nature portfolio

Corresponding author(s): Kohji Nishida

Last updated by author(s): Mar 31, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection in this study.
Data analysis	All data analysis was performed using publicly available software including Cell Ranger (v6.1.2), Seurat (v4.3.0), Harmony (v0.1.1), Monocle3, SCENIC (v1.3.1), SoupX (v1.6.2), STAR (v2.7.10b), RSEM (v1.3.3), and scVelo (v0.2.3). No custom code was developed.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The scRNA-seq and bulk RNA-seq datasets generated and analyzed during this study have been deposited in ArrayExpress under accession codes E-MTAB-14019 and E-MTAB-13992.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Not reported. No direct human participants were involved, sex and gender data were not collected or analyzed.
Reporting on race, ethnicity, or other socially relevant groupings	See above.
Population characteristics	See above.
Recruitment	See above.
Ethics oversight	See above.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	scRNA-seq: Sample size was n = 1 for each time point (2, 4, and 8 weeks), as described in the Methods section. scRNA-seq was performed on TP63-positive cells: 6,138 cells at 2 weeks, 6,541 cells at 4 weeks, 8,621 cells at 8 weeks, and 7,877 cells at 8 weeks with EGF treatment. Total cell counts were: 2,721 cells at 2 weeks, 6,177 cells at 4 weeks, and 6,772 cells at 8 weeks. Bulk RNA-seq (Figure 6e): n = 1 per condition. Figure 6b and 6d (qPCR and flow cytometry): n = 7, with statistical analysis described in the figure legend. Immunohistochemistry was conducted using samples from three independent biological replicates (N=3) for the main figures. For certain supplementary experiments, representative images from a single biological replicate (N=1) are shown, as indicated in the figure legends.
Data exclusions	No data were excluded from the analyses.
Replication	Immunostaining results were replicated using two independent hiPSC-derived colonies. For other assays, including RNA-seq and flow cytometry, biological replicates were not performed for all conditions; however, consistent results were obtained across different time points or cell fractions.
Randomization	Randomization was not applicable, as all samples were derived from standardized hiPSC differentiation protocols, and experimental groups were defined based on time points or fluorescence-based sorting.
Blinding	Blinding was not performed. Data collection and analysis were based on objective measurements (e.g., fluorescence intensity, gene expression), and group allocation was determined by marker expression or differentiation time points.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

The following antibodies were used in this study:
Mouse monoclonal anti-TP63(Santa Cruz Biotechnology,Cat# sc-8431; RRID: AB_628091; clone 4A4)
Rabbit polyclonal anti-PAX6(BioLegend,Cat# 901301, RRID:AB_2565003)
Rabbit polyclonal anti-KRT12 (Abcam,Cat# ab185627, RRID:AB_2889825)
Mouse monoclonal anti-TKT(Santa Cruz Biotechnology,Cat# sc-390179, RRID:AB_2925185; clone H-7)
Mouse monoclonal anti-KRT15(Santa Cruz Biotechnology,Cat# sc-47697, RRID:AB_627847; cloneLHK15)
Mouse monoclonal anti-KRT13(Abcam, Cat# ab16112; RRID: AB_302267; clone AE8)
Mouse monoclonal anti-MUC4(Abcam, Cat# ab52263; RRID: AB_881163; clone 8G-7)
Mouse monoclonal anti-MUC5AC(Santa Cruz Biotechnology Cat# sc-33667; RRID: AB_627973; clone CLH2)
Mouse monoclonal anti-KRT10(Santa Cruz Biotechnology, Cat# sc-52318, RRID:AB_629836; DE-K10)
Rabbit monoclonal anti-TFAP2A(Cell Signaling Technology, Cat# 3215, RRID:AB_2227429; clone C83E10)
Rabbit monoclonal anti-NANOG (Cell Signaling Technology, Cat# 4903, RRID:AB_10559205; clone D73G4)
Mouse monoclonal anti-SSEA4 (Cell Signaling Technology, Cat# 4755, RRID:AB_1264259; clone MC813)
Rabbit monoclonal anti-OCT4A (Cell Signaling Technology, Cat# 2840, RRID:AB_2167691; clone C30A3)
Mouse monoclonal anti-TRA-1-60(S) (Cell Signaling Technology, Cat# 4746, RRID:AB_2119059; clone TRA-1-60(S))
Goat polyclonal anti-TFAP2C(R and D Systems, Cat# AF5059, RRID:AB_2255891a)
Mouse monoclonal anti-MSX2(R and D Systems, Cat# MAB7917, clone 786607, AB_3096365)
Goat polyclonal anti-GATA3(R and D Systems Cat# AF2605, RRID:AB_2108571)
Rabbit monoclonal anti-CDH1(Abcam, Cat# ab40772, RRID:AB_731493; EP700Y)
Mouse monoclonal anti-KRT8(Santa Cruz Biotechnology,Cat# sc-8020, RRID:AB_627857; clone C51)
Rabbit polyclonal anti-FOXG1(Abcam, ab18259, RRID:AB_732415)
Rabbit polyclonal anti-TUBB3(Sigma-Aldrich, Cat# T2200, RRID:AB_262133)
Mouse monoclonal anti-VSX2 (Santa Cruz Biotechnology, Cat# sc365519; RRID:AB_10842442; clone E-12)
Rabbit monoclonal anti-MITF(Cell Signaling Technology Cat# 12590, RRID:AB_2616024; clone D5G7V)
Mouse monoclonal anti-RAX(Santa Cruz Biotechnology, Cat# sc-271889, RRID:AB_10708730; clone G-12)
Mouse monoclonal anti-PAX2 (Santa Cruz Biotechnology, Cat# sc-130387, RRID:AB_2236656; clone 60-P)
Mouse monoclonal anti-EGFP(Santa Cruz Biotechnology, Cat# sc-9996, RRID:AB_627695; clone B-2)
Rabbit polyclonal anti-tdTomato(Rockland, Cat# 600-401-379S, RRID:AB_11182807)

Validation

For each antibody, validation information has been provided by the manufacturer on their website.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

The study used the human iPSC line 201B7. This cell line is commercially available and was obtained from the Center for iPS Cell Research and Application (CiRA), Kyoto University. It was used as the parental line to generate PAX6-EGFP/TP63-tdTomato dual reporter iPSCs.

Authentication

The parental iPSC line 201B7 was obtained from a reputable source and has been previously characterized. The reporter iPSC line generated in this study was not further authenticated beyond the validation of successful gene knock-in.

Mycoplasma contamination

We routinely test for mycoplasma using MycoAlert (LONZA, https://www.lonzabio.jp/catalog/789/).

Commonly misidentified lines
(See ICLAC register)

None of the cell lines used in this study appear in the ICLAC Register of Misidentified Cell Lines.

Plants

Seed stocks — *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

Novel plant genotypes — *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

Authentication — *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

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

Instrument — Cell Sorter SH800S

Software — Data acquisition and analysis were performed using Sony Cell Sorter SH800S software.

Cell population abundance — Sorted populations were used to generate epithelial sheets with clearly distinct phenotypes, supporting the validity of the gating strategy.

Gating strategy — Two gating strategies were used depending on the experimental context: in some cases, tdTomato-positive cells were further subdivided based on EGFP fluorescence intensity, and in other cases, the entire tdTomato-positive population was sorted regardless of EGFP expression.
To exclude autofluorescent cells, control iPSCs without reporter knock-in were used to define the baseline fluorescence. A 99th percentile cutoff was applied to the control population, and cells in the experimental group exceeding this threshold were defined as reporter-positive.Additionally, FSC/SSC-based gating was used 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.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.