Spatiotemporal single cell analyses reveal a transient population of retinal progenitor cells in the ciliary margin of developing human retina

Majlinda Lako

majlinda.lako@ncl.ac.uk

Newcastle University https://orcid.org/0000-0003-1327-8573

Birthe Dorgau

newcastle university

Joseph Collin

Newcastle University, International Centre for Life, Newcastle upon Tyne, United Kingdom

Agata Rozanska

Newcastle University

Darin Zerti

Newcastle University

Moira Crossier

Newcastle University

Rafiqul Hussain

Newcastle University

Jonathan Coxhead

Newcastle Cancer Centre at the Northern Institute for Cancer Research

Tamil Dhanaseelan

Newcastle University

Aara Patel

University College London

Jane Sowden

UCL

David FitzPatrick

University of Edinburgh https://orcid.org/0000-0003-4861-969X

Rachel Queen

Newcastle University

Article

Keywords: developing human retina, single cell, RNA-Seq, ATAC-Seq, spatial transcriptomics, ciliary margin, GRNs, TEADs

Posted Date: July 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3160527/v1

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Communications on April 26th, 2024. See the published version at https://doi.org/10.1038/s41467-024-47933-x.

Abstract

The emergence of retinal progenitor cells (RPCs) and differentiation to various retinal cell types represent fundamental processes during retinal development. Herein, we provide a comprehensive single cell characterisation of transcriptional and chromatin accessibility changes that underline RPC specification and differentiation over the course of human retinal development up to midgestation. Our lineage trajectory data demonstrate the presence of early RPCs, which transit to late RPCs, and further to transient neurogenic progenitors, that give rise to all the retinal neurons. Combining single cell RNA-Seq with spatial transcriptomics of early eye samples, we demonstrate for the first time the transient presence of early RPCs in the ciliary margin zone with decreasing occurrence from 8 PCW of human development. In RPCs, we identified a significant enrichment for TEAD transcription factor binding motifs, which when inhibited led to loss of RPCs and retinal lamination, and inhibition of photoreceptor and retinal ganglion cell differentiation.

Introduction

Retina is the innermost, light-sensitive tissue that lines the back of the eye and is vital for light sensing and image processing. The retina is derived from a germinal zone in the optic vesicle in which neuroepithelial cells proliferate to give rise to the six types of retinal neuronal and one glial cell type, organized within three different layers. All these cell types derive from retinal progenitor cells (RPCs) in an orderly spatio-temporal manner that has been well studied in vertebrates [1]. Great progress has been achieved in the last five years providing both gene expression and epigenetic profiles of the developing human retina [2, 3, 4]. Single cell (sc) RNA-Seq studies have enabled molecular characterisation of all the retinal cell types and have provided the developmental trajectories that lead to retinal cell specification from RPCs during retinal development [5, 6, 7, 8, 9]. The transcriptomic studies have been complemented with sc assay for transposase-accessible chromatin sequencing (scATAC-seq), enabling identification of key transcription factors relevant to specific retinal cell fates and the gene regulatory networks (GRNs) that underline the cell-state changes [7, 10, 11, 12]. Notwithstanding, these studies have not provided as yet any information on the spatial resolution of RPCs or retinal neurons.

The retina of fish and amphibians continuously integrates new neurons generated from RPCs residing in the distal tip of the retina known as the ciliary margin zone (CMZ). In chicks the CMZ also contributes to retinal neurogenesis during development, but the adult chick CMZ cells have a more restricted potential as they only contribute to a small fraction of the retina and moreover do not participate in retinal regeneration following injury [13]. In mice, the ciliary body which arises from the CMZ was postulated to contain a population of pigmented cells that were able to form clonogenic spheres and could be differentiated to express marker genes found in photoreceptors, bipolar and Muller glia cells [14]. However subsequent studies demonstrated that these pigmented ciliary epithelial cells fail to differentiate into retinal neurones *in vitro* or *in vivo* [15]. Instead, progenitors distinct from the classical RPCs, characterised by Msx1 expression and located in the proximal CMZ, were shown to give rise to non-pigmented ciliary epithelial cells and multipotent neural RPCs [16]. Whether the human CMZ displays comparable properties during retinal development remains unclear.

Herein we generated a scRNA-Seq atlas of 24 samples spanning 13 time points from 7.5-21 PCW of human development, enabling identification of early and late RPCs and their transition to neurogenic progenitors and retinal neurons. Complementing scRNA-Seq with spatial transcriptomics (ST) we demonstrate the transient presence of early RPCs in the CMZ of developing human retina. Complementary scATAC-Seq identified key transcription factors and signalling pathways that underline RPCs proliferation and differentiation. Together, these comprehensive single cell analyses shed light on the molecular mechanisms governing human retinal development and provide guidance on the generation of RPCs and differentiated retinal cell types from human pluripotent stem cells (PSCs).

Results

scRNA-Seq atlas of human developing eyes and retinas

To better understand the emergence of RPCs, their heterogeneity and differentiation to retinal cell types, we performed scRNA-Seq of 11 embryonic and fetal eyes spanning 7.5-15 PCW and 13 retinal samples from 10-21 PCW (**Table S1**). Each sample was embedded using Uniform Manifold Approximation and Projection (UMAP) and clustered using Seurat graph-based clustering (**Figure S1, Table S1**). In our earlier study of human retinal development [2], we reported expression of RPC markers (e.g. VSX2) at the peripheral margin of developing human retina as early as 6.5 PCW, and at the outer neuroblastic zone of central retina at 7.8 PCW, marking this time period an important developmental window of human RPC divergence. To better understand this process at the single cell level, we first focused our scRNA-Seq analysis in the 7.5-8.5 PCW embryonic eyes, revealing the presence of retinal and non-retinal cell clusters (**Figure S1A**). At this stage of human development, the neuroepithelium-derived optic vesicle is surrounded by periocular mesenchyme (POM), which is derived from neural crest and mesoderm and contributes to anterior and non-neural ocular tissue development. In accordance, we identified neural crest cell

clusters adjacent to POM and corneal endothelial and corneal stromal cells [17] (**Figure S1A**). The POM cell clusters of 7.5-8.5 PCW eyes displayed high expression of characteristic markers described in other species such as collagen chains (*COL3A1, COL5A1, COL1A1*), proteoglycan decorin (*DCN*) [18] and latent Transforming Growth Factor Beta Binding Protein 1 (*LTBP1*) [19]. Other non-retinal cell clusters including extraocular muscle, ocular surface epithelium, red blood cells, monocytes and macrophages were also identified in most eye samples of this developmental window (**Figure S1A**, **Table S1**).

Amongst the retinal cell clusters, RPCs and retinal ganglion cells (RGCs) were identified in all 7.5 - 8.5 PCW eyes, consistent with our previous findings of RGC presence at the basal side of the inner neuroblastic zone at 8 PCW [2]. A recent scRNA-Seq study has documented the presence of early and late RPCs and their transcriptional signature [5]. In accordance, we found that RPCs clusters of 7.5 and 7.7 PCW eyes displayed high expression of characteristic early RPC markers such as *SFRP2, RAX, PAX6, VSX2, ZIC2, ZIC1, SIX3, SIX6* etc. [20] (**Table S1**). While the expression of these markers was maintained in RPCs found in the 8 and 8.5 PCW eye specimens, expression of late RPC markers such as *CCND1, ASCL1* and *HES6* [5] became prominent, suggesting the co-existence of early and late RPCs. A proliferating cone photoreceptor cluster marked by high expression of proliferation (*TOP2A, PCNA*), neurogenic progenitors (*HES6*) and cone markers (*RXRG, GNB3*) was detected in two eye specimens obtained at 7.7 and 8 PCW (**Figure S1A**) as well as one of the 10 PCW samples (**Figure S1B**). In the latter sample, the proliferating cone cluster was distinct to the mature cones. Notably a small cluster of horizontal cells was detected in two specimens of 8 and 8.5 PCW of development (**Figure S1A**), corroborating with the first reported emergence of immature horizontal cells at day 59 of human development [4]. A retinal pigment epithelium (RPE) cell cluster with high expression of characteristics markers namely *PMEL, TYRP1, MITF* and *GJA1* was identified in all eyes of this stage (**Figure S1A**, **Table S1**).

A recent scRNA-Seq study provided evidence of the existence of neurogenic RPCs in addition to the early and late RPCs, based on gene expression profiling [5]. Sridhar and colleagues [6] further defined the neurogenic into intermediate transitional T1, T2 and T3 cell populations with the capacity to give rise to defined types of retinal neurons at specific developmental windows. Our analysis of 10-14 PCW samples demonstrated the presence of some of these neurogenic transitional progenitors; however, a better definition was obtained from the analysis of retinal samples (> 10 PCW) due to a higher number of analysed cells within the retina *per se* (**Figure S1B**, **Table S1**). From 14 PCW, we were able to detect all types of retinal neurons apart from bipolar and Muller glia cells, which were detected from 14 and 16 PCW respectively (**Figure S1B,C, Table S1**). Published topographical studies report the highest density of microglia in the retinal periphery and retinal margin at ~ 8 PCW. Their presence in the central retina is not observed until 12 PCW (Diaz-Araya et al., 1995). In accordance with these studies, presence of cell clusters with high expression of microglia markers were observed from 14 PCW retinal samples (**Figure S1B, C, Table S1**).

Pseudo-temporal analyses reveal transition of RPCs to T1-T3 neurogenic cell clusters and retinal neurons

To identify gene regulatory networks (GRNs) that control RPCs specification and differentiation we integrated the transcriptomes of all retinal cells from 7.5-16 PCW human embryonic/fetal eyes with those of 10-21 PCW retinas. 48,856 cells were embedded using UMAP and clustered with Seurat (**Figure 1A**). Forty-three clusters were identified: of these, thirty-eight were composed of RPCs and retinal neurons, one of microglia, one of fibroblasts and two expressed markers of more than one cell type (**Table S2**). These last four clusters and an amacrine cell cluster with high expression of mitochondrial markers (cluster 41) were removed prior to further analysis. For illustrative purposes, all cell clusters defined as the same cell type, are annotated with the same color (**Figure 1A**).

Six RPC clusters were identified with enriched expression of typical markers including *SFRP2, DAPL1, HES1, HMGA1, PAX6, RAX* etc. Adjacent to the RPCs, a Muller glia cell cluster was found with high expression of *RLBP1, SLC1A3, APOE, VIM, CLU, SOX9* (**Table S2**). Eight clusters named proliferating RPCs were also found next to RPCs and were characterized by high expression of proliferation markers (e.g., *MKi67, TOP2A*) and RPCs markers (*SFRPP2, SOX2, HES1, ID3* etc.) (**Figure 1A**). All differentiated retinal cell types were easily identified based on the expression of characteristic cell markers (**Table S2**). Between the RPCs and retinal neurons, we identified a progenitor cell cluster with high expression of *ATOH7, HES6 and DLL3*, which was defined as the transitional T1 cluster. Two other cell clusters emanating from T1 were identified leading to the horizontal and amacrine cells clusters, and to photoreceptors and bipolar cells. These were defined as T2 and T3 based on the high expression of *PTF1A* and *PRDM13* and *FABP7* and *OTX2* respectively (**Table S2**).

To fully define the transitions from RPCs to T1, T2 and T3 neurogenic progenitors, we performed pseudo-temporal analysis of gene expression changes following cell cycle regression (**Figure 1B**). This analysis revealed bimodal densities of RPCs, mirroring the transition between early and late RPCs described by Lu and colleagues [5], followed by the transitional T1 and then T2 and T3 neurogenic progenitors (**Figure 1C, D**). Early and late RPCs showed overall similar expression patterns of markers, but also some differences in the expression level. For example, the inhibitor of Wnt pathway (*SFRP2*) was more highly expressed in the early RPCs, while Muller glia cell markers (*VIM, FOS*) were more prominent in the late RPCs (**Figure 1E, Table S2**). High expression of well described neurogenic markers such as *ATOH7, HES6* and *OLIG2* was characteristic of the T1 and to a lesser extent of the T3 progenitor cluster. In accordance with the position of T2 between horizontal and amacrine cells, high expression of key transcription factors (*ONECUT2, TFAP2C*) required for differentiation to these two

lineages was observed (**Table S2**). The transitional cluster T3 shared the expression of highly expressed T1 markers, and additionally displayed ethe expression of typical photoreceptor precursors (*RXRG, NRL*) and bipolar cell markers (*VSX1*), consistent with its positioning between the T1, and photoreceptor and bipolar cell clusters (**Figure 1A, C, Table S2**). To tease out the transcriptional machinery that controls the specification of each differentiated retinal cell type and their transition, separate pseudo-temporal analyses of RGCs, horizontal and amacrine, and photoreceptor and bipolar cells were conducted (**Figures S2, Table S2**). These analyses show that RGCs go through a T1 transition (data not shown), horizontal and amacrine cells through a T1-T2 (**Figure S2A**) and photoreceptor and bipolar cells through a T1-T3 transition state (**Figure S2B**), corroborating recently published scRNA-Seq data on few stages of human fetal retinal development [6].

During the course of these analyses many transcription factors highly expressed on defined progenitor subtypes or retinal neurons were identified. Human horizontal cells express the well-known markers such as *ONECUT1, ONECUT3, LHX1, PROX1* and show considerable overlap in gene expression with amacrine cells which are characterised by high expression of *TFAP2A, LHX9* and *MEIS2* (**Figure S2A, Table S2**). The cone and rod precursors were characterised by high expression of *THRB*, and *NRL* and *NR2E3* respectively, but interestingly also shared high expression of RGC and horizontal and amacrine cell markers (**Figure S2B, Table S2**). For example, high expression of RGC marker *SNCG* was found in the cone precursor cluster, while rod precursors displayed high expression of *PROX1*, a marker of retinal progenitors, horizontal and AII amacrine cells [21, 22]. These findings were further corroborated by immunofluorescence analysis showing co-expression of cone photoreceptor marker RXRG with RGC marker SNCG in the 8 and 11 PCW, which tailed off at 12 PCW (**Figure S3**). Notably, these co-expressing cells were in a proliferative state at 8 PCW, but not in the later stages of development.

ST analyses reveal the location of early RPCs in the CMZ of 8-13 PCW developing human eyes

Our pseudo-temporal analysis above and a recent scRNA-seq analysis of human retinal development identified clear transcriptional signatures of early and late RPCs [5], however their spatial location in the developing retina has not been defined to date. To investigate this in detail, we performed ST analysis on an 8 PCW eye sample (**Figure 2A-D**) revealing the presence of 12 spatially organised cell clusters (**Table S3**) including POM, lens, vitreous, corneal stroma, extraocular muscle and RPE (**Figure 2C**). Notably, the optic stalk (cluster 9, **Figure 2C**) was clearly defined by the ST analysis and characterised by high expression of *PAX2, ZIC1, HES1, SOX2, LHX2, THY1, PAX5* and *ID3* (**Table S3**). Histologically, a single cell layer was present at the peripheral retina, but outer and inner neuroblastic layers were present in the central retina (**Figure 2A**, **C**). The cells residing in the inner neuroblastic layer of central retina (cluster 5) were characterised by high expression of RGC markers such as *GAP43, PRPH, SNCG, INA, NEFL* and *NEFM*, corroborating our previous immunofluorescence findings [2].

High expression of genes typically marking the ciliary margin (e.g., *WNT2B* [23]), eye field (*RAX* and *PAX6),* earlyRPCs (*ID3, HMGA1, MDK, SFRP2)* and pigmented cell (*PMEL, TYR* and *TYRP1)* markers (**Table S3**) were characteristic of the cluster 4, which was defined as CMZ. Subclustering of cluster 4 resulted in two further cell subclusters (**Figure 2E, F**) with subcluster 0 expressing at high level marker genes associated with early RPCs (*FGF19, SFRP2, DAPL1, ZIC1, ID1, ID3, HMGA1, EEF1A1, TPT1, TMSB4X*), and subcluster 1 expressing at high level ciliary body (*KCNJ8, TPM2, ADGRA2A*) and iris pigment epithelium signature markers (*TYR, TRYP1, DCT, SILV, MLANA, PMEL)*. In contrast, the progenitor cells residing in the outer neuroblastic layer of the central retina (cluster 10, **Figure 2C**) were characterised by high expression of late RPC markers such as *RORB, CKB, HES1, HES5, ASCL1, HES6* and *NEUROD1* (**Table S3**). This led us to hypothesise that the early and late RPCs could be spatially located in different regions of the developing human retina. To investigate this further, we generated a gene expression signature for early and late RPCs from our pseudo-temporal analysis. Violin plots (**Figure 2G, I**) indicated the highest expression of early RPCs markers in the subcluster 0 of CMZ (cluster 4), while the highest expression of the late RPCs markers was observed in the outer neuroblastic layer of central retina (cluster 10). These findings were further corroborated by the spatial mapping, showing the predominant localisation of early RPCs in the CMZ, the late RPCs in the central retina (**Figure 2H, J**), and the transitional neurogenic clusters T1, T2 and T3 predominantly in the outer neuroblastic layer of the central retina (cluster 10, **Figure S4**). Interestingly, an early RPCs expression signature was also observed in and around the optic stalk (**Figure 2H**), albeit less intense than in the CMZ.

Similar cell clusters with enriched expression of RPCs and pigmented epithelial cell markers were consistently found in the ciliary margin zone of 10, 11 and 13 PCW eye specimens analysed with the ST approach (**Figure 3A-C**, **Table S3**); but the frequency of early RPCs was very much reduced compared to the 8 PCW sample. We were unable to fit larger size eyes to the current Visium ST slides, thus it was not possible to extend the ST analyses to fetal eyes older than 13 PCW. However, we utilised the scRNA-Seq data and plotted the ratio of early to late RPCs, showing a significant reduction in fraction of early RPCs from 8 PCW onwards (**Figure 3D**).

To complement the ST analyses and obtain more insights into early and late RPC localisation into the CMZ during retinal development, we performed RNA-Scope investigations using a marker of early RPCs (*ZIC1*), iris pigmented epithelium (*TFPI2*), late RPCs/neurogenic progenitors marker (*HES6*) and ciliary body (*OPTC*) in eye samples from 6.3-16 PCW. Consistently with data obtained above, we observed *ZIC1* expression in the CMZ as well as rest of retinal neuroepithelium of 6.6 PCW retinas (**Figure S5A, B**). In contrast, the expression of *HES6* was first seen in the central retina (**Figure S5B**) spreading to the periphery, up to CMZ, from 8 PCW until the last 16 PCW specimens

examined (**Figure S5D, G**, **Figure S6C, D, G**). Although *ZIC1* expression was still present in the CMZ of 8 PCW retinas (**Figures S5C, D**) it was reduced and by 10 PCW, it was localised in small patches at the posterior end of CMZ of less than 50% of eye sections assessed (**Figure S5E, F, G**). In 13 and 16 PCW, *ZIC1* expression was completely absent from the CMZ (**Figure S6A-F**), however its expression in the rest of retinal neuroepithelium persisted in a co-expressing pattern with *HES6* (**Figure S6C, D, G)**. Together these data corroborate localisation of early RPCs in the CMZ of developing human retinas with decreasing frequency from 8 PCW of development and reveal the propagation of neural retina differentiation from the centre to the periphery.

Characterisation of chromatin accessibility during human retinal development

To investigate chromatin accessibility during human retinal development, we performed scATAC-Seq analyses of two developing human eyes (8.5 PCW), and ten retinal samples dissected from 10-21 PCW (**Table S4**). 118,883 cells were captured using the 10XGenomics Chromium Single Cell ATAC Library and Gel Bead Kit. The corresponding scRNA-Seq datasets were used as reference maps to identify ATAC-Seq clusters for each sample (**Figure S7**). Similarly, to scRNA-Seq analysis, defined clusters of corneal stroma, epithelium, endothelium, and periocular mesenchyme, CMZ, extraocular muscle, red blood cells, microglia and optic nerve were found in the 8 PCW human eye samples alongside retinal clusters comprised of RPCs, RGCs, horizontal and amacrine cells and transient neurogenic cluster T1 (**Figure S7A**). All the retinal cell types were present between 10-14 PCW, albeit some at the precursor stage (for example rod precursors at 10 PCW), in addition to RPCs and the three types of transient neurogenic clusters T1, T2, T3 (**Figure S7A**). From 16 PCW, small clusters of microglia and the last-born cell type, Muller glia cells were evident (**Figure S7B**), corroborating the scRNA-Seq data and previously published sequential order of retinal cell emergence [2, 5, 6].

54,045 retinal cells from 12 samples and 9 developmental timepoints encompassing 8-21 PCW with a 3694,28 average median fragments per cell were integrated. The cells were clustered using the chromatin accessibility peaks near previously known marker genes resulting in 22 clusters (**Figure S7C, Table S4**). These included the abundant cell types (for example rods, cones) as well as the rarer cell types (e.g., microglia) and several types of amacrine cells (gabaergic, glycinergic and starburst). The DNA accessibility peaks were classified using annotation from cellranger and associated with promoters (if found within -1000 to +100 bp of the transcription start sites), exons, introns, distal (if found within 200 kb of the closest transcription start site), or intergenic regions (if not mapped to any genes) (**Figure 4A, Table S5**). This analysis enabled us to identify cell type specific regions of accessibility in RPCs, transient neurogenic progenitors T1, T2, T3 and retinal neurons (**Figure 4B**). All clusters displayed scATAC-Seq marker peak enrichment of cell type specific marker genes (**Figure 4C**).

We then went on to predict transcription factor (TF) binding motifs within the scATAC peaks using Signac (**Figure 5A, Table S6**), followed by foot printing validation analysis (**Figure 5B, C**). In RPCs, we identified binding motifs for TFs expressed in the optic stalk (VAX1, VAX2[24]), eye field (RAX[25]) and RPCs (LHX6, VSX2, SOX2[26, 27, 28]) as well as TF binding motifs that were shared with other cell types (for example FOS, NFI family members and SOX6 binding motifs were shared with Muller glia cells, **Figure 5A**). Notably we identified binding motifs for TFs not previously associated with RPCs, for example EN1, GBX2, LBX2, SHOX2 and GSX1 (**Figure 5A,B, Table S6**).

In accordance with sequential emergence of T2 and T3 neurogenic clusters from T1, we observed a set of shared TF binding motifs for NEUROG2, NEUROD1, HAND2, TAL1:TCF3, PITF1A, ATOH7, BHLHA15, MYF5 and ATOH1 (**Figures 5A**). The earliest born retinal cell types, RGCs displayed the POU4F and EBF family members binding motifs, whilst horizontal and amacrine cells were characterised by ONECUT and CUX, and TFAP2 and MEIS family members binding motifs respectively (**Figure 5A, C**). Rod and cone photoreceptors are derived from the T3 neurogenic progenitors; hence shared binding motifs were identified for TFs such as OTX2 [29], CRX [30] and DMBX1 [31] (**Figure 5A, C**), which are well described in the literature for their role in photoreceptor specification. Importantly, shared binding motifs in T3 and photoreceptors were also discovered for PITX1 and GSC TFs not associated previously with a role in rods or cones (**Figure 5A**). Bipolar cells are also derived from the T3 progenitors; hence enrichment of T3 TF binding motifs (such as OTX2) was observed in addition to less well characterised TFs such ZBTB18 (**Figure 5C**). Enrichment of NFI binding motif family members was observed for glia cells (Muller glia and microglia) in accordance with their role in regulating specification of the late-born cell types in the retina [32].

scATAC-Seq enables prediction of gene regulatory networks and novel TFs that govern retinal development

To identify GNRs governing the RPCs differentiation to transient neurogenic progenitors (T1, T2, T3) and the retinal neurons, the upstream regulator tool in IPA was used, combined with overlay analysis of DA peaks. A large number of upstream regulators including TGFβ, IGF-1, FGF-2, Sonic hedgehog were predicted to be activated in RPCs (**Table S7**). Their main common characteristic was predicted activation of key genes encoding proteins important for RPCs proliferation (*HES1* [33], *CCND1* [34], *ID3* [35]) and inhibition of transcription factors (*ATOH7* [36], *NEUROD4, NEUROD1* [37], *PTF1A* [38], and *HES6* [39]) (**Figures 6A, B**) that define the RPC competence to RGCs, horizontal and amacrine cell fates, and photoreceptor differentiation. Amongst the predicted inhibited upstream regulators, we found key transcription factors such as *PAX6* [40]and *ASCL1* [41], shown to control the timing and specificity of retinal neurogenesis (**Figures 6C, D**). Together these findings suggest the

presence of a finely tuned balance between activation of proliferation regulators and inhibition of neurogenic cues to maintain RPC self-renewal. Conversely in the transient neurogenic progenitors, we predicted the activation of upstream neurogenic regulators (e.g., ASCL1 in T1, FOXA2 in T2, GTF2IRD1 in T3), which control the expression of TF necessary for retinal cell type specific differentiation (**Figure S8A-C**) and inhibition of regulators (e.g. BMP4, LIN28A, IGF-1) that govern RPCs fate (**Figure S8D-F**).

Performing the same analysis on the differentiated retinal cells revealed some interesting insights and predicted novel upstream regulators (**Table S7**). For example, Eomesodermin (EOMES), a target gene of Pou4f2, required for RGC and optic nerve development in mouse [42], was predicted to be activated in horizontal cells (**Figure S9B**), resulting in activation of *LHX1* [26, 43], a transcription factor that specifies horizontal cells, while suppressing *LHX9*, a transcription factor required for amacrine cell subtype specification [26]. We identified a novel putative upstream regulator in RGCs, namely KLF2 (**Figure S9A**), which is predicted to inactivate Notch signalling, an important event required for RGC differentiation [44]. Our scRNA-Seq data have shown that the basic-helix-loop-helix PTF1A is highly expressed in the T2 progenitors, which give rise to amacrine and horizontal cells. A similar role for PTF1A has been demonstrated during mouse retinal development [38], albeit transient expression in all types of amacrine cells has been demonstrated in the zebrafish [45]. In accordance with these findings, the PTF1A upstream regulator was identified in amacrine cells, resulting in activation of TFs that promote amacrine cell specification and function *(**Figure S9C**)*.

Our analysis was also able to predict more complex upstream regulators, which are ligand-activated and able to play a role in the control of gene expression in a cell type specific manner. For example, the peroxisome proliferator-activated receptor *γ* (PPAR*G*), was identified as a putative novel upstream regulator in cone photoreceptors (**Figure S9D**) and predicted to activate amongst others *PRDM1*, which has been demonstrated to stabilise photoreceptor cell fate in OTX2+ progenitors by preventing bipolar cell induction [46]. In contrast, the upstream regulator TCF7 in bipolar cells is predicted to suppress *PRDM1* and activate expression of key genes important for bipolar cell function such as *GNAO1* and *TRPM1* [47] (**Figure S9F**). G-protein coupled receptors (GPCRs) play a significant role in many tissues by transducing complex signalling networks that coordinate gene expression. In accordance, one of the putative activated upstream regulators in rods, was the GPCR Rhodopsin (RHO), the most abundant protein in rods which functions as the primary photoreceptor molecule of vision (**Figure S9E**). Our data suggests that RHO may regulate the transcription of rod specific phosphodiesterases (e.g., *PDE6G, PDE6B*) and transducins (*GNAT1, GNGT1*).

scATAC-seq trajectory analysis reveals a role for TEAD transcription factors in retinal lamination and RPC, photoreceptor and RGCs specification

The TF binding motif analysis in RPCs (**Figure 5A**) revealed an enrichment for TEA domain (TEAD) TFs in RPCs (**Figure 7A**, **Table S6**). TEADs play an essential role in mediating YAP-dependent gene expression resulting in transcription of target genes responsible for cell proliferation, inhibition of apoptosis or retinal neurogenesis [48]. During retinal development, YAP's expression is restricted to the outer neuroblastic layer where RPCs reside [49]. The TEAD TFs display different expression patterns in the retina, with Tead2 being highly expressed in the proliferating cells located at the basal side of the outer neuroblastic layer of mouse retina and Tead3 being highly enriched in the inner neuroblastic layer (genepaint.org).

To assess if TEADs binding is essential for RPC development and differentiation, we treated PSCs-derived retinal organoids with TEAD palmitoylation inhibitor MGH-CP1, that blocks TEAD2/4-YAP interaction and suppresses the expression of their target genes in cancer cell lines [50]. Treatment was started at day 25, coinciding with immunofluorescence detection of RPCs in the retinal organoids. Three different doses (2.5, 5 and 10 μM) of MGH-CP1 were added every two days to the culture media for 14 days. Under control conditions (treated with vehicle) retinal organoids develop a bright phase retinal neuroepithelium by day 28 (day 3 of treatment) which expanded over time. By the end of treatment the majority of control retinal organoids (> 75%) had full coverage with bright phase retinal neuroepithelium, with a minority of organoids displayed partial coverage (15%) or no coverage (7%) at all (**Figure S10A**, **B**). In contrast, the great majority of retinal organoids treated with MGH-CP1, displayed either partial or full loss of bright phase neuroepithelium in a dose dependent manner. These findings were fully corroborated by immunofluorescence analyses, which showed the presence of partial retinal neuroepithelium harbouring VSX2+ RPCs in organoids treated with 5 μM MGH-CP1 and much reduced and mislocalised RPCs in the organoids treated with 10 μM MGH-CP1 (**Figure 7B, F**). The latter were also often characterised by the presence of SNCG+ RGCs in the apical layer of the organoids and internal rosettes with VSX2+ RPCs or Ki67 proliferating cells (**Figure 7C-E**). Quantitative analysis demonstrated attenuated RPCs, photoreceptor, and RGCs specification in a dose-dependent manner in retinal organoids treated with MGH-CP1 (**Figure 7f, H-J**). In accordance with TEADs role in activating transcription of genes important for cell proliferation, we noted a significant reduction in the fraction of Ki67+ cells in the retinal organoids treated with the highest dose of MGH-CP1 (**Figure 7G**), which also showed a small but significant increase in the percentage of apoptotic cells (**Figure S10C**).

Discussion

In this study we analyzed the transcriptional and DNA accessibility profiles of over 277,000 cells from 35 human eyes and retinal samples collected from 7.5-21 PCW of human development to delineate the diversification of RPCs and the GRNs that govern their differentiation. Using spatiotemporal single cell RNA-Seq analyses we demonstrate the transient localisation of early RPCs in the CMZ of developing human retinas and propagation of neural retina differentiation from the centre to the periphery. Single cell ATAC-Seq analysis revealed a significant enrichment of TEAD transcription factor binding motifs in RPCs, which when inhibited led to loss of retinal lamination, and attenuated RPCs, photoreceptor and RGCs differentiation.

The human eye is a heterogenous entity, comprised of diverse tissues derived from neuroectoderm, neural crest and mesenchymal cells. Single cell RNA-Seq studies have focused on the transcriptome of developing and/or adult cornea [51], iris, ciliary body [52], neural retina [5, 53 6], RPE and choroid [54 55]. However, a comprehensive single cell spatiotemporal profiling of human developing eyes has not been attempted before. Herein, we undertook scRNA-Seq analysis of 10 developing human eyes encompassing 7.5-10 PCW of development, revealing for the first-time transcriptional signature of neural crest, extraocular muscle and POM cells in addition to ciliary body, iris pigmented epithelium, ocular surface epithelium, stroma and endothelium, neural retina and RPE cells. In all cases, the neural crest cell clusters were closely associated with cell types derived therefrom including corneal stroma and endothelium, POM, periocular connective tissue and melanocytes, or those that require early signals from neural crest for their development such as extraocular muscle [56] and lens fibers [57]. Defects in neural crest formation are at the heart of several severe craniofacial and ocular anomalies including ocular coloboma, glaucoma etc. [17], hence a comprehensive understanding of transcriptome of neural crest cells and its derivatives as described herein, is of great importance for understanding the complexities underlying congenital eye diseases.

Taking advantage of ST, we were able to spatially locate various structures in the developing human eyes of 8-13 PCW, revealing for the first time the spatial and single cell transcriptome of the optic stalk, which forms the primitive connection between the retina and the brain. Importantly, the ST approach enabled us to analyse at single cell resolution the CMZ in developing human eyes of 8, 10, 11 and 13 PCW. Using PSCs-derived retinal organoids, Kuwahara and colleagues provided evidence of a putative CMZ, containing sphere-forming cells, able to generate *de novo* retinal cells [59]. However, whether the human CMZ is a source of RPCs during embryogenesis has been a long-standing question in the field. Using a combination of single cell RNA-Seq and ST, we show the presence and localisation of early RPCs in the CMZ of 8 PCW human eyes with decreasing frequency as development proceeds from 8 PCW onwards. Moreover, the CMZ located early RPCs were characterised by distinct expression signature compared to the ciliary body and iris pigmented epithelium, suggesting that they are a distinct entity to the putative stem cell ciliary body cells suggested by Gautam and colleagues [52]. Together our data provide for the first-time evidence on the transitional presence of early RPCs in the CMZ at a particular stage of human retinal development that needs to be functionally validated both *in vivo* using a larger number of early eye specimens and *in vitro* in retinal organoids by combining cell barcoding with single cell analyses. We were also able to locate some of the early RPCs in and around the optic stalk. The region around the optic stalk is initially continuous with the optic fissure margins. In the human eye fissure closure begins in the middle and most of the posterior part eventually forms part of the optic nerve head. Like the CMZ, the fissure margins are a zone of transition between the neural retina and RPE, thus it is likely that the optic stalk region may retain characteristics in common with the CMZ. In accordance, high expression of early RPC markers (*SOX2, HES1, ZIC1, NR2F1, LHX2*) and typical optic stalk marker PAX2, was observed in the optic stalk cluster (**Table S3**). The relationship between the early RPCs and optic stalk and how these contribute to retinogenesis, remains to be further investigated.

In addition to early and late RPCs, we were able to demonstrate the presence of transient neurogenic progenitors, named T1, T2, and T3 by both scRNA- and ATAC-seq, corroborating recently published data [6]. The pseudotime analyses demonstrated that RGCs transit through the T1 neurogenic progenitors, horizontal and amacrine cells via T1, T2 progenitors and bipolar cells and photoreceptors via T1, T3 progenitors. We were also able to demonstrate that both cone and rod photoreceptors develop from a precursor stage which precedes their maturation. Notably, we also observed at the early stages of development some “plasticity” regarding lineage transcription factor expression such as *SNCG*, a marker of RGCs was expressed in cone precursors and *PROX1*, a marker of horizontal and amacrine cells expressed in rod photoreceptors. Strikingly, Prox1 expression has been noted in rod progenitors as well as the Müller glia cells in the fish [60], while high co-expression of RGC markers was reported recently in the proliferating cones of retinoblastoma pluripotent stem cell derived organoids as well as patient retinoblastoma samples [61, 62]. Together these data suggest that although the expression of certain TFs maps to individual retinal cell trajectories, some of these may be reused temporary during the development of other retinal cell types or redeployed upon re-entry into the cell cycle and conversion to malignancy.

We complemented the scRNA-Seq with ATAC-Seq analysis of similarly staged developing human eyes and retinas. We found the scATAC-Seq to be more informative when it came to subclustering of amacrine cells into the three main subtypes namely gabaergic, glycinergic and starburst amacrine cells. Notably we also found scATAC-Seq to predict cell type clusters at earlier stages than scRNA-Seq analysis. For example, horizontal and amacrine cell precursors were identified at 8 PCW in the scATAC-Seq but 10 PCW in the scRNA-Seq. This could be due to DNA chromatin accessibility preceding key TF expression and cell fate determination. Similarly to recently published studies of

scATAC-Seq in the developing and adult retina [7, 10, 11, 12], we were able to reveal the accessible chromatin regions and putative TF binding motifs for the RPCs, transient neurogenic progenitors and all the retinal cells as they emerged during the course of development. This identified both well characterised TFs (e.g., POU4F in RGCs, ONECUT2 in horizontal cells, RAX in RPCs) as well as novel TFs (e.g. EN1, GBX2, LBX2, SHOX2, TEAD1-3 and GSX1 in RPCs), which deserve further functional validation in animal models and retinal organoids.

The scATAC-seq analyses enabled us to predict GRNs and novel TFs that govern retinal neurogenesis. In particular, enrichment of TEAD binding in RPCs was highlighted, suggesting a role for the Yap-Hippo signalling pathway during human retinal development. Evidence demonstrating the role of Hippo-Yap signalling during retinogenesis in mice has shown that developing retinas devoid of Yap display disrupted apical junctions, rosette formation and loss of laminar arrangements [49]. Our data fully corroborate these findings, showing the presence of rosettes within retinal organoids and disrupted laminar organisation when TEAD binding is disrupted. In addition to playing a role in retinal lamination, activated Yap has been shown to simultaneously interact with TEADs and Rx1 during zebrafish embryogenesis to drive the expression of proliferation-related genes and attenuate the trans-activation of photoreceptor genes respectively [63]. In our organoid model, we observed attenuated photoreceptor differentiation, which suggests that upon inhibition of TEAD binding, Yap may interact with the human ortholog of Rx1 (RAX) to suppress photoreceptor differentiation. Importantly we observed a significant reduction in RPCs and cellular proliferation at the highest dose of the inhibitor, but reduced RGCs and photoreceptor specification at both 5- and 10 μM doses, suggesting that the latter two events may be more sensitive to TEAD inhibition. While activation of Hippo-Yap signalling has been studies in the context of Muller glia cell activation [64, 65], our data provide for the first-time important insights into the function into Hippo-Yap and related TEAD signalling during the very early stages of human retinogenesis.

The data generated herein have been submitted to open access online resources adding valuable information to the currently existing scRNA and ATAC-Seq to increase the sample and read size, but most importantly the novel use of ST to reveal the localisation of RPCs and neurogenic progenitors provides important insights into the long-debated source of retinal progenitors during human development.

Methods

scRNA- and -ATAC-Seq

11 samples of developing human eyes and 14 samples of neural retina from 7.5-21 post-conception weeks (**Tables S1, S4**) were obtained from the Human Developmental Biology Resource under ethics permission 08/H0906/21+5 issued by the NorthEast Newcastle and North Tyneside 1 Research Ethics Committee. All samples were isolated and dissociated to single cells using a neurosphere dissociation kit (Miltenyi Biotech). Approximately 10,000 cells from each sample were captured, and sequencing libraries generated using the Chromium Single Cell 3’ Library & Gel Bead Kit (version 3, 10x Genomics). 10,000 of the subsequent nuclei were captured, and sequencing libraries generated using the Chromium Single Cell ATAC Library & Gel Bead Kit (version 1, 10x Genomics). Single cell RNA-Seq libraries were sequenced to 50,000 reads per cell and scATAC-Seq libraries were sequenced to 25,000 reads per nucleus on an Illumina NovaSeq 6000.

scRNA-Seq analysis

The BCL files were de-multiplexed using CellRanger mkfastq version 3.01 and then aligned and quantify them against the human reference genome GRCh38 using Cellranger count. We performed quality control checks for each sample in R for each sample and removed any cells with fewer than 1000 reads or 500 genes or greater than 10% mitochondrial reads. Any cells which expressed haemoglobin genes were also removed from the analysis. Doublets were predicted using DoubletFinder and filtered from the data. The Seurat R package (version 4.3.0) was then used to process the data prior to integration. Firstly, the raw data was normalised using the standard parameters. The FindVariableFeatures function used to select 2000 highly variable genes. The data was then scaled using ScaleData and the following variables were regressed out "percent.mt", "nCount_RNA", "nFeature_RNA". Principle component reduction with the 2000 highly variable genes selected, was applied to the scaled data using FindPCA function.

Harmony (version 0.1.1) was used to remove sample batch effects from the data. A Uniform Manifold Approximation and Projection (UMAP) reduction was than applied to the first 10 harmony corrected components. The Seurat graph-based method was used to cluster the data. Resolutions from 0.2 to 2.2 were tested. Differential expression analysis using the standard settings in the FindMarkers function from Seurat were used to identify markers genes within each cluster. Cell types were then assigned to these clusters (**Table S1**). We then performed pseudotime on 4 branches of the UMAP, namely: RPC-T1-T2-T3, RPC, RPC-T1-RGC-T2-HC-AC, RPC-T1-T2-HC-AC, RPC-T1-T2-RGCs-HCs-ACs and T1-T3-Cones-Rods-BC. The data was subset by these cell type groups we re-clustered the data using the method described in the previous section. The cell types were re-annotated to ensure robust assignment of the cell types. Monocle 3 was used to the order the cells and infer a pseudotime trajectory within the separate branches. We used Seurat FindAllMarkers to identify differentially expressed genes in

the different cell types within each branch. The top 10 genes for each cell type, with the cells ordered by pseudotime order, were visualised in heatmaps.

scATAC-ATAC analysis

Peaks were detected using Cellranger ATAC software (version 1.2) in each of the samples. A set of shared peaks was then defined using Bedtools merge (version 2.30) and the Cellranger ATAC reanalyse function was then used to call peaks using the shared peak set. The datasets were imported using Signac and quality control steps were performed to remove cells low quality cells. We excluded cells with fewer than 20% of reads in peak region fragments, or less than 3000 peak region fragments. Cells with a TSS enrichment score of less than 2 and Blacklist ratio greater than 0.05 or a nucleosome signal of less than 4 were also removed from downstream analysis.

Signac was then used to perform term frequency-inverse document frequency (TF-IDF) normalisation and singular value decomposition (SVD) dimension reduction for each individual sample. This was followed by UMAP reduction using components 2 to 30, and cluster analysis using Seurat. Signac was then used to generate a gene activity matrix based on open regions for each cell and FindAllMarkers from Seurat was used to predict upregulated genes for each cluster. These gene lists were used to assign cell type identity to the clusters. Retinal cell types were selected from each sample for integration and the normalisation, dimension reduction, and clustering steps, and cell type annotation steps described for the individual samples were applied to the combined dataset. We then identified differentially accessible peaks for each of the annotated cell types using the logistic regression (LR) test from the FindAllMarkers function. The average peak value for each cell type was calculated and differentially accessible peaks were plotted using the ComplexHeatmap package.

Chromvar was used to compute per cell motif activity scores for each cell and FindAllMarkers was used to compute enriched motifs for each cell type. The top enriched motifs ordered by average difference in z-score between each cell type are shown in a heatmap generated with the ComplexHeatmap package. Motif plots and Footprint plots were generated using Signac. The differentially accessible peaks were analysed using Qiagen Ingenuity Pathway Analysis (IPA). The lists of promoters were used to predict upstream regulators and motif data was overlaid onto the prediction to look for consensus in the predictions.

ST

Fresh frozen sections of four fetal retina samples of 8, 10, 11 and 13PCW were used for the spatial transcriptome analyses performed with the Visium Spatial Gene Expression kit from 10XGenomics. First the tissue optimisation was performed defining 30 minutes as the most optimal permeabilization time window. The ST procedure was performed according to manufacturer's instructions. Four tissue sections from each sample were carefully placed into the four capture areas, fixed, haematoxylin and eosin stained and images in order to preserve histological information. This makes it possible to overlay the cell tissue image and the gene expression data in a later step. After permeabilization, reverse transcription reagents were added on top of the tissues. The tissues were subsequently removed, leaving the cDNA coupled to the arrayed oligonucleotides on the slide. The cDNA-RNA hybrids were cleaved off the chip and the sequencing libraries were prepared. The sequencing depth varied between 25,000,000 and 200,000,000 million reads.

ST analysis

Spaceranger version 1.0 was used to demulitplex the data, generate gene expression matrices by aligning the FASTQs to GRCH38. The pipeline was also used to calculate spot co-ordinates using fiducial detection and to identify the spots covered by the tissue. The results from Spaceranger were imported into R using Spaniel (version 1.12). The data from each sample, which was derived from 4 consecutive sections within the tissue, was normalized and clustered, with a resolution of 0.5, using Seurat as described in the scRNA analysis section. FindAllMarkers was used to identify differentially expressed genes clusters. Spaniel was used to visualize the clusters overlaid on the tissue and generate spatial expression plots.

RNAscope

RNAscope *in situ* hybridization assay was used to determine the expression profile of *ZIC1, TFPI2, OPTC* and *HES6* during the development of the human retina. Formalin-fixed, paraffin embedded staged human fetal eyes were provided by the MRC/Wellcome Trust funded Human Developmental Biology Resource; www.hdbr.org). Tissue sections were taken on a microtome at 8µm intervals to SuperFrost microscope slides and baked for 1 h at 60°C before the paraffin was removed in xylene. The sections were first dehydrated in two changes of 100% ethanol before a target retrieval was performed by heating the sections for 8 min at 95°C and incubating with a protease enzyme cocktail (ACD- Cat. No. 322381) for 15 min at 40°C. RNAscope probes Hs-TFPI2 (ACD- Cat No. 470361), Hs-OPTC-C3 (ACD- Cat No. 1165211-C3), Hs-ZIC1-C2 (ACD- Cat No. 542991-C2), Hs-PAX2 (ACD- Cat No. 442541), Hs-HES6-C3 (ACD- Cat No. 521301-C3) were hybridised to the tissue for 2 h at 40°C followed by multiple rounds of signal amplification. Positive (ACD- Cat No. 320861) and negative (ACD- Cat No. 320871) control

probes were used to confirm specificity. The annealed probes were detected using Opal fluorophores OPAL 570 (c1) OPAL 650 (C2) and OPAL 520 (C3) and imaged using a Zeiss microscope and ZEN software.

Immunofluorescence analyses (IF)

Human fetal retinal tissue was fixed and IF performed on cryostat sections as previously described (Mellough et al., 2019). Sections were reacted against the following primary antibodies: CRX (1:200, Abnova, H00001406-MO2), Ki67 (1:200, Abcam, ab15580), Recoverin (1:1000, Millipore-Merck, ab5585), RXRγ (1:200, Santa Cruz Biotechnology, sc-555), SNCG (1:500, Antibodies.com, A121664) and VSX2 (1:100, Santa Cruz, sc-365519). Secondary antibodies were conjugated to Alexa488 (Jackson Immuno Research Laboratories), Cy3 (Jackson Immuno Research Laboratories) and Alexa 647 (Thermo Fisher). Antibody specificity was assessed by omitting the primary antibodies.Images were obtained using a Zeiss Axio Imager.Z1 microscope with ApoTome.2 accessory equipment and AxioVision or Zen software. Between 5 and 10 images were collected from each IF analyses. Images are displayed as a maximum projection and adjusted for brightness and contrast in Adobe Photoshop CS6 (Adobe Systems).

Retinal organoid differentiation

WT2 hiPSCs derived and characterised in our lab [66] were expanded in mTESR™1 (StemCell Technologies, 05850) on growth factor reduced Matrigel (BD Biosciences, San Jose, CA) coated plates at 37°C and 5% CO_2. Retinal organoids were generated as followed: hiPSCs were dissociated into single cells using Accutase (Gibco, A1110501) and seeded at a density of 7,000 cells/well onto U-bottom 96-well plates (Helena, 92697T) prior coated with Lipidure (AMSbio, AMS.52000011GB1G) in mTeSR™1 with 10 µM Y-27632 ROCK inhibitor (Chemdea). After 2 days, 200 µl of differentiation medium as described in Dorgau et al. [67] was added. Half of the differentiation medium was changed every 2 days until day 18 of differentiation. Then, the media was supplemented with 10% Fetal Calf Serum (FCS; Life Technologies, UK), Taurine (Sigma-Aldrich) and T3 (Sigma-Aldrich) and retinal organoids were transferred to 6-well low attachment plates (Corning, 3471). Retinoic Acid (RA; 0.5 µM; Sigma-Aldrich) was added from day 90 to day 120 of differentiation. Media was changed every 2-3 days. MGH-CP1 (Sigma-Aldrich) at different concentrations (2.5µM, 5µM and 10µM) or the vehicle control (DMSO) were added to the culture media and refreshed at every media change. Retinal organoids were incubated with MGH-CP1 or DMSO for 2 weeks, starting at day 25 of differentiation and were collected after the incubation period at day39 of differentiation.

Declarations

Data availability: All the single cell data have been deposited to GEO under the following accession numbers: GSE234971. The analyses scripts are described in: https://github.com/RachelQueen1/BBSRC_Retina.

To review GEO accession GSE234971:

Go to https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.ncbi.nlm.nih.gov%2Fgeo%2Fquery%2Facc.cgi%3Facc%3DGSE234971&data=05%7C01%7Cmajlinda.lako%40newcastle.ac.uk%7Ccd09a768126d4d363a7c08db7ec68cf5%7C9c5012c9b61644c2a91766814fbe3e87%7C1%7C0%7C638243163457404185%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=JD3QxqfhcOj2emgoLK2sq4nfFWsR3%2BSfYDKamJcU8p8%3D&reserved=0

Enter token yvspieyszvivhen into the box

Acknowledgements: The authors are grateful for funding received from BBSRC UK (BB/T004460/1) and MRC UK (MR/S035826/1, MR/S036237/1). The human embryonic and fetal material was provided by the Joint MRC / Wellcome (MR/R006237/1, MR/X008304/1 and 226202/Z/22/Z) Human Developmental Biology Resource (https://www.hdbr.org/). The authors would like to acknowledge the Bioimaging Unit (Newcastle University), especially Dr Veronika Boczonadi, for their support and assistance in this work, and Prof. Roy Quinlan (University of Durham) for help with the definition of lens specific clusters in the ST analyses.

Competing interests: No competing interests declared.

Author contributions:

BD –performed experiments, data collection and analyses, figure preparation

JC – performed experiments, data collection, fund raising

AR, DZ, MC, RH, JC, TD - performed experiments, data collection

AP, JS, DRF – provided key reagents, data discussion and fund raising

RQ – bioinformatic data analyses, data submission, figure preparation, contributed to manuscript writing and fund raising

ML – experimental design, data analyses, figure preparation, manuscript writing and fund raising

References

1. Marquardt T, Gruss P. Generating neuronal diversity in the retina: one for nearly all. *Trends Neurosci* **25**, 32-38 (2002).
2. Mellough CB*, et al.* An integrated transcriptional analysis of the developing human retina. *Development* **146**, (2019).
3. Aldiri I*, et al.* The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron* **94**, 550-568 e510 (2017).
4. Hoshino A*, et al.* Molecular Anatomy of the Developing Human Retina. *Dev Cell* **43**, 763-779 e764 (2017).
5. Lu Y*, et al.* Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. *Dev Cell* **53**, 473-491 e479 (2020).
6. 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 e1644 (2020).
7. Wang SK*, et al.* Single-cell multiome of the human retina and deep learning nominate causal variants in complex eye diseases. *Cell Genom* **2**, (2022).
8. 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).
9. Singh RK, Winkler PA, Binette F, Petersen-Jones SM, Nasonkin IO. Comparison of Developmental Dynamics in Human Fetal Retina and Human Pluripotent Stem Cell-Derived Retinal Tissue. *Stem Cells Dev* **30**, 399-417 (2021).
10. Lyu P*, et al.* Gene regulatory networks controlling temporal patterning, neurogenesis, and cell-fate specification in mammalian retina. *Cell Rep* **37**, 109994 (2021).
11. Xie HH*, et al.* Chromatin accessibility analysis reveals regulatory dynamics of developing human retina and hiPSC-derived retinal organoids. *Sci Adv* **6**, (2020).
12. Finkbeiner C, Ortuno-Lizaran I, Sridhar A, Hooper M, Petter S, Reh TA. Single-cell ATAC-seq of fetal human retina and stem-cell-derived retinal organoids shows changing chromatin landscapes during cell fate acquisition. *Cell Reports* **38**, (2022).
13. Fischer AJ, Reh TA. Identification of a proliferating marginal zone of retinal progenitors in postnatal chickens. *Dev Biol* **220**, 197-210 (2000).
14. Tropepe V*, et al.* Retinal stem cells in the adult mammalian eye. *Science* **287**, 2032-2036 (2000).
15. Gualdoni S*, et al.* Adult Ciliary Epithelial Cells, Previously Identified as Retinal Stem Cells with Potential for Retinal Repair, Fail to Differentiate into New Rod Photoreceptors. *Stem Cells* **28**, 1048-1059 (2010).
16. Belanger MC, Robert B, Cayouette M. Msx1-Positive Progenitors in the Retinal Ciliary Margin Give Rise to Both Neural and Non-neural Progenies in Mammals. *Developmental Cell* **40**, 137-150 (2017).
17. Williams AL, Bohnsack BL. The Ocular Neural Crest: Specification, Migration, and Then What? *Front Cell Dev Biol* **8**, (2020).
18. Doane KJ, Ting WH, McLaughlin JS, Birk DE. Spatial and temporal variations in extracellular matrix of periocular and corneal regions during corneal stromal development. *Experimental Eye Research* **62**, 271-283 (1996).
19. Bi L, Lwigale P. Transcriptomic analysis of differential gene expression during chick periocular neural crest differentiation into corneal cells. *Dev Dyn* **248**, 583-602 (2019).
20. Bassett EA, Wallace VA. Cell fate determination in the vertebrate retina. *Trends Neurosci* **35**, 565-573 (2012).
21. Dyer MA, Livesey FJ, Cepko CL, Oliver G. Prox1 function controls progenitor cell proliferation and horizontal cell genesis in the mammalian retina. *Nat Genet* **34**, 53-58 (2003).
22. Perez de Sevilla Muller L, Azar SS, de Los Santos J, Brecha NC. Prox1 Is a Marker for All Amacrine Cells in the Mouse Retina. *Front Neuroanat* **11**, 39 (2017).
23. Kubo F, Takeichi M, Nakagawa S. Wnt2b controls retinal cell differentiation at the ciliary marginal zone. *Development* **130**, 587-598 (2003).
24. Mui SH, Kim JW, Lemke G, Bertuzzi S. Vax genes ventralize the embryonic eye. *Genes Dev* **19**, 1249-1259 (2005).

25. Furukawa T, Kozak CA, Cepko CL. rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina. *P Natl Acad Sci USA* **94**, 3088-3093 (1997).
26. Balasubramanian R, Bui A, Ding Q, Gan L. Expression of LIM-homeodomain transcription factors in the developing and mature mouse retina. *Gene Expr Patterns* **14**, 1-8 (2014).
27. Zou CJ, Levine EM. Vsx2 Controls Eye Organogenesis and Retinal Progenitor Identity Via Homeodomain and Non-Homeodomain Residues Required for High Affinity DNA Binding. *Plos Genet* **8**, (2012).
28. Taranova OV*, et al.* SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Gene Dev* **20**, 1187-1202 (2006).
29. Nishida A*, et al.* Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nat Neurosci* **6**, 1255-1263 (2003).
30. Furukawa T, Morrow EM, Cepko CL. Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* **91**, 531-541 (1997).
31. Wong L, Weadick CJ, Kuo C, Chang BSW, Tropepe V. Duplicate dmbx1 genes regulate progenitor cell cycle and differentiation during zebrafish midbrain and retinal development. *Bmc Dev Biol* **10**, (2010).
32. Clark BS*, 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-+ (2019).
33. Wall DS*, et al.* Progenitor cell proliferation in the retina is dependent on Notch-independent Sonic hedgehog/Hes1 activity. *Journal of Cell Biology* **184**, 101-112 (2009).
34. Das G, Choi Y, Sicinski P, Levine EM. Cyclin D1 fine-tunes the neurogenic output of embryonic retinal progenitor cells. *Neural Dev* **4**, 15 (2009).
35. Du Y, Yip HK. The Expression and Roles of Inhibitor of DNA Binding Helix-Loop-Helix Proteins in the Developing and Adult Mouse Retina. *Neuroscience* **175**, 367-379 (2011).
36. Wu FG*, et al.* Single cell transcriptomics reveals lineage trajectory of retinal ganglion cells in wild-type and Atoh7-null retinas. *Nat Commun* **12**, (2021).
37. Cherry TJ, Wang S, Bormuth I, Schwab M, Olson J, Cepko CL. NeuroD Factors Regulate Cell Fate and Neurite Stratification in the Developing Retina. *J Neurosci* **31**, 7365-7379 (2011).
38. Fujitani Y*, et al.* Ptf1a determines horizontal and amacrine cell fates during mouse retinal development. *Development* **133**, 4439-4450 (2006).
39. Bae S, Bessho Y, Hojo M, Kageyama R. The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development* **127**, 2933-2943 (2000).
40. Philips GT*, et al.* Precocious retinal neurons: Pax6 controls timing of differentiation and determination of cell type. *Developmental Biology* **279**, 308-321 (2005).
41. Nelson BR, Hartman BH, Ray CA, Hayashi T, Bermingham-McDonogh O, Reh TA. Acheate-scute like 1 (Ascl1) is required for normal delta-like (Dll) gene expression and notch signaling during retinal development. *Dev Dyn* **238**, 2163-2178 (2009).
42. Mao CA, Kiyama T, Pan P, Furuta Y, Hadjantonakis AK, Klein WH. Eomesodermin, a target gene of Pou4f2, is required for retinal ganglion cell and optic nerve development in the mouse. *Development* **135**, 271-280 (2008).
43. Poche RA, Kwan KM, Raven MA, Furuta Y, Reese BE, Behringer RR. Lim1 is essential for the correct Laminar positioning of retinal horizontal cells. *J Neurosci* **27**, 14099-14107 (2007).
44. Nelson BR, Gumuscu B, Hartman BH, Reh TA. Notch activity is downregulated just prior to retinal ganglion cell differentiation. *Dev Neurosci-Basel* **28**, 128-141 (2006).
45. Jusuf PR, Harris WA. Ptf1a is expressed transiently in all types of amacrine cells in the embryonic zebrafish retina. *Neural Development* **4**, (2009).
46. Brzezinski JA, Lamba DA, Reh TA. Blimp1 controls photoreceptor versus bipolar cell fate choice during retinal development. *Development* **137**, 619-629 (2010).
47. Morgans CW*, et al.* TRPM1 is required for the depolarizing light response in retinal ON-bipolar cells. *P Natl Acad Sci USA* **106**, 19174-19178 (2009).
48. Lee M, Goraya N, Kim S, Cho SH. Hippo-yap signaling in ocular development and disease. *Dev Dynam* **247**, 794-806 (2018).
49. Kim JY*, et al.* Yap is essential for retinal progenitor cell cycle progression and RPE cell fate acquisition in the developing mouse eye. *Developmental Biology* **419**, 336-347 (2016).
50. Sun Y*, et al.* Pharmacological blockade of TEAD-YAP reveals its therapeutic limitation in cancer cells. *Nat Commun* **13**, (2022).

51. Collin J*, et al.* A single cell atlas of human cornea that defines its development, limbal progenitor cells and their interactions with the immune cells. *Ocul Surf* **21**, 279-298 (2021).
52. Gautam P*, et al.* Multi-species single-cell transcriptomic analysis of ocular compartment regulons. *Nat Commun* **12**, 5675 (2021).
53. Lukowski SW*, et al.* A single-cell transcriptome atlas of the adult human retina. *EMBO J* **38**, e100811 (2019).
54. Collin J*, et al.* Single cell RNA sequencing reveals transcriptional changes of human choroidal and retinal pigment epithelium cells during fetal development, in healthy adult and intermediate age-related macular degeneration. *Hum Mol Genet*, (2023).
55. Voigt AP*, et al.* Single-cell transcriptomics of the human retinal pigment epithelium and choroid in health and macular degeneration. *Proc Natl Acad Sci U S A* **116**, 24100-24107 (2019).
56. Bohnsack BL*, et al.* Development of Extraocular Muscles Requires Early Signals From Periocular Neural Crest and the Developing Eye. *Arch Ophthalmol-Chic* **129**, 1030-1041 (2011).
57. Grocott T, Johnson S, Bailey AP, Streit A. Neural crest cells organize the eye via TGF-beta and canonical Wnt signalling. *Nat Commun* **2**, 265 (2011).
58. Harris WA, Perron M. Molecular recapitulation: the growth of the vertebrate retina. *Int J Dev Biol* **42**, 299-304 (1998).
59. Kuwahara A, Ozone C, Nakano T, Saito K, Eiraku M, Sasai Y. Generation of a ciliary margin-like stem cell niche from self-organizing human retinal tissue. *Nat Commun* **6**, (2015).
60. Cid E*, et al.* Prox1 expression in rod precursors and Muller cells. *Exp Eye Res* **90**, 267-276 (2010).
61. Rozanska A*, et al.* pRB-Depleted Pluripotent Stem Cell Retinal Organoids Recapitulate Cell State Transitions of Retinoblastoma Development and Suggest an Important Role for pRB in Retinal Cell Differentiation. *Stem Cells Transl Med* **11**, 415-433 (2022).
62. Liu J*, et al.* A high-risk retinoblastoma subtype with stemness features, dedifferentiated cone states and neuronal/ganglion cell gene expression. *Nat Commun* **12**, 5578 (2021).
63. Asaoka Y, Hata S, Namae M, Furutani-Seiki M, Nishina H. The Hippo Pathway Controls a Switch between Retinal Progenitor Cell Proliferation and Photoreceptor Cell Differentiation in Zebrafish. *Plos One* **9**, (2014).
64. Kastan N*, et al.* Small-molecule inhibition of Lats kinases may promote Yap-dependent proliferation in postmitotic mammalian tissues. *Nat Commun* **12**, (2021).
65. Kastan NR*, et al.* Development of an improved inhibitor of Lats kinases to promote regeneration of mammalian organs. *P Natl Acad Sci USA* **119**, (2022).
66. Melguizo-Sanchis D*, et al.* iPSC modeling of severe aplastic anemia reveals impaired differentiation and telomere shortening in blood progenitors. *Cell Death Dis* **9**, (2018).
67. Dorgau B*, et al.* Laminin gamma3 plays an important role in retinal lamination, photoreceptor organisation and ganglion cell differentiation. *Cell Death Dis* **9**, 615 (2018).

Figures

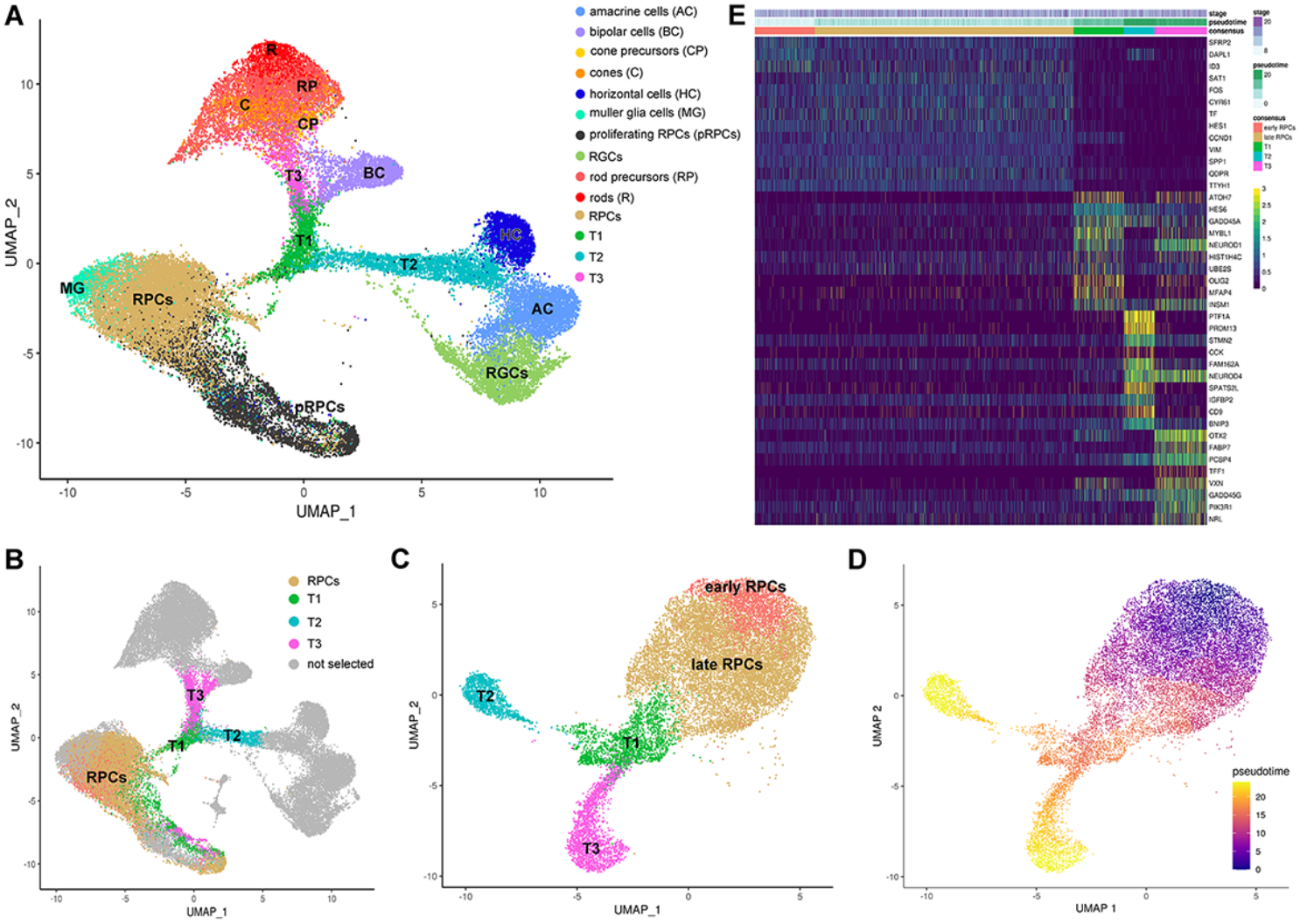

Figure 1

RPCs identification and their developmental trajectories during human retinal development. A) UMAP plot of integrated scRNA-Seq fetal retina cells. Each cluster was identified based on expression of retinal specific cell markers. Highly expressed markers for each clusters are shown in Table S2. **B**) RPCs and transient neurogenic progenitors named T1, T2 and T3 were identified. Highly expressed markers for each cluster along the pseudotime trajectory are shown in Table S2. **C, D**) Pseudotime analysis demonstrating transition from early to late RPCs, and to T1 progenitors, which further commit to either T2 or T3 transient neurogenic progenitors. **E**) Gene expression heatmap showing similarities in gene expression patterns between early and late RPCs, but distinct gene expression signatures in T1, T2 and T3 neurogenic progenitors.

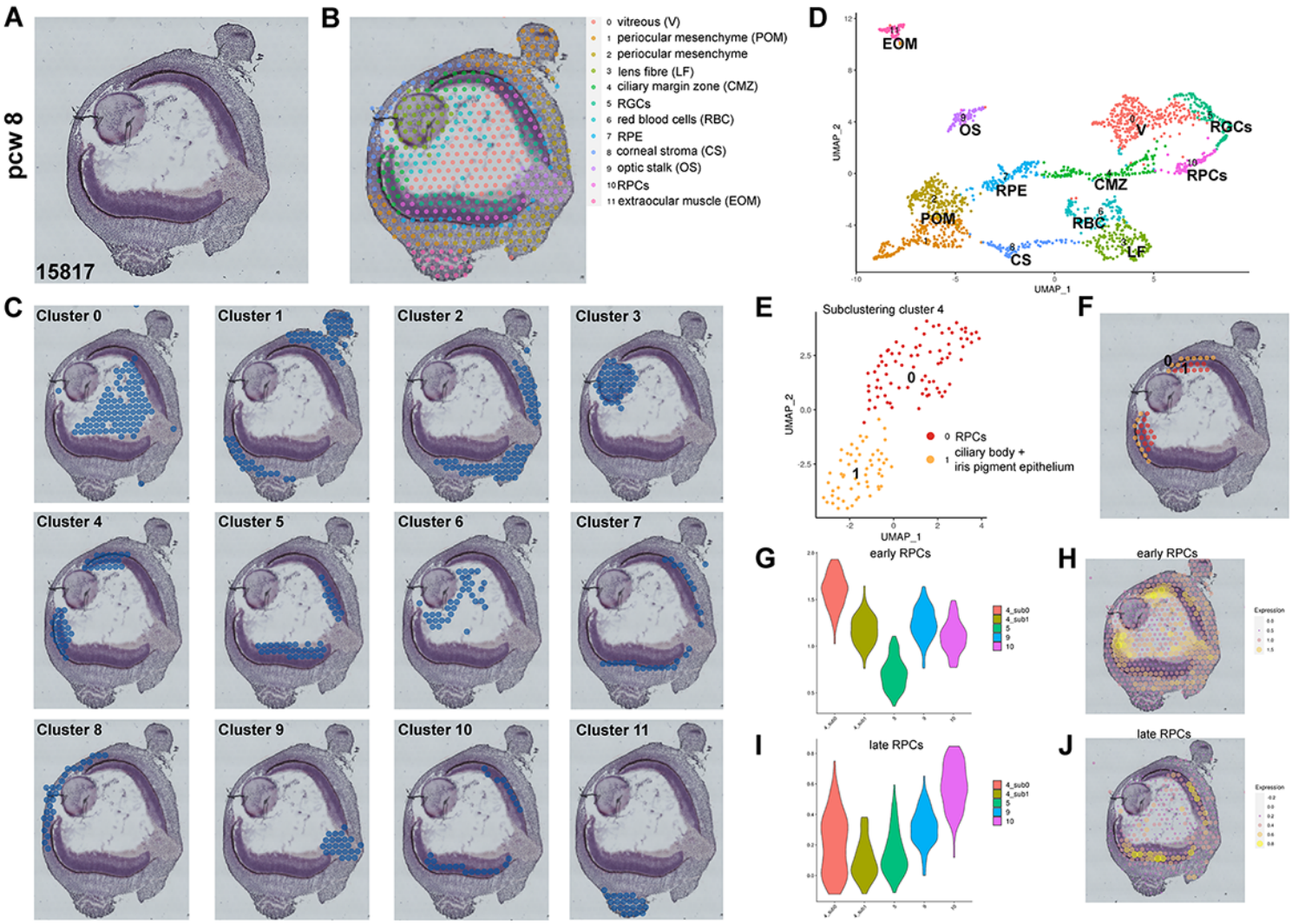

Figure 2

ST analysis of 8 PCW human eye sections reveals the location of early RPCs in the CMZ. A) Representative histological staining of the 8 PCW fresh frozen human eye section. **B, C**) Spatial localisation of the 12 clusters identified from the ST analysis. Highly expressed markers for each cluster are shown in Table S3. **D**) UMAP of spatial transcriptomics scRNA-Seq data. **E**) Subclustering of ciliary margin zone (cluster 4) reveals the presence of two subclusters namely RPCs and ciliary body, and iris pigmented epithelial cells. **G** and **I**) Expression violin plots showing the expression of early RPCs in the peripheral retina (CMZ) and late RPCs in the central retina, respectively. **H** and **J**) Spatial localisation of early and late RPCs in the CMZ and central retina, respectively.

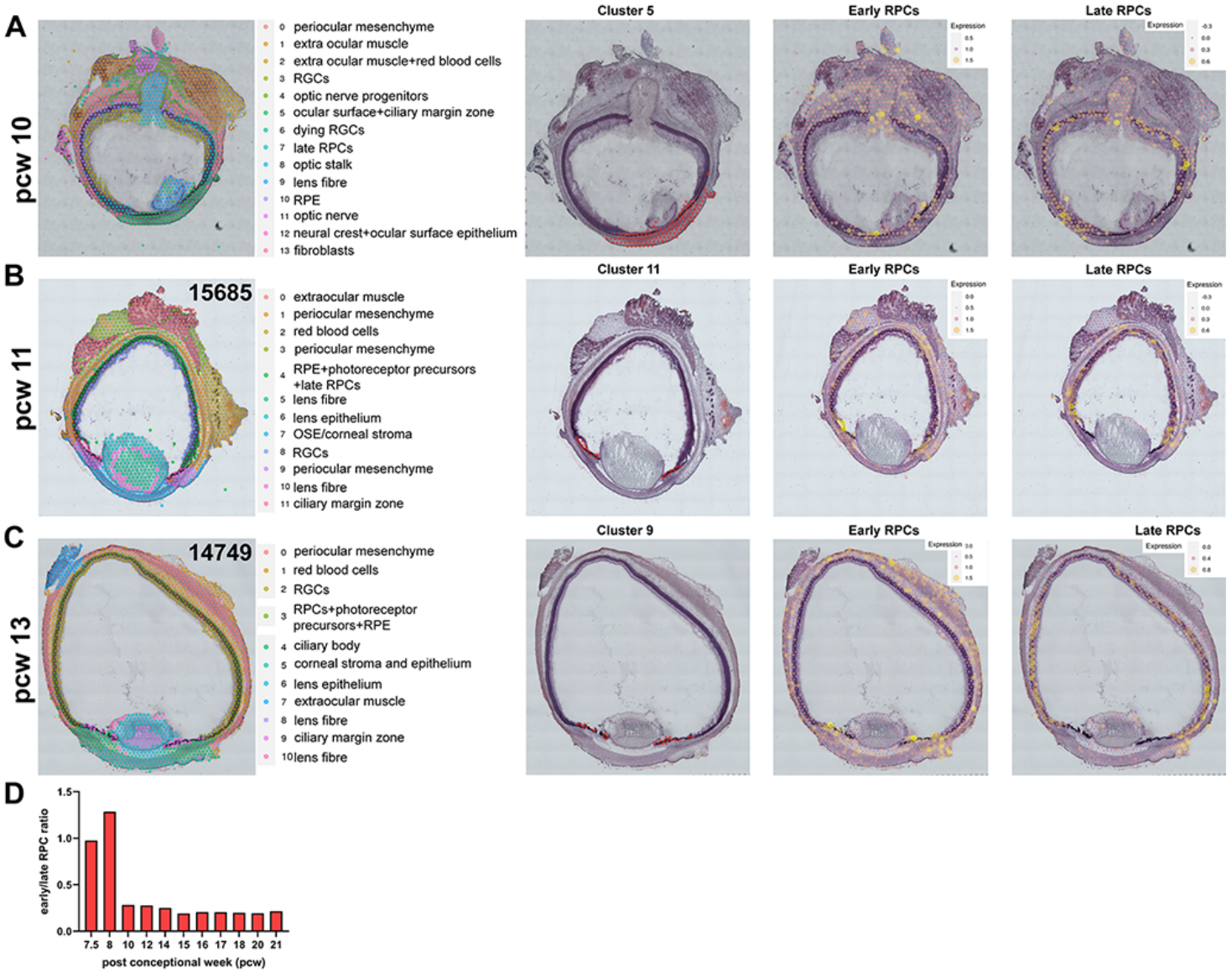

Figure 3

ST analyses demonstrate decreased presence of early RPCs in the ciliary margin zone as development proceeds from 10 to 13 PCW. A-C) Spatial localisation of all cell clusters identified from the ST analyses (left hand side). The CMZ, early and late RPCs cluster localisations are shown on the right-hand side. Highly expressed markers for each cluster are shown in Table S3. Note the scarce presence of early RPCs in the CMZ of 13 PCW human eye. **D**) Early RPCs reach a peak at 7.5-8 PCW and then decline from 10 PCW onwards in the human developing retina. The ratio of early to late RPCs is inferred from the scRNA-Seq data.

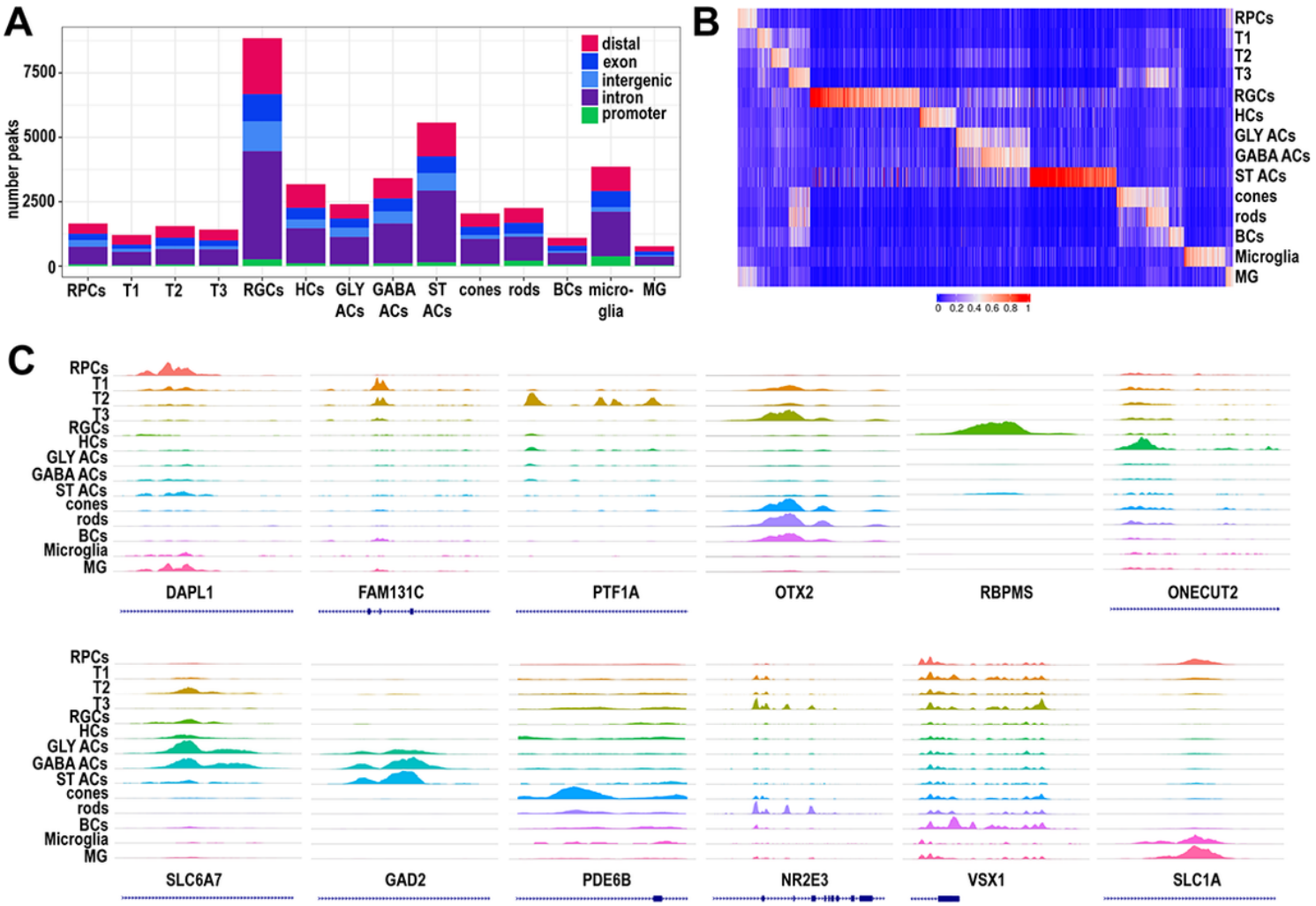

Figure 4

Single cell ATAC-Seq analysis of developing retina samples reveals cell type specific chromatin accessibility profiles. A) The number and type of chromatin accessibility profiles for each cell type. **B**) Heatmap showing differentially accessible of chromatin accessibility peaks (columns) for each cell type (rows). **C**) Representative examples of chromatin accessibility peaks for retinal cell specific marker genes. Each track represents the aggregate scATAC signal of all cells from the given cell type normalized by the total number of reads in TSS regions.

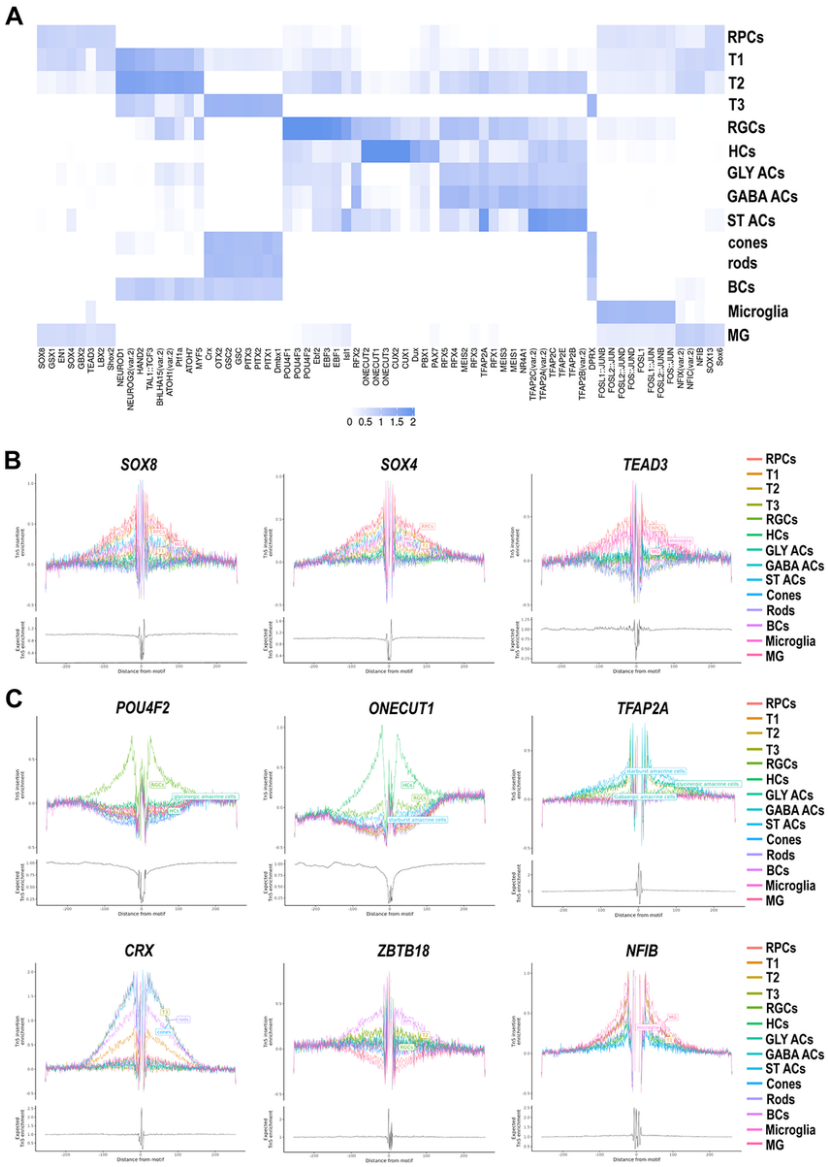

Figure 5

Motif analysis of accessible DNA peaks predicts cell type specific TFs in the developing human retina. A) Heatmap of transcription factor binding motifs enriched in each cell type. More significant enrichment is indicated by the darker colours. **B**) Footprinting analysis of selected TFs predicted to show a significant enrichment in RPCs. **C**) Footprinting analysis of selected TFs predicted to show a significant enrichment in transient neurogenic progenitors and retinal neurons. Additional abbreviations to those mentioned in the main text: Gly ACs- glycinergic amacrine cells, GABA ACs – gabaergic amacrine cells, ST ACs – starburst amacrine cells**,** HCs – horizontal cells, MG- Muller glia cells.

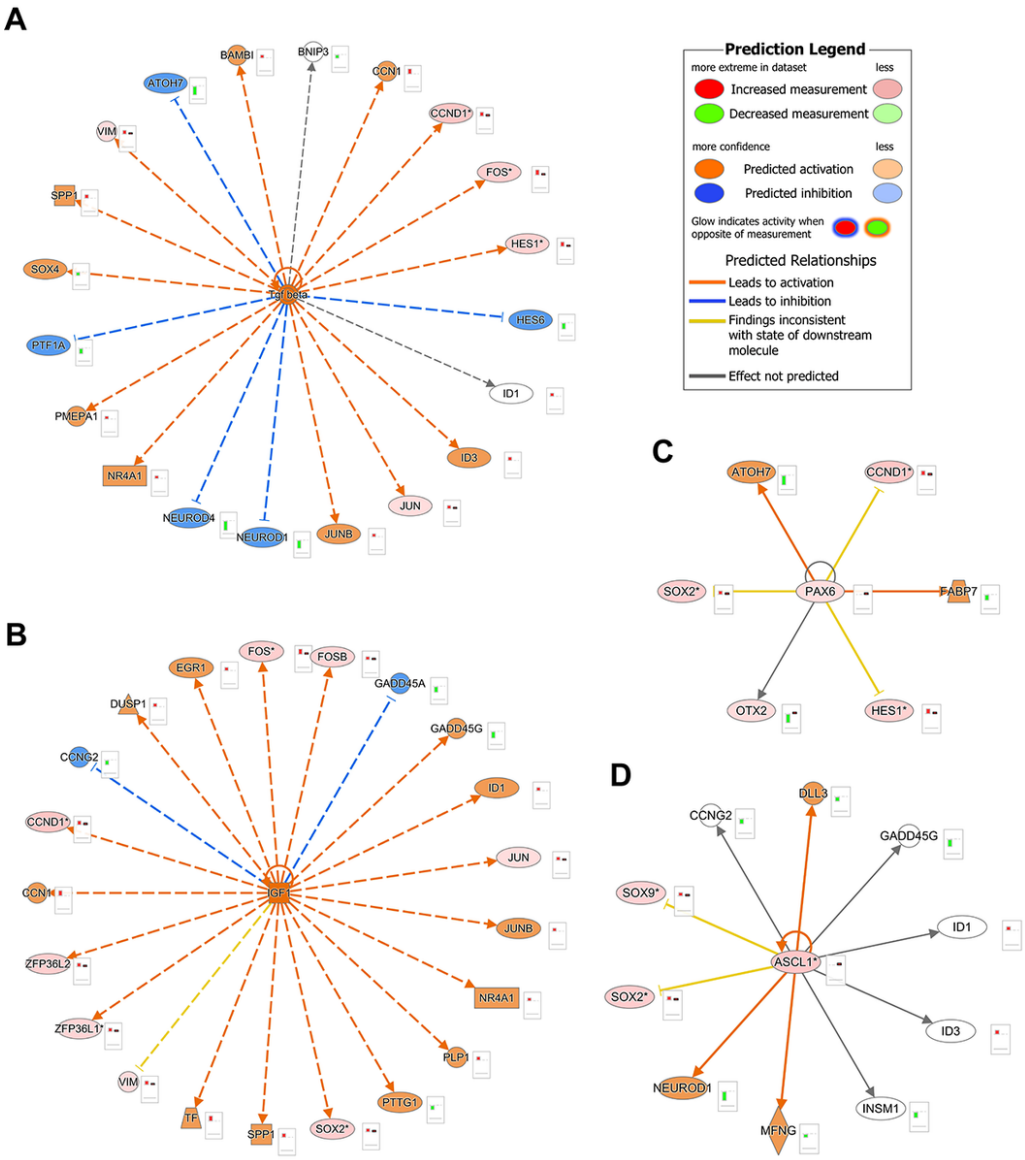

Figure 6

Representative gene regulatory networks in RPCs depicting activated upstream regulators (A, B) and inhibited upstream regulators (C, D) and their target genes. Upstream regulatory networks were generated with IPA using differentially expressed genes from the scRNA-Seq data and differential accessibility analysis in the scATAC-Seq data. The networks show predictions of upstream regulators which might be activated or inhibited to explain observed upregulation/downregulations in the data. The barplots next to each molecule represent the relative expression in the sRNA-Seq (column 1) and scATAC-Seq datasets (column 2). The colours for the network nodes/barplots indicate observed upregulation/ increased chromatin accessibility (red), predicted upregulation/increased chromatin accessibility (orange), observed downregulation (green) and predicted downregulation/ decreased chromatin accessibility (blue). The colour of the edges represents the relationships between the molecules; orange = prediction and observation are consistent with activation; blue = prediction and observation are consistent with downregulation; yellow = prediction and observation are inconsistent; and grey relationship between the molecules is available in the IPA knowledge database. *- indicates duplicates in scATAC-Seq dataset.

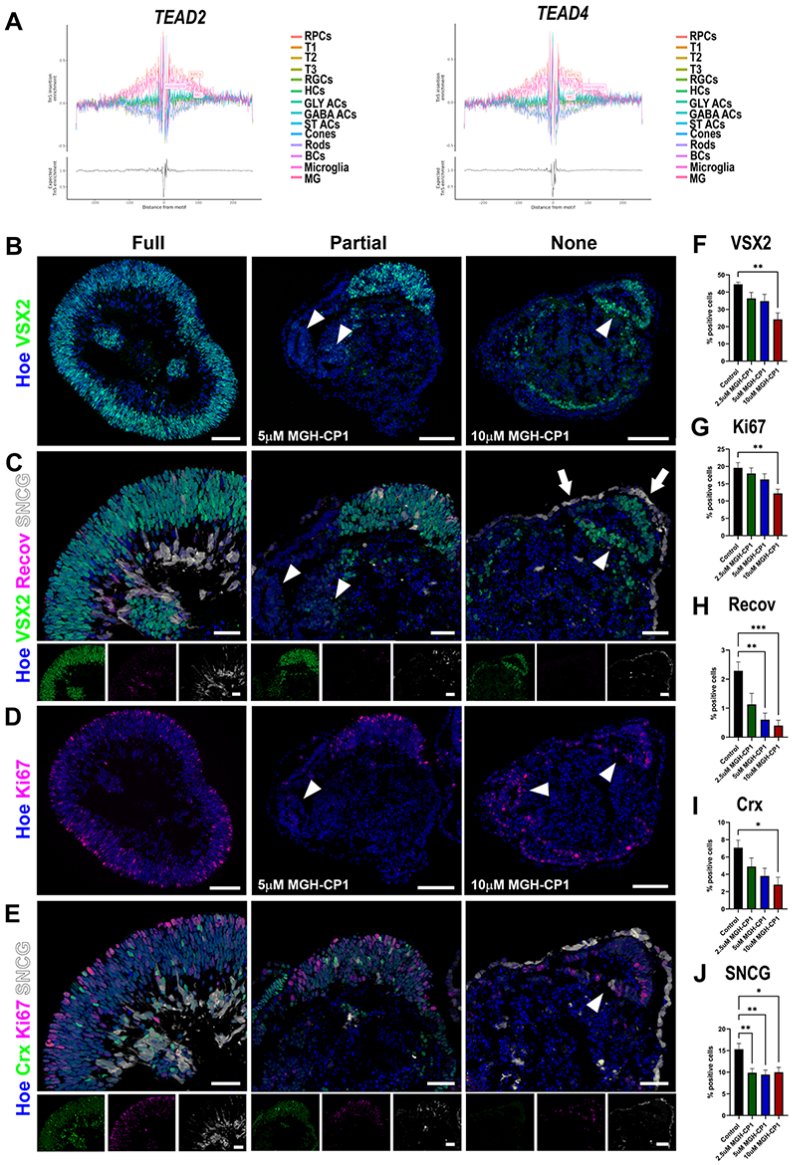

Figure 7

TEAD binding plays a significant role in retinal lamination and RPCs, RGCs and photoreceptor specification. **A**) Footprinting analysis of TEAD2 and TEAD4 showing a significant enrichment in RPCs. Additional abbreviations to those mentioned in the main text: Gly ACs- glycinergic amacrine cells, GABA ACs – gabaergic amacrine cells, ST ACs – starburst amacrine cells, HCs – horizontal cells, MG- Muller glia cells. **B-J**) Quantitative immunofluorescence analyses for the presence of $VSX2^+$ RPCs, $Ki67^+$ proliferating cells, $SCNG^+$ RGCs and $Recoverin^+$ and CRX^+ photoreceptor precursors reveal loss of RPCs, disturbed retinal lamination, and attenuation of photoreceptor and RGCs specification. White arrowheads show the presence of rosettes comprised of RPCs or $Ki67^+$ proliferating cells. Scale bars 100 μM for B-E and 50 μM for the bottom panel insets.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx

- TableS6.xlsx
- TableS7.xlsx
- SupplementcombinedNatCom.pdf