METHODS

Human pathology samples. Corneal epithelium squamous metaplasia and all other tissues were obtained as de-identified surgical specimens, fixed in 5% formalin, embedded into paraffin, sectioned and stained for immunofluorescence studies.

Isolation and culture of limbal stem cells and skin epidermal stem cells. Post-mortem human eyeballs were obtained from eye banks and limbus regions were taken and washed in cold PBS with 100 international units (IU) penicillin and 100 μg ml^{-1} streptomycin, and cut into small pieces. Cell clusters were obtained by 0.2% collagenase IV digestion at 37 °C for 2 h, single cells were obtained by further digestion with 0.25% trypsin-EDTA at 37 °C for 15 min. Primary cells were seeded on plastic plates coated with 2% growth factor reduced Matrigel (354230, BD Biosciences). Limbal stem cells from GFP-labelled rats and rabbits were isolated and cultured using the same method as for human LSCs.

Human epidermis was obtained from donor skin biopsy of eye lids, and hair follicles were removed under microscope. Primary human and rabbit epidermal stem cells were isolated from interfollicular epidermis using the same method as described for human limbal stem cells. Culture medium was as follows: DMEM/F12 and DMEM (1:1) with 1/100 penicillin–streptomycin, 10% fetal bovine serum, 10 ng ml^{-1} EGF, 5 μg ml^{-1} insulin, 0.4 ug ml^{-1} hydrocortisone, 10^{-10} M cholera toxin and 2×10^{-9} M 3,3′,5-triiodo-L-thyronine.

All cells used in the current manuscript are from primary cultured cells made in our laboratories, and mycoplasma contamination tests were routinely carried out and were negative.

***In vitro* three-dimensional differentiation protocol.** Three-dimensional differentiation was performed on a 24-well plate or an 8-well chamber. In brief, dissociated single stem cells were embedded in matrigel at 2×10^4 cells per 50 μl gel. Three-dimensional structures were formed after 14–18 days culture in a differentiation medium CnT-30 (limbal stem cell differentiation) or CnT-02 (skin epidermal stem cell differentiation) (Cellntec).

Immunofluorescence and laser confocal microscopy. To detect the localization of proteins in cultured cells, cells were fixed with 4% paraformaldehyde for 20 min, then permeablized with 0.3% Triton X-100-PBS for 5 min twice and blocked in PBS solution containing 5% bovine serum albumin and 0.3% TritonX-100%, followed by an overnight incubation in primary antibodies at 4 °C. After three washes in PBS, cells were incubated with secondary antibody. Cell nuclei were counterstained with DAPI (4′,6-diamidino-2-phenylindole). For immunofluorescence of paraffin-embedded tissue sections, de-paraffinization was performed, followed by the same immunofluorescence protocol described above.

The following antibodies were used: mouse anti-p63 monoclonal antibody, rabbit anti-K5 monoclonal antibody, mouse anti-K10 monoclonal antibody, mouse anti-K14 monoclonal antibody with biotin labelled, mouse anti-K19 monoclonal antibody, (MA1-21871, RM2106S0, MS611P0, MS115B0, MS1902P0, Thermo Fisher Scientific), rabbit anti-PAX6 polyclonal antibody (PRB-278P, Covance), mouse anti-K1 monoclonal antibody (sc-376224, Santa Cruz), Rabbit anti-WNT7A polyclonal antibody, mouse anti-K3/K12 monoclonal antibody, rabbit anti-K12 monoclonal antibody (ab100792, ab68260, ab124975, Abcam), mouse anti-Ki67 monoclonal antibody (550609, BD BioSciences), anti-GFP rabbit monoclonal antibody and anti-GFP mouse monoclonal antibody (G10362, A11120, Invitrogen). The secondary antibodies, AlexaFluor-488- or 568-conjugated anti-mouse or rabbit immunoglobulin-G (IgG) (Invitrogen) were used at a dilution of 1:500. Images were obtained using an Olympus FV1000 confocal microscope.

Quantitative PCR. RNA was isolated using an RNeasy kit (Qiagen) and subjected to on-column DNase digestion. Complementary DNA synthesis was performed using a superscript III reverse transcriptase kit according to the manufacturer's instructions (Invitrogen). qPCR was performed by 40-cycle amplification using gene-specific primers (Extended Data Table 1; top) and a Power SYBR Green PCR Master Mix on a 7500 Real Time PCR System (Applied Biosystems). Measurements were performed in triplicates and normalized to endogenous GAPDH levels. Relative fold change in expression was calculated using the ΔΔCT method (cycle threshold (CT) values < 30). Data are shown as mean ± s.d. based on three replicates.

Genome-wide gene expression microarray and data analysis. Total RNA was isolated from LSCs, SESCs and differentiated CECs from three-dimensional differentiation assay. Gene expression microarray analysis was performed using an Illumina human genome microarray system, with each sample in biological replicate ($n = 2$ per group; Human HT-12 v4 Expression BeadChip; Illumina, San Diego, California). Raw data were deposited into the GEO database under accession number GSE32145. Expression-level data were generated by the Illumina BeadStudio version 3.4.0 and normalized using quartile normalization. Probes whose expression level exceeded a threshold value of 64 in at least one sample were considered detected. The threshold value was found by inspection from the distribution plots of $\log_2$ expression levels. Detected probes were sorted according to their q value, which is the smallest false discovery rate (FDR) at which the probe is called significant. FDR was evaluated using significance analysis of microarrays and its implementation in the official statistical package sam[19]. To avoid false positive calls due to spuriously small variances, the percentile of standard deviation values used for the exchangeability factor $s0$ in the regularized t-statistic was set to 50. We combined the LESC and CEC samples into one group of four samples, and looked for differentially expressed genes between this group and SESCs samples. The top 100 significant genes in this comparison are presented in Extended Data Fig. 2. All genes in this figure are significant at the FDR level of 0.01 or less. A heatmap was created using in-house hierarchical clustering software, and colours qualitatively correspond to fold changes.

RNA-seq and hierarchical cluster analysis. Total RNA was purified by a Picropure RNA isolation kit (Life Technology). RNA-seq was performed as described previously[20]. In brief, 600 ng of total RNA was first converted to cDNA by superscript III first strand synthesis kit with primer Biotin-B-T. The cDNA was purified by NucleoSpin Gel and PCR Clean-Up Kit column (Clontech) to remove free primers and enzyme. Then terminal transferase (NEB) was applied to block the terminal of a cDNA 3′ end. Streptavidin-coaged magnetic beads (Life Technology) were further applied to isolate cDNAs. After RNA degradation by sodium hydroxide, second-strand cDNA was synthesized by random priming with primer A-N8. The second strand cDNA was eluted from beads by heat denaturing. The cDNA was then used as template to construct libraries by amplifies with barcode primers and primer PB. The sequencing was done on Hiseq 2000 system.

Hierarchical cluster analysis was performed with cluster and Java TreeView[21]. The raw data were first filtered using default parameters provided by the program Cluster. The filtered data were further adjusted by log transformation, genes and arrays were centred by median, and then both gene and array were hierarchically clustered with euclidean and average linkage. The hierarchical trees and gene matrix were visualized and generated by Java Treeview.

Lentiviral RNA interference and *PAX6* transduction. Lentiviral shRNAs targeting *PAX6, WNT7A* and *FZD5* genes were either cloned into pLKO.1 plasmid between Age I and EcoR I or purchased directly from Sigma. shRNAs targeting sequences for gene-specific knockdowns were as follows: *PAX6*, CGTCCATCTT TGCTTGGGAAA and AGTTTGAGAGAACCCATTATC; *WNT7A*, CGTGCT CAAGGACAAGTACAA and GCGTTCACCTACGCCATCATT; *FZD5*, CGCG AGCCCTTCGTGCCCATT and TCCTAAGGTTGGCGTTGTAAT. We used a lentiviral pLKO.1-puro Non-Target shRNA control plasmid encoding a shRNA that did not target any known genes from any species as a negative control in all gene knockdown experiments (Sigma).

Lentiviral shRNA particles were prepared according to a previous described protocol[22]. In brief, replication-incompetent lentiviral particles were packaged in 293T cells by co-transfection of shRNA constructs with packaging mix (pCMV-dR8.2 and pCMV-VSVG at a 9:1 ratio). Virus was collected two times at 48 h and 72 h post transfection.

For transduction, *PAX6a* open reading frame (ORF) was PCR amplified from cDNAs purchased from Thermo Scientific (MHS6278-202756612) and inserted into pLenti CMV-GFP Puro vector between BamH1 and BsrG1. *PAX6b* was generated by PCR-mediated point mutation strategy with primers PAX6 InF and PAX6 InR (Extended Data Table 1; top). For GFP labelling, pLenti CMV-GFP Hygro (656-4) purchased from Addgene was used. The lentiviral particles were packaged by co-transfection with packaging plasmids psPax2 and pMD2.G.

For lentiviral infection, cells were infected for 16–20 h with fresh media containing individual virus and polybrene at a final concentration of 8 μg ml^{-1}. The infected cells were further selected by 2 μg ml^{-1} uromycin for 48 h or 200 μg ml^{-1} hygromycin for 72 h.

Western blot analysis and co-immunoprecipitation. For western blot analysis, cells were washed once with PBS and then collected in cell lysis buffer (50 mM Tris-HCl, pH 6.8; 2% SDS; 10% Glycerol; 100 mM DTT). Protein concentration was quantified by Nanodrop and Bromophenol blue was added to a final concentration of 0.1%, then 25 μg of total lysate was fractionated on a 4–12% NUPAGE gel (Life Technology). Proteins were transferred to a nitrocellulose membrane at 100 V for 1 h. The membrane was blocked with 5% milk and probed with relevant antibodies and mouse anti-β-actin monoclonal antibody (A5316, Sigma).

To detect interaction between FZD5 and WNT7A, a 10-cm dish of limbal stem cells at 90% confluence was collected; the cell pellet was resuspended in 700 μl of co-immunoprecipitation (Co-IP) buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 2.5 mM $MgCl_2$, 0.5% NP-40, 1× proteinase inhibitor) and incubated on ice for 20 min, then centrifuged at 13,000 r.p.m. at 4 °C for 20 min. The 600 μl of supernatant were aliquoted into two pre-chilled Eppendorf tubes, 5 μg of rabbit anti-FZD5 monoclonal antibody (#5266, Cell Signaling) or WNT7A antibodies was added to each tube and incubated at 4 °C overnight. Protein A/G magnetic beads (50 μl, Thermo Fisher) were added to each tube, and incubated at 4 °C for 2 h, washed with a Co-IP buffer and eluted in 1× SDS sample buffer (Life Technology) at 70 °C. The input and elutes were fractionated on 4–12% NUPAGE gel and blotted with FZD5 and WNT7A antibodies.

Cell transplantation. All animal studies were performed in full accordance with the Association for Research in Vision and Ophthalmology (ARVO) statement, Use of Animals in Ophthalmic and Vision Research, and approvals were obtained from Institutional Animal Care Committees.

New Zealand white rabbits (2.0 kg to 2.5 kg, male) were used in the study. Rabbits were anaesthetized with intramuscular injection of xylazine hydrochloride (2.5 mg ml^{-1}) and ketamine hydrochloride (37.5 mg ml^{-1}). To create a limbal stem cells deficiency model (Extended Data Fig. 7f), corneal and limbal epithelium was removed by 360-degree conjuntival peritomy and lamellar dissection to remove anterior scleral and corneal stromal tissues, 2 mm posterior from limbus towards the centre of the cornea. This dissection ensured removal of LSC and the entire corneal epithelium. Rabbit GFP-labelled LSCs (5×10^5), $PAX6^+$ SESCs or shPAX6 LSCs cells were mixed with fibrin (25 mg ml^{-1}) and thrombin (25 U per ml) and seeded onto the exposed stromal bed of a recipient cornea and limbal area; the surface was then covered by a human amniotic membrane (Bio-tissue), which is secured with 10.0 VICRYL sutures (Ethicon) (Extended Data Table; bottom). As a negative control, only amniotic membrane was applied to the denuded cornea. Antibiotics (levofloxacin) and steroids (betamethasone) were applied to both eyes immediately after the cell transplant procedures, and were administered three times a day for 2 weeks. Animals were randomly assigned into each experimental group. The investigator who performed cell transplantation was blind to the identity of cells used. Another investigator carried out assessment of the effect of corneal epithelial repair in rabbits and was again blind to the identity of cells used in the transplantation. For analysis, we exclude only animals that died of post-operative complications such as infection, as they did not reach the end point for assessment of cell transplantation effect; this criterion is pre-established.

19. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98,** 5116–5121 (2001).
20. Fox-Walsh, K., Davis-Turak, J., Zhou, Y., Li, H. & Fu, X. D. A multiplex RNA-seq strategy to profile poly(A^+) RNA: application to analysis of transcription response and 3′ end formation. *Genomics* **98,** 266–271 (2011).
21. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA* **95,** 14863–14868 (1998).
22. Xue, Y. *et al.* Direct conversion of fibroblasts to neurons by reprogramming PTB- regulated microRNA circuits. *Cell* **152,** 82–95 (2013).