**Supplementary Information. Supplementary Fig. 1-3**

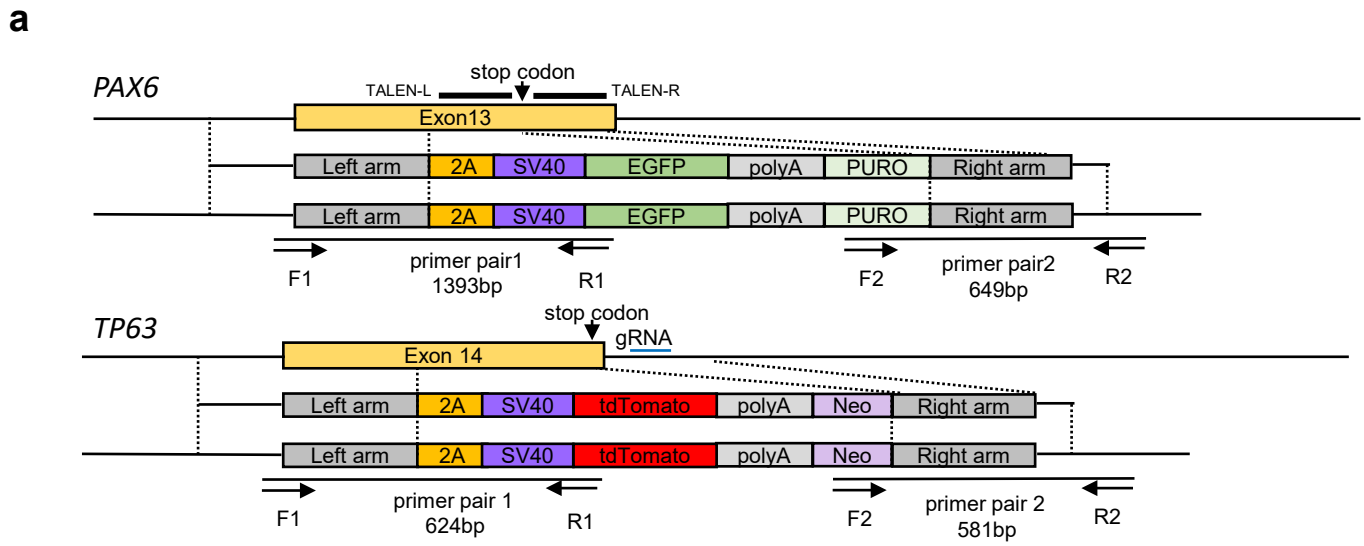
**Supplementary Note 1.** Flow cytometry methodology and gating strategy.

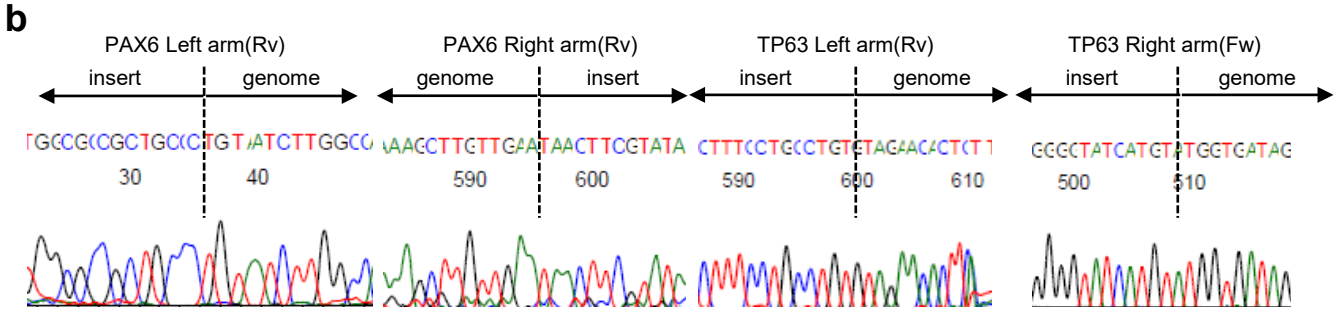
**Supplementary Data 1.** Excel file containing lists of cluster-specific genes related to Fig.s 2, 3, 4, Supplementary Fig.2, and 3.

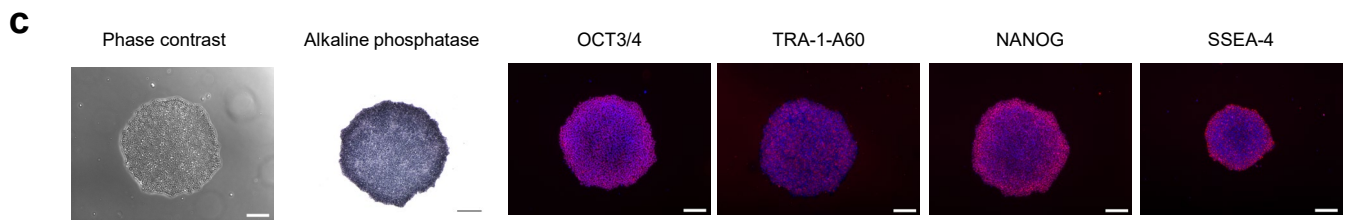
**Supplementary Data 2.** Excel file containing lists of cluster-specific regulons, related to Fig.s 2, 3, 4, and Supplementary Fig. 2.

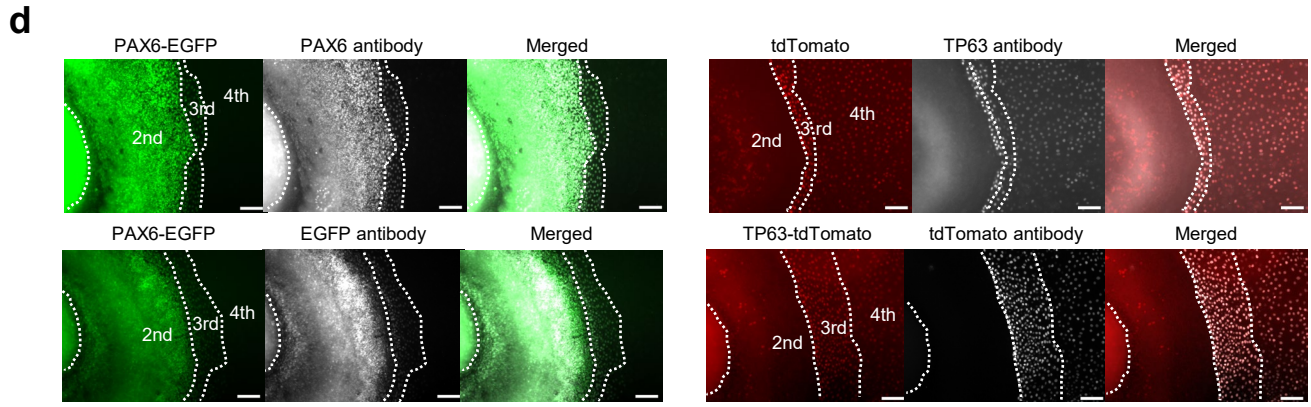
**Supplementary Data 3.** Excel file containing lists of PAX6-regulated genes identified by SCENIC analysis, classified into PAX6-extended and high-confidence targets. Related to Fig. 2, 3, 4, and Supplementary Fig. 2.

**Supplementary Data 4.** The source data behind the graphs in the paper. Related to Fig. 6b and 6d.




Supplementary Fig. 1

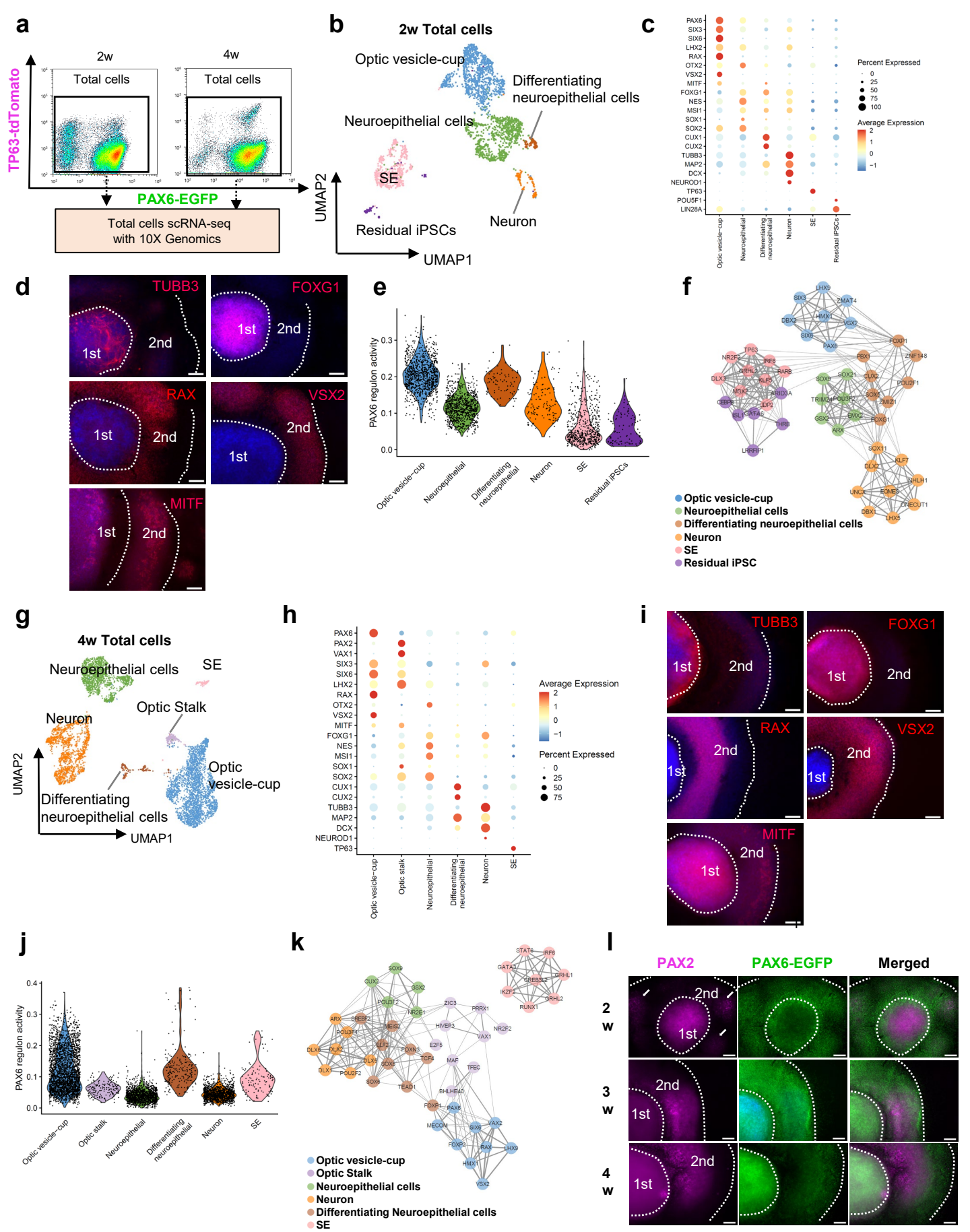
**Supplementary Fig 1. Generation and validation of dual reporter hiPSCs expressing PAX6-EGFP and TP63-tdTomato.**

(a) Schematic representation of gene targeting strategies for knock-in of EGFP at the PAX6 locus and tdTomato at the TP63 locus. Primer positions used for genotyping are indicated.

(b) Genomic sequencing confirming precise integration of the reporter constructs at the PAX6 and TP63 loci.

(c) Characterization of undifferentiated reporter hiPSC colonies by phase-contrast imaging, alkaline phosphatase staining, and immunofluorescence for pluripotency markers OCT3/4, TRA-1-60, NANOG, and SSEA-4.

(d) Immunofluorescence analysis at 4 weeks of differentiation showing co-localization of PAX6 with EGFP and TP63 with tdTomato across SEAM zones**.** Representative images are from one biological replicate . Scale bars: 100 μm.




Supplementary Fig. 2

**Supplementary Fig 2. Ocular development analysis combining SEAM and scRNA-seq**

(a) Flow cytometry analysis of PAX6-EGFP and TP63-tdTomato expression in total cells at 2 and 4 weeks post differentiation. These cells were used for scRNA-seq.

(b) UMAP plot of 2-week scRNA-seq data showing clusters including optic vesicle cup, neuroepithelial cells, differentiating neuroepithelial cells, neurons, residual iPSCs, and SE.

(c) Bubble plot showing expression levels of representative genes in each cluster at 2 weeks.

(d) Immunofluorescence staining of SEAM zones (1st: zone 1, 2nd: zone 2, 3rd/4th: zone 3/4), showing marker localization across zones using TUBB3, FOXG1, RAX, VSX2, and MITF. Scale bar: 100 μm. Representative images are from one biological replicate using two independent colonies.

(e) Violin plot depicting the PAX6 regulon activity in each cell cluster of Total cells at 2 weeks.

(f) Network diagram from SCENIC analysis, illustrating the TF regulon activities specific to each cluster within the Total cells at 2 weeks. Node colors and edge thickness represent clusters and coefficients, respectively. Abbreviations: SE, Surface Ectoderm; w, week(s).

(g) UMAP plot of 4-week scRNA-seq data showing clusters including optic vesicle cup, neuroepithelial cells, differentiating neuroepithelial cells, neurons, optic stalk, and SE.

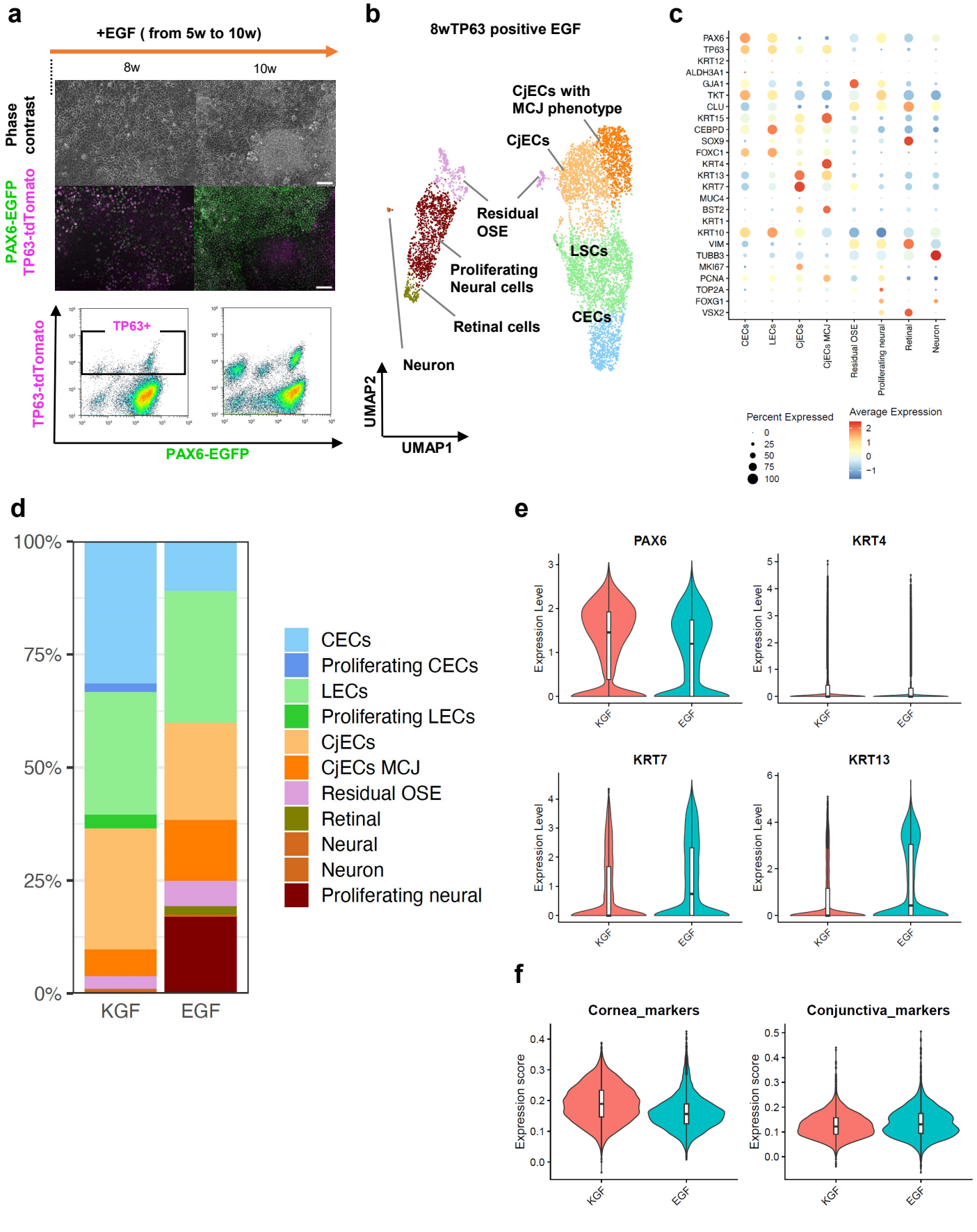
(h) Bubble plot showing expression levels of representative genes in each cluster at 4 weeks.

(i) Immunofluorescence staining of SEAM zones (1st: zone 1, 2nd: zone 2) showing marker localization across zones using TUBB3, FOXG1, RAX, VSX2, and MITF. Scale bar: 100 μm. Representative images are from one biological replicate.

(j) Violin plot depicting the PAX6 regulon activity in each cell cluster of total cells at 4 weeks.

(k) Network diagram from SCENIC analysis illustrating the TF regulon activities specific to each cluster within the total cells at 4 weeks. Node colors and edge thickness represent clusters and correlation coefficients, respectively.

(l) Immunofluorescence staining showing dynamic expression patterns of PAX2 and PAX6-EGFP in SEAM at 2, 3, and 4 weeks. Co-expression at 2 weeks becomes spatially separated by 4 weeks. Scale bar: 100 μm. Representative images are from one biological replicate using two independent colonies. scRNA-seq analyses (b, c, e, f, g, h, j, k) were performed using one biological replicate.




Supplementary Fig. 3

**Supplementary Figure 3. Single-cell transcriptomic comparison of TP63-positive cells induced by EGF.**

(a) Sequential imaging of hiPSC differentiation into ocular surface epithelial cells across 8–10 weeks with EGF treatment, with PAX6-EGFP and TP63-tdTomato fluorescence. Flow cytometry plots indicate TP63-positive cell frequency during EGF induction. Representative images are from three independent biological replicates as applicable. Scale bars: 100 μm

(b) UMAP plot of TP63-positive cells at 8 weeks after EGF treatment showing clusters of CECs, LECs, CjECs, CjECs with MCJ phenotype, residual OSE, and retinal, neural, and proliferating neuroepithelial cells.

(c) Bubble plot showing expression levels of representative genes across clusters.

(d) Bar plot comparing the proportions of cell types in TP63-positive cells between KGF- and EGF-treated conditions.

(e) Violin plots of PAX6, KRT4, KRT7, and KRT13 expression in KGF- and EGF-treated cells.

(f) Violin plots comparing corneal and conjunctival marker scores between KGF- and EGF-treated TP63-positive cells.

scRNA-seq analyses (b, c) were performed using one biological replicate.

**Supplementary Note 1. Flow cytometry methodology and gating strategy**

**Sample preparation**

Flow cytometry was performed on differentiated PAX6-EGFP/TP63-tdTomato hiPSCs at 2, 4, 8, and 10 weeks. Cells were dissociated using StemPro™ Accutase™ (Thermo Fisher Scientific), and analyzed and sorted using a SH800S cell sorter (Sony, Tokyo, Japan). Data acquisition and analysis were performed using the instrument's dedicated software.

**Instrument**

Cell sorter: SH800S (Sony, Tokyo, Japan)

**Software**

Data acquisition and analysis were conducted using Sony Cell Sorter SH800S software.

**Cell population abundance**

Sorted cell populations were used to generate epithelial sheets exhibiting clearly distinct phenotypes, supporting the validity of the gating strategy.

**Gating strategy**

Two gating strategies were applied depending on the experimental context: in some cases, tdTomato-positive cells were further subdivided based on EGFP fluorescence intensity, while in others, the entire tdTomato-positive population was sorted regardless of EGFP expression.

To exclude autofluorescent cells, control iPSCs lacking reporter knock-ins were used to define baseline fluorescence. A 99th percentile cutoff was applied to the control population, and cells in the experimental groups exceeding this threshold were defined as reporter-positive.

Additionally, FSC/SSC-based gating was employed to exclude debris and cell aggregates. Further gating was applied to eliminate diagonally distributed populations likely attributable to autofluorescence, ensuring accurate detection of specific fluorescent signals.