor pools was > 30% (**Table 2**).

The number of eQTL identified per cell type ranged from ten to 456 (**Table S4**), indicating consistent power to detect eQTL and also a similarity of cell types in each population as expected. We assessed the overlap of eQTL between cell types, and found that the majority of cis-eQTL are cell type-specific (**Figure 2A**). Cell-type ubiquitous (shared across cell cell-types) eQTL were rare, with only one eGene (*RPS26*) being common to all cell types. However, we identified a greater degree of shared eQTL amongst related cell types, such as subpopulations belonging to the RGC lineage, specifically RGC1, RGC2, RGC3 and RPC9, which share eQTL for 58 eGenes (**Figure 2B, C**).

One potential explanation for the cell type-specific eQTL detection is that the gene is only expressed in one cell type, and therefore, we would not expect to observe an eQTL in the other cell types where the gene is not expressed. To evaluate this, we correlated the expression of each gene that had a significant cell type-specific eQTL effect, with its expression levels in each of the other cell types (**Figure 2C**). These results indicate that cell type-specific eQTL are not a function of cell type-specific gene expression, showing high levels of correlation in almost all instances. Another possible explanation for the cell type-specific eQTL is low statistical power to detect eQTL in multiple cell types. To assess this hypothesis, we implemented an empirical framework to test the rank of the test statistics for eGene SNP effects across the non-significant cell types for each cell type-specific eQTL. In almost all instances we observed none, or very limited, enrichment of the test statistic across cell types.

Among the RGC eQTL identified, a number are directly involved in neurogenesis or neurodegeneration. For example, *CNOT6L* is a poly(A) specific mRNA deadenylase that controls transcription, and reduced expression has been found to impair retinal differentiation in *D. melanogaster*<sup>42</sup>. *DNAJA1* belongs to a large family of chaperones, and has been shown to prevent neurodegeneration by decreasing  $\alpha$ -synuclein aggregates<sup>43</sup>, whilst *TSPAN2*, is known to support myelination<sup>44</sup>.

To evaluate the context-specific relationship on the eQTL allelic effect due to disease status, we tested for a SNP-by-disease state interaction and identified 2,175



lead eQTLs (**Table S5**). A total of 290 eQTLs were identified where there was a significant interaction between the SNP allelic effect and POAG disease status. Fifty-four of these eQTLs were specific to RGC lineage subpopulations (**Table S6**). These eQTL are of particular interest, as the data suggest that the allelic effect of the SNP differs due to disease. Interestingly, *rs28368130* at chromosome 9p21, a locus which has been definitively associated with POAG <sup>8,45</sup>, was found to influence *CDKN2B* expression in the RGC1 cell population with a disease-status FDR  $p = 4.67 \times 10^{-4}$ . An eQTL for *SPP1* was found across all RGC lineages (**Figure 3A**). *SPP1* encodes for osteopontin, which is known to have reduced expression in eyes from patients with POAG <sup>46</sup>, and is thought to be neuroprotective to the retina <sup>47,48</sup>. *IGFBPL1* regulates axonal growth in RGCs <sup>49</sup> and was also found to have a statistically significant disease-state interaction eQTL in the RPC9, RGC1 and RGC2 subpopulation of cells (**Figure 3B**). Similarly, *SAR1A*, which is involved in transport between the endoplasmic reticulum to the Golgi apparatus and is associated with axonal growth <sup>50</sup>, has an eQTL identified by *rs4746023*, in RGC1 cell types. In patients with POAG, carrying each additional copy of the A allele causes an increase by an average of 1.4 transcripts per cell, which is approximately 2 times higher than in healthy controls (**Figure 3C**).

### ***Disease-specific differential expression of genes across cell types identifies altered transthyretin expression in POAG RGCs.***

We next sought to evaluate the relationship between disease status and regulation of the transcriptome and a cellular level, testing for differences in the expression levels of genes in each cell population. In total, after Bonferroni correction we identified 3,118 genes whose expression was either up- or down- regulated in POAG cases relative to the controls (**Table S7**). We can be confident that these results are due to the genetic effects underlying POAG risk, as at all steps from iPSC generation, differentiation, cell capture, and library preparation, the cell lines were either managed in either shared conditions, or randomized with respect to disease status (**Methods**). Further, no firm environmental factors have been found to definitively predispose to POAG risk, and are



unlikely to account to a difference in gene expression in differentiated cells, given the epigenetic profile of fibroblast-derived iPSCs is reset during reprogramming<sup>51,52</sup>. Focusing on the three RCG populations, we identified 144 genes differentially expressed between POAG cases and controls (**Figure 4A**). Consistent with our observations of cell type-specific eQTL, 68.06% of genes were identified as differentially expressed in only one cell type, reinforcing the conclusions that the genetic effects of POAG are highly cell type-specific. Interestingly, *TTR* was found to be differentially expressed between POAG cases and healthy controls across all RGC subpopulations (**Figure 4A**). Coding variants in *TTR* are known to cause familial amyloidotic polyneuropathy, which is frequently associated with glaucoma<sup>53</sup>.

### ***Transcriptome-wide association study (TWAS) identifies novel and refine known genetic associations for glaucoma***

We leveraged this iPSC-derived retinal organoid single cell eQTL data with our recently reported multitrait glaucoma GWAS summary statistics to prioritize glaucoma risk genes in a TWAS<sup>8</sup>. In the single-cell TWAS, we identified seven genes associated with POAG after Bonferroni correction (**Figure 4B**). Of the five genes identified in the RGC1 subpopulation, one is located at a locus (chromosome 17q21) recently associated with POAG<sup>54</sup>. Here, we implicate *KANSL1-AS1*, which was also identified as a major eGene for RGCs (**Figure 2B**). *KANSL1-AS1* was also found by TWAS to be associated with POAG in the RPC9 subpopulation.

The TWAS results also helped fine-map potential causal genes at known GWAS loci<sup>55,56</sup>. Most of the identified genes map to loci that have previously been associated with POAG<sup>54</sup>. Briefly, the five TWAS genes on chromosome 2 are located near GWAS reported gene *BRE* and share the same GWAS variant rs6741499 (or in strong LD)<sup>8</sup>, which is also associated with IOP (Gao et al., 2018). Of these, the top TWAS hit *MPV17* encodes a mitochondrial inner membrane protein involved in the metabolism of reactive oxygen species<sup>57</sup> and has been found to play an important role in the pathogenesis of

RGC damage <sup>58</sup>. The gene *CTD-307407.5* on chromosome 11 from RGC2 subpopulation is near *MALAT1*, which is also associated with VCDR <sup>59</sup>.

We then compared the TWAS results based on scRNA-seq data to bulk RNA-seq data. The bulk retinal transcriptome data were described previously <sup>60</sup>. Of the three genes with available bulk TWAS results, only *KANSL1-AS1* was significant after Bonferroni correction ( $P_{\text{bulk TWAS}} = 2.95 \times 10^{-4}$ ).

## DISCUSSION

Here we present a large-scale scRNA-seq analysis of iPSC-derived RGCs. We generated over 100 patient-specific iPSC lines and differentiated them into RGCs using retinal organoids. Following the capture of over 330,000 cells, we analysed 258,071 cells from 112 individuals. We identified a total of 2,235 eQTL across all cell types, and found 58 eGenes that are specific to the RGC lineage. POAG culminates in loss of RGCs, and in iPSC-derived RGCs we identified disease-associated loci. Analysis, conditional on disease status, implicated 54 statistically significant RGC eQTLs, and single cell TWAS identified seven genes at loci previously associated with POAG.

Several recent studies have employed scRNA-seq to characterize transcriptomic changes during human retinal development using fetal tissue <sup>38,61</sup>, retinal organoids <sup>62–64</sup> or both <sup>30,31</sup>. Our results complement these findings and are in concordance with those of Sridhar *et al.* <sup>30</sup> who also observed RPCs as the major cell type in early organoids followed by RGCs, photoreceptors and interneurons. Similar observations were made by Lu and colleagues <sup>31</sup>. The absence of glial cells within the retinal organoids in our dataset is not surprising given these are generally the last retinal cells to develop, <sup>65</sup> and emerged in older retinal organoids <sup>30</sup>. Furthermore, we did not observe statistically significant differences in cellular composition of organoids derived from healthy controls and patients with POAG, which suggests a high level of consistency across differentiation batches.

It is becoming widely recognized that the pharmaceutical pipeline for drug development has stalled, and that there is a pressing need for human models of disease to improve our molecular understanding of common, complex diseases and facilitate preclinical trials<sup>66</sup>. We investigated the impact of genetic background and disease status on gene expression through the eQTL mapping. Highlighting the power of large-scale iPSC studies to uncover disease-specific profiles, this work lays the foundation for context-specific drug screening and underscores the efficiency of using stem cell models for dissecting complex disease.

## **ACKNOWLEDGMENTS**

We thank Vikrant Singh, Anthony Cook, Alison Conquest and Dominik Kaczorowski for providing technical support. This research was supported by National Health and Medical Research Council Practitioner Fellowships (AWH, JEC, DAM), Senior Research Fellowship (AP, 1154389; SM), Career Development Fellowship (JEP, 1107599) and Investigator Fellowship (JEP, 117578), an Australian Research Council Future Fellowship (AP, FT140100047), an International Postgraduate Research Scholarship & Research Training Program Scholarship (MD), and by grants from the National Health and Medical Research Council (1150144, 1143163) and Australian Research Council (180101405), the Joan and Peter Clemenger Foundation, the Goodridge Foundation, the Ophthalmic Research Institute of Australia, the BrightFocus Foundation, the Philip Neal bequest, Stem Cells Australia – the Australian Research Council Special Research Initiative in Stem Cell Science, the University of Melbourne and Operational Infrastructure Support from the Victorian Government.

## **AUTHOR CONTRIBUTIONS**

Conceptualization, J.E.P., A.P., A.W.H.; Methodology, M.S.D., A.S., X.H., S.M., J.E.P., A.P., A.W.H.; Investigation, M.S.D., H.H.L., X.H., G.E.L., D.H., L.R., P.S., J.E., CL., U.N., R.M., V.G., S.Y.L., J.L., L.G.; Resources, H.H.L., E.S., S.L.G., L.C., L.K., D.A.M., J.E.C., J.E.P., A.P., A.W.H.; Data analysis, M.S.D., A.S., X.H., S.M., J.E.P., A.P., A.W.H.; Writing - original draft, M.S.D., A.S., X.H., S.M., J.E.P., A.P., A.W.H.; Writing - review & editing, M.S.D., A.S., X.H., G.E.L., D.H., D.A.M., J.E.C., S.M., J.E.P., A.P., A.W.H.; Supervision and project administration, J.E.P., A.P., A.W.H.; Funding Acquisition, D.A.M., J.E.C., S.M., J.E.P., A.P., A.W.H.

## **DECLARATION OF INTEREST**

The authors declare no competing interests.

## FIGURE LEGENDS

**Figure 1. Generation of retinal organoids and identification and characterisation of cell subpopulations.** (A) Retinal organoids were generated from iPSCs over a period of 49 days. iPSCs were differentiated as a monolayer for the first 28 days, and then cultured in 3D as a suspension for 7 days. Resulting organoids were then plated onto Matrigel and grown until retinal ganglion cells started to project out of the organoid at 49 days. These were harvested for scRNA-seq. (B) Uniform Manifold Approximation and Projection (UMAP) representation of cells grouped into 23 subpopulations, as identified by Louvain clustering. (C) UMAP plot of the cell types and lineages, as determined by analysis of differentially expressed genes of individual subpopulations and trajectory analysis. RGC clusters form one branch of the trajectory. Other cell types - RPE, interneurons, cones and lens - arise from another branch of the trajectory. The last main branch consists of differentiating RPC subpopulations. RPC: retinal progenitor cell. RGC: retinal ganglion cell. RPE: retinal pigmented epithelium. (D) Heatmap of the mean expression of cell type-specific gene markers across in each subpopulation. Expression values have been scaled and converted to z-scores and genes have been grouped by cell type. AC, amacrine cell; BC, bipolar cell; HC, horizontal cell; MG, Müller glia; PR, photoreceptor; RPC, retinal progenitor cell; RPE, retinal pigment epithelium; RGC, retinal ganglion cell.

**Figure 2. Cell subpopulation-specific eQTLs.** Summary of eQTLs specific to cell subpopulations. (A) Minimal overlap of genes with significant eQTLs (eGenes) show they are predominantly subpopulation-specific. A small proportion of these eGenes were model-specific. (B) Chromosomal map of significant loci in RGC subpopulations RGC1, RGC2, RGC3 and RPE. Loci were labelled as significant if  $FDR < 5 \times 10^{-8}$ . (C) Relationships between genotype and mean gene expression at some of the most significant shared eQTL loci in RGC subpopulations - *CNOT6L*, *DNAJA1* and *TSPAN2*, in comparison to corresponding loci within RPE subpopulation.

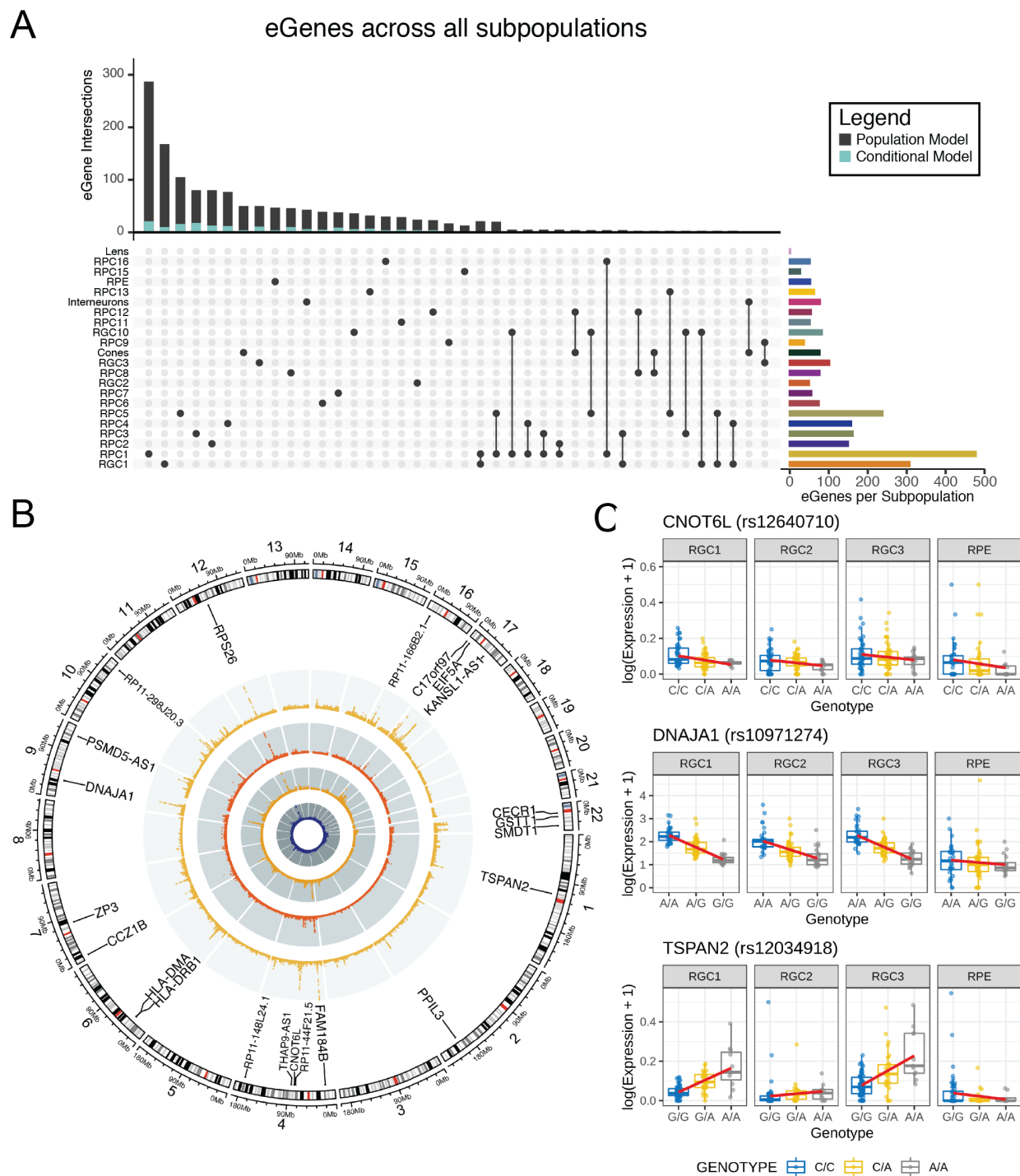
**Figure 3. Disease-associated cis-eQTLs specific to cell subpopulations.** SNP-by-disease interactions were tested for all SNPs within 1 MBP of each gene for each cell subpopulation. (A) The mean gene expression of the following genes: *SPP1*, *IGFBPL1* and *SAR1A* were grouped by disease and genotype and plotted as boxplots. Regression lines for each condition represent the relationship between genotype and expression. (B) Locus zoom plots focus on regions  $\pm 30,000$ bp around locus and genes plotted in the previous panel. FDR values for the significance of eQTL within these regions are plotted by location, alongside the ideograms of genes within this region.

**Figure 4. Prioritization of glaucoma risk genes.** (A) Volcano plot displaying genes differentially expressed between POAG cases and healthy controls (B) Manhattan plot displaying the transcriptome-wide association analysis (TWAS) from the scRNA-seq data.

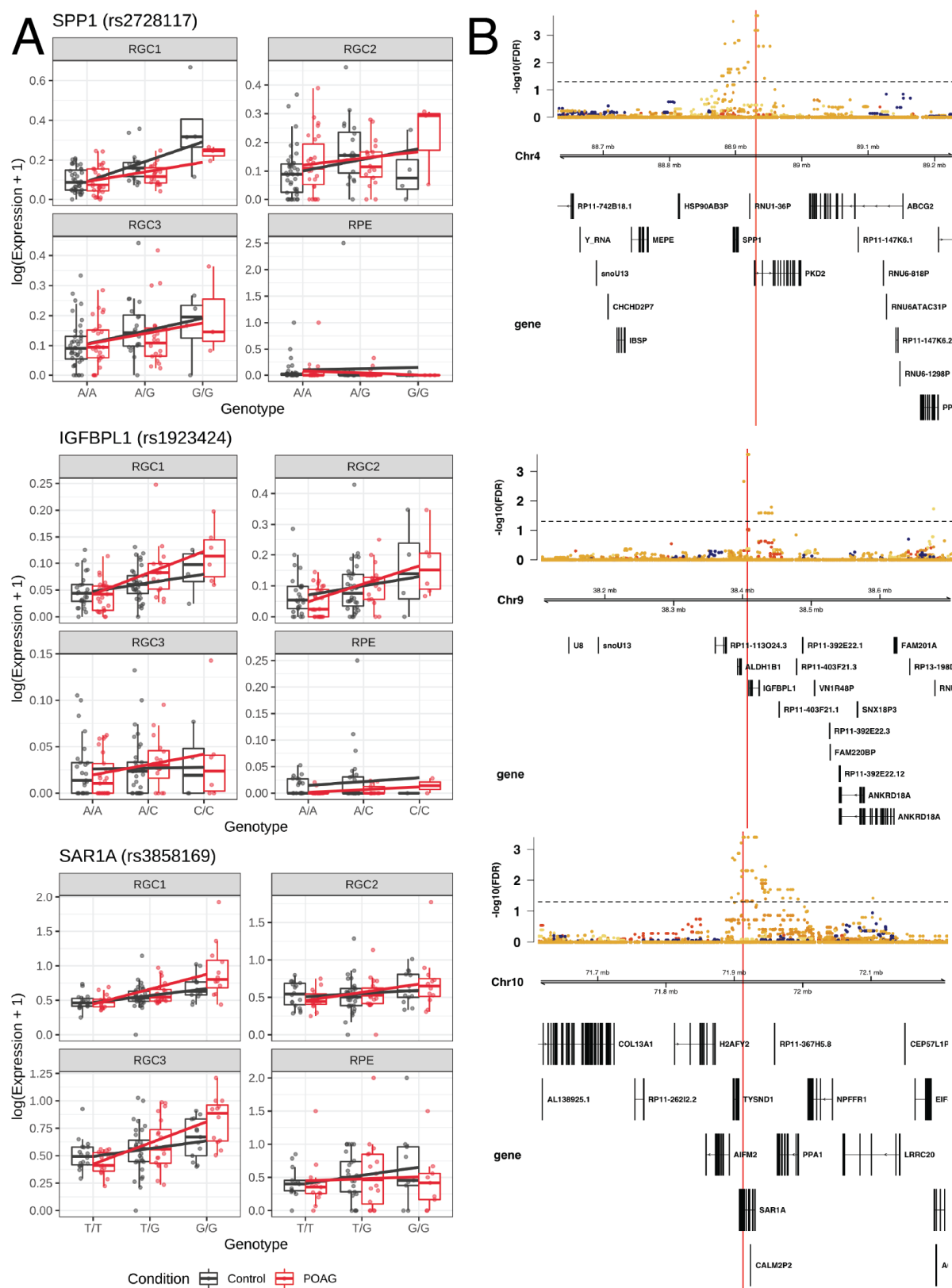




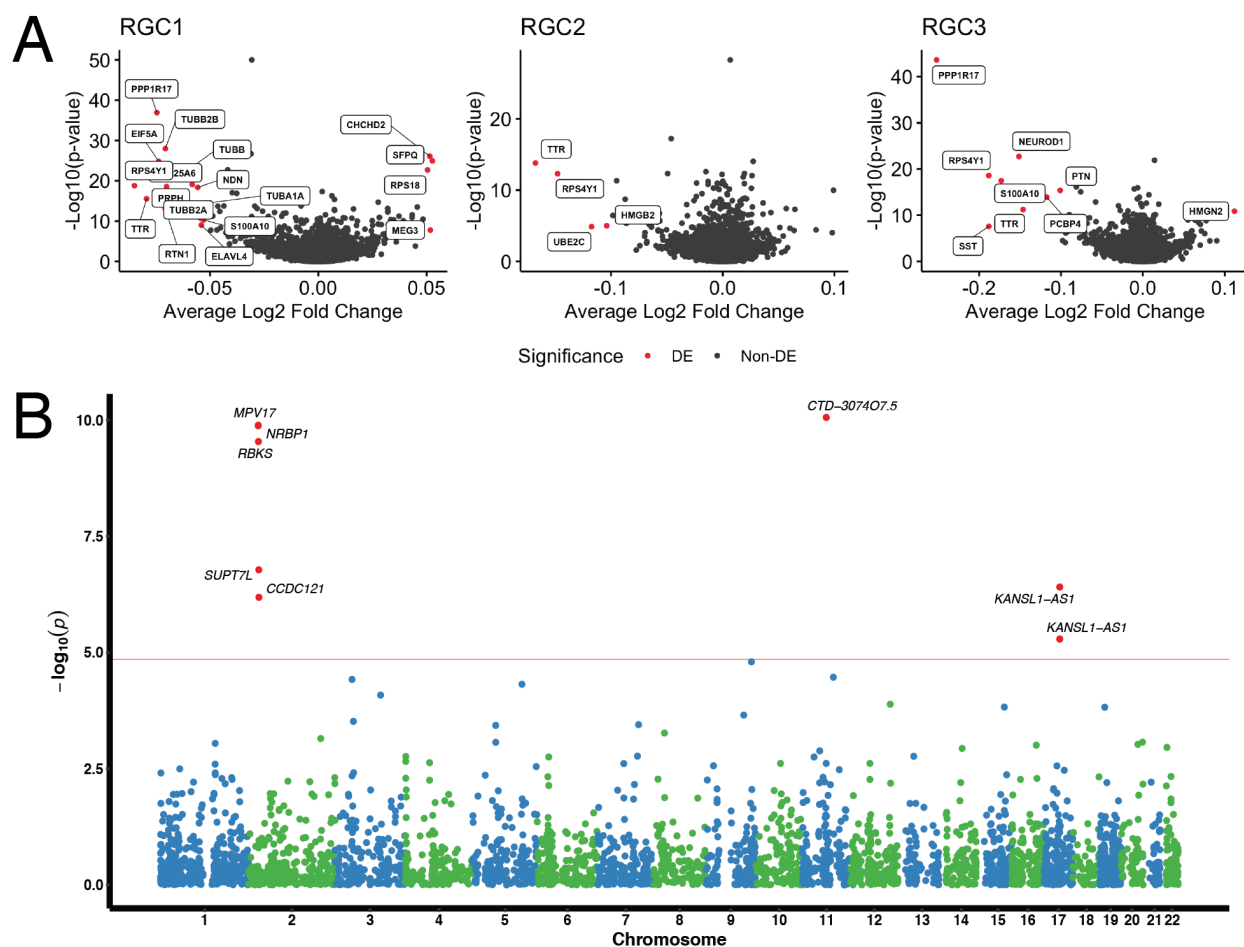
**Figure 2**



**Figure 3**



**Figure 4**



# TABLES

**Table 1. Percentage breakdown of captured cell population.** Cells were initially grouped into subpopulations based on Louvain-based clustering. These subpopulations were then annotated using gene markers associated with each known cell type associated with human optic cups.

Cell types	Proportion (%)
Retinal progenitor cells	77.4
Retinal ganglion cells	17.0
Retinal pigmented epithelium cells	1.27
Photoreceptors/ bipolar cells	2.56
Interneurons	1.66
Lens cells	0.04

**Table 2. Breakdown of significant cis-eQTLs detected in the full cohort and by disease-status conditional tests.** The relationship between genotype and expression was tested at loci within 1MB of each gene, using four different models. Population and disease models were tested using mean expression of all donors, and the population model was used to test control and POAG donors separately. eQTL were labelled significant using the threshold of FDR < 0.05 and gene expression in at least 30% of the donors tested.

Model	Number of significant eQTLs	Number of significant eGenes	Number of significant eSNPs
Population	2,235	1,447	1,861
Disease	2,175	1,395	1,809
Control only population	1,025	804	903
POAG only population	874	693	770

# METHODS

## Participant recruitment.

All participants gave informed written consent<sup>78</sup>. All experimental work was approved by the Human Research Ethics committees of the Royal Victorian Eye and Ear Hospital (11/1031H, 13/1151H-004), University of Melbourne (1545394), University of Tasmania (H0014124) and the University of Western Australia (RA/4/1/5255) as per the requirements of the National Health & Medical Research Council of Australia (NHMRC) and in accordance with the Declarations of Helsinki. We recruited a large cohort of patients with POAG and sex-, ethnically, age-matched individuals, through the Glaucoma Inheritance Study in Tasmania and the Australian and New Zealand Registry of Advanced Glaucoma, local ophthalmic clinics and adjunct studies (mean  $\pm$  SD age: 59.7  $\pm$  10.3 years at diagnosis for case subjects; 70.1  $\pm$  8.8 years at recruitment for controls). To refine translational significance, recruitment of patients with POAG focused on a) end-stage disease, which carries the greatest visual morbidity: virtually the entire optic nerve function is lost through RGC degeneration; b) patients designated as having Normal Tension Glaucoma, as they have RGCs with an increased susceptibility to degeneration compared to cases with trabecular dysfunction or a very high intraocular pressure; c) “super selection” of the most severe POAG cases. Case-inclusion criteria were: in the worst eye a vertical cup:disc ratio  $>0.95$  and a best-corrected visual acuity worse than 6/60 due to POAG or on a reliable Humphrey Visual Field a mean deviation of  $\leq -22$ dB; or at least 2 out of 4 central squares involved with a Pattern Standard Deviation of  $<0.5\%$ . The maximum documented pre-treatment IOP, measured by Goldmann applanation tonometry, was recorded. To fulfil a standard clinical diagnosis of Normal Tension POAG the maximum-recorded IOP was required to be  $<22$  mmHg. Clinical-exclusion criteria were signs of secondary or syndromic glaucoma. For each control participant, a complete ophthalmologic evaluation (incorporating automated visual field testing, fundus and optic disc imaging, corneal pachymetry) was performed. An ophthalmic history was obtained, with questions centred on age at diagnosis, family history, surgical intervention for glaucoma or cataract, macular degeneration, retinal

detachment, and refractive surgery. Control subjects, with no known family history of glaucoma, and who had normal IOP, optic discs (optical coherence tomography retinal nerve fibre layer analysis within age-matched normal limits), and visual fields, were selected for analysis.

## **Fibroblast culture**

Skin biopsies were obtained from non-sun exposed regions using a 3mm<sup>2</sup> dermal punch. Fibroblasts were expanded, cultured and banked in DMEM with high glucose, 10% fetal bovine serum (FBS), L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin (all from Thermo Fisher Scientific, USA) as described previously<sup>79</sup>. All cell lines were mycoplasma-free. Fibroblasts at passage (p) 2 were used for reprogramming.

## **Generation, selection and iPSC maintenance**

A TECAN liquid-handling platform was used to maintain and passage iPSCs, as described in<sup>26</sup>. iPSCs were generated by nucleofection with episomal vectors expressing *OCT-4*, *SOX2*, *KLF4*, *L-MYC*, *LIN28* and shRNA against *p53*<sup>80</sup> in feeder- and serum-free conditions using TeSR™-E7™ medium as described previously<sup>27</sup>. The reprogrammed cells were maintained on the automated platform using TeSR™-E7™ medium, with daily medium change. Pluripotent cells were selected by sorting anti-human TRA-1-60 Microbeads<sup>27</sup>. Cell number was determined; cells were subsequently plated onto vitronectin XF™-coated plates and in StemFlex™ medium. Subsequent culturing was performed on the automated platform using StemFlex™ medium, which was changed every 2-3 days. Passaging was performed weekly on the automated platform using ReLeSR™ onto vitronectin XF™-coated plates as described in<sup>27</sup>. TRA-1-60 quantifications were performed on a MACSQuant® Analyzer 10 immediately prior to passaging to fresh plates as described<sup>27</sup>. Pluripotency was assessed by expression of the markers OCT3/4 (sc-5279, Santa Cruz

Biotechnology) and TRA-1-60 (MA1-023-PE, Thermo Fisher Scientific, USA) by immunocytochemistry. Virtual karyotyping was undertaken using Illumina Human Core Exome or UK Biobank Axiom™ Arrays, as described in <sup>26</sup>. The iPSC lines FSA0001, FSA0002, FSA0005, FSA0006, IST2168, IST2607, MBE1006, MBE2900, MBE2906, MBE2909, TOB0199, TOB0224, TOB0412, TOB0421, TOB0431, TOB0435, TOB0474, WAB0450, WAB0069 were characterised in <sup>26</sup>.

### **Differentiation of iPSCs into retinal organoids**

Retinal organoids were generated following <sup>17</sup> with modifications. Briefly, on day 28 formed retinal organoids with surrounding tissue were excised using a 21G needle. They were maintained in suspension culture for 7 days in PRO medium (DMEM/F12, 1:1, L-glutamine, 1% Non-essential amino acids, Penicillin-Streptomycin 10,000 U/ml) supplemented with B27 and FGF2 (10 ng/ml). On day 35, organoids were transferred to Matrigel-coated 24-well tissue culture plates and maintained for 14 days in NDM medium (Neurobasal, 1% MEM non-essential amino acids, 1% GlutaMAX, 1% glucose (45%), Penicillin-Streptomycin 10,000 U/ml) with 2% B27 and 1% N2 added fresh. Medium was changed every 2-3 days. Optic cups were dissociated with Papain Dissociation System following the manufacturer's instructions. Briefly, cells were harvested with papain (20 U/mL) and DNase I (2,000 U/mL) for 30 minutes at 37 °C. Subsequently, NDM medium was added at a 1:1 ratio and cells were gently triturated with a P1000 pipette followed by centrifugation (5 minutes, 300 g, 4 °C). Cells were resuspended in 1 mL of 0.1% BSA in PBS solution. Subsequently, cells were counted and assessed for viability with Trypan Blue using a Countess II automated counter, then pooled at a concentration of 1000 cells/  $\mu$ L ( $1 \times 10^6$  cells/mL). Final cell viability estimates ranged between 79-99%.

### **Transcriptome profiling of single cells from retinal organoids and cell-based quality control**



Multi-donor single-cell suspensions were prepared for scRNA-seq using the Chromium Single Cell 3' Library & Gel bead kit (10x Genomics; PN-120237). Each pool was loaded onto individual wells of 10x Genomics Single Cell A Chip along with the reverse transcription (RT) master mix to generate single-cell gel beads in emulsion (GEMs). Reverse transcription was performed using a C1000 Touch Thermal Cycler with a Deep Well Reaction Module (Bio-Rad) as follows: 45 min at 53 °C; 5 min at 85 °C; hold 4 °C. cDNA was recovered and purified with DynaBeads MyOne Silane Beads (Thermo Fisher Scientific; catalog no. 37002D). Subsequently, it was amplified as follows: for 3 min at 98°C; 12× (for 15 sec at 98 °C; for 20 sec at 67 °C; for 60 sec at 72°C); for 60 sec at 72 °C; hold 4 °C followed recommended cycle number based on targeted cell number. Amplified cDNA was purified with SPRIselect beads (Beckman Coulter; catalog no. B23318) and underwent quality control following manufacturer's instructions. Sequencing libraries for each pool were labelled with unique sample indices (SI) and combined for sequencing across two 2 x 150 cycle flow cells on an Illumina NovaSeq 6000 (NovaSeq Control Software v1.6) using S4 Reagent kit (catalog no. 20039236). Raw base calls from the sequencer then underwent demultiplexing, quality control, mapping and quantification with the Cell Ranger Single Cell Software Suite 3.1.0 by 10x Genomics (<https://www.10xgenomics.com/>). Processed reads from the sequencer were mapped to the *Homo sapiens* reference *hg19/GRCh37* from ENSEMBL (release 75), and the pipeline was run using the estimated cell count value of 20,000. scRNA-seq data from each pool underwent quality control separately in R using the *ascend* package<sup>69</sup>. Cells were removed if they did not meet thresholds calculated from 3 Median Absolute Deviations (MAD) of the following statistics: total Unique Molecular Identifier (UMI) counts, number of detected genes, and fraction of mitochondrial and ribosomal transcripts to total expression.

### **SNP genotype analysis and imputation**

DNA was extracted from cell pellets using QIAamp DNA Mini Kit (QIAGEN, 51306) following the manufacturer's instructions. DNA concentration was determined with

SimpliNano spectrophotometer (GE Life Sciences), diluted to a final concentration of 10-15 ng/μl and genotyped on UK Biobank Axiom™ Arrays at the Ramaciotti Centre for Genomics, Sydney, Australia. Samples previously screened using Illumina arrays, were re-genotyped on UK Biobank Axiom™ Arrays. Genotype data were exported into PLINK data format using GenomeStudio PLINK Input Report Plug-in v2.1.4 and screened for SNP and individual call rates (<0.97), HWE failure ( $P < 10^{-6}$ ), and MAF (<0.01)<sup>81</sup>. Samples with excess autosomal heterozygosity or with sex-mismatch were excluded. In addition, a genetic relationship matrix from all the autosomal SNPs were generated using the GCTA tool and one of any pair of individuals with estimated relatedness larger than 0.125 were removed from the analysis<sup>82</sup>. Individuals with non-European ancestry were excluded outside of an “acceptable” box of +/- 6SD from the European mean in PC1 and PC2 in a SMARTPCA analysis. The 1000G Phase 3 population was used to define the axes, and the samples were projected onto those axes. Imputation was performed on each autosomal chromosome using the Michigan Imputation Server with the Haplotype Reference Consortium panel (HRC r1.1 2016) and run using Minimac3 and Eagle v2.3<sup>67,68</sup>. Only SNPs with INFO > 0.8 were retained.

### **Demultiplexing of cell pools into individual donors**

*demuxlet* v1.0<sup>70</sup> and *scrublet* v0.20<sup>71</sup> were used to demultiplex cells from mixed-donor pools using transcriptome and genotype data. Each pool was demultiplexed separately. *demuxlet* was run with exon-only SNPs and the following arguments: “*--field GP, --geno.error = 0.01, --min-mac 1, --min-callrate 0.5, alpha = 0.5, doublet-prior = 0.5*”. Cells were initially assigned a putative donor based on the maximum likelihood of reads from scRNA-seq overlapping sets of unique variants (SNPs) mapped by genotyping. *scrublet* was then used to confirm if a cell was a neotypic doublet by building a simulation of doublets based on sampled transcriptome data and scoring the cell based on its neighbors in k-means nearest neighbor graph. A cell was designated a singlet if *scrublet* agrees, and if the posterior probability of it being a singlet in *demuxlet* is greater

than 0.99. Unassigned donors, doublets and cells with ambiguous assignments were omitted from downstream analyses.

### **Aggregation, normalization and dimensionality reduction of scRNA-seq datasets**

The unfiltered count matrices of all batches were combined into one dataset using the *cellranger aggr* pipeline. This pipeline equalized the read depth of all batches by downsampling reads from higher-depth libraries to match the lowest depth library<sup>83</sup>. Cells that had been removed from single-batch analyses due to low quality, being labeled as a doublet, or with conflicting assignments, were also removed from the combined expression matrix. The SCTransform function from Seurat (v3.0.2) was applied to the filtered count matrix to perform cell-cell and batch normalization<sup>84</sup>. The fraction of mitochondrial and ribosomal expression of total expression was regressed out as a part of this step, and the top 3000 most variable genes were used to calculate residuals. The residuals were then reduced to 30 principal components (PCs) using PC Analysis (PCA). These PCs were reduced further into two dimensions, for visualization and clustering via Uniform Manifold Approximation Projection (UMAP)<sup>75</sup>.

### **Identification and annotation of cell subpopulations**

Graph-based clustering via the Louvain algorithm that was implemented in Seurat was used to identify cell subpopulations<sup>85,86</sup>. First, cell-cell Euclidean distances calculated from PCs were projected onto a K-means nearest-neighbor graph (KNN). Next, the Louvain algorithm was implemented at resolutions between 0 and 1, at 10 equal intervals of 0.1. Finally, the movement of cells between subpopulations at these resolutions were visualized on to a *clustree* plot (Figure S2A) as implemented in the *clustree* R package<sup>87</sup>. Regions of stability were identified from the plot, and the resolution where this region began - 0.4, was selected as the optimal resolution. Cells were divided into 22 subpopulations at this resolution. Cells that could not be assigned to a subpopulation - singletons, were assigned to a group designated 'Cluster 0'. To annotate each subpopulation, the combined Likelihood Ratio Test (LRT) as described

by <sup>88</sup> was applied to normalized, log-transformed UMI counts. Disease-specific markers within subpopulations were identified using the Wilcoxon Rank Sum test <sup>89</sup>. Genes were classified as markers if they were differentially expressed, based on the thresholds of average  $\log_2$  fold change  $> |0.25|$  and Bonferroni-corrected p-value  $< 0.01$ .

### **Identification of differentiation lineages via pseudotime analysis**

Differentiation lineages were identified using the *slingshot* R package <sup>76</sup>. Singletons were excluded from the trajectory, and the UMAP matrix was used as input for the 'slingshot' function. RPC1 was selected as the root of the trajectories due to the expression of the proliferative marker MKI67, and the end-points of trajectories were defined as the relatively mature cell subpopulations: RGC1, Cones, Interneurons, RPE and Lens.

### **Identification of cis-eQTL using transcriptome and genotype data**

We investigated cis-eQTL associated with each subpopulation using the R package Matrix eQTL <sup>90</sup>. To prepare scRNA-seq data for this stage of analysis, the normalized, corrected UMI count matrix was split by subpopulation. The mean expression of each gene, for all cells within the subpopulation from a donor, was used to build a gene-donor count matrix. Genes that were expressed in less than 30% of the donor population were excluded, and remaining mean counts were transformed ( $\log(x+1)$ ). SNPs with a Minor Allele Frequency (MAF) of less than 10% were filtered from the genotype dataset. SNPs within one mega-base pairs of a gene were tested. To identify lead eQTL based on population, Matrix eQTL was run with an additive linear model using sex, age and the top six genotype principal components as covariates. To identify eQTL specific to POAG, disease status was included in the model. eQTL were determined to be significant based on a threshold of FDR  $< 0.05$ . Lead eQTLs identified using the disease model underwent additional testing using the *lm* function from R to measure the size and significance of interactions between genotype and disease status.

## Transcriptome wide association study analysis

Transcriptome-wide association study analysis (TWAS) was performed using summary statistics generated by eQTL analysis. We customized our prediction models based on the online tutorial (available through <https://github.com/hakyimlab/MetaXcan>) to evaluate the association between predicted gene expression and glaucoma risk using GWAS summary statistics. The gene-expression prediction models were constructed from the single-cell eQTL data using the following models: “blup”, “lasso”, “top1” and “enet”. The GWAS summary statistics from a recent multitrait meta-analysis of glaucoma were used <sup>8</sup>. S-PrediXcan was used to harmonize the GWAS summary statistics and to compute the gene-level association results using each subpopulation of single-cell expression data and glaucoma GWAS summary statistics <sup>91</sup>. A total of 3,573 genes (across tested subpopulations) were tested for glaucoma. To correct for multiple testing, we adjusted the gene-level association P values using the Bonferroni correction method ( $0.05 / (\text{total number of tests across all subpopulations})$ ). We then compared the transcriptome-wide association results based on scRNA-seq data to bulk RNA-seq data. The bulk retinal transcriptome data were described previously <sup>60</sup>. Briefly, 406 controls and age-related macular degeneration cases that passed RNA-seq and genotyping quality control were modeled with mixed models, LASSO, or elastic net according to Gusev et al <sup>92</sup>. The effect sizes from these models were used as weights to calculate the gene-trait associations.

# **SUPPLEMENTAL INFORMATION TITLES AND LEGENDS**

**Figure S1. Identification and characterisation of cell subpopulations.**

**Figure S2. Variation in cell differentiation between control and POAG samples.**

**Figure S3. Pseudotime analysis across lineages and sub-lineages.**

**Figure S4. Boxplots displaying the relationships between genotype and mean gene expression for all eQTLs in RGC subpopulations with a  $FDR < 5 \times 10^{-8}$ .**

**Table S1. Summary of single cell RNA-sequencing metrics.**

**Table S2. Cell-donor deconvolution summary metrics.**

**Table S3. Subpopulation-specific markers identified by differential expression gene analysis.**

**Table S4. Number of significant cis-eQTLs across cell subpopulations.**

**Table S5. Lead cis-eQTLs from RGC subpopulations.**

**Table S6. Statistically significant SNP-by-disease state interaction eQTLs in the RGC lineage specific subpopulations.**

**Table S7. Disease-specific differentially-expressed genes in retinal ganglion cells.**

**Supplemental Results. Description of delineated cell subpopulations.**

## REFERENCES

1. Tham, Y.-C. *et al.* Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040. *Ophthalmology* vol. 121 2081–2090 (2014).
2. Quigley, H. A. Open-angle glaucoma. *N. Engl. J. Med.* **328**, 1097–1106 (1993).
3. Garway-Heath, D. F. *et al.* Latanoprost for open-angle glaucoma (UKGTS): a randomised, multicentre, placebo-controlled trial. *Lancet* **385**, 1295–1304 (2015).
4. Polubriaginof, F. C. G. *et al.* Disease Heritability Inferred from Familial Relationships Reported in Medical Records. *Cell* **173**, 1692–1704.e11 (2018).
5. Wang, K., Gaitsch, H., Poon, H., Cox, N. J. & Rzhetsky, A. Classification of common human diseases derived from shared genetic and environmental determinants. *Nature Genetics* vol. 49 1319–1325 (2017).
6. Stone, E. M. *et al.* Identification of a gene that causes primary open angle glaucoma. *Science* **275**, 668–670 (1997).
7. Rezaie, T. *et al.* Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science* **295**, 1077–1079 (2002).
8. Craig, J. E. *et al.* Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nat. Genet.* **52**, 160–166 (2020).
9. Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93**, 779–797 (2013).
10. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts



- by defined factors. *Cell* **131**, 861–872 (2007).
11. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
  12. Gill, K. P. *et al.* Enriched retinal ganglion cells derived from human embryonic stem cells. *Sci. Rep.* **6**, 30552 (2016).
  13. Ohlemacher, S. K. *et al.* Stepwise Differentiation of Retinal Ganglion Cells from Human Pluripotent Stem Cells Enables Analysis of Glaucomatous Neurodegeneration. *Stem Cells* **34**, 1553–1562 (2016).
  14. Gill, K. P., Hewitt, A. W., Davidson, K. C., Pébay, A. & Wong, R. C. B. Methods of Retinal Ganglion Cell Differentiation From Pluripotent Stem Cells. *Transl. Vis. Sci. Technol.* **3**, 7 (2014).
  15. Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10**, 771–785 (2012).
  16. Zhong, X. *et al.* Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs. *Nat. Commun.* **5**, 4047 (2014).
  17. Reichman, S. *et al.* Generation of Storable Retinal Organoids and Retinal Pigmented Epithelium from Adherent Human iPS Cells in Xeno-Free and Feeder-Free Conditions. *Stem Cells* **35**, 1176–1188 (2017).
  18. Fligor, C. M. *et al.* Three-Dimensional Retinal Organoids Facilitate the Investigation of Retinal Ganglion Cell Development, Organization and Neurite Outgrowth from Human Pluripotent Stem Cells. *Sci. Rep.* **8**, 14520 (2018).
  19. Capowski, E. E. *et al.* Reproducibility and staging of 3D human retinal organoids

- p across multiple pluripotent stem cell lines.
- Development*
- 146**
- , (2019).
20. Völkner, M. *et al.* Retinal Organoids from Pluripotent Stem Cells Efficiently Recapitulate Retinogenesis. *Stem Cell Reports* **6**, 525–538 (2016).
  21. Arno, G. *et al.* Mutations in REEP6 Cause Autosomal-Recessive Retinitis Pigmentosa. *Am. J. Hum. Genet.* **99**, 1305–1315 (2016).
  22. Tucker, B. A. *et al.* Duplication of TBK1 Stimulates Autophagy in iPSC-derived Retinal Cells from a Patient with Normal Tension Glaucoma. *J. Stem Cell Res. Ther.* **3**, 161 (2014).
  23. Li, K. *et al.* HiPSC-derived retinal ganglion cells grow dendritic arbors and functional axons on a tissue-engineered scaffold. *Acta Biomater.* **54**, 117–127 (2017).
  24. Wills, Q. F. *et al.* Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nat. Biotechnol.* **31**, 748–752 (2013).
  25. Kolodziejczyk, A. A. *et al.* Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell Stem Cell* **17**, 471–485 (2015).
  26. Daniszewski, M. *et al.* Single-Cell Profiling Identifies Key Pathways Expressed by iPSCs Cultured in Different Commercial Media. *iScience* **7**, 30–39 (2018).
  27. Crombie, D. E. *et al.* Development of a Modular Automated System for Maintenance and Differentiation of Adherent Human Pluripotent Stem Cells. *SLAS Discov* **22**, 1016–1025 (2017).
  28. Hoshino, A. *et al.* Molecular Anatomy of the Developing Human Retina. *Developmental Cell* vol. 43 763–779.e4 (2017).

29. Clark, B. S. *et al.* Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification. *Neuron* **102**, 1111–1126.e5 (2019).
30. Sridhar, A. *et al.* Single-Cell Transcriptomic Comparison of Human Fetal Retina, hPSC-Derived Retinal Organoids, and Long-Term Retinal Cultures. *Cell Rep.* **30**, 1644–1659.e4 (2020).
31. Lu, Y. *et al.* Single-cell analysis of human retina identifies evolutionarily conserved and species-specific mechanisms controlling development. *Neuroscience* **121** (2019).
32. Ericson, J. *et al.* Pax6 Controls Progenitor Cell Identity and Neuronal Fate in Response to Graded Shh Signaling. *Cell* **90**, 169–180 (1997).
33. Graham, V., Khudyakov, J., Ellis, P. & Pevny, L. SOX2 functions to maintain neural progenitor identity. *Neuron* **39**, 749–765 (2003).
34. Zibetti, C., Liu, S., Wan, J., Qian, J. & Blackshaw, S. Epigenomic profiling of retinal progenitors reveals LHX2 is required for developmental regulation of open chromatin. *Commun Biol* **2**, 142 (2019).
35. de Melo, J. *et al.* Lhx2 Is an Essential Factor for Retinal Gliogenesis and Notch Signaling. *J. Neurosci.* **36**, 2391–2405 (2016).
36. Langer, K. B. *et al.* Retinal Ganglion Cell Diversity and Subtype Specification from Human Pluripotent Stem Cells. *Stem Cell Reports* **10**, 1282–1293 (2018).
37. Rheaume, B. A. *et al.* Single cell transcriptome profiling of retinal ganglion cells identifies cellular subtypes. *Nat. Commun.* **9**, 2759 (2018).

38. Hu, Y. *et al.* Dissecting the transcriptome landscape of the human fetal neural retina and retinal pigment epithelium by single-cell RNA-seq analysis. *PLoS Biol.* **17**, e3000365 (2019).
39. Mu, X., Fu, X., Beremand, P. D., Thomas, T. L. & Klein, W. H. Gene regulation logic in retinal ganglion cell development: Isl1 defines a critical branch distinct from but overlapping with Pou4f2. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6942–6947 (2008).
40. Pan, L., Deng, M., Xie, X. & Gan, L. ISL1 and BRN3B co-regulate the differentiation of murine retinal ganglion cells. *Development* **135**, 1981–1990 (2008).
41. Wu, F. *et al.* Two transcription factors, Pou4f2 and Isl1, are sufficient to specify the retinal ganglion cell fate. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1559–68 (2015).
42. Vicente, C. *et al.* The CCR4-NOT complex is a tumor suppressor in *Drosophila melanogaster* eye cancer models. *J. Hematol. Oncol.* **11**, 108 (2018).
43. Pemberton, S. *et al.* Hsc70 protein interaction with soluble and fibrillar alpha-synuclein. *J. Biol. Chem.* **286**, 34690–34699 (2011).
44. de Monasterio-Schrader, P. *et al.* Uncoupling of neuroinflammation from axonal degeneration in mice lacking the myelin protein tetraspanin-2. *Glia* **61**, 1832–1847 (2013).
45. 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).
46. Chowdhury, U. R., Jea, S.-Y., Oh, D.-J., Rhee, D. J. & Fautsch, M. P. Expression profile of the extracellular matrix protein osteopontin in primary open-angle glaucoma and

- the normal human eye. *Invest. Ophthalmol. Vis. Sci.* **52**, 6443–6451 (2011).
47. Ruzafa, N., Pereiro, X., Aspichueta, P., Araiz, J. & Vecino, E. The Retina of Osteopontin deficient Mice in Aging. *Mol. Neurobiol.* **55**, 213–221 (2018).
  48. Chang, S. W., Kim, H. I., Kim, G. H., Park, S. J. & Kim, I.-B. Increased Expression of Osteopontin in Retinal Degeneration Induced by Blue Light-Emitting Diode Exposure in Mice. *Front. Mol. Neurosci.* **9**, (2016).
  49. Guo, C. *et al.* IGFBPL1 Regulates Axon Growth through IGF-1-mediated Signaling Cascades. *Sci. Rep.* **8**, 1–13 (2018).
  50. Aridor, M. & Fish, K. N. Selective targeting of ER exit sites supports axon development. *Traffic* **10**, 1669–1684 (2009).
  51. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
  52. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151**, 1617–1632 (2012).
  53. Reynolds, M. M. *et al.* Ocular Manifestations of Familial Transthyretin Amyloidosis. *Am. J. Ophthalmol.* **183**, 156–162 (2017).
  54. Gharahkhani, P., Jorgenson, E., Hysi, P. & Khawaja, A. P. A large cross-ancestry meta-analysis of genome-wide association studies identifies 69 novel risk loci for primary open-angle glaucoma and includes a genetic link .... *BioRxiv* (2020).
  55. Wainberg, M. *et al.* Opportunities and challenges for transcriptome-wide association studies. *Nat. Genet.* **51**, 592–599 (2019).
  56. Gusev, A. *et al.* Transcriptome-wide association study of schizophrenia and

- chromatin activity yields mechanistic disease insights. *Nat. Genet.* **50**, 538–548 (2018).
57. Spinazzola, A. *et al.* MPV17 encodes an inner mitochondrial membrane protein and is mutated in infantile hepatic mitochondrial DNA depletion. *Nat. Genet.* **38**, 570–575 (2006).
  58. McMonnies, C. Reactive oxygen species, oxidative stress, glaucoma and hyperbaric oxygen therapy. *J. Optom.* **11**, 3–9 (2018).
  59. Ramdas, W. D. *et al.* A genome-wide association study of optic disc parameters. *PLoS Genet.* **6**, e1000978 (2010).
  60. Ratnapriya, R. *et al.* Retinal transcriptome and eQTL analyses identify genes associated with age-related macular degeneration. *Nat. Genet.* **51**, 606–610 (2019).
  61. Aldiri, I. *et al.* The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron* **94**, 550–568.e10 (2017).
  62. Mao, X. *et al.* Single-Cell RNA Sequencing of hESC-Derived 3D Retinal Organoids Reveals Novel Genes Regulating RPC Commitment in Early Human Retinogenesis. *Stem Cell Reports* **13**, 747–760 (2019).
  63. Kim, S. *et al.* Generation, transcriptome profiling, and functional validation of cone-rich human retinal organoids. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 10824–10833 (2019).
  64. Collin, J. *et al.* Deconstructing Retinal Organoids: Single Cell RNA-Seq Reveals the

- Cellular Components of Human Pluripotent Stem Cell-Derived Retina. *Stem Cells* **37**, 593–598 (2019).
65. Vecino, E. & Acera, A. Development and programmed cell death in the mammalian eye. *Int. J. Dev. Biol.* **59**, 63–71 (2015).
  66. Cook, D. *et al.* Lessons learned from the fate of AstraZeneca’s drug pipeline: a five-dimensional framework. *Nat. Rev. Drug Discov.* **13**, 419–431 (2014).
  67. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
  68. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
  69. Senabouth, A. *et al.* ascend: R package for analysis of single-cell RNA-seq data. *Gigascience* **8**, (2019).
  70. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).
  71. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281–291.e9 (2019).
  72. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
  73. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
  74. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**,



- 1888–1902.e21 (2019).
75. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* (2018) doi:10.1038/nbt.4314.
76. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
77. Van den Berge, K. *et al.* Trajectory-based differential expression analysis for single-cell sequencing data. doi:10.1101/623397.
78. McCaughey, T. *et al.* An Interactive Multimedia Approach to Improving Informed Consent for Induced Pluripotent Stem Cell Research. *Cell Stem Cell* **18**, 307–308 (2016).
79. Aasen, T. *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* **26**, 1276–1284 (2008).
80. Okita, K. *et al.* A more efficient method to generate integration-free human iPS cells. *Nat. Methods* **8**, 409–412 (2011).
81. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
82. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
83. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *bioRxiv* 065912 (2016) doi:10.1101/065912.
84. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**,

296 (2019).

85. Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
86. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
87. Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, (2018).
88. McDavid, A. *et al.* Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* **29**, 461–467 (2013).
89. Hollander, M., Wolfe, D. A. & Chicken, E. *Nonparametric Statistical Methods*. (John Wiley & Sons, 2013).
90. Shabalin, A. A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).
91. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).
92. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).