Cell Genomics Transparent Peer Review Record

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

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

Summary

Initial submission: Received : September 19, 2020

Scientific editor: Orli Bahcall and Rita Gemayel

First round of review: Number of reviewers: 3
Revision invited : December 17, 2020
Revision received : March 8, 2021

Second round of review: Number of reviewers: 2
Revision invited : January 26, 2022
Revision received : January 28, 2022

Third round of review: Number of reviewers: 1
Accepted : May 11, 2022

Data freely available: Yes

Code freely available: Yes

Referees' reports, first round of review

Reviewer #1 (Comments to authors)

The authors have performed a technical tour-de-force to try to understand differences between glaucoma cases and controls gene expression, using iPSCs given the unavailability of appropriate human tissue, and are to be congratulated. While they have compressed a huge amount of work into a readable manuscript, at times it was not easy to follow through, and I hope my comments below will improve the manuscript. In particular, the results for each set of experiments do not particularly reference previous results- for example CDKN2B has a SNP allelic effect with respect to glaucoma status, but is not mentioned again- so I assume has no differential expression? Similarly the TWAS results are mildly disappointing- in terms of refining known loci there are only 3 loci identified in the end, and none of the better-known loci are discussed. Maybe a table of the authors' multitrait GWAS significant hits and what was found here for future researchers to refer to? It would be good to have a better flow-through of information.

1. I would like clarification about the recruitment. The advanced glaucoma is clearly a great phenotype, but it is unclear from the description whether the subjects recruited ALL had end-stage disease, all had normal tension glaucoma- using the term recruitment "focused" on a) end stage b) NTG makes it unclear. Similarly the age of diagnosis of cases is given, and the age at recruitment of controls, but not the age of recruitment of cases (which may have bearing on later comments)
2. I would like some clarification of methods- in the first paragraph of results, the authors talk of batches, pooling, 160 cell lines, SNPs from 162 individuals, 23 failures, 110 donor cell lines (and 55+57 human subjects). It is unclear to me how these all relate and where 162 individuals come from and exactly what was pooled. How many cases and controls were left? Could this be in some sort of flow chart?
3. There were 17% RGCs, 77% RPCs and fewer other cells, as expected. Were these proportions similar in cases and controls?
4. The study identified 58 e-genes and refers to Fig 2B/C where there are fewer names of genes shown. Are these results tabulated anywhere? And the circos plot is not fully explained in the legend as to what the different colors mean.
5. The authors on page 7 give two possible explanations for the cell type-specific eQTL detection but then discount these. Cell type specificity and correlation properties are well known (see GTEX Consortium Science 2020). Any other possible reasons? Could another possibility be random clonal differentiation with aberrant translational patterns? Apart from a few gene-markers that prove that these cells are RGC cells of some sort, how differentiated they were and how similar to actual cells (for example in animals)?
6. Given the "poster-boy" status of CDKN2B in NTG, it would have been helpful to include this in the figure. This is probably an ubiquitous effect, give Burdon et al showing allele-specific methylation of this promoter.
7. How was the FDR for the 54 eQTLs specific to RGCs actually calculated? I don't think I saw this and some have 10-2 FDRs in the table.
8. Bottom page 7 fig 3C is referenced but I assume you mean 3A? And you refer to A alleles but the figure illustrates T/G alleles.
9. Table S7 is referenced after 3118 genes but later in same paragraph 144 genes are differentially expressed (and table looks to have around 156 lines so I assume relates to the 144 genes?). Suggest move reference to correct place.
10. Regarding differential expression, while I understand the randomisation should avoid false positives, differing ages might have some effect? Even during the natural conception 10-20% of CpG sites remain methylated. Epigenetic memory and age: there is talk about differential pluripotency depending on the age of the donor. How well age-matched were the cases and controls at skin biopsy?
11. Further comment on the TWAS results might be helpful- TWAS infers expression but this study actually has expression data- and how these relate to the authors' previous publication cited that used non-RGC expression.
12. While RGCs are the end-organ damaged in glaucoma, some comment on the importance of IOP and the fact that previous GWAS have found mainly IOP genes would be helpful to the non-expert reader.

Reviewer #2 (Comments to authors)

In this study, Daniszewski et al. use organoid technology to study the genetics of primary open angle glaucoma (POAG). The authors established iPSC lines from POAG patients and controls, then used them to derive retinal organoids, which were subjected to scRNA-seq. Transcriptional analysis enabled the authors to identify disease specific differentially expressed genes and retinal ganglion cell specific eQTLs.

The authors present a large dataset that they are depositing in the Human Cell Atlas. Though a number of sc-RNAseq studies now have been performed on retinal organoids, this dataset will be a useful resource for the field. The study would benefit from a better description of the overall rationale. Specifically, what classes of genes identified in POAG GWAS would be expected to show differential expression in retinal ganglion cells, as opposed to other putative target cell types i.e. trabecular meshwork cells. Why would gene expression in immature (fetal) RGC) be expected to correlate with that in mature cells undergoing pathology or repair? These are questions that pertain to many disease modeling studies with iPSC, particularly for degenerative diseases, but the authors should address them. It might be important that the authors focused on normal tension glaucoma (as they explain in the methods), where RGC susceptibility may be more central, and if so, this should be explained up front. The study does not permit conclusions regarding the role of newly identified RGC eQTL or DEG in the disease, but the information may be useful in conjunction with other data, and the brief discussion does not claim too much from the data.

Specific comments:

1. Page 5-authors should comment of the diversity of the population studied with respect to sex and ethnicity. How many clones from each individual were studied? How many replicate differentiations were performed? What was the overall reproducibility? Were failures to differentiate equally distributed between patients and controls?
2. Page 5-what was the rationale for harvest of organoids at the two-week timepoint. Would additional maturation have provided better information (most cells are still progenitors at this point).
3. Page 5 in the end, how many cells from how many control and patient lines were assessed
4. Page 5-Figure S2 shows very substantial variation in the proportion of RGC from patient to patient (though not in a disease related fashion). How did this variation impact on gene expression profiles?
5. How convincing is Figure 2c-two cell types and only three genes does not seem sufficient to draw general conclusions. In some cases, effects of genotype across RGC subsets are not that striking. In S4, many eQTL also operate in RPE, including KANSL1-AS1
6. Page 7-what is the basis for the statement that a number of RGC eQTL are directly involved in neurogenesis or neurodegeneration. You might expect the three examples cited to turn up by chance, given the diversity of processes (and apparently species) surveyed. How about actual gene ontology analysis?
7. Pge 7 and Figure 3a-unless I have misunderstood, these eQTL in genes previously associated with POAG can certainly be detected in (most subsets of) RGC, but there seem to be no consistent differences in patient versus control cells. I suppose that this makes sense, but how is one to identify important eQTL without knowledge a priori if this is the case. I could not find the Figure 3c referred to in the text in my copy of the manuscript.
8. Page 8-what exactly is the link postulated between TTR, familial amyloidotic polyneuropathy, and POAG. This does not seem convincing.
9. Page 8-explain clearly how your multitrait glaucoma GWAS should relate to POAG? Is POAG a subset of the diseases examined in the GWAS? How exactly did you use the GWAS statistics?

Reviewer #3 (Comments to authors)

Daniszewski et al. used single cell RNA-seq (scRNA-seq), combined with eQTL analysis and GWAS, to study the associations between genetic variation and primary open-angle glaucoma (POAG). They reprogrammed fibroblasts to iPSCs from >50 POAG patients and matched controls and differentiated iPSC-derived retinal organoids. They performed scRNA-seq on ~250,000 cells and found that the vast majority of cells were retinal progenitor cells (RPCs). No difference in the cell type proportions between cases and controls was found. To determine the associations between genetic variation and gene expression, they performed an eQTL analysis on all cells and on each cell type. They found 2,235 eGenes across all cells, but they had very little power to detect eQTLs on each cell type (10-456 eGenes/cell type). To determine associations between genetic variation, gene expression and POAG, they performed TWAS, and found seven genes associated with disease, most of which were at known GWAS loci.
While this study represents a great resource to study the transcriptome of retinal organoids, genetic analyses are weak and need to be strengthened.

Major comments:
A- The authors use terminology in a confusing manner:
a. Page 4-5: "Using an additive linear model, a total of 54,786 eQTLs were found to be associated with 21,512 SNPs". Because of LD structure, reporting the number of variants whose genotype is significantly associated with gene expression is meaningless. It would be more useful to report the number of eGenes. Additionally, these numbers are not described anywhere else in the text.
b. Page 6: eQTLs underlie association, not causality. The authors cannot refer to eQTL analysis as "To

explore cell type-specific genetic control of gene expression".

B- Many of the analyses need to be substantially improved and the Methods section lacks important details:
a. Batch effects are a common and real problem in Genomic studies. In the Results section on page 5 first paragraph it is unclear how the data were generated to avoid batch effects. Sentences from this paragraph: 1. The cohort had 57 healthy and 55 POAG individuals; 2. Twenty-three lines did not differentiate; 3. 330,569 cells from 160 individual cell lines; 4. SNPs from 162 individuals; 5. scRNA-seq data assigned to 128 donors; and 6. Data from 110 donor iPSC lines used for subsequent analyses. This is impossible to follow and the methods and Table S1 do not clarify. Were cases and controls differentiated at the same time and mixed for scRNA-seq generation? If not, how were the data generated? The supplemental table should include a differentiation number and which samples were combined for scRNA-seq generation. On page 8 this is stated, "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)." But I don't see this information provided in the Methods.

b. Page 5, section "Identification and characterisation of 23 subpopulations from 253,107 cells". The authors identify 23 clusters. While this is technically correct, it would be extremely surprising that retinal organoids include 24 cell types. Indeed, they are combined into six distinct cell types. Can the authors find a clustering resolution that allows them to identify these biologically relevant cell populations? If not, what would the reason be?

c. The differential expression analysis could be more informative, considering the resources generated in this study. With ~250,000 individual cells differential expression analysis could provide invaluable information beyond the observed 3,118 genes that were differentially expressed. With single cell technologies, there are a few methods that would provide better insights into the transcriptomic differences between cases and controls. The authors should explore a few of these to strengthen their analysis:
i. Pseudotime: the authors do not find differences between cell composition of cases and controls, but this does not exclude that, overall, cases and controls may display differences in maturity. Pseudotime may be tested for eQTLs too (see Cuomo et al., Nat Comm 2020 for an example)
ii. Differential expression analysis cell type by cell type. Are there differentially expressed genes in each cell type? Is there a cell type that accounts for the most differences between cases and controls?

d. Page 6, "The genetic control of gene expression is highly cell type-specific": "We identified a total of 2,235 eQTL across all cell types, which surpassed a study-wide significance threshold of FDR < 0.05". FDR correction for the eQTL analysis should be explained. There are methods that account for LD when performing gene by gene FDR correction (eigenMT) and studies that suggest to perform a two- or three-step approach to perform FDR correction (Huang, NAR 2018: doi: 10.1093/nar/gky780). The authors do not describe how they performed FDR correction.

e. Page 6, "The genetic control of gene expression is highly cell type-specific": The authors find 2,235 eQTLs and 1,447 eGenes across all cell populations. I do not understand what the authors intend by "eQTLs". Are these all independent eQTL signals (primary and conditional eQTLs)? If so, this should be clearly stated and the methods should describe how conditional eQTLs were performed. If not, it would be very surprising to find only 1.5 eVariants/eGene, considering the LD structure of the genome. It is also surprising that the authors used a very stringent MAF cutoff for eQTLs (10%, when the usual threshold for studies with a similar size is 5%). Also, these numbers do not match those described in the Introduction (54,786 eQTLs and 21,512 SNPs).

f. Page 6, "The genetic control of gene expression is highly cell type-specific": finding that almost all eQTLs are cell type-specific is the opposite of what has been found in previous eQTL studies. For example, the latest GTEx papers show that many eQTLs are shared between different tissues, while other papers (Kim-Hellmuth, Science 2020; Donovan, Nat. Comm. 2020; Cuomo, Nat Comm. 2020; Jerber, bioRxiv 2020) have shown that only a subset of eQTLs are cell type-specific. This needs to be discussed in discussion as to why the results from this study are so different.

g. Page 7, second paragraph: "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)". The figure shows three examples of eQTL signals in four cell types, of which one (DNAJA1) supports the cell type-specificity of eQTLs, one (CNOT6L) shows an eQTL shared between RGCs and RPE, and one (TSPAN2) does not have good agreement between RGCs. Based on these examples, I do not understand how the authors can conclude that "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". The authors also provide an alternative hypothesis: "Another possible explanation for the cell type-specific eQTL is low statistical power to detect eQTL in multiple cell types" and describe a convoluted analysis that, in their opinion, does not support the hypothesis, but they do not provide any supporting data.

h. Page 7, last paragraph and its corresponding Methods section (page 26-27: "Identification of cis-eQTL using transcriptome and genotype data"): "To identify eQTL specific to POAG, disease status was included in the model". I interpret this as adding the disease status as a covariate in the eQTL analysis. If this is the correct interpretation (please, explain how this analysis was performed), adding disease status as a covariate in the eQTL analysis would remove the effects of the "disease" variable from the analysis. If the authors aim at finding variants that are associated with gene expression only in disease, a better approach would be to add an interaction term (genotype:disease) to the linear regression model and test its significance. For more details on this approach, please look at Alasoo, Nat Genet 2018.

i. Page 24: "Participant recruitment". The cohort needs to be better described. It is indicated that the cohort of POAG and patients and cases were sex-, ethnically, and age-matched. Only the mean ages are given: cases 59.7 and controls 70.1. Complete information on each individual needs to be provided. This is important given how the SNP imputation was conducted (page 25) and the requirement to take this information into consideration as covariates in calling the eQTLs. The authors should describe the individuals included in this study at least in a supplemental table.

j. page 26-27: "Identification of cis-eQTL using transcriptome and genotype data". Considering the number of samples tested, the number of eGenes should be higher. By changing the normalization step and adding covariates, the power should be improved:
i. Gene expression is usually quantile-normalized to be used as input for eQTL analysis. Having normalized data allows to compare effect sizes across genes that are expressed at different levels. Log-transformation risks to flatten the differences for genes expressed at high levels.
ii. Using PEER factors has been shown to increase the number of detected eQTLs, as PEER factors reduce the noise derived from unknown sources of variability in the data. Optimizing the number of PEER factors to use can greatly improve the number of eQTLs detected.

Authors' response to the first round of review

Reviewer #1:

The authors have performed a technical tour-de-force to try to understand differences between glaucoma cases and controls gene expression, using iPSCs given the unavailability of appropriate human tissue, and are to be congratulated. While they have compressed a huge amount of work into a readable manuscript, at times it was not easy to follow through, and I hope my comments below will improve the manuscript. In particular, the results for each set of experiments do not particularly reference previous results- for example CDKN2B has a SNP allelic effect with respect to glaucoma status, but is not mentioned again - so I assume it has no differential expression? Similarly the TWAS results are mildly disappointing- in terms of refining known loci there are only 3 loci identified in the end, and none of the better-known loci are discussed. Maybe a table of the authors' multitrait GWAS significant hits and what was found here for future researchers to refer to? It would be good to have a better flow-through of information.

1. I would like clarification about the recruitment. The advanced glaucoma is clearly a great phenotype, but it is unclear from the description whether the subjects recruited ALL had end-stage disease, all had normal tension glaucoma- using the term recruitment "focused" on a) end stage b) NTG makes it unclear. Similarly the age of diagnosis of cases is given, and the age at recruitment of controls, but not the age of recruitment of cases (which may have bearing on later comments)

Action. We have 'tightened' our description of the cohort, such that it now reads:

"POAG patients required a clinical diagnosis of advanced normal tension glaucoma, …" In addition, we have included specific details regarding the age at collection for both the case and controls subjects: "(mean ± SD age: 69.1 ± 14.4 years at biopsy for case subjects; 68.1 ± 8.2 years at biopsy for controls)." Given that this information could be useful for future epigenetic work, these details for each individual cell line has been included in **Table S1**.

2. I would like some clarification of methods - in the first paragraph of results, the authors talk of batches, pooling, 160 cell lines, SNPs from 162 individuals, 23 failures, 110 donor cell lines (and 55+57 human subjects). It is unclear to me how these all relate and where 162 individuals come from and exactly what was pooled. How many cases and controls were left? Could this be in some sort of flow chart?

Action. Thank you for this suggestion. We have now inserted a flowchart that describes progression of numbers (**Figure S1**). The first section of results was rewritten to clarify the experimental flow, as follows:

"***Large-scale generation of patient iPSCs, differentiation into retinal organoids and scRNA-seq.*** We recruited a large cohort of 183 individuals, which included healthy (n=92, of which 50 were female) and patients with advanced POAG (n=91, of which 50 were female). The mean ± SD age at biopsy for controls was : 68.1 ± 8.2 years, and 69.1 ± 14.4 years for case subjects. Participants underwent skin biopsy and their cultured fibroblasts were reprogrammed to iPSCs using episomal vectors as

we previously described(Crombie et al. 2017). Genotyping data were also generated from participants and after quality control and imputation, yielded 7,691,208 autosomal SNPs at a minor allele frequency (MAF) above 0.01. iPSC lines were differentiated in batches (25 batches, 6-8 lines with equal numbers of control and POAG lines per batch) to neural retina for 28 days in adherent cultures. 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**). This timeline was based on work by others, which described RGC emergence by day 35 of retinal organoid differentiation (Reichman et al. 2017), and RGC neurite extension following plating of dissociated organoids by day 40 (Fligor et al. 2018). Twenty-two lines did not differentiate to retinal organoids and were discarded (healthy: 11 lines, of which 1 was female; POAG: 11 lines of which 5 were female). Cells from the remaining 161 individual cell lines were harvested and divided into 25 batches for scRNA-seq, with each batch containing cells from 6-8 cell lines and had a targeted capture of 2,000 cells per line. A total number of 330,569 cells were captured via scRNA-seq and sequenced to a mean read depth of 41,020 per cell (**Table S1**). Individual cells were traced back to their cell line donor using a combination of transcriptome and genotype-based methods. Lines were removed based on the following criteria: failed genotype and virtual karyotyping QC, monogenic POAG, non-European background and low cell capture numbers. Individual cells were removed based on scRNA-seq metrics as described in the methodology. 247,50 cells (Healthy: 128,175, POAG: 119,345) from 110 iPSC lines (Healthy: 56 of which 35 were female, mean ± SD age of samples 67.5 ± 7.8 years; POAG: 54 of which 33 were female, 71.8 ± 11.5 years) were retained for subsequent analyses."

3. There were 17% RGCs, 77% RPCs and fewer other cells, as expected. Were these proportions similar in cases and controls?

Action. Yes the proportions were similar and the following sentence was included in the manuscript: "We compared the percentage of cell types between patients with POAG and healthy controls, and observed no statistically significant differences between the groups (**Table S5**, **Figure S2**)."

4. The study identified 58 e-genes and refers to Fig 2B/C where there are fewer names of genes shown. Are these results tabulated anywhere? And the circos plot is not fully explained in the legend as to what the different colors mean.

Action. The number of eQTL results has been tabulated in Table 2, whereby the 58 eGenes mentioned in the paragraph were shared between members of a lineage. To clarify this, we have included an additional supplementary table containing these results, and the text in the figure caption has been altered to now read:

"(B) Chromosomal map of significant loci in RGC subpopulations RGC1 (light orange), RGC2 (red), RGC3 (dark orange) and RPE (blue). Loci were labelled as significant if FDR < 5×10^{-8}. **Table S6** contains full details of significant loci."

5. The authors on page 7 give two possible explanations for the cell type-specific eQTL detection but then discount these. Cell type specificity and correlation properties are well known (see GTEX Consortium Science 2020). Any other possible reasons? Could another possibility be random clonal differentiation with aberrant translational patterns? Apart from a few gene-markers that prove that these cells are RGC cells of some sort, how differentiated they were and how similar to actual cells (for example in animals)?

Action. Thank you for this suggestion. To assess the similarity of our iPSC-derived retinal cell types, we used *scPred* - an unbiased gene-marker free cell classification method to classify our cells based on the dataset released by (Yan et al. 2020), who obtained retinal specimens from adult patients. The comparison indicates that the cell generated with our retinal organoid differentiation method cluster with scPred cell classifications, thus indicating a close relationship to native cells. In particular, the *in vitro* RGCs correlate with the *in vivo* RGC population as shown on the figure below. We have now inserted this additional analysis as panel C into **Figure S3**.

Direct inspection of specific canonical genes, also revealed a similar pattern between our iPSC-derived retinal organdies and the human retinal dataset released by (Yan et al. 2020). Please see the figure below, which we have not currently included in our revised manuscript due to space considerations.

6. Given the "poster-boy" status of CDKN2B in NTG, it would have been helpful to include this in the figure. This is probably an ubiquitous effect, given Burdon et al showing allele-specific methylation of this promoter.

Action. We have further investigated the role of genetic variation on *CDKN2B* expression, and sought support of possible methylation influences. In our original submission we identified an interaction effect between the eQTL and disease status in the RGC1 cell population. In further analysis we identified that there was variation in the number of donors with a non-zero expression of the *CDKN2B* gene. We then tested if there was a relationship between the percentage of donors with non-zero expression and genotype classes for POAG and Controls. The analysis identified a significant association between disease status and genotype class for gene silencing. To highlight this we have included information relating to the expression frequency in

Figure 4B, and the following text has been inserted:
"In further analysis we identified that there was variation in the number of donors with a non-zero expression of the *CDKN2B* gene, supportive of previous work showing allele-specific methylation of this promoter region [67]. We then tested if there was a relationship between the percentage of donors with non-zero expression and genotype classes for cases and controls, and identified a significant association across all RGC subtypes (χ_2, RGC1 $p=5.17\times10^{-11}$; RGC2 $p= 9.42\times10^{-10}$, RGC3 $p= 1.57\times10^{-14}$)(**Figure 4B**)."
7. How was the FDR for the 54 eQTLs specific to RGCs actually calculated? I don't think I saw this and some have 10-2 FDRs in the table.
Action. The following sentence has now been inserted into the Methods section: "The FDR was calculated using the Benjamini–Hochberg procedure for all tested gene-SNP, as applied using MatrixEQTL (Shabalin 2012)."
8. Bottom page 7 fig 3C is referenced but I assume you mean 3A? And you refer to A alleles but the figure illustrates T/G alleles.
Action. Thank you for this astute observation. We verified this and the figure has been fully updated, and is now Figure 4.
9. Table S7 is referenced after 3118 genes but later in the same paragraph 144 genes are differentially expressed (and table looks to have around 156 lines so I assume it relates to the 144 genes?). Suggest move reference to correct place.
Action. Thank you for highlighting this, and the reference to this Table has been moved.
10. Regarding differential expression, while I understand the randomisation should avoid false positives, differing ages might have some effect? Even during the natural conception 10-20% of CpG sites remain methylated. Epigenetic memory and age: there is talk about differential pluripotency depending on the age of the donor. How well age-matched were the cases and controls at skin biopsy?
Action. We have now included additional details regarding the age of the case and control participants.
At recruitment: Healthy: n=92, of which 50 were female; mean ± SD age: 68.1 ± 8.2 years at biopsy. POAG: n=91, of which 50 were female; 69.1 ± 14.4 years at biopsy
At analysis: Healthy: n=56 of which 35 were female; 67.5 ± 7.8 years. POAG: n=54 of which 33 were female; 71.8 ± 11.5 years.
This information is now provided within an updated result section, methods, Table S1 and in a flowchart presented as Figure S1.
11. Further comment on the TWAS results might be helpful- TWAS infers expression but this study actually has expression data- and how these relate to the authors' previous publication cited that used non-RGC expression.
Action. We have now inserted the following text into the Discussion Section:
In the TWAS framework, the gene expression data (association between SNPs and genes) were used to train prediction models to determine gene expression levels by genetic variants (genetically regulated gene expression, GReX) (Gamazon et al. 2015; Gusev et al. 2016). The prediction models were used to impute gene expression levels in the GWAS dataset based on the trained weights from multiple-SNP prediction models, which could be further used to evaluate the association between imputed gene expression levels and the GWAS phenotype (that is glaucoma), and to identify

genes associated with disease traits. In this study, we performed the first glaucoma TWAS based on cell type-specific expression profiling. The gene expression profiling between different tissues could be quite different, eg, bulk retinas versus subpopulations belonging to the RGC lineage. In our TWAS analysis, the single cell gene expression data from different subpopulations were used to train prediction models, and then to impute the gene expression levels in GWAS dataset based on summary statistics (Barbeira et al. 2018). The single-cell level resolution TWAS can provide new insights into the potential causal genes for glaucoma in specific cell types.

12. While RGCs are the end-organ damaged in glaucoma, some comment on the importance of IOP and the fact that previous GWAS have found mainly IOP genes would be helpful to the non-expert reader.

Action. The following text was added to the first paragraph of the introduction: "Elevated IOP was long considered a distinguishing feature of POAG; however, it is now clear that it is not a direct determinant of disease development [4]. Patients with elevated IOP may not develop glaucomatous optic neuropathy, while those with IOP within the normal population range may sustain significant RGC loss [5–7]."

Reviewer #2:

In this study, Daniszewski et al. use organoid technology to study the genetics of primary open angle glaucoma (POAG). The authors established iPSC lines from POAG patients and controls, then used them to derive retinal organoids, which were subjected to scRNA-seq. Transcriptional analysis enabled the authors to identify disease specific differentially expressed genes and retinal ganglion cell specific eQTLs.

1. Page 5-authors should comment of the diversity of the population studied with respect to sex and ethnicity. How many clones from each individual were studied? How many replicate differentiations were performed? What was the overall reproducibility? Were failures to differentiate equally distributed between patients and controls?

Action. Our cohort was age and sex matched. This information is now added to the result section: "We recruited a large cohort of 183 sex- and age- matched individuals,which included healthy (n=92, of which 50 were female) and patients with advanced POAG (n=91, of which 50 were female)". The cohort exclusively focused on individuals of European descent. The selection of individuals based on ethnicity was performed as described in the methods section. "... 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 (Yang et al. 2011). 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 (**Figure S5**)."

This plot has now been included as Figure S5. Two individuals did not cluster with the European (EUR) supercluster, and were excluded from analysis. Following reprogramming, polyclonal iPSC lines were generated for each patient as described in the methods section. Reproducibility of differentiation was similar between control and POAG iPSC lines with 81 control iPSC lines and 80 POAG iPSC lines successfully

used for differentiation. The following information was added to the results section: "... Twenty-two lines did not differentiate to retinal organoids and were discarded (healthy: 11 lines, of which 1 was female; POAG: 11 lines of which 5 were female). Cells from the remaining 161 individual cell lines were harvested and divided into 25 batches for scRNA-seq, with each batch containing cells from 6-8 cell lines and had a targeted capture of 2,000 cells per line." The information on replicate differentiations is presented in Table S1 (donor quality control, number of pools). The following clarification was added to the result section: "... 247,520 cells (Healthy: 128,175; POAG: 119,345) from 110 iPSC lines (Healthy: 56 of which 35 were female, mean ± SD age of samples 67.5 ± 7.8 years; POAG: 54 of which 33 were female, 71.8 ± 11.5 years) were retained for subsequent analyses."

2. Page 5-what was the rationale for harvest of organoids at the two-week time point. Would additional maturation have provided better information (most cells are still progenitors at this point).

Action. We apologise for this confusion. RGCs were actually harvested after 49 days of differentiation, as based on work from Reichman and colleagues (2017, Stem Cells) which showed the emergence of RGC markers BRN3A/B by day 35, and a peak expression of these markers (mRNA and proteins) between days 42 and 84 of retinal organoid differentiation. The differentiation timeframe (49-day-differentiation) we used is also consistent with the thorough work of Fligor *et al* (2018, Sci Rep) which assessed RGC emergence during retinal organoid differentiation from hPSCs and showed that RGCs are present within the organoids from day 30, with a plating of dissociated organoids at day 40 allowing for RGC neurite extension. In our differentiation protocol, iPSC lines were first differentiated into retinal organoid for 28 days in adherent cultures, followed by 7 days in suspension culture (day 35) and plated onto Matrigel for an additional 2-weeks (day 49) to allow RGC neuronal outgrowth. Hence cells were harvested after 49 days days of retinal differentiation. The last 2-week-time-point of plated organoids was chosen following optimisation with the hESC line H9 (WiCell) and a BRN3B reporter hESC line (A81-H7, Fligor *et al* 2018, Sci Rep) which showed high expression of BRN3B in the plated organoids, with BRN3B positive neuronal outgrowth observed from 3 days of plating and length of these projections increased over time reaching higher levels by 14 days after plating whilst keeping cell viability high. For clarity the following sentences were inserted into the result section on page 5: "This timeline was based on work by others, which described RGC emergence by day 35 of retinal organoid differentiation (Reichman et al. 2017), and RGC neurite extension following plating of dissociated organoids by day 40 (Fligor et al. 2018)."

3. Page 5 in the end, how many cells from how many control and patient lines were assessed.

Action. The results section was updated to reflect the number of control and patient lines retained for analysis: "Lines were removed based on the following criteria: failed genotype and virtual karyotyping QC, monogenic POAG, non-European background and low cell capture numbers. Individual cells were removed based on scRNA-seq metrics as described in the methodology. 247,520 cells (Healthy: 128,175, POAG: 119,345) from 110 iPSC lines (Healthy: 56 of which 35 were female, mean ± SD age of samples 67.5 ± 7.8 years; POAG: 54 of which 33 were female, 71.8 ± 11.5 years) were retained for subsequent analyses."

4. Page 5-Figure S2 shows very substantial variation in the proportion of RGC from patient to patient (though not in a disease related fashion). How did this variation impact on gene expression profiles?

Action. We thank the reviewer for this interesting suggestion. Following their advice, we assessed the effect of RGC proportion on both differentially expressed genes in all cell types, and the detection of eQTL in RGC subpopulations. We included the proportion of RGCs as a latent variable in DE analysis using MAST and compared the results to the baseline and found it had the biggest impact on the interneuron subpopulation (please see graph below). Enrichment of differentially expressed genes related to % RGC showed the upregulation of apoptosis pathways, growth pathways and developmental pathways (Table S9). As RGCs are one of the earliest cell types to arise during retinal development, they can influence the behaviour of surrounding cells, including progenitors and differentiating cells. It is already established that RGCs influence retinogenesis by interactions with progenitor cells and other retinal neurons (D'Souza and Lang 2020). The variation we observed in the DE analysis thus suggests that RGCs influence the differentiation and survival of early interneurons, by upregulated pathways associated with neuronal death, cell growth and differentiation. This statement was added to the Supplemental Results together with a new panel in Figure S3: "Comparing gene expression of samples with high levels of RGCs the other samples revealed an effect of RGC proportion on both differentially expressed genes in all cell types (Fig S3C), and the detection of eQTL in RGC subpopulations. In particular, RGC numbers mainly influenced gene expression in interneurons, with an upregulation of pathways linked to apoptosis growth and development (**Table S9**). As RGCs are one of the earliest cell types to arise during retinal development, they can influence retinogenesis by interactions with progenitor cells and other retinal neurons (D'Souza and Lang 2020). The variation in gene expression observed in interneurons thus suggests that RGCs influence the differentiation and survival of early interneurons, by upregulated pathways associated with neuronal death, cell growth and differentiation".

5. How convincing is Figure 2c-two cell types and only three genes does not seem sufficient to draw general conclusions. In some cases, effects of genotype across RGC subsets are not that striking. In S4, many eQTL also operate in RPE, including KANSL1-AS1

Action. We now have included details of all significant loci as a supplementary table (Table S6) and have revised the text to reiterate this point. As displayed in the upset plot (Figure 3A), it is noteworthy that the majority of eQTLs were not common across cell types. Figure 2C are just examples of genes that had cell type-specific eQTL, and notably - KANSL1-AS1 did not have any significant interactions in the RPE cells of this study.

6. Page 7-what is the basis for the statement that a number of RGC eQTL are directly involved in neurogenesis or neurodegeneration. You might expect the three examples cited to turn up by chance, given the diversity of processes (and apparently species) surveyed. How about actual gene ontology analysis?

Action. We have now undertaken gene ontology analysis on eGenes associated with each RGC subpopulation and for comparison, RPE. No significant results were found using an Over Representation Analysis (ORA) (Boyle et al. 2004) and Gene Set

Enrichment Analysis (GSEA) (Subramanian et al. 2005) (p-value threshold <0.05). We were able to perform a group Gene Ontology (GO) using clusterProfiler (Yu et al. 2012) which showed some groups of eGenes from the RGC subpopulations - such as those related to detoxification and biomineralization, were absent in eGenes from the RPE subpopulation.

7. Pge 7 and Figure 3a-unless I have misunderstood, these eQTL in genes previously associated with POAG can certainly be detected in (most subsets of) RGC, but there seem to be no consistent differences in patient versus control cells. I suppose that this makes sense, but how is one to identify important eQTL without knowledge a priori if this is the case. I could not find the Figure 3c referred to in the text in my copy of the manuscript.

Action. We apologise for this confusion. We approached the problem of identifying genes and eQTL with differing effects on disease status in several ways. We first tested for eQTL using all genes and SNPs that passed quality control in each cell using data from all individuals. For those genes that pass the study-wide FDR threshold and expression in at least 30% of donors, we tested if there was any interaction between the allelic effect and disease status, but fitting the following model (*i.e.* gene expression = SNP + SNP:disease + residual). The significance of the interaction term was also tested against a study-wide FDR threshold. We have revised Figure 4 to graphically display these results and have also updated the text to clarify this analysis strategy: "To identify eQTL that had alternative allelic effects under different disease statuses, we included an interaction term (SNP:disease status) in the original linear model for each eQTL identified by the first round of analysis. eQTL with interacting effects were determined to be significant based on a threshold of FDR < 0.05 of the interaction term."

8. Page 8-what exactly is the link postulated between TTR, familial amyloidotic polyneuropathy, and POAG. This does not seem convincing.

Action. Individuals carrying specific mutations in TTR leading to Transthyretin V30Mrelated
familial amyloidosis with polyneuropathy (FAP TTRV30M) often show ocular manifestations including glaucoma. A second reference supporting this point was added to this statement (Beirao *et al*, 2015).

9. Page 8-explain clearly how your multitrait glaucoma GWAS should relate to POAG? Is POAG a subset of the diseases examined in the GWAS? How exactly did you use the GWAS statistics?

Action. The multitrait glaucoma GWAS is currently the largest GWAS for this disease, and thus, presents the best power for our TWAS analysis. We only used the glaucoma GWAS statistic and the following sentence regarding multitrait GWAS has been inserted: "In previous work, we combined GWAS for multiple genetically correlated traits, and here used the glaucoma-specific effect size estimates and P-values for SNPs across the genome."

Reviewer #3:

Daniszewski et al. used single cell RNA-seq (scRNA-seq), combined with eQTL analysis and GWAS, to study the associations between genetic variation and primary open-angle glaucoma (POAG). They reprogrammed fibroblasts to iPSCs from >50

POAG patients and matched controls and differentiated iPSC-derived retinal organoids. They performed scRNA-seq on ~250,000 cells and found that the vast majority of cells were retinal progenitor cells (RPCs). No difference in the cell type proportions between cases and controls was found. To determine the associations between genetic variation and gene expression, they performed an eQTL analysis on all cells and on each cell type. They found 2,235 eGenes across all cells, but they had very little power to detect eQTLs on each cell type (10-456 eGenes/cell type). To determine associations between genetic variation, gene expression and POAG, they performed TWAS, and found seven genes associated with disease, most of which were at known GWAS loci. While this study represents a great resource to study the transcriptome of retinal organoids, genetic analyses are weak and need to be strengthened.

We thank the reviewer for providing helpful and informative comments and suggestions. We have addressed each of these, and replied in detail below. In particular, we have improved on both the genetic analysis and description of the results and methods throughout.

Major comments:

A- The authors use terminology in a confusing manner:

a. Page 4-5: "Using an additive linear model, a total of 54,786 eQTLs were found to be associated with 21,512 SNPs". Because of LD structure, reporting the number of variants whose genotype is significantly associated with gene expression is meaningless. It would be more useful to report the number of eGenes. Additionally, these numbers are not described anywhere else in the text.

Action. We apologise for this confusion, and agree that it would be more useful to report the number of eGenes. Reference to the number of eQTL associated with eSNPs has been removed and replaced with reference to the number of eGenes.

b. Page 6: eQTLs underlie association, not causality. The authors cannot refer to eQTL analysis as "To explore cell type-specific genetic control of gene expression".

Action. We agree that eQTLs identified are associated and not necessarily causal and as such we have removed this phrase.

B- Many of the analyses need to be substantially improved and the Methods section lacks important details:

a. Batch effects are a common and real problem in Genomic studies. In the Results section on page 5 first paragraph it is unclear how the data were generated to avoid batch effects. Sentences from this paragraph:

1. The cohort had 57 healthy and 55 POAG individuals;
2. Twenty-three lines did not differentiate;
3. 330,569 cells from 160 individual cell lines;
4. SNPs from 162 individuals;
5. scRNA-seq data assigned to 128 donors; and
6. Data from 110 donor iPSC lines used for subsequent analyses.

This is impossible to follow and the methods and Table S1 do not clarify. Were cases and controls differentiated at the same time and mixed for scRNA-seq generation? If not, how were the data generated? The supplemental table should include a differentiation number and which samples were combined for scRNA-seq generation. On page 8 this is stated, "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)." But I don't see this information provided in the Methods.

Action. We apologize for the lack of clarity of the paragraph. Cases and controls were differentiated simultaneously and mixed for scRNA-Seq. Differentiation number and sample identity per batch are now added Table S4. The first section of results was rewritten to clarify the experimental flow, please see response to Reviewer 1 point 2: "iPSC lines were differentiated in batches (25 batches, 6-8 lines with equal numbers of control and POAG lines per batch) to neural retina for 28 days in adherent cultures. 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**). Cell lines were harvested in 25 batches with 6-8 lines per batch and multiplexed for scRNA-Seq with a targeted capture of 2,000 cells per line".

Information on shared or randomized conditions have been included in Table S1 and the method section: page 25 for Differentiation of iPSCs into retinal organoids: "Cell lines were harvested in 25 batches (6-8 lines per batch) and multiplexed for scRNASeq with a targeted capture of 2,000 cells per line".

Computational methods were also used to correct for batch effects, as described in the methods section: "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 (Zheng et al. 2016). The SCTransform function from Seurat (v3.0.2) was applied to the filtered count matrix to perform cell-cell and batch normalization (Hafemeister and Satija 2019). "

b. Page 5, section "Identification and characterisation of 23 subpopulations from 253,107 cells". The authors identify 23 clusters. While this is technically correct, it would be extremely surprising that retinal organoids include 24 cell types. Indeed, they are combined into six distinct cell types. Can the authors find a clustering resolution that allows them to identify these biologically relevant cell populations? If not, what would the reason be?

Action. As shown in Figure S2A, clustering was performed at resolutions ranging from 0 to 1 in steps of 0.1. Even at the lowest resolution, there were 13 subpopulations. Considering we are working with stem cell-derived cultures, the cells exist on a continuum between the progenitor state and the final cell state, which is supported by our trajectory analysis. These subpopulations should therefore be considered cell states, rather than just cell types.

c. The differential expression analysis could be more informative, considering the resources generated in this study. With ~250,000 individual cells differential expression analysis could provide invaluable information beyond the observed 3,118 genes that were differentially expressed. With single cell technologies, there are a few methods that would provide better insights into the transcriptomic differences between cases and controls. The authors should explore a few of these to strengthen their analysis:

i. Pseudotime: the authors do not find differences between cell composition of cases

and controls, but this does not exclude that, overall, cases and controls may display differences in maturity. Pseudotime may be tested for eQTLs too (see Cuomo et al., Nat Comm 2020 for an example)

Action: We thank the reviewer for their interesting suggestion of applying pseudotime to eQTL analysis and have updated the analysis to include this information:

"Trajectory analysis of subpopulations reveals disease-specific mechanisms in RGC lineages

We studied the ordering of the subpopulations across pseudotime by performing trajectory inference using the *slingshot* package [56], as described in **Methods**. Trajectory inference revealed a complex, branching trajectory that consisted of 12 lineages (**Figure 1C, 2A**). Three of these lineages (6, 7 and 9) comprise of RGC subpopulations branching off from RPC9 **(Figure 2A, 2B)**. We examined these lineages in greater detail by studying gene expression patterns related to disease and pseudotime. Gene ontology analysis of the RGC lineages revealed an overrepresentation of genes involved in neurogenesis (**Figure 2C**). Further to this, there was a significant difference in the distribution of cells across pseudotime in the lineage terminating with RGC3 (Lineage 7), based on disease status (Kolmogorov-Smirnov Test: p-value - 0.028).

We then used *tradeSeq [57]* to investigate the nature of this lineage, and determine if disease status affected gene expression patterns across the trajectory. We identified 1,471 genes that were differentially expressed between the conditions, across pseudotime (Benjamini & Hochberg FDR < 0.05) (**Table S6**). Disease ontology of these genes was performed with Gene Set Enrichment Analysis (GSEA) [58], and revealed associated with four disease pathways - schizophrenia, psychotic disorders, disease of mental health and cognitive disorder, as annotated by the Disease Ontology database [59]. We also applied pseudotime to cis-eQTL mapping to determine if this had a significant interaction between genotype and POAG. This uncovered one new eGene - *HMGB1*, which had pseudotime as a significant interaction term ($p=1.76 \times 10^{-7}$) at SNP rs9578147 (**Figure 2D**). Interestingly, *HMGB1* is involved with nucleosome stabilization, and is released from injured cells and to induce an inflammatory response [60]. It has been shown to induce RGC death in NMDA-mediated retinal neurodegeneration [61,62], and is present in glaucomatous retina [63]."

ii. Differential expression analysis cell type by cell type. Are there differentially expressed genes in each cell type? Is there a cell type that accounts for the most differences between cases and controls?

Action: In addition to canonical markers, we assigned cell types to subpopulations using differentially expressed genes. The table for all differentially expressed genes for each subpopulation can be found in Table S4.

d. Page 6, "The genetic control of gene expression is highly cell type-specific": "We identified a total of 2,235 eQTL across all cell types, which surpassed a study-wide significance threshold of FDR < 0.05". FDR correction for the eQTL analysis should be explained. There are methods that account for LD when performing gene by gene FDR correction (eigenMT) and studies that suggest a two- or three-step approach to perform FDR correction (Huang, NAR 2018: doi: 10.1093/nar/gky780). The authors do not describe how they performed FDR correction.

Action. We apologise for not making this clear in our initial submission. The following

sentence has now been inserted into the Methods section: “The FDR was calculated using the Benjamini–Hochberg procedure for all gene-SNP interactions, as applied using MatrixEQTL (Shabalin 2012).”

e. Page 6, "The genetic control of gene expression is highly cell type-specific": The authors find 2,235 eQTLs and 1,447 eGenes across all cell populations. I do not understand what the authors intend by "eQTLs". Are these all independent eQTL signals (primary and conditional eQTLs)? If so, this should be clearly stated and the methods should describe how conditional eQTLs were performed. If not, it would be very surprising to find only 1.5 eVariants/eGene, considering the LD structure of the genome. It is also surprising that the authors used a very stringent MAF cutoff for eQTLs (10%, when the usual threshold for studies with a similar size is 5%). Also, these numbers do not match those described in the Introduction (54,786 eQTLs and 21,512 SNPs).

Response. We apologise for the confusion here. We defefine eQTL as an independent genetic association between SNPs and a gene (termed an eGene if there is 1 or more signifuicant eQTL). An eQTL typically contains many eVariants (eSNPs) and to test for secondary independent associations we take the top eSNP from the 1st round of eQTL analysis and include it as a conditional effects in the regression model and test a second round of eQTL tests. This accounts for LD between SNPs as any SNPs in LD with the original top eSNP will not have a significant effect in the second round of analysis. However, unlinked (and associated) secondary eQTL can be identified. In total we identify eQTL for 1,447 eGenes, of which 788 have a significant secondary eQTL. We have updated the text throughout with a clearer explanation of terms.

On the MAF, we decided to use a MAF cut-off of 0.1 to ensure that we (on average) had enough individuals in the rare homozygous genotype class to accurately estimate the mean effect for that genotype class.

f. Page 6, "The genetic control of gene expression is highly cell type-specific": finding that almost all eQTLs are cell type-specific is the opposite of what has been found in previous eQTL studies. For example, the latest GTEx papers show that many eQTLs are shared between different tissues, while other papers (Kim-Hellmuth, Science 2020; Donovan, Nat. Comm. 2020; Cuomo, Nat Comm. 2020; Jerber, bioRxiv 2020) have shown that only a subset of eQTLs are cell type-specific. This needs to be discussed in discussion as to why the results from this study are so different.

Action. We appreciate this point, as the apparent difference in the conclusions could mislead readers. When testing for overlap of eQTL between cell-types, we tested if the top eSNPs were shared in common. This is in contrast to the papers mentioned above which tested for overlap in eGene. i.e. does the same gene have an eQTL between cell types. While this analysis strategy is valid, we were concerned that it would mask cell-type specific effects when independent eQTL acting on the same gene in different cell types. In other work (Yazar et al. Under Revision at Science) we have found that these types of effects are more common than true shared eQTL. we have added some new text in the discussion to make this clear.

g. Page 7, second paragraph: "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)". The figure shows three examples of eQTL

signals in four cell types, of which one (DNAJA1) supports the cell type-specificity of eQTLs, one (CNOT6L) shows an eQTL shared between RGCs and RPE, and one (TSPAN2) does not have good agreement between RGCs. Based on these examples, I do not understand how the authors can conclude that "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". The authors also provide an alternative hypothesis: "Another possible explanation for the cell type-specific eQTL is low statistical power to detect eQTL in multiple cell types" and describe a convoluted analysis that, in their opinion, does not support the hypothesis, but they do not provide any supporting data.

Action. To ensure clarity of manuscript we have removed this section of text and simplified the message, such that it now reads:

"Whilst some eGenes are identified across multiple cell types (such as *RPS26;* **Figure 3B, 3C**), the three RGC subpopulations share 17 eGenes, such as *DNAJA1* and *TSPAN2*, that are absent from other subpopulations (**Figure 3B**). *DNAJA1* belongs to a large family of chaperones, and has been shown to prevent neurodegeneration by decreasing α-synuclein aggregates [64], whilst *TSPAN2*, is known to support myelination [65]."

h. Page 7, last paragraph and its corresponding Methods section (page 26-27: "Identification of cis-eQTL using transcriptome and genotype data"): "To identify eQTL specific to POAG, disease status was included in the model". I interpret this as adding the disease status as a covariate in the eQTL analysis. If this is the correct interpretation (please, explain how this analysis was performed), adding disease status as a covariate in the eQTL analysis would remove the effects of the "disease" variable from the analysis. If the authors aim at finding variants that are associated with gene expression only in disease, a better approach would be to add an interaction term (genotype:disease) to the linear regression model and test its significance. For more details on this approach, please look at Alasoo, Nat Genet 2018.

Action. We apologise for the confusion over the statistical model that we used here. As highlighted above, we did apply a test for the interaction between SNP and Disease status, rather than just fitting disease status in the model. We have edited this section to provide a clearer explanation of our approach:

"To identify eQTL that had alternative allelic effects under different disease statuses, we included an interaction term (SNP:disease status) in the original linear model for each eQTL identified by the first round of analysis. eQTL with interacting effects were determined to be significant based on a threshold of FDR < 0.05 of the interaction term."

i. Page 24: "Participant recruitment". The cohort needs to be better described. It is indicated that the cohort of POAG and patients and cases were sex-, ethnically, and age-matched. Only the mean ages are given: cases 59.7 and controls 70.1. Complete information on each individual needs to be provided. This is important given how the SNP imputation was conducted (page 25) and the requirement to take this information into consideration as covariates in calling the eQTLs. The authors should describe the individuals included in this study at least in a supplemental table.

Action. The control and POAG cohorts were matched for age and sex. These matchings were respected at collection and in the subsequent samples used for the

sequencing and analysis.
At recruitment: Healthy: n=92, of which 50 were female; mean ± SD age: 68.1 ± 8.2 years at biopsy. POAG: n=91, of which 50 were female; 69.1 ± 14.4 years at biopsy
At analysis: Healthy: n=56 of which 35 were female; 67.5 ± 7.8 years. POAG: n=54 of which 33 were female; 71.8 ± 11.5 years.
This information is now provided within an updated result section, methods, Table S4 and in a flowchart presented in a revised **Figure S1**.
j. page 26-27: "Identification of cis-eQTL using transcriptome and genotype data". Considering the number of samples tested, the number of eGenes should be higher. By changing the normalization step and adding covariates, the power should be improved:
Action: We have now performed quantile-normalisation, and as suggested this did improve our power for discovery of eQTL. As such, the text (results & method) and figures have been updated with this revised analysis.
Our revised breakdown of results are displayed here:
Model Number of eQTL Number of eGenes Number of eSNPs
Population 4,484 3,102 3,892
Disease 4,443 3,091 3,860
Control 2,985 2,394 2,492
POAG 2,460 2,090 2,136
i. Gene expression is usually quantile-normalized to be used as input for eQTL analysis. Having normalized data allows to compare effect sizes across genes that are expressed at different levels. Log-transformation risks to flatten the differences for genes expressed at high levels.
Action. Thank you for this suggestion. As outlined in the point above, we have now performed quantile-normalisation, and revised our results accordingly.
ii. Using PEER factors has been shown to increase the number of detected eQTLs, as PEER factors reduce the noise derived from unknown sources of variability in the data. Optimizing the number of PEER factors to use can greatly improve the number of eQTLs detected.
Response. Again, thank you for this suggestion. To directly investigate this, we applied PEER factors to a subset of cells (those designated in the RGC1 cluster). Specifically, PEER factor residuals were applied to our gene expression matrix. UMI logcounts were quantile_normalised using limma, and results filtered with FDR threshold of 0.05. Interestingly, when compared to quantile normalisation, we identified a lower number of eQTLs PEER factors were applied (please see the figure below), and as such have not included these results into the manuscript.__

Referees' report, second round of review

Reviewer #2 (Comments to authors)

In their rebuttal and revised manuscript, the authors have clarified a number of significant issues raised in three in depth reviews. The clarity of the manuscript has been improved. The data will be a great resource for workers in the field.

Reviewer #3 (Comments to authors)

The approach described by the authors likely results in an underestimation of the number of the overlapping eQTL signals between cell types, as it does not take LD structure into account: i.e. if two variants are in perfect LD but one is found as the lead eSNP for cell type A and the other is found as the lead eSNP for cell type B, the authors would not describe these two eQTL signals as "shared". To address this issue, colocalization methods (see https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004383) would provide a more meaningful estimation of the number of cell type-specific eQTL signals.

The justification that the authors provide in the Discussion "This strategy, as opposed to testing for eGene overlap, helps ensure that the estimated proportion of shared eQTL is not inflated due to occurrences where the same gene's expression is associated with two independent eQTL in different cell types" should be reformulated once the differences between cell types have been addressed taking LD structure into account.

Extent they have addressed concerns, advance and suitability for publication in Cell Genomics or another journal.

Overall the authors addressed most concerns except for the comment above. If this concern is addressed, the publication would be suitable for Cell Genomics.

Authors' response to the second round of review

Reviewer #2: In their rebuttal and revised manuscript, the authors have clarified a number of significant issues raised in three in depth reviews. The clarity of the manuscript has been improved. The data will be a great resource for workers in the field.

We thank the reviewer for their valuable suggestions and feedback.

Reviewer #3: The approach described by the authors likely results in an underestimation of the number of the overlapping eQTL signals between cell types, as it does not take LD structure into account: i.e. if two variants are in perfect LD but one is found as the lead eSNP for cell type A and the other is found as the lead eSNP for cell type B, the authors would not describe these two eQTL signals as "shared". To address this issue, colocalization methods (see https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004383) would provide a more meaningful estimation of the number of cell type-specific eQTL signals.
The justification that the authors provide in the Discussion "This strategy, as opposed to testing for eGene overlap, helps ensure that the estimated proportion of shared eQTL is not inflated due to occurrences where the same gene's expression is associated with two independent eQTL in different cell types" should be reformulated once the differences between cell types have been addressed taking LD structure into account.

We thank the reviewer for the suggestion on how to address the involvement of LD structure in overlapping eQTL signals between cell types. We have now further refined our approach by performing conditional eQTL analysis on eGenes with an eQTL in more than one subpopulation, and have updated the manuscript to accommodate these new findings.

We implemented the following methodology and have inserted the following text into the manuscript:

To identify eGenes with overlapping eQTL signals in more than one subpopulation, we performed multi-directional conditional analysis on pairs of subpopulations. If a gene had an eQTL in subpopulation A and subpopulation B, we tested whether the allelic effects of $eSNP_A$ and $eSNP_B$ were dependent on each other by including $eSNP_A$

as a fixed covariate in the linear model for subpopulation B, and $eSNP_B$ in the linear model for subpopulation A. eSNPs were independent if the association remained significant. To determine if eSNPs tag the same causal variant in both subpopulations or were in LD, we tested the change in allelic effect between this model and the original model for significance. If the change was not significant, then the eSNPs tag independent causal variants for the same gene in different cell subpopulations.

The following text has been inserted into the results section of the manuscript:

647 out of 3091 genes with an eQTL - eGenes, were detected in more than one cell type, and only 215 of these eGenes had an eQTL observed in two or more cell types (**Figure 3B, 3C, Table S8**). RPE and RPC14 did not have any overlapping eQTL with any other cell types, while RPC1 and RGC1 had the greatest number of overlapping eQTL (27) (Pearson's correlation: r2 = 0.46, p value = 0.02) **(Figure 3B)**. As the majority subpopulations are retinal progenitors, non-RPC subpopulations share more eQTL signals with RPC subpopulations than each other. Two genes had an eQTL in all subpopulations but RPC2 - *RPS26* and *GSTT1* (**Figure 3D***).* Only *GSTT1* had overlapping eQTL in 16 subpopulations that indicates the variants associated with this eQTL are either in linkage disequilibrium with each other or are targeting the same causal variant. Approximately half of the eGenes detected in the retinal ganglion cell subpopulations (RGC1: 46.9%, RGC2: 58.9% and RGC3: 51.1%) were exclusive to this cell type, and only seven of these eGenes - *PPP1R17*, *RASD1*, *NXPH1*, *IGFBPL1*, *SAPCD2*, *KRTAP5-AS1* and *TK1* were found in at least one RGC subpopulation **(Figure 3D)**.

Referees' report, third round of review

Reviewer #3 (Comments to authors)

The authors addressed my concern about taking LD structure into account when examining overlapping eQTL signals between cell types.