Single cell transcriptomic analysis reveals the impact of elevating neurogenic factor expression on human retinal organoid development

Xiangmei Zhang[1], Igor Mandric[2], Kevin H. Nguyen[1], Thao T. T. Nguyen[1], Matteo Pellegrini[2], James C. R. Grove[1], Steven Barnes[1,3], *Xian-Jie Yang[1,4]

[1]Department of Ophthalmology, Stein Eye Institute
[2]Department of Molecular, Cell, and Developmental Biology
[3]Doheny Eye Institute, [4]Molecular Biology Institute,
University of California, Los Angeles, CA, USA

Corresponding author: Xian-Jie Yang

Address: Stein Eye Institute
UCLA School of Medicine
100 Stein Plaza
Los Angeles, CA 90095, USA

Tel: 1-310-825-7020
Fax: 1-310-825-6992
Email: yang@jsei.ucla.edu

Running title:

Sc RNA-seq analysis of neurogenic factor function in human retinal organoids

Abstract

Transcription factors containing the bHLH motif play critical roles in retinogenesis. However, their precise functions in uncommitted retinal progenitors remain to be fully elucidated. Here, we investigate the roles of ATOH7 and Neurog2 in human ES cell-derived 3D retinal organoids by elevating their expression at the onset of neurogenesis. Using single cell transcriptome analyses, we have identified three progenitor states in retinal organoids: neural stem cell-like progenitors, neurogenic progenitors, and cell cycle-exiting progenitors. The exiting progenitors feed into a postmitotic neuroblast pool that serves as the common source for three early born neuronal lineages. Both ATOH7 and Neurog2 accelerate the transition from the stem cell state to the neurogenic state, and expand the exiting progenitor and neuroblast populations. Despite similar enhancement of retinal ganglion cell production, ATOH7 and Neurog2 show differential promotion of neuronal types with distinct molecular characteristics. Moreover, single cell transcriptome analyses reveal new evidence that ATOH7 and Neurog2 impact uncommitted progenitors through modulation of an interactive gene network, thus providing novel insight into human retinal development.

Introduction

As an integral component of the central nervous system, the vertebrate neural retina retains a highly conserved laminar structure that senses, processes, and delivers visual information to the brain. Classic cell birth dating and lineage tracing studies have established that the seven major neuronal cell types constituting the retinal network are generated in a temporal order from a common ocular progenitor pool during development [1-3]. The ensuing research has ruled out a rigid deterministic cell fate specification mechanism, but instead supports a view that multipotent progenitors progressively evolve through different competence states to enable the sequential production of distinct cell types [1,4]. Cumulative molecular genetic studies have uncovered important roles of cell intrinsic factors involved in retinal development [5]. Among these, transcription factors containing the basic helix-loop-helix (bHLH) motif have emerged as important players regulating the production and differentiation of various retinal cell types [6]. Multiple bHLH factors are expressed during retinogenesis either in proliferating progenitors or in postmitotic neurons; however, their dynamic regulation and function in specific cellular contexts remain to be fully elucidated.

The bHLH factors Atoh7 and Neurog2 are both expressed early in the developing retinal epithelium. Atoh7 plays a critical role in the development of an early born retinal neuronal type, the retinal ganglion cells (RGCs), which project axons through the optic nerve to multiple higher visual centers [2,7]. Both Atoh7 and Neurog2 mRNAs are expressed by subsets of early retinal progenitors, with some co-expression at the protein level during certain time windows [8-10]. Loss of Atoh7 function in mouse, zebrafish, and humans results in a severe reduction of RGCs leading to a diminished optic nerve and blindness [11-15]. Atoh7 deficiency also causes a minor abnormality in the production of cone photoreceptors, another early-born cell type in the retina [11]. In contrast, genetic ablation of Neurog2 yields a transient stall of neurogenesis but without severe lasting deficits [16]. In the mouse retina, Atoh7 protein is not detected in fully differentiated RGCs [9,10], suggesting that its main biological activity is transiently required in uncommitted early progenitors. Ectopic expression of Atoh7 in different late stage retinal progenitors either redirects progenitors towards an RGC fate [17] or fails to specify the RGC fate [18]. Therefore, Atoh7 is thought to confer a competent state of progenitors to adopt early cell fates [19]. In the absence of Atoh7, co-expression of two downstream transcription factors Islet1 and Pou4f2 is sufficient to rescue the RGC production deficit and ensure full execution of the RGC differentiation program in the mouse retina [20-24]. We have shown previously that viral mediated expression of human ATOH7 in the developing chicken retina results in precocious neurogenesis and a significant increase in RGC production [25], supporting a hypothesis that a critically high threshold of ATOH7 expression triggers uncommitted early progenitors to exit the cell cycle and predominantly adopt the RGC fate [26].

In the human retina, RGC development occurs during the first trimester and remains a minor cell population in the mature retina [27-29]. This scarcity of human RGCs has hindered research on human RGC development as well as blinding diseases caused by RGC loss. The advancements of pluripotent stem cell technologies in the preceding decade have led to robust stem cell based retinal organoid culture systems [30-33], which provide an excellent opportunity to produce and study human RGCs in vitro. A series of recent studies have applied powerful single cell transcriptomic analyses to characterize the molecular features of cell types in the mature retina and to analyze developmental processes of mammalian, including human retinas [34-39]. In this study, we have used human embryonic stem cell (ESC)-derived 3D retinal organoids to test whether elevating ATOH7 expression levels enhances human RGC production. In addition, we have performed single cell transcriptomic analysis to investigate the effects of modulating neurogenic factor ATOH7 and Neurog2 expression levels on early human retinogenesis. Results of this study further validate that expression levels of these neurogenic factors critically impact human retinal cell production. Moreover, our findings provide new insights into the

function of ATOH7 and Neurog2 in uncommitted progenitors to modulate an interactive gene network that plays important roles during cell cycle withdrawal and cell fate determination.

Results

Viral mediated bHLH factor expression affects progenitor proliferation in 3D human retinal organoids.

To investigate the roles of bHLH neurogenic factors during development of the human retina, we established H9 embryonic stem cell (ESC)-derived 3D retinal organoid cultures [30,40]. These human retinal organoids showed typical morphology of the retinal neural epithelium and co-expressed PAX6 and VSX2 transcription factors (Supplementary Fig. 1), a characteristic feature specific to retinal progenitors. To regulate gene expression during retinogenesis, we constructed lentiviral vectors that encode the doxycycline (Dox) inducible TetO promoter upstream of the human ATOH7 cDNA fused with the Flag epitope tag (LV-ATOH7f) or co-expressing EGFP and puromycin resistant genes (LV-AEP) (Fig. 1a). We also produced a previously described lentiviral TetO vector co-expressing the mouse Neurog2, EGFP, and puromycin resistant genes (LV-NEP) [41](Fig. 1a). Immunohistochemistry confirmed Dox-induced bHLH protein expression following in vitro co-infections of H9 ESCs with LV-rtTA and either LV-ATOH7f, LV-AEP, or LV-NEP (Fig. 1b). We next carried out co-infection of retinal organoids with LV-rtTA and either LV-AEP or LV-NEP at the onset of retinogenesis followed by Dox inductions (Fig. 1c). Live imaging of the EGFP reporter showed that at 24 hours post Dox induction, significant numbers of LV-AEP or LV-NEP infected cells had already migrated towards the inner retina compared to cells from the control LV-GFP infected organoids. This trend continued and became more obvious as the Dox induction times lengthened (Supplementary Fig. 2). By 6-days after Dox induction, the majority of LV-AEP and LV-NEP infected cells were located in the inner layer of the organoids (Fig. 1d, 1e), indicating that viral mediated ATOH7 and Neurog2 expression impacted retinal organoid development.

To investigate whether virally expressed bHLH factors affected retinal progenitor proliferation, we performed BrdU pulse-labeling and examined progenitor marker expression by immunohistochemistry. In the control LV-GFP infected retinal organoids, BrdU-labeled and the proliferating cell nuclear antigen (PCNA)-labeled progenitors occupied the ventricular zone, whereas phospho-histone 3 (PH3) labeled M phase cells were detected at the ventricular surface of the organoids (Fig. 2a, 2b, 2c). Furthermore, PCNA showed co-labeling with PAX6-expressing progenitors in the ventricular zone, but was absent from the $PAX6^+$ cells located in the inner layer, where postmitotic $POU4F^+$ RGCs resided (Fig. 2b, 2c). In contrast to LV-GFP infected retinal organoids, in which GFP^+ cells were distributed throughout the neural epithelium, the majority of GFP^+ cells in LV-AEP and LV-NEP infected organoids were located in the inner retinal layer and devoid of co-labeling with BrdU or PCNA (Fig. 2d, 2e). Quantification of dissociated retinal organoids confirmed that percentages of GFP^+PCNA^+ double labeled cells were reduced significantly from 52.7 ± 8.0 % for LV-GFP infection to 9.0 ± 3.8 % and 2.0 ± 1.9 % for LV-AEP and LV-NEP infections, respectively (Fig. 2f). These results demonstrate that viral driven ATOH7 or Neurog2 expression promoted cell cycle exit.

Elevated neurogenic factor expression promotes RGC production in 3D retinal organoids.

The retinal projection neurons (RGCs) are among the earliest neuronal cells produced during retinogenesis [2,28]. In our human retinal organoid cultures, RGC genesis was detected as early as Day 25 and continued through Day 60 (Fig. 2c). By Day 40, retinal organoid-derived neurons exhibited voltage-gated Na^+, K^+ and Ca^{2+} channels as well as spontaneous and provoked electrophysiological excitability in vitro during whole cell patch clamp recording, characteristic of native RGCs (Supplementary Fig. 3). TTX-sensitive Na^+ channels, TEA-sensitive K^+ channels, and Cd^{2+}-sensitive Ca^{2+} channel currents were recorded. Average Na^+ current amplitude was 219 ± 76 pA (n=12) but in cells with large Na^+ currents (~1

nA), multiple action potentials were observed, while cells expressing smaller Na^+ currents (200-400 pA) typically produced a single spike. Average K^+ current amplitude at +40 mV was 438 ± 49 pA (n=40).

To determine whether viral mediated ATOH7 or Neurog2 expression promoted RGC production, we performed immunohistochemical analyses of retinal organoid sections using known RGC markers (Fig. 3). In LV-ATOH7f infected retinal organoids, signals of the viral reporter Flag closely correlated with POU4F-expressing RGCs in the inner retina as detected by a pan-POU4F/BRN3 antibody (Fig. 3a). Compared to control LV-GFP infected organoids, both LV-AEP and LV-NEP infected cells showed increased co-labeling with the RGC markers NF145, NeuN, and DCX in the inner retina (Fig. 3b). Similarly, in attached organoid cultures that displayed extensive neurite outgrowth, the ATOH7f-expressing cells showed extensive co-labeling for the RGC markers POU4F, NF145, and RBPMS (Fig. 4a).

We also performed flow cytometry analyses of dissociated retinal organoid cells to quantify the effects of viral mediated ATOH7 or Neurog2 expression on RGC genesis (Fig. 4b, 4c, 4d). After 4 days of Dox induction at Day 39, ATOH7 expression led to a 2-fold increase of $POU4F^+$ cells, from 6.6 ± 0.68 % of total cells in LV-GFP infected organoids to 13.7 ± 0.59 % of total cells in LV-ATOH7f infected organoids (Fig. 4c). Among LV-ATOH7f infected cells, 57.1 ± 1.9 % were $POU4F^+$ compared to 6.1 ± 1.1 % among LV-GFP infected cells (Fig. 4d). However, LV-ATOH7f did not increase expression of the RGC marker ISLET1 at this stage (Fig. 4c). In the parallel analysis, LV-NEP infection did not significantly promote $POU4F^+$ cells, but instead increased $ISLET^+$ cells from 5.9 ± 0.95 % to 11.1 ± 0.89 % of total cells (Fig. 4c), and from 4.4 ± 1.1 % to 18.4 ± 1.6 % among viral infected cells (Fig. 4d). Similar analyses at Day 47 following a 7-day Dox induction showed significantly increased co-labeling with the RGC markers ISLET1 (> 10-fold of 6.3%), DCX (>4-fold of 21%), and NeuN (>20-fold of 4%) of LV-AEP and LV-NEP infected cells compared with the control LV-GFP infected cells (Fig. 4e). These results demonstrate that elevated ATOH7 or Neurog2 expression promoted human RGC production in retinal organoids.

Single cell RNA-sequencing analysis reveals accelerated neurogenesis.

To elucidate the influences asserted by viral driven ATOH7 and Neurog2 expression on transcriptome, we performed single cell RNA-sequencing (sc RNA-seq) analysis. We first used fluorescent activated cell sorting to enrich for LV-GFP, LV-AEP, and LV-NEP infected retinal organoid cells between Day 45 to Day 48 (Supplementary Fig.4), followed by 10X Genomics automated single-cell capture, mRNA barcoding, and cDNA library preparation. The high throughput DNA sequencing resulted in 192,932, 153,154, and 132,651 mean readings per cell for LV-GFP, LV-AEP, and LV-NEP samples, respectively. After aligning to the reference human genome and eliminating poor quality cells, the final single cell datasets had a mean gene range between 2935-3079 per cell, and consisted of 3004 cells for LV-GFP, 2063 cells for LV-AEP, and 3909 cells for LV-NEP infected samples.

The sc RNA-seq datasets were subjected to Seurat cell clustering analysis [42], which resolved into 12-15 clusters, and visualized as 2-dimensional UMAPs (Fig. 5a). We applied dot plot analysis using known genes to assign various cell clusters into seven categories or states; each was represented with one or two highly or uniquely expressed genes (Fig.5b; Supplementary Fig. 5). The neural stem cell, neurogenic progenitor, and cell cycle-exiting progenitor categories all expressed CCDN1, encoding cyclin D1, indicating that they all belonged to proliferative cells. However, the neural stem cell clusters showed distinctively high levels of PLOD2 transcripts (procollagen-lysine,2-oxoglutarate 5-dioxygenase 2), whereas the exiting progenitors exhibited upregulation of GADD45A (growth arrest and DNA damage inducible alpha) and downregulation of the retinal progenitor marker VSX2. The exiting progenitors and postmitotic neuroblasts were the main populations expressing ATOH7 mRNA (see Fig. 9). Markers identifying the early born retinal neuronal types included ISL1 and POU4F2 for RGCs, CRX and RCVRN for photoreceptors, and PRDM13 and TFAP2A for horizontal and amacrine cells (HC/AC) (Fig. 5b; Supplementary Fig. 5). These sc RNA-seq analyses thus not only allowed identification of the early

postmitotic neuronal types in the human retinal organoids at the time of analysis, but also revealed sub-classes of neural progenitors with distinct molecular signatures, and transitional states including cell cycle exiting progenitors and postmitotic neuroblasts.

Next, we combined different cell clusters assigned to each of the seven cell categories (Fig. 6a) and performed a pseudotime trajectory analysis (Fig. 6b). Without defining the starting and ending points, the trajectory of LV-GFP infected cells revealed that the neural stem cells were closely related to the neurogenic progenitors, which in turn produced exiting progenitors that developed into the postmitotic neuroblasts and a single trajectory including three types of retinal neurons. In both LV-AEP and LV-NEP infected samples, the pseudotime trajectory displayed a more clearly defined progression from the neurogenic progenitors toward the exiting progenitors, which in turn gave rise to neuroblasts. The postmitotic neuroblasts showed a single node for bifurcated trajectories separating the RGC and HC/AC branch from the photoreceptor branch (Fig. 6b). Quantification of the different cell categories/states further demonstrated that LV-AEP and LV-NEP infection significantly reduced the percentage cells in the neural stem cell category from 25.2 % to less than 4% (Fig. 6c). Neurog2 expression also caused a significant reduction of neurogenic progenitors from 45 % to 26.9 %. Concomitantly, ATOH7f and Neurog2 induction increased the percentage of exiting progenitors from 2.2 % to 4.5 % as well as the percentage of neuroblasts from 5.0 % to 14.9 % and 20.2 %, respectively. Moreover, consistent with marker analysis, elevated ATOH7f and Neurog2 expression resulted in increased proportion of RGCs from 7.0 % to 22.1 % and 27.7 % of total cells, respectively (Fig. 6c). Interestingly, Neurog2, but not ATOH7f, enhanced the photoreceptor population from 3.4 % to 9.8 %; whereas ATOH7f, but not Neurog2, promoted HC/AC proportions from 3.1 % to 4.6 %; (Fig. 6c). These sc RNA-seq data demonstrated that elevating ATOH7f and Neurog2 expression promoted transitions from the pre-neurogenic to the neurogenic state and enhanced neurogenesis of early retinal cell types in retinal organoids.

Neurogenic factors promote transitions of distinct developmental states.

Since viral mediated expression of ATOH7 and Neurog2 affected transitions between developmental states, we explored the key characteristics of each cell state. We compiled differentially expressed genes (DEGs, adjusted p values <0.05) of each cell category for LV-GFP, LV-AEP and LV-NEP samples (Supplement Tables 1-3), and constructed heatmaps for the top 10 DEGs displaying more than Log 1.5-fold change of expression levels (Fig. 7a) (Supplement Tables 4-6). In the control LV-GFP infected sample, the top 10 DEGs in the neural stem cell category included SLC2A1, encoding glucose transporter protein type 1 (GLUT1), and GPI, glucose-6 phosphate isomerase, and were quite distinct from those of the neurogenic progenitors. In LV-AEP and LV-NEP infected retinal organoids, the neural stem cells were not only reduced, but also shared genes, such as SFRP2, IFITM3, and VIM with the neurogenic progenitors. In all three virus infected samples, the exiting progenitors shared top DEGs, including HES6, a suppressor of HES1 [43], and HMGB2, a member of the chromosomal high mobility group of proteins [44], indicating common processes involved in cell cycle exit. The heatmaps also revealed that neuroblasts in each virus infected sample expressed some genes associated with the RGC, photoreceptor, and HC/AC cell lineages, suggesting that the newly postmitotic neuroblasts were poised at an intermediate developmental state and in the process of committing to specific cell fates.

To further decipher the changes of gene expression and the key biological processes that occur during normal developmental transitions, we constructed volcano plots (Fig. 7b) and performed gene ontology (GO) analyses (Fig. 7c) using significant DEGs from the LV-GFP infected cell categories. The neural stem cell state was associated with higher levels of transcripts for several glycolytic pathway enzymes such as SLC2A1, GPI, ENO1, and PGK1, as well as low levels of mitochondrial respiratory chain components and proteins required for rapid cell proliferation such as PCNA and TOP2A (Fig. 7b). This

was consistent with the dominant GO pathways for the neural stem cell state (Fig. 7c). In contrast, cells in the neurogenic progenitor state expressed high levels of SFRP2, FGF19, and SOX2 (Fig. 7b), correlating with the biological responses to growth stimulation and the mitotic cell cycle (Fig. 7c). The dominant GO terms associated with the exiting progenitors included cell cycle arrest and Notch signaling (Fig. 7c). The exiting progenitor to neuroblast transition was marked by elevated expression of the Notch signaling inhibitor HES6 and the antiproliferation factors BTG1 [45], BTG2, and GADD45A [46] (Fig. 7b), consistent with the GO pathway analysis (Fig. 7c). The predominant GO pathways for the three neuronal types reflected their phenotypic differentiation with high axonal growth and light transduction processes associated with RGCs and photoreceptors, respectively (Fig. 7c).

ATOH7 and Neurog2 assert differential effects on cell cycle and neuronal differentiation.

To examine the potential impact of neurogenic factor expression on the cell cycle, we combined the LV-GFP, LV-AEP, and LV-NEP sc-RNA-seq datasets and carried out Seurat cell clustering analysis (Fig. 8a). Based on feature plots of known genes (Supplementary Fig. 6), cell clusters in the combined UMAP were identified as distinct cell categories. Furthermore, the progenitor cell clusters were assigned to different phases of the cell cycle (Fig. 8a) based on known functions of cell cycle genes, including CDK2 associated with G1/S phase [47], MCM4 as a DNA replication licensing factor [48], CCNB2 encoding M phase cyclin B2 [49], and PLK1 involved in spindle assembly and cytokinesis [50] (Supplementary Fig. 6). Slingshot analysis performed for the combined dataset illustrated the developmental trajectory of cell clusters from the neural stem cell state through the different phases of the cell cycle toward cell cycle exit (Fig.8a). Although cell clusters 5 and 6 both expressed M phase genes, they clearly had distinct transcript profiles since cluster 6 was poised to exit the cell cycle (Fig. 8a). As expected, individual UMAPs of ATOH7 and Neurog2 infected samples displayed significantly reduced neural stem cell populations (Fig. 8b; Fig. 6c). In addition, quantitative analyses showed that ATOH7 and Neurog2 infection altered cell cycle distributions among progenitors compared to controls (Fig. 8c). For example, LV-AEP increased G1 distribution by 5 % while decreasing S phase cells by 3.7 %, whereas LV-NEP caused a 7 % S phase cell reduction and a 4.5% increase of M phase cells, which included the exiting progenitors (Fig. 8c).

The Slingshot analysis demonstrated that the postmitotic neuroblast pool, cluster 7, served as the root source giving rise to three neuronal cell lineages (Fig. 8a). Interestingly, UMAPs of individual virus infected retinal organoids revealed differential effects of ATOH7 and Neurog2 on neuronal fate specification. ATOH7 elevation resulted in clear enhancement of the RGC cluster 8, without affecting the photoreceptor cluster 10 (Fig. 8b, also see Fig. 6c). In contrast, Neurog2 overexpression not only significantly enhanced the production of photoreceptor cluster 10 and RGC cluster 8, but also generated a distinct RGC cluster 9, which was largely absent in LV-GFP and LV-AEP infected retinal organoids (Fig. 8b, also see Fig. 6c). Comparison analyses revealed that the two RGC clusters consisted of cells with differential gene expression levels. For example, GAP43 and NSG1 were expressed by cells in both clusters, however, they were frequently detected at higher levels in cluster 8 (Fig. 8d, 8e). Although both clusters expressed ISL1 and POU4F2 (Supplementary Fig. 6), we often observed lower levels of POU4F2 and SNCG in many cluster 9 cells (Fig. 8d). These results suggest that ATOH7 and Neurog2 might differentially influence neuroblast fate specification and/or neuronal differentiation.

Neurogenic factors modulate early retinal gene network.

Next, we examined the effects of viral mediated ATOH7f and Neurog2 elevation on endogenous bHLH gene expression. Since the lentiviral vectors did not contain poly-A sequences associated with ATOH7 and Neurog2 cDNAs, and the single cell cDNA libraries were constructed using oligo dT priming, we were able to analyze expression of the endogenous genes using the sc RNA-seq datasets, while

excluding genes encoded by the viruses. We first examined expression of ASCL1, an early onset bHLH neurogenic factor. Consistent with previous studies [34,36], feature plots showed that ASCL1 was predominantly expressed in neural stem cells and progenitors (Fig. 9a). Viral mediated ATOH7f and Neurog2 expression resulted in significant reduction of $ASCL1^+$ cell from 29.7 % to 17.1 % and 13.2 %, respectively (Fig. 9c); but neither affected the median levels of ASCL1 expression (Fig. 9b). In retinal organoids, ATOH7 and NEUROG2 were detected in exiting progenitors, neuroblasts, and some postmitotic neurons (Fig. 9a). Both virally expressed ATOH7f and Neurog2 increased the number of endogenous ATOH7-expressing cells from 9.1 % to 22.0 % among total cells (Fig. 9c), and LV-NEP infection also increased the median ATOH7 expression level (Fig. 9b). Interestingly, LV-NEP but not LV-AEP infection caused a 2-fold increase of endogenous NEUROG2-expressing cells as well as elevated the median expression level (Fig. 9b, 9c), suggesting that Neurog2 positively regulates endogenous NEUROG 2. In addition, elevated levels of ATOH7f and Neurog2 both correlated with the increased percentages of cells expressing OLIG2, another bHLH gene in retinal organoids [51] (Fig. 9c).

In addition to ASCL1, ATOH7, NEUROG2, and OLIG2, sc RNA-seq analysis identified additional Notch signaling pathway genes and bHLH genes affected by ATOH7f and Neurog2 expression (Supplementary Table 1-3; Supplementary Fig. 7). Violin plots and quantification showed that the Notch ligands DLL1 and DLL4 were upregulated, especially by Neurog2 (Fig. 9d, 9e). In addition, the percentage of cells expressing the Notch signaling effector HES1 was significantly reduced compared to the control LV-GFP infected cells (Fig. 9e), while changes in median expression levels were mild (Fig. 9d). In contrast, not only the percentages of HES6 and HEY1 expressing cells were increased (Fig. 9e), the median HES6 expression levels were significantly elevated by bot LV-AEP and LV-NEP (Fig. 9d). Since HES1 was predominantly expressed by progenitors, whereas HES6 and HEY1 were upregulated in exiting progenitors and neuroblasts (Supplementary Fig. 7), these changes reflected the trend of enhanced neurogenesis. Data from sc RNA-seq also showed that viral expression of ATOH7f or Neurog2 caused substantial increases of several bHLH genes involved in neuronal fate specification or differentiation, including NEUROD1, NEUROD4, NHLH1 and NHLH2 (Fig. 9d, 9e; Supplementary Fig. 7). In all cases, elevated Neurog2 strongly impacted both the number of cells expressing these genes as well as their expression levels.

Together, these results suggested that viral mediated ATOH7f and Neurog2 expression in the developing retinal organoids affected an interactive gene network. We therefore performed STRING network analysis by inputting bHLH genes and a few selected genes known to be involved in RGC development. The resulting network model included previously reported molecular interactions as well as novel regulatory relationships revealed in this study (Fig. 9f). We also summarize the observed temporal progression of gene expression as retinal organoid cells advanced through the developmental states as defined by sc RNA-seq analysis (Fig. 9g).

Discussion

In this study we have used human ESC-derived organoid cultures as a model system to investigate embryonic retinal development. The 3D retinal organoids retain the unique molecular signature and correct tissue polarity of the embryonic retinal epithelium in vivo, and generate the expected early retinal cell lineages and functional neurons. By transducing human retinal organoids with inducible viral vectors, we were able to determine whether elevating ATOH7 or Neurog2 affected neuronal production, and examine the biological processes and the gene network involving these neurogenic factors.

We have identified three classes of retinal progenitors by performing single cell transcriptome analysis. The neural stem cells express PAX6, VSX2, LHX2, and SOX2, and thus clearly possess the

identity of the retinal primordium. Their main GO pathways show cellular metabolic characteristics common to stem cells [52], with prominently featured DEGs of the glycolytic pathway, including SLC2A1, GPI, PGK1, and PLOD2, and low mitochondrial respiratory chain genes. This state likely represents a naïve pre-neurogenic status found in the retinal primordium before the onset of neurogenesis or later at the ciliary margin of the retina. These neural stem cells express cyclin D1, but have relatively low levels of PCNA and TOP2A, suggesting that they might be slow cycling cells. The neural stem cell category was significantly reduced in LV-AEP and LV-NEP transduced retinal organoids, indicating that expression of these neurogenic factors promotes the transition from the pre-neurogenic to the neurogenic state. This finding is consistent with our previous observation in the chicken retina, where ectopic ATOH7 expression in the pre-neurogenic peripheral retina induces precocious neurogenesis ahead of the neurogenic wave front [25]. These results indicate that the onset of neurogenic factor expression among naïve retinal progenitors can trigger or accelerate the transition into a neurogenic state in the context of the ocular primordium. Noticeably, the transition from the stem cell state to the neurogenic state is also accompanied by upregulation of the chromosomal high mobility group genes, such as HMGB2 and HMGN2, reflecting the underlying epigenetic changes accompanying this transition.

Our sc RNA-seq analysis also identified a distinct progenitor state among the proliferating neurogenic progenitor population, the cell cycle exiting progenitors. This group of cells still express the characteristic M phase genes such as CDC20 that interacts with the mitosis complex, and MZT1 and PLK1 which organize the mitotic spindles. However, the exiting progenitors show significant upregulation of the Notch ligands DLL1 and DLL4, as well as the bHLH genes ATOH7 and NEUROD1. Both viral mediated ATOH7 and Neurog2 elevation caused expansion of the exiting progenitor state. The developmental trajectory analysis points to a progression from the exiting progenitors to the postmitotic neuroblasts, which exhibit a transcriptome profile partially overlapping with all three early neuronal lineages, including VSX1, NHLH1, ATOH7, ONECUT2, NEUROD1, and the growth arrest gene GADD45A. Thus, the neuroblasts appear to serve as a transitional cell pool poised for terminal cell fate choices and differentiation. Interestingly, we have observed that relative to ATOH7, elevating Neurog2 asserted more potent effects in promoting cell cycle withdrawal, neuroblast expansion, and neuronal differentiation in retinal organoids.

Among the early bHLH factors detected by sc RNA-seq in retinal organoids, a higher percentage of cells expressed ATOH7 (9%) than NEUROG2 (3.2%) endogenously. Although elevating either ATOH7 or Neurog2 caused enhanced neurogenesis, we also detected differential effects of these two factors. For example, at Day 39, ATOH7 promoted POU4F$^+$ cells, whereas Neurog2 increased ISL1$^+$ cells; but both factors enhanced ISL1$^+$ cells by Day 47. Furthermore, the RGC cluster enhanced by ATOH7 elevation shared transcriptomic signatures with RGCs from the control organoids. In contrast, Neurog2 induction promoted two distinct groups of RGCs, but only one of the two matched the expression profile of RGCs in the control organoids. Further analyses showed that although both RGC groups expressed ISLET1 and POU4F2, they exhibited different levels of known RGC markers such as SNCG, POU4F2, and GAP43. The promotion of the two RGC subtypes may reflect the more potent neurogenic effect of Neurog2, which might have accelerated the RGC differentiation program to acquire additional subtype features. In addition, the effect of elevated Neurog2 to induce ISL1$^+$ neurons preferentially earlier on may also influence outcomes of neuronal differentiation. Alternatively, the viral induced exogenous Neurog2 could have promoted the production of a hybrid neuronal cell type, which does not naturally occur during retinogenesis. Future research using long term cultures may provide data to address these possibilities. In addition to the RGC phenotypes, ATOH7 increased the HA/AC lineage without impacting cone photoreceptor genesis, consistent with the developmental trajectory. In contrast, Neurog2 significantly enhanced photoreceptor production without affecting the HC/AC lineage, likely reflecting the effect of Neurog2 induced NEUROD1 upregulation.

Expression of cell intrinsic factors can be profoundly influenced by extrinsic signaling events during neural development [53]. Notch mediated cell-to-cell signaling plays important roles during retinogenesis, and disruption of Notch signaling can influence RGC and cone photoreceptor development [6,54-57]. In addition, differentiated RGCs have been shown to produce secreted signals that modulate progenitor behaviors. Among the signaling molecules released by differentiated RGCs are Shh, GDF11, and VEGF, all of which promote retinal progenitor proliferation while simultaneously suppressing RGC production [58-61]. Accumulating evidence indicate that the Notch signaling downstream effector Hes1 may serve as an important integration node for distinct signaling pathways that converge upon retinal progenitors to influence cell proliferation and control neurogenesis. Hes1 is known to directly control cell proliferation by repressing the CDK inhibitor p27(Kip1) [43]. The negative feedback of extrinsic signals on RGC genesis is in part mediated by Hes1 suppression of ATOH7 [62-64]. At the present time, the temporal expression sequence and the complex regulatory relationships among the various early bHLH factors in the retina are not fully understood but their elucidation is likely crucial for a better understanding of cell fate selection and the progressive changes in progenitor competent states. The single cell RNA-seq technology has provided a tool to decipher gene regulatory networks at the transcription level on a global scale.

As demonstrated by sc RNA-seq data, HES1 is predominantly expressed by neural stem cells and neurogenic progenitors, and is down regulated among exiting progenitors. The sc RNA-seq data also show that ASCL1 is present at low levels in retinal organoid neural stem cells, and is mainly expressed by neurogenic progenitors, suggesting that it could play a role in establishing competence for the neurogenic state. This is consistent with the potent regeneration effects of ASCL1 in mature retinal Muller glial cells which promote reentry into the cell cycle and the generation of new neurons [65-67]. Intriguingly, viral mediated ATOH7 and Neurog2 expression decrease both HES1 and ASCL1 expression. Concomitantly, ATOH7 and Neurog2 significantly upregulate the expression of the HES1 inhibitor HES6 [43] among exiting progenitors. Interestingly, Atoh7 has been shown to promote another Notch effector Hes5.3 [68,69]. These data suggest that ATOH7 and NEUROG2 comprise a select group of neurogenic factors that can dampen the positive effect of HES1 on cell proliferation and relieve the inhibitory effect of HES1 on neurogenesis. Moreover, viral mediated ATOH7 and Neurog2 expression significantly increase NEUROD1, which initiates its expression in exiting progenitors, as well as several other factors including NEUROD4, NHLH1 and NHLH4, which are expressed predominantly by postmitotic neurons. We summarize our analysis of these newly uncovered regulatory relationships together with previously validated data in a gene network model focusing on RGC development (Fig. 9f, 9g). Future investigations are necessary to determine whether the regulatory effects asserted by ATOH7 and Neurog2 are direct or indirect, and if the gene network is indeed interactive as proposed.

Expression of the HES1 gene in developing cortical progenitors is oscillatory due to autorepression by its own product, and as a consequence, the levels of its target Neurog2 also oscillate in cortical progenitors [70-73]. However, it has not yet been demonstrated if autorepression induced HES1 oscillations occur in the developing retina, and the modes of regulation between HES1 and bHLH neurogenic genes in the retina remain largely unknown. Results from this study support the hypothesis that high levels of AOTH7 or NEUROG2 trigger a withdrawal from the cell cycle which leads to the birth of a neuroblast. Critical questions remain regarding how the level of ATOH7 is regulated in neurogenic progenitors to result in a stochastic expression pattern. It has been shown that the ATOH7 gene contains enhancers that mediate direct positive regulation by Pax6 [74]. In Hes1 mutants, Atoh7 is precociously expressed along with the formation of RGC and HC/AC [75], supporting that Hes1 negatively regulates Atoh7 expression. Our sc RNA-seq analysis reveals that viral mediated ATOH7 expression elevates endogenous ATOH7 without affecting endogenous NEUROG2 expression, whereas Neurog2 induction leads to increases in both ATOH7 and NEUROG2 expression. These results suggest positive

autoregulation of ATOH7 and NEUROG2, as well as cross-regulation of ATOH7 by NEUROG2. This conclusion is consistent with previous reports that suggest cross regulation between Atoh7 and Neurog2 [16,76]. Since only a fraction of ATOH7 protein expressing cells appear to co-express Neurog2 protein during a given window of time [9], progenitors that have coincidental expression of both factors are more likely to reach or exceed a threshold level of Atoh7, which may effectively counteract the negative impact of Notch and secreted signals, and lead to a terminal cell division and neurogenic event. We therefore propose that integrated cell extrinsic signals and interacting cell-intrinsic factors could converge, resulting in stochastic expression of the ATOH7 gene, and thus enabling a limited subset of progenitors that have reached this threshold to exit the cell cycle and initiate downstream RGC and/or cone photoreceptor differentiation programs. Similar explanations can be extended to the observed stochastic expression patterns of multiple neurogenic bHLH factors among neurogenic progenitors [77,78]. The sc RNA-seq approach enables us to survey multiple genes simultaneously, thus allowing us to construct a gene regulatory network model (Fig. 9f) that integrates previous findings and can be further interrogated. This model focuses on bHLH factors expressed during early retinogenesis, but does not exclude other genes involved in the developmental process. In fact, RGC and other neuronal fate determinations are known to be regulated by multiple transcription factors [79-85], which together coordinate epigenetic changes and orchestrate transcriptomic responses required for neuronal differentiation.

In summary, we have shown that elevating neurogenic factor expression in ESC-derived retinal organoids significantly enhances retinogenesis, which can serve as a useful approach to produce authentic human RGCs for studying development and degenerative diseases. Our single cell transcriptome analysis provides novel insights into the interactive network of neurogenic factors and their functions during human retinogenesis.

Methods

Lentiviral construction and production

The inducible lentiviral vector plasmids LV-GFP and LV-NEP (YS-TetO-FUW-Ng2-P2A-EGFP-T2A-Puro) were generous gifts from Dr. Thomas Sudhof [41]. The human ATOH7 cDNA was obtained as described previously [25]. The LV-ATOH7f vector was constructed by replacing the EGFP gene in the LV-GFP vector with ATOH7 cDNA fused to the Flag epitope tag at the c-terminus. The LV-AEP vector was constructed by custom synthesizing the continuous open reading frame of ATOH7f-P2A-EGFP-T2A-Puro by GenScript, and replacing the EGFP gene in the LV-GFP vector. All lentiviral vector plasmids were verified by DNA sequencing. The LV-rtTA lentiviral vector (FUW-M2rtTA) with the Ubi promoter was obtained from AddGene (Plasmid #20342).

Lentiviral stocks were produced by co-transfection of HEK 293T cells with a given viral vector DNA and the third-generation lentiviral helper plasmids with VSVG pseudotyping as described [86-88]. The 293T cell medium DMEM containing 10% fetal bovine serum (SigmaAldrich, 12103C) was changed to serum free CD293 (ThermoFisher, 11913) one day post transfection, and viral supernatants were harvested every 24 hours. Combined viral stocks were concentrated by ultracentrifugation as previously described [88].

Human retinal organoid derivation

Human H9 ES cells were cultured and passaged on Matrigel (Corning, 356231) coated dishes in mTeSR1 medium (Stemcell Technologies, 05850). Retinal organoids were generated based on a previously described protocol [30] with modifications. At the start of the culture (Day 0), H9 ES cells (at 80-90% confluency) were enzymatic detached using dispase (1 mg/ml, Stemcell Technologies, 07923).

Detached cells were transferred into medium at 3:1 ratio of mTeSR1 to Neural Induction Medium (NIM) that consists of DMEM/F12 with 1x N2 supplement (ThermoFisher, 17502048), 1x non-essential amino acid (NEAA; ThermoFisher, 11140050), 2 µg/ml Heparin (ThermoFisher, H7482), and 1x Antibiotic-Antimycotic (Anti-Anti; ThermoFisher, 15240112) in low-attachment plates (Corning, 3471) to allow the formation of embryonic bodies (EBs). During the next three days, the medium was replaced daily with the ratio of NIM to mTeSR1 increased to 100%. At Day 6, human BMP4 (R&D Systems, 314-BP-010) was added to the NIM medium to a final concentration of 55 ng/ml [40]. At Day 7, EBs were collected and seeded in NIM as adherent cultures in 6-well dishes (Corning, 3516) till Day 16. At Day 16, the visible neural rosettes formed from attached EBs were manually lifted, collected, and further cultured as suspensions in Retinal Differentiation Medium (RDM) consisting of DMEM to F12 at 3:1, 1x B27 supplement (ThermoFisher,1754044), 1x NEAA, and 1x Anti-Anti. During the 7 days after lifting neural rosettes, 5 µM SU-5402 (SigmaMillipore, SML0443) and 3 µM GSK inhibitor CHIR99021 (Stemgent, 04-0004) were added to RDM. At Day 20, the translucent optic vesicle-like structures were manually separated from the rest of the suspension culture, collected, and cultured as retinal organoids in RDM. From Day 24, 10 % FBS (SigmaAldrich, 12103C), 100 µM Taurine (SigmaAldrich, 0625), and 500 µM retinoic acid (SigmaAldrich, R2625) were added, and the medium was changed twice a week.

Lentiviral Infection and transgene induction

Retinal organoids were infected by different lentiviruses in conjunction with LV-rtTA three times between Day 23 and Day 40. TetO promoter induction was carried out by adding doxycycline (Dox) to a final concentration of 2 µg/ml (SigmaAldrich, D3072) according to experimental designs as indicated in the results.

Attached and dissociated retinal organoid cultures

Retinal organoids between Day 30-35 were cut into small pieces (0.1- 0.5 mm) and plated on Matrigel coated culture dishes, or dissociated with Trypsin (SigmaAldrich, T9935) to single cells and plated on poly-D-lysine and laminin coated glass coverslips (Corning, 354087). After attachment, retinal organoid cells were cultured in RDM or BrainPhys neuronal medium (Stemcell Technologies, 05790) with SM1 (Stemcell Technologies, 05711) and N2 supplements (ThermoFisher, 17502048), 20 ng/ml BDNF (PeproTech, 450-02), 20 ng/ml GDNF (Stemcell Technologies, 78058), 1 mM dibutyryl cyclic-AMP (Stemcell Technologies, 73882), and 200 nM ascorbic acid (Stemcell Technologies,72132) till desired time, followed by immunofluorescent labeling or electrophysiological recordings.

Electrophysiological recording

Electrophysiological recordings were performed using dissociated retinal organoid cells cultured as a monolayer on glass coverslips between Day 40-45 using previously described methodologies [89-92]. Whole cell patch clamp was performed at room temperature using an Axopatch 200B amplifier controlled by pClamp 11 data acquisition software (Molecular Devices). The pipette solution contained 20 mM KCl, 120 mM K-gluconate 0.1 mM $CaCl_2$, 1 mM EGTA, 10 mM HEPES, 3 mM Mg-ATP, 0.2 mM Li-GTP, and 8 mM phosphocreatine, at pH7.2. The bathing solution (for recording K^+ currents, Na^+ currents and for current clamp) contained125 mM NaCl, 3mM KCl, 2 mM $CaCl_2$, 1.25 mM NaH_2PO_4, 1 mM $MgCl_2$, 25 mM $NaHCO_3$ and 10 mM glucose bubbled continuously with 95% O_2 – 5% CO_2. A high barium external solution for Ca^{2+} channel current recordings containing 110 mM choline chloride, 5 mM KCl, 1 mM $MgCl_2$ 7 mM $BaCl_2$, 15 mM TEACl, 0.1 mM 4-aminopyridine, 20 mM glucose, 10 mM HEPES and 1 µM tetrodotoxin, adjusted to pH 7.4, was used with a CsCl intracellular solution containing 140 mM CsCl, 1 mM $CaCl_2$, 11 mM EGTA, 2 mM $MgCl_2$, and 10 mM HEPES, at pH7.2.

Immunohistochemistry and imaging

For live imaging of whole mounts retinal organoids, EGFP signals were first captured using a Leica MZ10F fluorescent dissecting microscope, followed by image acquisition using an Olympus Flowview FV1000-IX81 (inverted) scanning laser confocal microscope. For whole mount immunolabeling, retinal organoids were fixed with 4% paraformaldehyde (PFA) in PBS for 30 minutes, followed by primary and secondary antibody incubations overnight at 4°C, with extensive washes in between [25]. Cryosections (14 μm) or attached cultures were processed for immunolabeling as previously described [25].

The following primary antibodies were used: mouse anti-GFP(1:200; Millipore, MAB#3580); goat anti-GFP (1:200; Rockland Inc, 600-101-215); rabbit anti-GFP (1:200; Rockland Inc, 400-401-215); mouse anti-FLAG (Clone M2)(1:500; SigmaAldrich, F3165); rabbit anti-ATOH7 (1:100; NovusBio, NBP1-88639); goat anti-Ngn2 (1:200; Santa Cruz Biotechnology, sc-19233); mouse anti-BrdU (1:1; GE Healthcare, RPN202); mouse anti-PCNA (1:500; SigmaAldrich, P3825); rabbit anti-phospho-histone 3 (ser10) (1:3000; Upstate Biotechnology, 06-570); rabbit anti-Pax6 (1:200; Chemicon, ab5409); goat anti-CHX10/VSX2 (N-18) (1:50; Santa Cruz Biotechnology, sc-21690); goat anti-BRN3a (1:100; Santa Cruz Biotechnology, sc-31984); goat anti-pan BRN3 (1:50; Santa Cruz Biotechnology, sc-6026); mouse anti-Islet1 (1:10; Developmental Study Hybridoma Bank, 39.4D5); mouse anti-doublecortin (E-6) (1:50; Santa Cruz Biotechnology, sc-271390); rabbit anti-NeuN (1:200; Abcam, ab177487); rabbit anti-NF145 (1:750; Millipore, AB1987); rabbit anti-RBPMS (1:200; [93]. Secondary antibodies used were: Alexa Fluor 488 donkey anti-mouse (1:500; ThermoFisher, A32766); Alexa Fluor 488 donkey anti-rabbit (1:500; ThermoFisher, A32790); Alexa Fluor 488 donkey anti-goat (1:500; ThermoFisher, A32814); Alexa Fluor 594 donkey anti-mouse (1:500; ThermoFisher, A32744); Alexa Fluor 594 donkey anti-rabbit (1:500; ThermoFisher, A32754); Alexa Fluor 594 donkey anti-goat (1:500; ThermoFisher, A32758); Alexa Fluor 647 donkey anti-mouse (1:500; ThermoFisher, A32787); Alexa Fluor 647 donkey anti-rabbit (1:500; ThermoFisher, A32795); Alexa Fluor 647 donkey anti-goat (1:500; ThermoFisher, A32849).

Slides were mounted with Fluoro-Gel (Electron Microscopy Sciences, 17985-10) after staining with DAPI (SigmaAldrich, D9542). Confocal images were acquired using an Olympus Flowview FV1000-BX61 (upright) scanning laser microscope with Plan-APO objectives. Images were arranged using Adobe Photoshop.

Cell marker quantification and statistics

Lentivirus infected retinal organoids at Day 39 were pooled (10-15 organoids in each sample n), dissociated into single cell suspensions using tyrpsin (SigmaAldrich, T-9935), and immunolabeled as previously described [60]. Flow cytometry was performed using LSRII Analytic Flow Cytometer for cell marker analyses. Quantification of FACS data was performed using FlowJo software (Tree Star, Inc.). In addition, pooled retinal organoids at Day 47 were dissociated with trypsin and plated as a monolayer for 3 hours followed by immunolabeling for the various cell markers listed above. Monolayer cell quantification was performed by counting marker-positive cells in multiple fields of independent samples (n=3) using captured confocal images. Bar graphs were constructed using Prism (Graphic Pad). Ordinary one-way ANOVA with Tukey's multiple comparison test was used for statistical analysis of cell marker quantifications, with p value < 0.05 considered significant.

Single cell cDNA library preparation and sequencing

Distinct pools of H9 ES cell-derived retinal organoids (12-20 retinal organoids/pool) co-infected by LV-rtTA and LV-EGFP, LV-AEP, or LV-NEP were induced by Dox and dissociated between Day 45 and Day 48 using trypsin and manual trituration. Dissociated cell suspensions were subjected to fluorescence activated cell sorting using FACSAriaII (BD Biosciences). Non-infected retinal organoid cells

were used to set thresholds for selecting EGFP-positive cells. Sorted EGFP-positive cells were collected in HBSS without Ca^{2+} and Mg^{2+} (ThermoFisher, 14170-112) containing 1% FBS and 0.4% BSA. The cells were washed with PBS containing 0.04% BSA, then counted with Countess II Cell Counter (ThermoFisher).

Automated single-cell capture, barcoding, and cDNA library preparation were carried out using 10X Genomics Chromium Controller with Chromium Single Cell 3' Library & Gel Bead Kit v2 reagents, with 12 cycles of cDNA amplification and 12 cycles of library amplification, following the manufacturer's instructions. Qubit dsDNA Assay kit (Life Technologies) and TapeStation 4200 (Agilent) were used to assess the quality and concentration of the libraries. Illumina NovaSeq6000 S2 paired-end 2x50bp mode was used to sequence the libraries.

Single cell RNA-sequencing data processing and quality control

10X Genomics Cell Ranger version 2.1.1 was used to demultiplex the raw base calls into FASTQ files (cellranger mkfastq). Spliced Transcripts Alignment to a Reference (STAR) version 2.5.1b (cellranger count) was used to perform sequence alignments to the reference human genome (GRCh38), barcode counts, and UMI counts to yield summary reports and t-Stochastic Neighboring Embedding (t-SNE) dimensionality reduction. For downstream analyses, cells with a number of unique molecular identifiers (UMI) > 2500 per cell and < 0.1 % mitochondrial gene expression were used. For LV-GFP, LV-AEP, LV-NEP samples, the mean reads per cell ranged from 139,000-195,000, with mean gene per cell ranging from 2935-3079. The resulting total single cell counts used for analysis were 3004 for LV-EGFP, 2063 for LV-AEP, and 3909 for LV-NEP infected samples.

Single cell RNA-sequencing data analysis and visualization

The analysis of sc RNA-seq data was performed using Seurat R package (https://satijalab.org/seurat/v2.2) [42,94]. Clustering of cells was performed by using Seurat FindCluster function (top 20 principal components, resolution 0.8) that implements the shared nearest neighbor modularity optimization algorithm. Nonlinear dimensionality reduction using UMAP (Uniform Manifold Approximation and Projection) was applied for the visualization of cells in two-dimensional space. Feature plots of known genes were used to designate clusters observed in the UMAP space into six major cell categories/states. Cell counts of each category were obtained using custom R code.

Pseudotime developmental progression of cell states was obtained by using the Monocle R package (version 2) to process the datasets with cell labels corresponding to the six cell categories and visualized as UMAPs. Pseudotime cell cycle progression and cell fate adoption analysis was performed using Slingshot R (version1.6.1) by combining the LV-GFP, LV-AEP, and LV-NEP sc RNA-seq datasets and assigning the start point as neural stem cells and end points as differentiated neuronal cell types.

Differentially expressed genes (DEGs) were identified using edgeR with significantly enriched genes in each cell category defined as those with adjusted *p* values < 0.05 for LV-GFP, LV-AEP, and LV-NEP data sets (Supplementary Tables 1-3). The top 10 enriched DEGs in each cell category were defined as those with adjusted *p*-value < 0.05, and log fold change >1.5 (Supplementary Tables 4-6). Heatmaps of the top 10 DEGs were generated using the Seurat package. Volcano plots were generated using EnhancedVolcano R package to show *p* values and fold changes of DEGs between two datasets. Gene Ontology (GO) Enrichment analysis was performed using ShinyGO v0.61 (http://bioinformatics.sdstate.edu/go) [95] and the Homo Sapiens background using p-value (FDR) cutoff at 0.05. The top 25 DEGs from each cell category within the LV-GFP dataset were used as inputs, and the redundancy of the output biological processes was manually reduced to the most predominant GO terms.

Feature plots of individual gene expression patterns in different cell clusters were presented as UMAPs. Violin plots for individual genes in all cell clusters were constructed to show expression levels and cell distributions. Kruskal-Wallis one-way ANOVA rank sum test and Tukey-Kramer-Nemenyi all-pairs test were used for statistical analysis, taken into consideration of both gene expression levels and cell numbers between different samples, with p value < 0.05 considered significant. The statistical tests were performed on R Studio using 'PMCMRPlus' [96] and 'FSA' [97] packages. Data were plotted using R Studio 'ggplot2' [98] and 'ggsignif' [99] packages. Cells with gene expression level < 0.2 were exclude from the violin plots and statistical analyses.

STRING analysis exploring protein-protein association network was performed using human protein database (version 11.0; https://string-db.org/) by inputting relevant genes involved in retinal development. The schematic network model shows known molecular interactions reported previously and new regulatory relationships described in this study.

Data availability

The authors declare that all data supporting the findings of this study described within the article and its supplementary information files are available from the corresponding author upon reasonable request.

Code availability

Source code to reproduce data analysis is available on GitHub [https://github.com/XXXXXXXX/Human_Epi].

Acknowledgments

This work was supported by NIH grant R01EY026319 to XJY, NIH core grant P30EY000331, and an unrestricted grant from the Research to Prevent Blindness to the Department of Ophthalmology at University of California Los Angeles.

Author Contributions

X-J Y. conceived the project and designed lentiviral vectors; X.Z. carried out human retinal organoid and neuronal cultures, performed immunohistochemistry and imaging, and FACS analysis; X.Z. and K.H.N. prepared viral stocks and sc RNA-seq libraries; T.T.T.N, I.M., and M.P. performed bioinformatic analysis; T.T.T.N. performed statistical analysis; J.C.R.G and S.B. performed electrophysiological recordings; X.Z, T.T.T.N., and X-J Y. prepared the figures, and analyzed the data; X-J Y. wrote the manuscript. All authors reviewed and commented on the manuscript.

Competing Interests

The authors declare no competing interests.

Supplementary information

Description of additional supplementary files.

References

1 Livesey, F. J. & Cepko, C. L. Vertebrate neural cell-fate determination: lessons from the retina. *Nature reviews. Neuroscience* **2**, 109-118, doi:10.1038/35053522 (2001).

2 Young, R. W. Cell differentiation in the retina of the mouse. *Anat Rec* **212**, 199-205 (1985).

3 Wong, L. L. & Rapaport, D. H. Defining retinal progenitor cell competence in Xenopus laevis by clonal analysis. *Development* **136**, 1707-1715, doi:10.1242/dev.027607 (2009).

4 Cepko, C. L., Austin, C. P., Yang, X., Alexiades, M. & Ezzeddine, D. Cell fate determination in the vertebrate retina. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 589-595, doi:10.1073/pnas.93.2.589 (1996).

5 Xiang, M. Intrinsic control of mammalian retinogenesis. *Cell Mol Life Sci* **70**, 2519-2532, doi:10.1007/s00018-012-1183-2 (2013).

6 Akagi, T. *et al.* Requirement of multiple basic helix-loop-helix genes for retinal neuronal subtype specification. *The Journal of biological chemistry* **279**, 28492-28498, doi:10.1074/jbc.M400871200 (2004).

7 Hoon, M., Okawa, H., Della Santina, L. & Wong, R. O. Functional architecture of the retina: development and disease. *Progress in retinal and eye research* **42**, 44-84, doi:10.1016/j.preteyeres.2014.06.003 (2014).

8 Brown, N. L. *et al.* Math5 encodes a murine basic helix-loop-helix transcription factor expressed during early stages of retinal neurogenesis. *Development* **125**, 4821-4833 (1998).

9 Miesfeld, J. B., Glaser, T. & Brown, N. L. The dynamics of native Atoh7 protein expression during mouse retinal histogenesis, revealed with a new antibody. *Gene expression patterns : GEP* **27**, 114-121, doi:10.1016/j.gep.2017.11.006 (2018).

10 Fu, X. *et al.* Epitope-tagging Math5 and Pou4f2: new tools to study retinal ganglion cell development in the mouse. *Developmental dynamics : an official publication of the American Association of Anatomists* **238**, 2309-2317, doi:10.1002/dvdy.21974 (2009).

11 Brown, N. L., Patel, S., Brzezinski, J. & Glaser, T. Math5 is required for retinal ganglion cell and optic nerve formation. *Development* **128**, 2497-2508 (2001).

12 Wang, S. W. *et al.* Requirement for math5 in the development of retinal ganglion cells. *Genes Dev* **15**, 24-29 (2001).

13 Kay, J. N., Finger-Baier, K. C., Roeser, T., Staub, W. & Baier, H. Retinal ganglion cell genesis requires lakritz, a Zebrafish atonal Homolog. *Neuron* **30**, 725-736, doi:10.1016/s0896-6273(01)00312-9 (2001).

14 Ghiasvand, N. M. *et al.* Deletion of a remote enhancer near ATOH7 disrupts retinal neurogenesis, causing NCRNA disease. *Nature neuroscience* **14**, 578-586, doi:10.1038/nn.2798 (2011).

15 Miesfeld, J. B. *et al.* The Atoh7 remote enhancer provides transcriptional robustness during retinal ganglion cell development. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 21690-21700, doi:10.1073/pnas.2006888117 (2020).

16 Hufnagel, R. B., Le, T. T., Riesenberg, A. L. & Brown, N. L. Neurog2 controls the leading edge of neurogenesis in the mammalian retina. *Developmental biology* **340**, 490-503, doi:10.1016/j.ydbio.2010.02.002 (2010).

17 Mao, C. A. *et al.* Reprogramming amacrine and photoreceptor progenitors into retinal ganglion cells by replacing Neurod1 with Atoh7. *Development* **140**, 541-551, doi:10.1242/dev.085886 (2013).

18 Prasov, L. & Glaser, T. Pushing the envelope of retinal ganglion cell genesis: context dependent function of Math5 (Atoh7). *Developmental biology* **368**, 214-230, doi:10.1016/j.ydbio.2012.05.005 (2012).

19 Brzezinski, J. A. t., Prasov, L. & Glaser, T. Math5 defines the ganglion cell competence state in a subpopulation of retinal progenitor cells exiting the cell cycle. *Developmental biology* **365**, 395-413, doi:10.1016/j.ydbio.2012.03.006 (2012).

20 Wu, F. *et al.* Two transcription factors, Pou4f2 and Isl1, are sufficient to specify the retinal ganglion cell fate. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1559-1568, doi:10.1073/pnas.1421535112 (2015).

21 Mu, X., Fu, X., Beremand, P. D., Thomas, T. L. & Klein, W. H. Gene regulation logic in retinal ganglion cell development: Isl1 defines a critical branch distinct from but overlapping with Pou4f2. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6942-6947, doi:10.1073/pnas.0802627105 (2008).

22 Liu, W., Mo, Z. & Xiang, M. The Ath5 proneural genes function upstream of Brn3 POU domain transcription factor genes to promote retinal ganglion cell development. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 1649-1654, doi:10.1073/pnas.98.4.1649 (2001).

23 Pan, L., Deng, M., Xie, X. & Gan, L. ISL1 and BRN3B co-regulate the differentiation of murine retinal ganglion cells. *Development* **135**, 1981-1990, doi:10.1242/dev.010751 (2008).

24 Li, R. *et al.* Isl1 and Pou4f2 form a complex to regulate target genes in developing retinal ganglion cells. *PloS one* **9**, e92105, doi:10.1371/journal.pone.0092105 (2014).

25 Zhang, X. M., Hashimoto, T., Tang, R. & Yang, X. J. Elevated expression of human bHLH factor ATOH7 accelerates cell cycle progression of progenitors and enhances production of avian retinal ganglion cells. *Scientific reports* **8**, 6823, doi:10.1038/s41598-018-25188-z (2018).

26 Prasov, L., Nagy, M., Rudolph, D. D. & Glaser, T. Math5 (Atoh7) gene dosage limits retinal ganglion cell genesis. *Neuroreport* **23**, 631-634, doi:10.1097/WNR.0b013e328355f260 (2012).

27 Hendrickson, A. Development of Retinal Layers in Prenatal Human Retina. *American journal of ophthalmology* **161**, 29-35 e21, doi:10.1016/j.ajo.2015.09.023 (2016).

28 Hoshino, A. *et al.* Molecular Anatomy of the Developing Human Retina. *Developmental cell* **43**, 763-779 e764, doi:10.1016/j.devcel.2017.10.029 (2017).

29 Mellough, C. B. *et al.* An integrated transcriptional analysis of the developing human retina. *Development* **146**, doi:10.1242/dev.169474 (2019).

30 Ohlemacher, S. K. *et al.* Stepwise Differentiation of Retinal Ganglion Cells from Human Pluripotent Stem Cells Enables Analysis of Glaucomatous Neurodegeneration. *Stem cells* **34**, 1553-1562, doi:10.1002/stem.2356 (2016).

31 Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell stem cell* **10**, 771-785, doi:10.1016/j.stem.2012.05.009 (2012).

32 Reichman, S. *et al.* From confluent human iPS cells to self-forming neural retina and retinal pigmented epithelium. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8518-8523, doi:10.1073/pnas.1324212111 (2014).

33 Meyer, J. S. *et al.* Optic vesicle-like structures derived from human pluripotent stem cells facilitate a customized approach to retinal disease treatment. *Stem cells* **29**, 1206-1218, doi:10.1002/stem.674 (2011).

34 Lu, Y. *et al.* Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. *Developmental cell* **53**, 473-491 e479, doi:10.1016/j.devcel.2020.04.009 (2020).

35 Lukowski, S. W. *et al.* A single-cell transcriptome atlas of the adult human retina. *EMBO J* **38**, e100811, doi:10.15252/embj.2018100811 (2019).

36 Sridhar, A. *et al.* Single-Cell Transcriptomic Comparison of Human Fetal Retina, hPSC-Derived Retinal Organoids, and Long-Term Retinal Cultures. *Cell reports* **30**, 1644-1659 e1644, doi:10.1016/j.celrep.2020.01.007 (2020).
37 Menon, M. *et al.* Single-cell transcriptomic atlas of the human retina identifies cell types associated with age-related macular degeneration. *Nature communications* **10**, 4902, doi:10.1038/s41467-019-12780-8 (2019).
38 Lo Giudice, Q., Leleu, M., La Manno, G. & Fabre, P. J. Single-cell transcriptional logic of cell-fate specification and axon guidance in early-born retinal neurons. *Development* **146**, doi:10.1242/dev.178103 (2019).
39 Hoang, T. *et al.* Gene regulatory networks controlling vertebrate retinal regeneration. *Science*, doi:10.1126/science.abb8598 (2020).
40 Kuwahara, A. *et al.* Generation of a ciliary margin-like stem cell niche from self-organizing human retinal tissue. *Nature communications* **6**, 6286, doi:10.1038/ncomms7286 (2015).
41 Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78**, 785-798, doi:10.1016/j.neuron.2013.05.029 (2013).
42 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology* **36**, 411-420, doi:10.1038/nbt.4096 (2018).
43 Bae, S., Bessho, Y., Hojo, M. & Kageyama, R. The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development* **127**, 2933-2943 (2000).
44 Bianchi, M. E. & Agresti, A. HMG proteins: dynamic players in gene regulation and differentiation. *Current opinion in genetics & development* **15**, 496-506, doi:10.1016/j.gde.2005.08.007 (2005).
45 Liu, C. *et al.* BTG1 potentiates apoptosis and suppresses proliferation in renal cell carcinoma by interacting with PRMT1. *Oncol Lett* **10**, 619-624, doi:10.3892/ol.2015.3293 (2015).
46 Barreto, G. *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445**, 671-675, doi:10.1038/nature05515 (2007).
47 Horiuchi, D. *et al.* Chemical-genetic analysis of cyclin dependent kinase 2 function reveals an important role in cellular transformation by multiple oncogenic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E1019-1027, doi:10.1073/pnas.1111317109 (2012).
48 Yabuta, N. *et al.* Mammalian Mcm2/4/6/7 complex forms a toroidal structure. *Genes Cells* **8**, 413-421, doi:10.1046/j.1365-2443.2003.00645.x (2003).
49 Liu, J. H. *et al.* Functional association of TGF-beta receptor II with cyclin B. *Oncogene* **18**, 269-275, doi:10.1038/sj.onc.1202263 (1999).
50 Lee, K. S., Oh, D. Y., Kang, Y. H. & Park, J. E. Self-regulated mechanism of Plk1 localization to kinetochores: lessons from the Plk1-PBIP1 interaction. *Cell Div* **3**, 4, doi:10.1186/1747-1028-3-4 (2008).
51 Hafler, B. P. *et al.* Transcription factor Olig2 defines subpopulations of retinal progenitor cells biased toward specific cell fates. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7882-7887, doi:10.1073/pnas.1203138109 (2012).
52 Gu, W. *et al.* Glycolytic Metabolism Plays a Functional Role in Regulating Human Pluripotent Stem Cell State. *Cell stem cell* **19**, 476-490, doi:10.1016/j.stem.2016.08.008 (2016).
53 Gouti, M., Metzis, V. & Briscoe, J. The route to spinal cord cell types: a tale of signals and switches. *Trends Genet* **31**, 282-289, doi:10.1016/j.tig.2015.03.001 (2015).
54 Jadhav, A. P., Cho, S. H. & Cepko, C. L. Notch activity permits retinal cells to progress through multiple progenitor states and acquire a stem cell property. *Proceedings of the National*

Academy of Sciences of the United States of America **103**, 18998-19003, doi:10.1073/pnas.0608155103 (2006).

55 Jadhav, A. P., Mason, H. A. & Cepko, C. L. Notch 1 inhibits photoreceptor production in the developing mammalian retina. *Development* **133**, 913-923, doi:10.1242/dev.02245 (2006).

56 Riesenberg, A. N., Liu, Z., Kopan, R. & Brown, N. L. Rbpj cell autonomous regulation of retinal ganglion cell and cone photoreceptor fates in the mouse retina. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 12865-12877, doi:10.1523/JNEUROSCI.3382-09.2009 (2009).

57 Yaron, O., Farhy, C., Marquardt, T., Applebury, M. & Ashery-Padan, R. Notch1 functions to suppress cone-photoreceptor fate specification in the developing mouse retina. *Development* **133**, 1367-1378, doi:10.1242/dev.02311 (2006).

58 Wang, Y., Dakubo, G. D., Thurig, S., Mazerolle, C. J. & Wallace, V. A. Retinal ganglion cell-derived sonic hedgehog locally controls proliferation and the timing of RGC development in the embryonic mouse retina. *Development* **132**, 5103-5113, doi:10.1242/dev.02096 (2005).

59 Zhang, X. M. & Yang, X. J. Regulation of retinal ganglion cell production by Sonic hedgehog. *Development* **128**, 943-957 (2001).

60 Sakagami, K., Gan, L. & Yang, X. J. Distinct effects of Hedgehog signaling on neuronal fate specification and cell cycle progression in the embryonic mouse retina. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 6932-6944, doi:10.1523/JNEUROSCI.0289-09.2009 (2009).

61 Kim, J. *et al.* GDF11 controls the timing of progenitor cell competence in developing retina. *Science* **308**, 1927-1930, doi:10.1126/science.1110175 (2005).

62 Miesfeld, J. B. *et al.* Rbpj direct regulation of Atoh7 transcription in the embryonic mouse retina. *Scientific reports* **8**, 10195, doi:10.1038/s41598-018-28420-y (2018).

63 Hashimoto, T., Zhang, X. M., Chen, B. Y. & Yang, X. J. VEGF activates divergent intracellular signaling components to regulate retinal progenitor cell proliferation and neuronal differentiation. *Development* **133**, 2201-2210, doi:10.1242/dev.02385 (2006).

64 Maurer, K. A., Riesenberg, A. N. & Brown, N. L. Notch signaling differentially regulates Atoh7 and Neurog2 in the distal mouse retina. *Development* **141**, 3243-3254, doi:10.1242/dev.106245 (2014).

65 Pollak, J. *et al.* ASCL1 reprograms mouse Muller glia into neurogenic retinal progenitors. *Development* **140**, 2619-2631, doi:10.1242/dev.091355 (2013).

66 Ramachandran, R., Fausett, B. V. & Goldman, D. Ascl1a regulates Muller glia dedifferentiation and retinal regeneration through a Lin-28-dependent, let-7 microRNA signalling pathway. *Nature cell biology* **17**, 532, doi:10.1038/ncb3144 (2015).

67 Jorstad, N. L. *et al.* Stimulation of functional neuronal regeneration from Muller glia in adult mice. *Nature* **548**, 103-107, doi:10.1038/nature23283 (2017).

68 Chiodini, F. *et al.* A positive feedback loop between ATOH7 and a Notch effector regulates cell-cycle progression and neurogenesis in the retina. *Cell reports* **3**, 796-807, doi:10.1016/j.celrep.2013.01.035 (2013).

69 Akazawa, C., Sasai, Y., Nakanishi, S. & Kageyama, R. Molecular characterization of a rat negative regulator with a basic helix-loop-helix structure predominantly expressed in the developing nervous system. *The Journal of biological chemistry* **267**, 21879-21885 (1992).

70 Kageyama, R., Shimojo, H. & Isomura, A. Oscillatory Control of Notch Signaling in Development. *Advances in experimental medicine and biology* **1066**, 265-277, doi:10.1007/978-3-319-89512-3_13 (2018).

71 Kageyama, R., Shimojo, H. & Ohtsuka, T. Dynamic control of neural stem cells by bHLH factors. *Neurosci Res* **138**, 12-18, doi:10.1016/j.neures.2018.09.005 (2019).

72 Imayoshi, I. *et al.* Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* **342**, 1203-1208, doi:10.1126/science.1242366 (2013).

73 Ochi, S., Imaizumi, Y., Shimojo, H., Miyachi, H. & Kageyama, R. Oscillatory expression of Hes1 regulates cell proliferation and neuronal differentiation in the embryonic brain. *Development* **147**, doi:10.1242/dev.182204 (2020).

74 Riesenberg, A. N. *et al.* Pax6 regulation of Math5 during mouse retinal neurogenesis. *Genesis* **47**, 175-187, doi:10.1002/dvg.20479 (2009).

75 Lee, H. Y. *et al.* Multiple requirements for Hes 1 during early eye formation. *Developmental biology* **284**, 464-478, doi:10.1016/j.ydbio.2005.06.010 (2005).

76 Matter-Sadzinski, L., Matter, J. M., Ong, M. T., Hernandez, J. & Ballivet, M. Specification of neurotransmitter receptor identity in developing retina: the chick ATH5 promoter integrates the positive and negative effects of several bHLH proteins. *Development* **128**, 217-231 (2001).

77 Boije, H., Rulands, S., Dudczig, S., Simons, B. D. & Harris, W. A. The Independent Probabilistic Firing of Transcription Factors: A Paradigm for Clonal Variability in the Zebrafish Retina. *Developmental cell* **34**, 532-543, doi:10.1016/j.devcel.2015.08.011 (2015).

78 He, J. *et al.* How variable clones build an invariant retina. *Neuron* **75**, 786-798, doi:10.1016/j.neuron.2012.06.033 (2012).

79 Jiang, Y. *et al.* Transcription factors SOX4 and SOX11 function redundantly to regulate the development of mouse retinal ganglion cells. *The Journal of biological chemistry* **288**, 18429-18438, doi:10.1074/jbc.M113.478503 (2013).

80 Chang, K. C. *et al.* Novel Regulatory Mechanisms for the SoxC Transcriptional Network Required for Visual Pathway Development. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **37**, 4967-4981, doi:10.1523/JNEUROSCI.3430-13.2017 (2017).

81 Koike, C. *et al.* Functional roles of Otx2 transcription factor in postnatal mouse retinal development. *Molecular and cellular biology* **27**, 8318-8329, doi:10.1128/MCB.01209-07 (2007).

82 Yamamoto, H., Kon, T., Omori, Y. & Furukawa, T. Functional and Evolutionary Diversification of Otx2 and Crx in Vertebrate Retinal Photoreceptor and Bipolar Cell Development. *Cell reports* **30**, 658-671 e655, doi:10.1016/j.celrep.2019.12.072 (2020).

83 Rehemtulla, A. *et al.* The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 191-195, doi:10.1073/pnas.93.1.191 (1996).

84 Oh, E. C. *et al.* Transformation of cone precursors to functional rod photoreceptors by bZIP transcription factor NRL. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1679-1684, doi:10.1073/pnas.0605934104 (2007).

85 Wu, F. *et al.* Onecut1 is essential for horizontal cell genesis and retinal integrity. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **33**, 13053-13065, 13065a, doi:10.1523/JNEUROSCI.0116-13.2013 (2013).

86 Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *Journal of virology* **72**, 8463-8471 (1998).

87 Zufferey, R. *et al.* Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *Journal of virology* **72**, 9873-9880 (1998).

88 Hashimoto, T. *et al.* Lentiviral gene replacement therapy of retinas in a mouse model for Usher syndrome type 1B. *Gene Ther* **14**, 584-594, doi:10.1038/sj.gt.3302897 (2007).

89 Hartwick, A. T., Lalonde, M. R., Barnes, S. & Baldridge, W. H. Adenosine A1-receptor modulation of glutamate-induced calcium influx in rat retinal ganglion cells. *Investigative ophthalmology & visual science* **45**, 3740-3748, doi:10.1167/iovs.04-0214 (2004).

90 Hirooka, K. *et al.* T-Type calcium channel alpha1G and alpha1H subunits in human retinoblastoma cells and their loss after differentiation. *Journal of neurophysiology* **88**, 196-205 (2002).

91 Lalonde, M. R., Jollimore, C. A., Stevens, K., Barnes, S. & Kelly, M. E. Cannabinoid receptor-mediated inhibition of calcium signaling in rat retinal ganglion cells. *Molecular vision* **12**, 1160-1166 (2006).

92 Sargoy, A., Sun, X., Barnes, S. & Brecha, N. C. Differential calcium signaling mediated by voltage-gated calcium channels in rat retinal ganglion cells and their unmyelinated axons. *PloS one* **9**, e84507, doi:10.1371/journal.pone.0084507 (2014).

93 Ye, L., Gu, L., Caprioli, J. & Piri, N. RNA-binding protein Rbpms is represented in human retinas by isoforms A and C and its transcriptional regulation involves Sp1-binding site. *Mol Genet Genomics* **293**, 819-830, doi:10.1007/s00438-018-1423-8 (2018).

94 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).

95 Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628-2629, doi:10.1093/bioinformatics/btz931 (2020).

96 Kessler, S., Pohlert, T., Breitung, V., Wilcsek, K. & Bierl, R. Comparative evaluation of four suspended particulate matter (SPM) sampling devices and their use for monitoring SPM quality. *Environ Sci Pollut Res Int* **27**, 5993-6008, doi:10.1007/s11356-019-07314-0 (2020).

97 FSA: Fisheries Stock Analysis v. R package version 0.8.30 (2020).

98 Hadley, W. *Ggplot2*. (Springer Science+Business Media, LLC, 2016).

99 ggsignif: Significane Brackets for 'ggplot2' v. R package version 0.6.0 (2019).

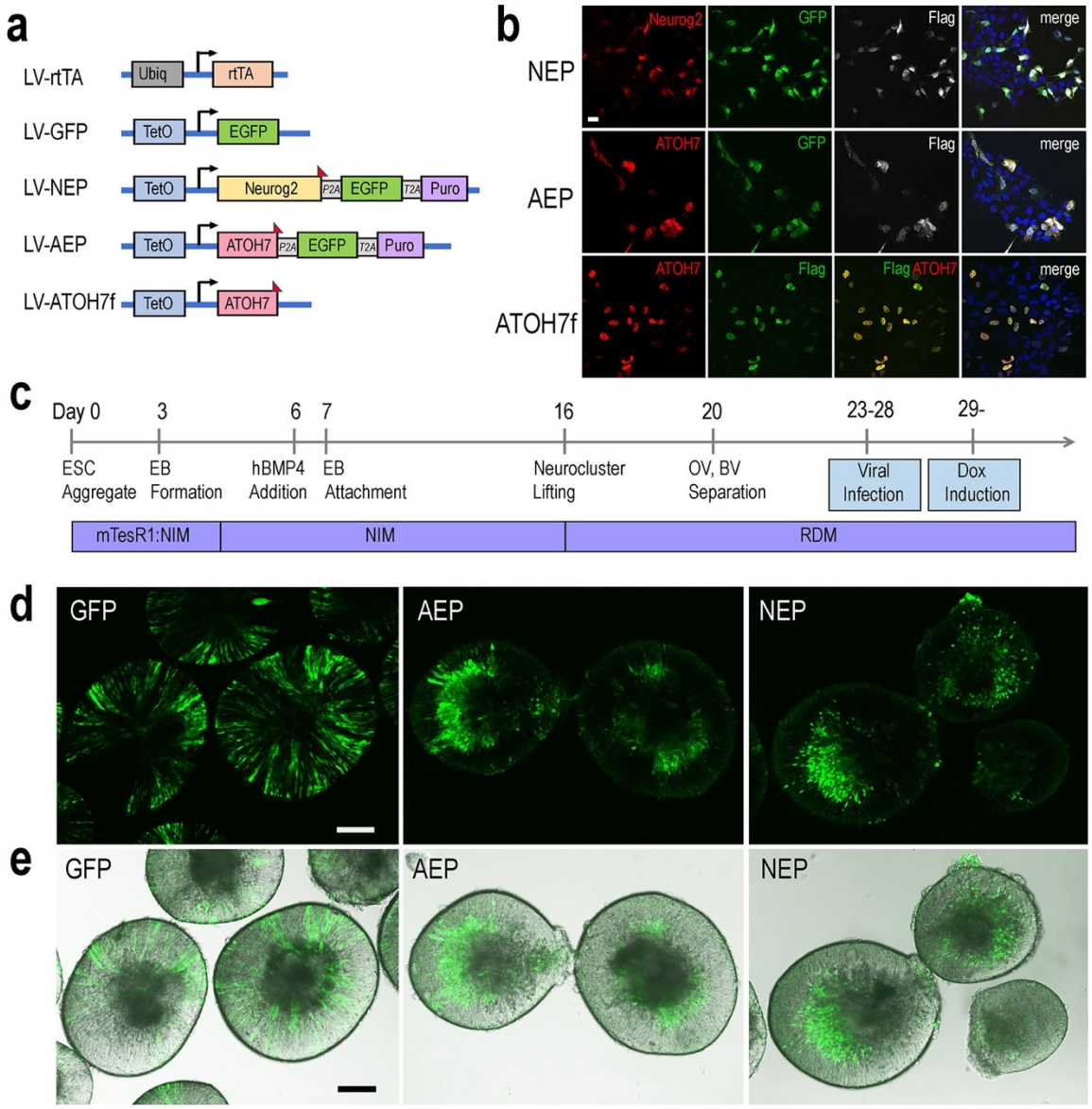

Fig. 1 Inducible lentivirus-mediated gene expression in human ES cell-derived retinal organoids.

a Schematics of lentiviral vectors used in the study. A vector expressing rtTA under the control of the constitutive Ubiq promoter was used to co-infect human ES cells with lentiviral vectors encoding the inducible TetO promoter. Puro, puromycin-resistant gene; *P2A* and *T2A*, intervening 2T sequences in frame with the coding sequences of the other genes in LV-NEP and LV-AEP. The Flag epitope tag was fused to the c-terminus of Neurog2 and ATOH7 coding regions. **b** Expression of Neurog2 and ATOH7 proteins as detected by immunofluorescent microscopy in H9 ES cells co-infected with LV-rtTA and either LV-NEP, LV-AEP, or LV-ATOH7f after 36 hours of Dox induction. Scale Bar for all, 20μm. **c** Experimental time course of human retinal organoid derivation and lentiviral infection. **d, e** Confocal images (**d**) and merged confocal and bright field images (**e**) show distribution of viral infected GFP^+ cells in live retinal organoids at Day 36 after 6 days of Dox induction. Names of lentiviruses are abbreviated as: GFP for LV-GFP, NEP for LV-NEP, AEP for LV-AEP. Scale Bars, 100 μm.

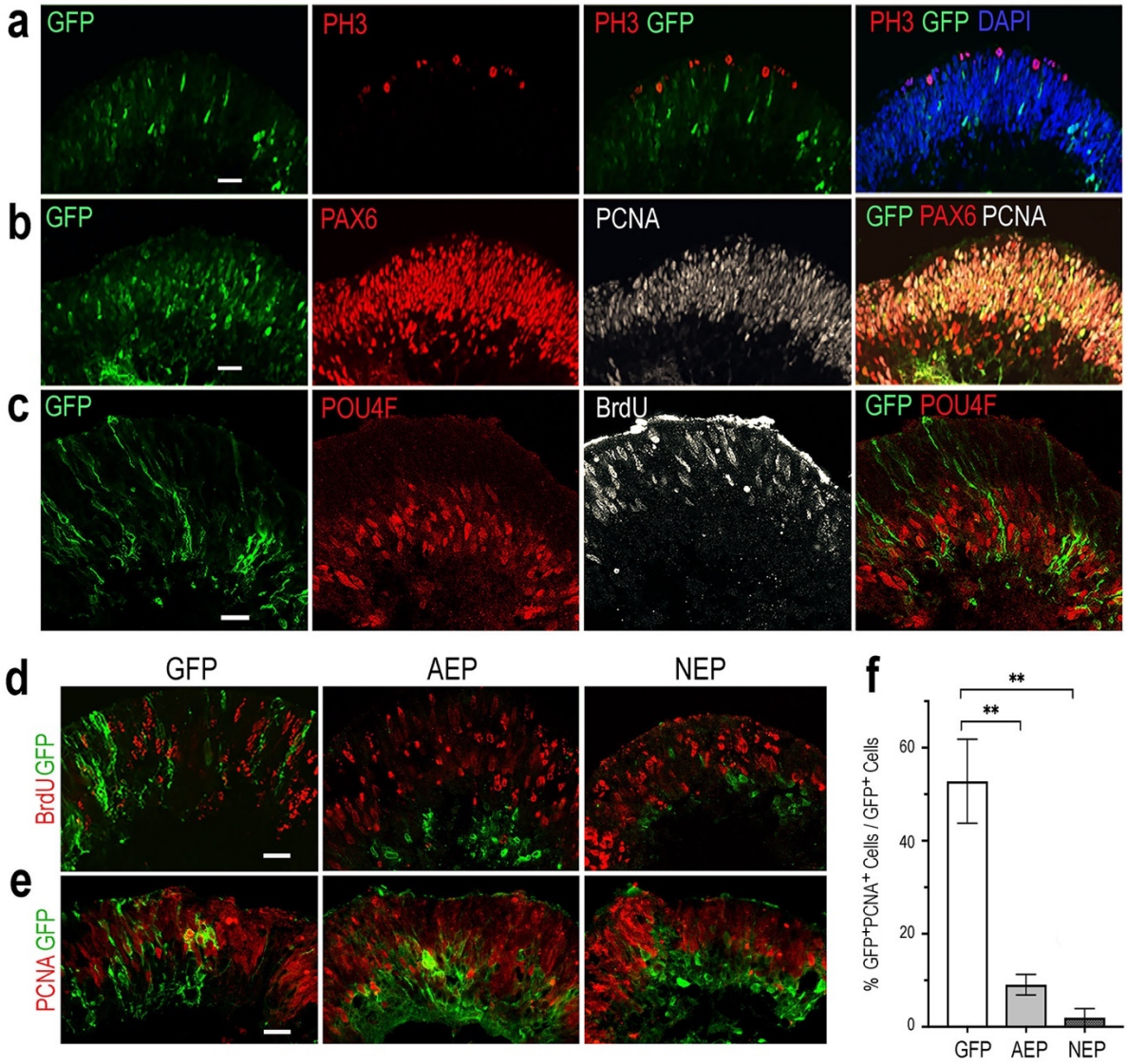

Fig. 2 Influences of viral mediated neurogenic factor expression on cell proliferation.

a-c Immunofluorescence labeling of cross sections from retinal organoids infected with LV-GFP. At Day 34, $PH3^+$ dividing cells are located at the ventricular surface (**a**), while $PCNA^+$ progenitors occupy the ventricular zone and are co-labeled with PAX6-expressing progenitors (**b**). Note that $PAX6^+PCNA^-$ postmitotic neurons are located to the inner layer of the retinal organoid. At Day 40, proliferating progenitors labeled with BrdU are distributed in the ventricular zone, whereas $POU4F^+$ postmitotic retinal ganglion cells (RGCs) reside in the inner layer of the retinal organoid (**c**). Scale bars, 20 μm. **d, e** Confocal images of cross sections from retinal organoids infected by LV-GFP, LV-AEP, or LV-NEP after 6-day Dox induction. Co-staining of GFP with BrdU at Day 36 (**d**, 3-hour labeling) or PCNA at Day 40 (**e**) shows that most LV-AEP and LV-NEP infected cells are postmitotic. Scale bars, 20 μm. **f** Quantification of double labeled $PCNA^+GFP^+$ progenitor cells among total viral infected GFP^+ cells at Day 47. Mean ± s.e.m. of n= 3 independent samples. One-way ANOVA, ** $p< 0.01$.

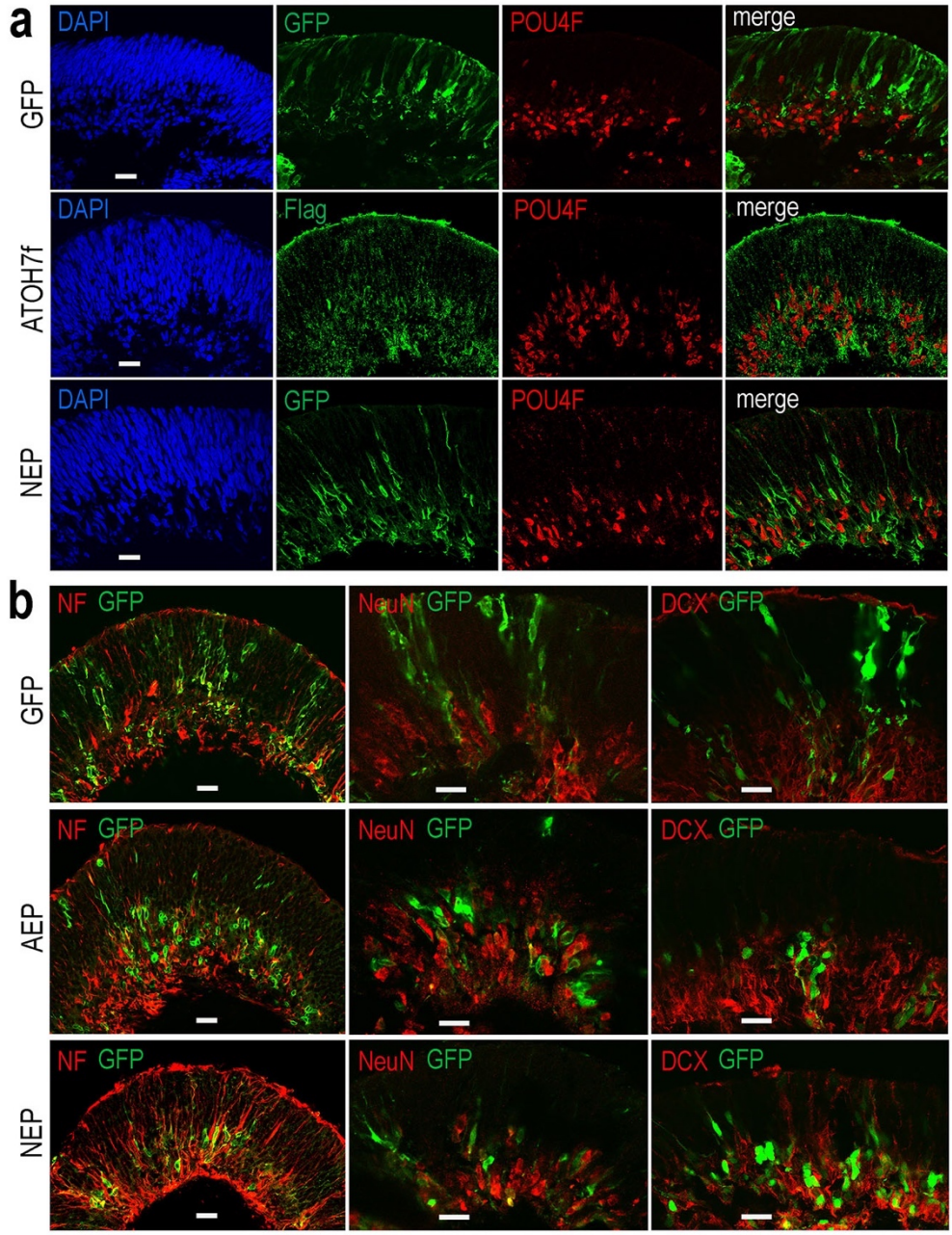

Fig. 3 Effects of viral mediated neurogenic factor expression on retinal ganglion cell genesis.

Confocal images of retinal organoid sections immunolabeled with retinal ganglion cell (RGC) markers. **a** Co-labeling of viral markers with a pan-POU4F antibody in LV-NEP and LV-ATOH7f infected retinal organoids at Day 43. **b** Merged images of viral marker GFP co-labeling with NF145 and NeuN at Day 36, and DCX at Day 46. Scale Bars, 20 μm.

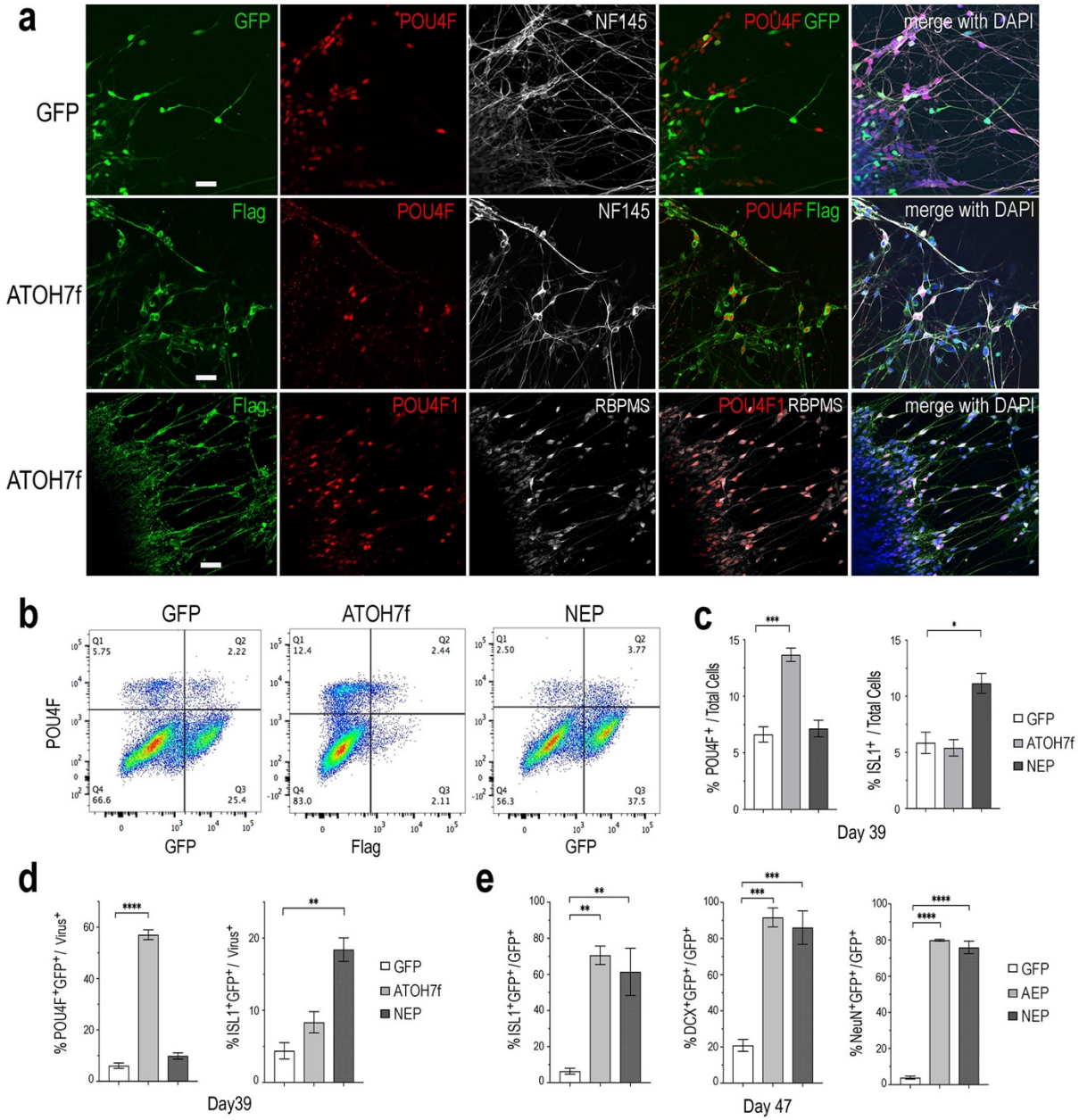

Fig. 4 Impacts of elevating neurogenic factor expression on retinal ganglion cell genesis.

a Immunofluorescent images of attached retinal organoid cultures infected with LV-GFP or LV-ATOH7f after 4-day Dox induction at Day 39. Co-labeling for viral vector markers with POU4F and NF145 at Day 38, or POU4F1 and RBPMS at Day 41 shows high correspondence of LV-ATOH7f infection and RGC marker expression. Scale bars, 20 μm. **b** Representative flow cytometry profiles of dissociated cells at Day 39 from LV-GFP, LV-ATOH7f, and LV-NEP infected retinal organoids after 4-day Dox induction. Cells were co-labeled for RGC marker POU4F and viral marker Flag for LV-ATOH7f, or GFP for LV-GFP and LV-NEP, respectively. **c, d** Bar graphs show flow cytometry quantification of RGC markers POU4F and ISL1 among total cells (**c**) and viral infected cells (**d**) at Day 39. **e** Quantification of cultured monolayer cells from LV-GFP, LV-AEP, and LV-NEP infected retinal organoids at Day 47 after 7-day Dox induction. Percentages of RGC marker ISL1, DCX, and NeuN positive cells among viral infected GFP^+ cells are shown. For **c, d, e** Mean ± s.e.m. of n= 3 independent samples. One-way ANOVA, **** $p< 0.0001$, *** $p< 0.001$, **$p< 0.01$.

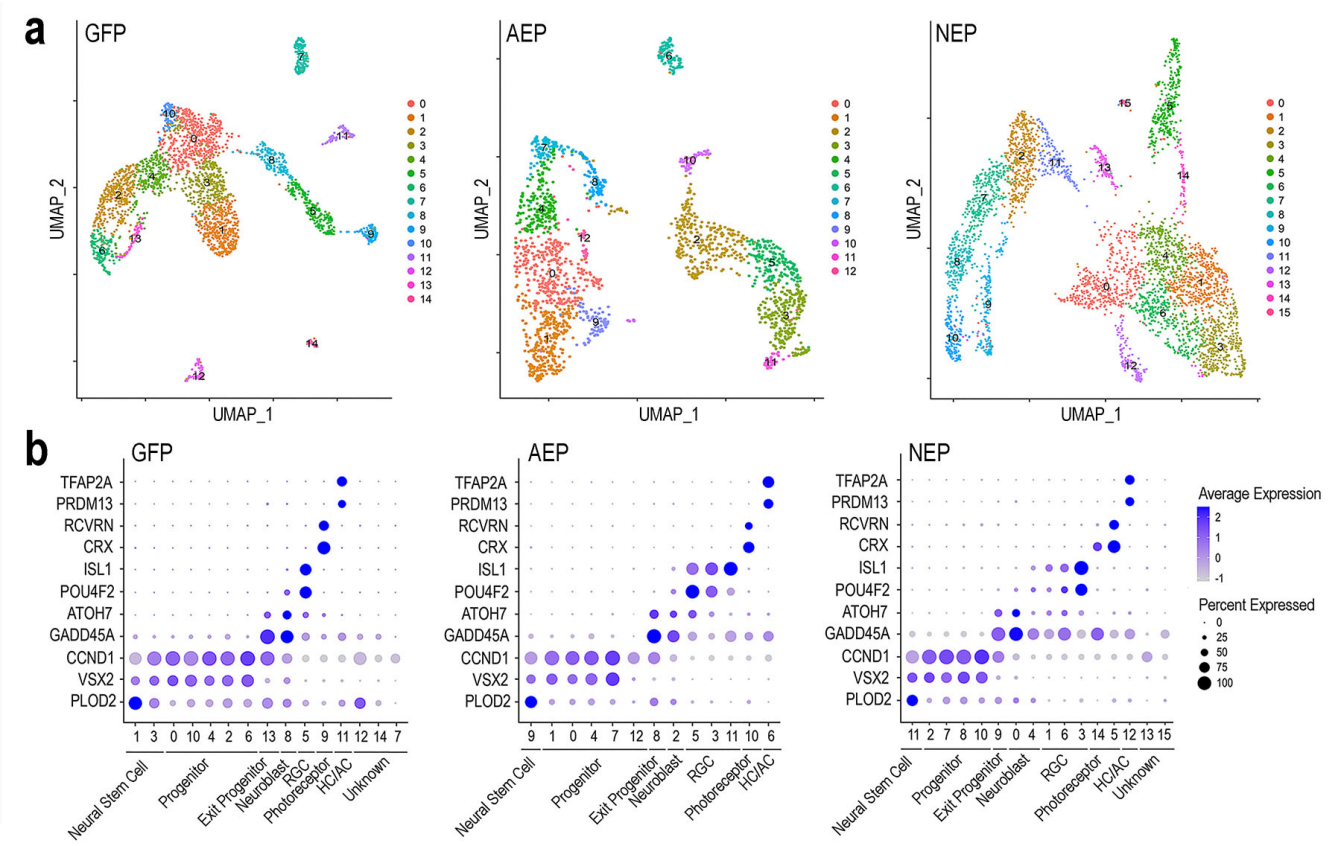

Fig.5 Characterization of lentivirus-infected human retinal organoids by single cell RNA-seq.

a UMAP visualization of cell clustering based on sequencing of single cell cDNA libraries from lentivirus infected retinal organoids between culture Day 45-48 after 8-day Dox induction. The numbers of single cell that passed quality control and used for downstream analysis are 3004 for LV-GFP, 2063 for LV-AEP, and 3909 for LV-NEP samples. **b** Dot plots show assignments of UMAP clusters into seven major cell categories based on expression of known genes for different cell states or types during early retinal development. Minor cell clusters in LV-GFP and LV-NEP infected retinal organoids that do not show characteristic retinal gene expression are designated as “unknown”. The average expression levels are represented by the intensity bar, and the percentages of cells expressing a given gene in a cluster are indicated as the dot size.

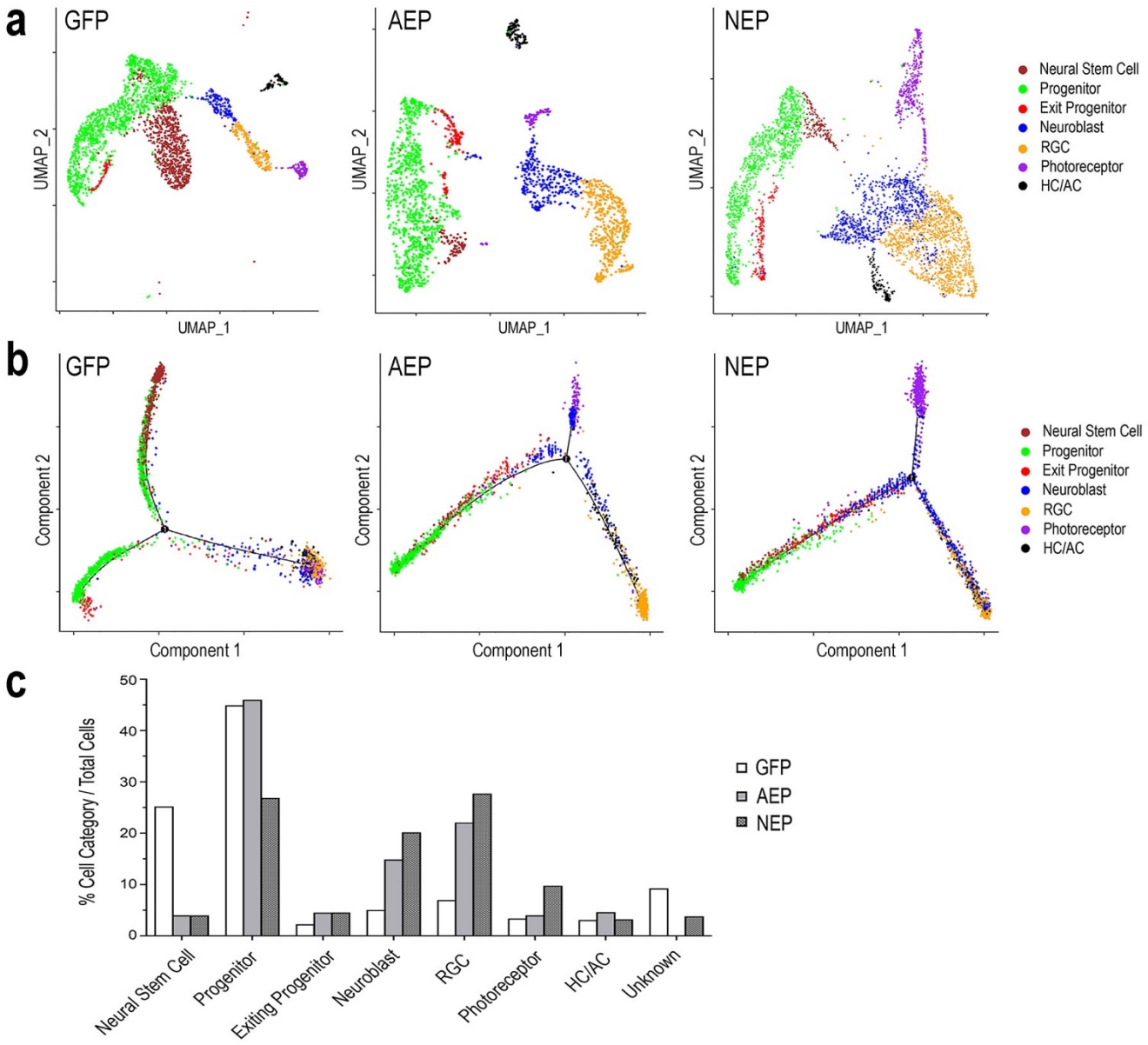

Fig. 6 Distinct cell categories based on single cell RNA-seq profiles and their developmental trajectory.

a UMAP presentations based on cell category assignments for single cell clusters from LV-GFP, LV-AEP, and LV-NEP infected retinal organoids. Each cell category is represented by a single color. The side legend shows the color codes used to represent distinct cell categories. **b** Pseudotime trajectory analysis using Monocle of different cell categories from LV-GFP, LV-AEP, and LV-NEP infected retinal organoid. **c** Bar graph shows the percentages of different cell categories among total cells in LV-GFP, LV-AEP, and LV-NEP infected retinal organoid cells.

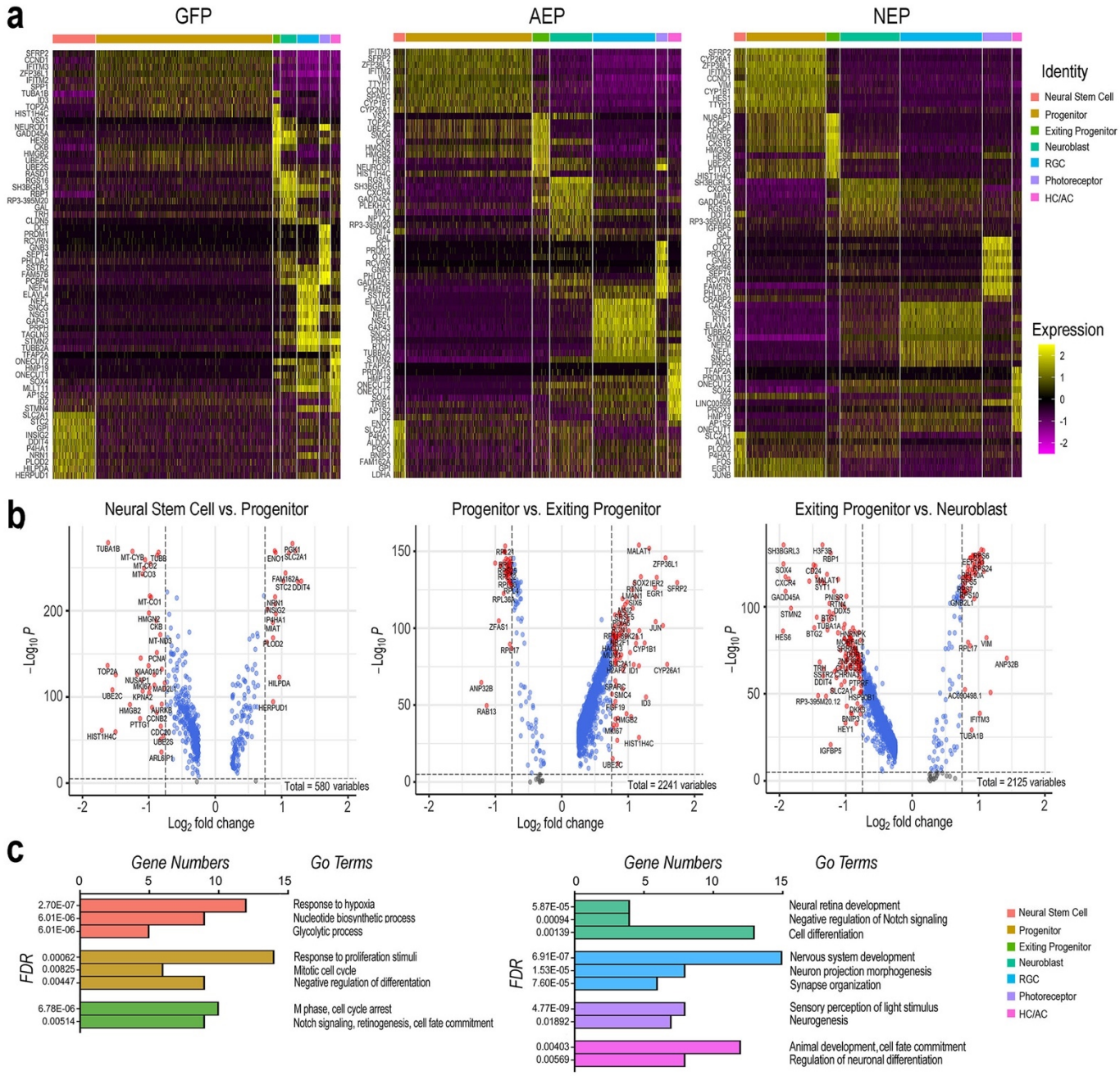

Fig. 7 Effects of elevating neurogenic factor expression on developmental transitions.

a Heatmaps of the top ten DEGs in each cell type category from cells of LV-GFP, LV-AEP, and LV-NEP infected retinal organoids. The side legend indicates the color codes for cell category identities atop the heatmaps. **b** Volcano plots show significant DEGs in LV-GFP infected retinal organoids during the three transitions: from the neural stem cell to the neurogenic progenitor state, from the cycling progenitor to the exiting progenitor state, and from the exiting progenitor to the neuroblast states. Genes with *p* value < 0.01 and fold change > Log_2 0.75 are shown as red dots. **c** Bar graphs show the predominant biological processes using GO pathway analysis for each cell category of LV-GFP infected retinal organoids. The false positive rates (FDR) and the gene numbers among the top 25 DEGs associated with a given GO term are shown. The right-side legend shows the same color codes corresponding to each cell category as in the heatmaps.

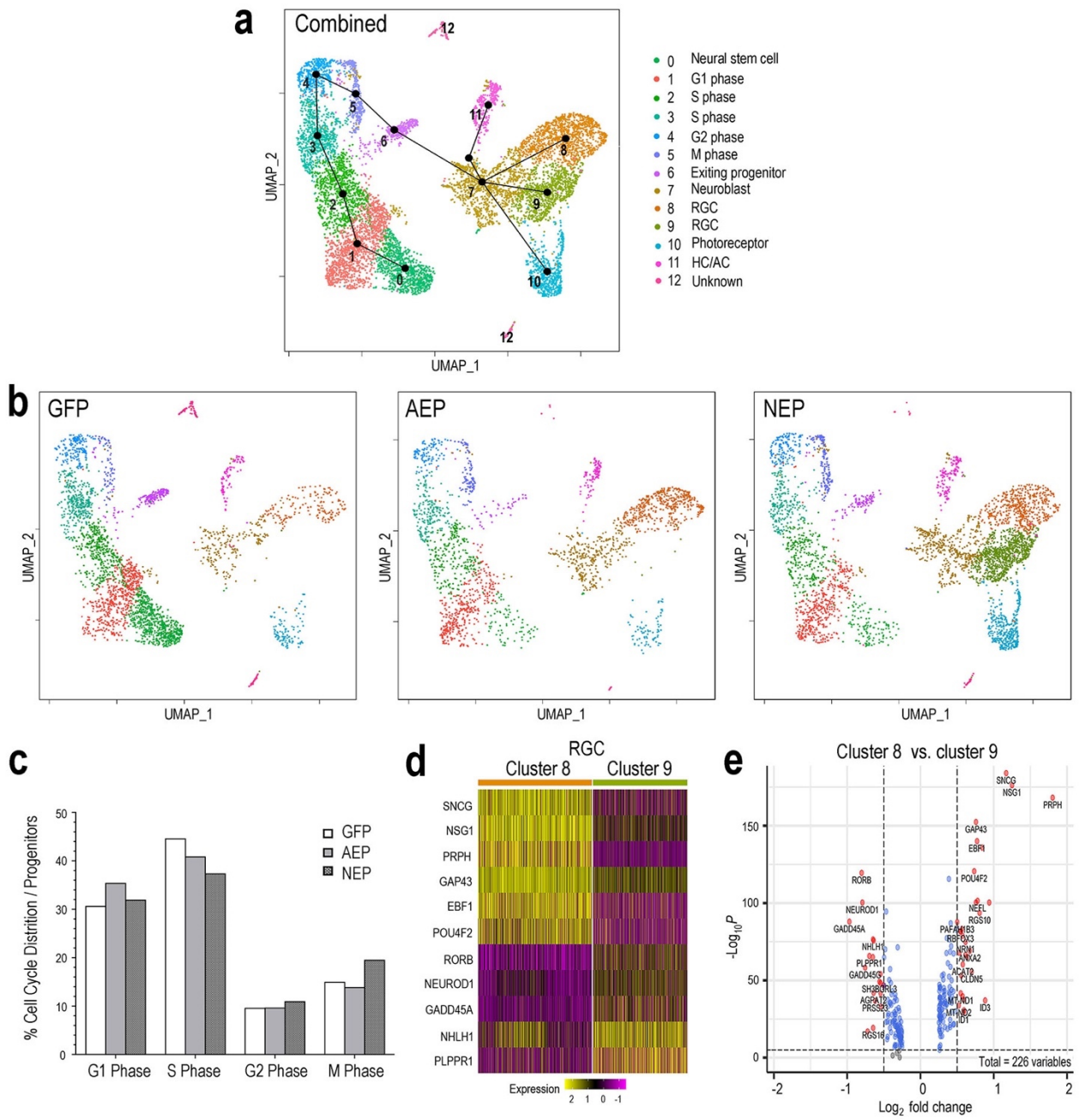

Fig. 8 Influence of neurogenic factors on cell cycle progression and retinal ganglion cell differentiation.

a UMAP of the combined sc RNA-seq dataset from LV-GFP, LV-AEP, and LV-NEP infected retinal organoids. Each cell cluster is assigned to an identity as shown on the side color legend based of feature plots of known genes. Line edges connecting individual clusters are generated by SLINGSHOT analysis to show the predicted developmental trajectory of the entire cell cohort. **b** UMAPs show cell clusters for individual samples from LV-GFP, LV-AEP, and LV-NEP infected retinal organoids with the same cell identity assignments as in the combined UMAP. **c** Distribution of cells in different phases of the cell cycle as percentages of total proliferative progenitor cells. The M phase cells include the exiting progenitors (cluster #6). **d** Heatmap shows differential expression levels of genes in the two RGC clusters 8 and 9. **e** Volcano plot shows DEGs between the two RGC clusters 8 and 9. Genes with *p* value < 0.01 and fold change > Log_2 0.5 are shown as red dots.

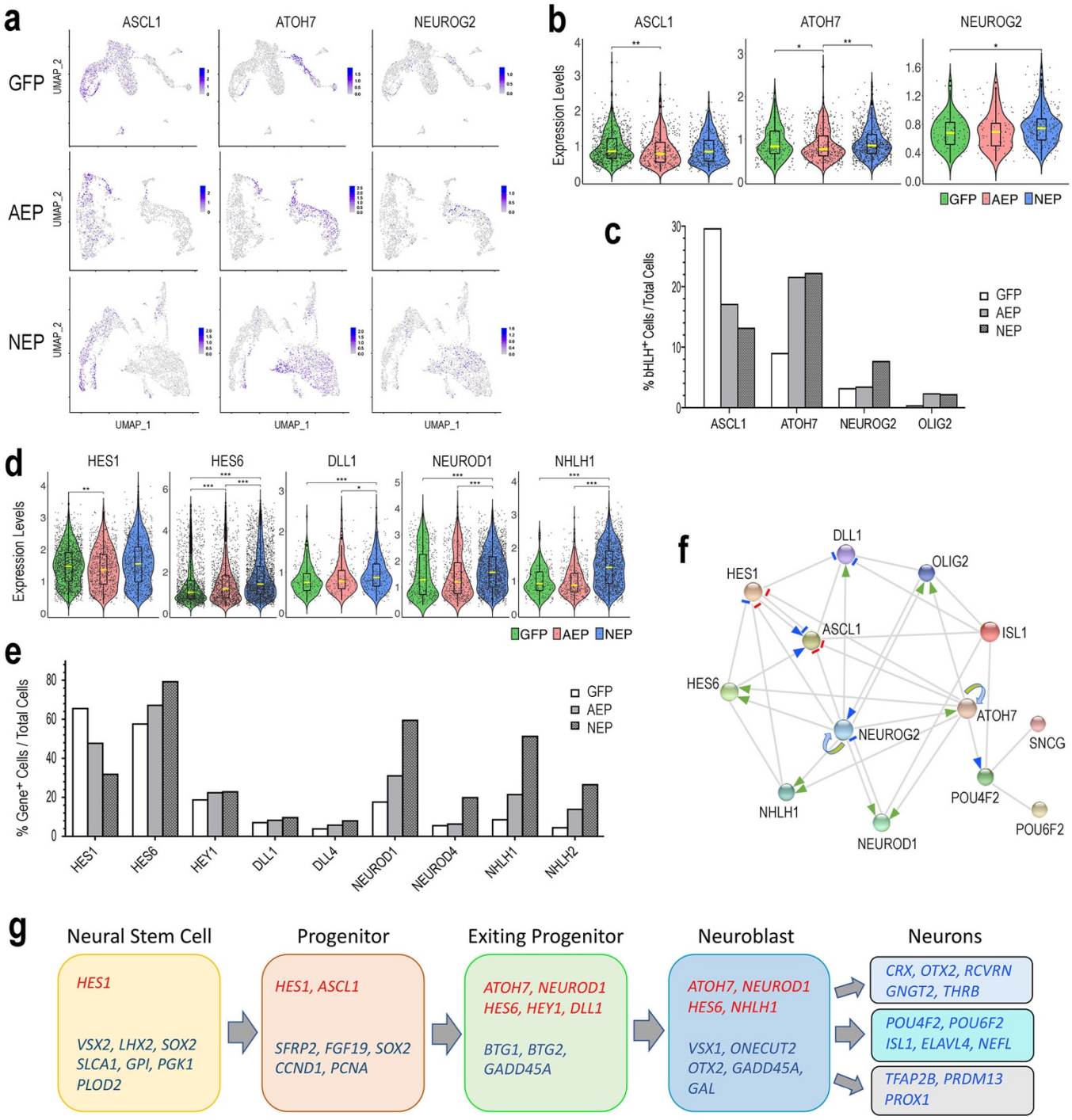

Fig. 9 Regulatory relationship among genes playing roles in early human retinogenesis.

a Feature plots show distribution of ASCL1, ATOH7, and NEUROG2 expressing cells among cell clusters in LV-GFP, LV-AEP, and LV-NEP infected retinal organoids. **b, d** Violin plots show comparisons of LV-GFP (green) versus LV-AEP (red) or LV-NEP (blue) induction on expression levels of endogenous genes in the retinal organoids. The box within the violin represents the middle 50% of the data, and the yellow line within the box indicates the median expression level. Statistical analysis generated P-values were based

on both the cell counts and expression levels. ***p<0.001, ** p< 0.01, *p< 0.05. **c, e** Bar graphs show percentages of cells expressing endogenous genes among total viral infected cells in LV-GFP, LV-AEP, and LV-NEP samples. **f** Schematic model based on STRING analysis summarizes a gene network involved in cell cycle exit and early retinogenesis. Grey edges represent protein-protein associations. Positive and negative regulatory relationships are indicated as arrows and short bars, respectively. Previously known molecular interactions from curated databases are shown as blue arrows and short bars. Effects of lentiviral induced ATOH7f and Neurog2 on endogenous human gene expression in retinal organoids are indicated as green arrows and red bars. Looping arrows indicate positive auto-regulation. **g** Summary of sequential expression of bHLH genes (red) and other key genes (blue) during early human retinal organoid development.