SI Appendix

A trajectory for human pluripotent stem cells toward the early otic lineage

Ealy M, Ellwanger DC, Kosaric N, Stapper AP, and Heller S

Contents:

SI Materials and Methods

Cell culture

H9 human embryonic stem cells (hESC, Wicell Research Institute) were maintained on γ-irradiated mouse embryonic fibroblast (CF-1 strain) feeder cells in KnockOut DMEM/F12 (Invitrogen) supplemented with 20% KnockOut serum replacement (Invitrogen), 1% non-essential amino acids (Invitrogen), 1 mM L-glutamine (Invitrogen), 0.1 mM EmbryoMax 2-mercaptoethanol (Millipore), 100 ug/ml penicillin-streptomycin (Invitrogen), and 10 ng/ml recombinant human basic FGF (FGF2, R&D Systems). All experiments were conducted with cell passages 28-40.

The human induced pluripotent stem cell line SUCI-0002 (referred to as iPSC throughout) was derived from a study participant with no reported hereditary hearing loss. Fibroblasts were isolated with a skin punch biopsy and maintained in Alpha MEM Earle's Salts (Irvine), Change Basal media (Irvine), 6% Chang C frozen supplement (Irvine), 20% FBS, 1mM L-glutamine, 100 ug/ml gentamicin, and 125 ng/ml Fungizone (Cellgro). Fibroblasts were reprogrammed using lentivirus-containing EF1a-hSTEMCCA-loxp (a gift from Dr. G. Mostoslavsky, Boston University). Reprogrammed colonies were identified with StainAlive TRA-1-60 antibody (DyLight 488) (Stemgent) and propagated on irradiated mouse embryonic fibroblasts. Pluripotency was confirmed by immunostaining for OCT4, NANOG, SSEA-1, SSEA-4, and TRA-1-81; and qRT-PCR for *DMNT3B*, *FGF4*, *LIN28A*, *NANOG*, *POU5F1*, *PRMT5*, *SALL4*, and *SOX2* (Fig. S12). Stanford University's Institutional Review Board approved all

methods for human subjects research. Stanford University's Stem Cell Research Oversight Panel approved all human stem cell research. Stanford University's Administrative Panel on Biosafety approved all work involving biohazardous agents.

For differentiation, H9 hESCs or iPSCs were dissociated to single cells using 1x Accutase cell detachment solution (EMD Millipore) and plated on matrigel (Corning) coated plates at a density of $5.26x10^{-4}$ cells/cm^2 in mouse embryonic fibroblast conditioned media supplemented with 10 ng/ml FGF2 and 10 μM Y-27632 (Enzo Life Sciences). FGF2 was used because it activates all known FGF receptors, thereby circumventing uncertainties from not knowing which specific FGFs are essential for human otic induction (see also (1)). Media was replaced with induction media once cells were 80% confluent, typically 2 days after plating single cells. Induction media consisted of 15% KnockOut serum replacement media supplemented with the below factors as described in the results. Guidance towards mesendoderm and intermediate mesoderm was conducted as previously described (2, 3).

Final concentrations for media supplements are as follows: 10 μM SB431542 (TGFß inhibitor, Selleck Bio), 1 μM FH535 (WNT inhibitor, Millipore), 20 ng/ml recombinant human BMP4 (R&D Systems), 1 μM LDN193189 (BMP inhibitor, Selleck Bio), 25 ng/ml FGF2 (R&D Systems), 25 ng/ml heparin sulfate (Sigma), 10 μM retinoic acid (Sigma), 3 μM CHIR99021 (WNT agonist, R&D Systems), 100 ng/ml recombinant human/mouse/rat Activin A (ActA, R&D Systems), and 1 μM purmorphamine (Shh agonist, Tocris Bioscience).

Gene expression analysis of bulk cultured cells

RNA was isolated from cells using the Qiagen RNeasy Plus Micro kit. 250ng of total RNA was used for cDNA synthesis using the Applied Biosystem's High Capacity Reverse Transcription Kit. cDNA was diluted 10-fold in DNA suspension buffer (Teknova) and used for the Fluidigm pre-amplification step for Dynamic Array gene expression analysis: cDNA was pre-amplified for 14 cycles with 500 nM DELTAgene pooled primer mix using 2x Taqman PreAmp Master Mix (Invitrogen), followed by Exo1 treatment (NEB). 5-fold diluted Exo1 treated pre-amplified cDNA was used for loading the 96.96 Dynamic Array chip on the Fluidigm Biomark HD. See Table S1 for a comprehensive list of primers used for gene expression analysis. Statistical significance for relative fold change values was determined using Student's t-test.

Quantitative RT-PCR primer validation

To validate primers, a 2-fold dilution series spanning 15 different concentrations was performed on bulk RNA from presumptive otic progenitors from a hESC embryoid body-based otic induction protocol (4). Six technical replicates were performed for each dilution to determine the threshold for reliable amplicon detection based on melting temperature curves for each primer pair. Primer efficiencies and R^2 values were determined based on the standard curve of the dilution series (Fig. S13). The limit of detection (LoD) value for each assay was determined by taking the mean C_t value for the most dilute sample in which positive amplification plots were detected in all six replicates with a standard

deviation <1. The overall LoD for the panel of assays was determined by taking the median of the LoDs for each assay, equal to a C_t of 22.

Single cell qRT-PCR

Single cells from the monolayer cell culture experiments were collected using the Fluidigm C1 Single Cell Autoprep system, following the manufacturer's protocol. Cultures were dissociated to single cells using Accutase treatment for 5-30 minutes, depending on culture condition, and passed through a 40 μM cell strainer. Approximately 500 cells were loaded onto the C1 10-17 μm chip (Fluidigm). Cell lysis, cDNA synthesis, and pre-amplification were performed on the C1 chip using validated amplicon-specific DELTAgene Assays (Fluidigm), and products harvested to run on the 96.96 Dynamic Array chip on the Fluidigm Biomark HD. In total, 856 cells were used to perform qRT-PCR analysis.

Single cell gene expression data processing

The raw qRT-PCR data was processed as matrix $\boldsymbol{X} \rightarrow \mathbb{R}^{l \times m \times n}$ listing expression values of n genes (features) captured in m cells (samples) under l biological conditions. The median expression intensity of each cell was normalized to the global median ϕ $(\phi = \text{median}\big(\text{median}_i(X_{.i.})\big),\ i = 1 \ldots m)$. Resulting C_t values < 0 were classified absent and set to $^{N}/_{A}$. 192 samples with no detectable reference gene expression (*ACTB* or *GAPDH*) or expressing < 2% of assay genes were removed from the analysis. Additionally, 12 samples were classified as doublets based on high reference gene expression levels above twice the standard deviation of the median gene expression. Removal of 204

samples after these quality control steps left 652 cells for data processing. Log_2 gene expression values (Log2Ex) were determined by the C_t value offset to the LoD, i.e. $\Delta\boldsymbol{X} = X_{hij} - \mathrm{LoD}, \text{with } \mathrm{LoD} = 22$. To enable a better appreciation of the true biological cell-to-cell variation, we minimize technical noise in Log2Ex values by adjusting to an invariant endogenous normalization factor. For this purpose, we computed the geNorm M value (5) for 11 genes expressed in at least 95% of cells in the sample panel to estimate the stability of assayed reference genes. We found *ACTB* to be the most stably expressed (Fig. S11*A*), whereas *GAPDH* showed reasonable variance $(M = 1.9)$. Each sample feature expression was normalized to the *ACTB* expression value, respectively. Log2Ex values were scaled to obtain a nonnegative $\boldsymbol{X}$ with a minimum value of $\varepsilon = 10^{-3}$, i.e. $\boldsymbol{X} = X_{hij} + |\min \boldsymbol{X}| + \varepsilon$. To account for non-random non-detected expression signals (drop-outs), i.e. measurements failing to produce a minimum amount of signal due to intrinsic cellular conditions (e.g. low RNA concentration) or technical noise (e.g. primer dimer), we implemented an iterative *k*-nearest neighbor based model. For each non-detect ($X_{hij} = {}^{\mathrm{N}}/_{\mathrm{A}}$), the respective gene expression $X_{h.j}$ found in the *k* most similar cells was analyzed. Cellular similarity was defined by the Euclidean metric ($d_{X_{h1j}X_{h2j}} = \sqrt{\sum_{j=1}^{3}(X_{h1j} - X_{h2j})^2}$) in the reduced three-dimensional space obtained by independent component analysis. Non-random absent values $(P < 0.05)$ were identified using the probability of observing x non-detects in the *k*-neighborhood by chance by means of the Binomial distribution $\mathfrak{B}(x; k, p) = \binom{k}{x} p^x (1-p)^{k-x}$. Missing values were imputed as follows: stochastic

absent expression signals were set to 0; dropped-out values (4.48%) were estimated fitting a generalized additive model (6) of gene expression as a cubic smoothed function of cell ordering of the *k*-neighborhood in the first independent component (Fig. S11*B*). The procedure was iterated until convergence, i.e. all values $X_{hij} \in \mathbb{R}_0^+$ (Fig. S11*B'*). Finally, we filtered 641 cells having less than 15% predicted drop-outs in 90 genes (Fig. S11*C*-*M*).

Multivariate data analysis

The gene expression profile of each cell was encoded by a highly dimensional vector in $\mathbb{R}^n$ (with $n = 90$ genes). Reducing the dimensionality can lower the variability in pairwise cell-to-cell distances, reduces redundant information from correlated features and, ultimately, simplifies the mining and interpretation of patterns in the data. We assumed that observed expression values, $\boldsymbol{X} \rightarrow \mathbb{R}^{l \times m \times n}$, are determined by a mixture of latent abstract sources (e.g. cell differentiation over time). As such, we applied the fastICA algorithm (7) to separate the n mixed signals into three independent components, $\boldsymbol{Y} \rightarrow \mathbb{R}^{l \times m \times 3}$. We performed multidimensional scaling using Barnes-Hut t-distributed stochastic neighbor embedding (tSNE) (8) to visualize the three-dimensional ordination of the cells on a generic two-dimensional XY plot. This allowed us to project quantitative information of individual gene expression levels across cells by means of fitted smooth contours using penalized splines (9).

Cells from other germ layer control cultures were compared to pNNE cells. Overall differences in gene expression were determined between cell

populations. Results of each comparison for all genes analyzed are presented in Table S2.

Cell subpopulations were determined using hierarchical clustering with Ward's minimum variance criterion (10) and Euclidean distance. To estimate the number of clusters in the data, we used the gap statistic (11). We calculated the difference of the expected within-cluster dispersion under a null reference distribution (500 Monte Carlo bootstrap samples) and the actual dispersion, i.e. a higher gap value denotes a superior goodness of clustering. The minimal number of clusters k was defined according to Tibshirani *et al.* (2001) (11) such that $\mathrm{gap}(k)$ is at most one standard deviation σ distant from the first local maximum, i.e. $\mathrm{gap}(k) \geq \mathrm{gap}(k+1) - \sigma(k+1)$.

The heterogeneity within each cell population was estimated using the sum of squares, i.e. $SS_h = \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n}(X_{hij} - \bar{X}_{h.j})^2$, where $\bar{X}_{h.j}$ is the average expression of gene j in cell population h of size m.

Temporal trajectory reconstruction

We aimed to reconstruct and compare the induced H9 and iPSC linages based on single-cell transcriptome profiles of our gene panel. First, we applied a type III two-factorial analysis of variance (ANOVA, Fig. S8*A*) to assess the effect of the independent variables *stage* and *cell line* on the gene expression profile. We removed genes of which the factor *cell line* masked the temporal variance, i.e. cell line specific genes (false-discovery rate corrected *F*-test < 10^{-3}). Then, for each cell line we determined those genes exhibiting a significant fraction of

variance explained by the manifest variable *stage* using a type III one-way ANOVA ($n_{H9} = 33$, $n_{iPSC} = 32$).

To order the cell samples (day 0, day 6 putative non-neural ectoderm, day 12, and day 18 for H9 ESCs and iPSCs) by progress, we constructed an undirected complete graph $G = (V, E)$, weighted by the Euclidean distance between cells in the three-dimensional independent component space. The weight of all edges connecting a pair of cells collected at non-adjacent time points was set to $w_{xy} = \infty$. We applied Prim's algorithm (12) to find the weighted shortest connection subgraph, i.e. the minimum spanning tree (MST), in G. The backbone of the trajectory was defined as the longest shortest path in the MST. We passed those parameters to the Monocle algorithm (13), which computed the optimal pseudo-temporal ordering of the cells and the identification of branches along the trajectory by structuring the data as a PQ tree.

Gene expression dynamics over pseudo-time were calculated by fitting a smoothed Gaussian likelihood based regression model using generalized additive models provided through the VGAM package (14). To correlate gene expression profiles in temporal terms, we standardized the feature variables to *Z*-scores. As the trajectories are of different length (i.e. varying number of collected cells), we partitioned cells along the trajectory around 10 medoids using the PAM algorithm (15) (Fig. S9). This enabled us to compute the Pearson's correlation coefficient of each gene between cell lines by using the average cluster expression intensity.

Estimating transcriptional resemblance to native mouse E10.5 otic cells

To quantify the similarity of induced H9 hESC and iPSC transcriptional profiles with those of known otic cells, we obtained the qPCR Log2Ex values matrix $\boldsymbol{X} \rightarrow \mathbb{R}^{m \times n}$ of $n = 96$ genes and $m = 382$ cells from the otocyst and the immediate neighboring neuroblast population of E10.5 mouse embryos from Durruthy-Durruthy *et al.* (2014) (16). To control for the effect of cofounding batch or species covariates, we processed the Log2Ex values analogously to the samples measured in this study (Supplementary Methods: Single cell gene expression data processing). We identified 37 homologous genes overlapping between both assays. The reference gene (*ACTB*) as well as 13 genes exhibiting a significant fraction of cell line specific expression variance (Fig. S8*A*) were removed from the subsequent analysis. To validate the capability of this gene subset to discriminate otic cells from neuroblasts (labeled non-otic for this purpose), we conducted hierarchical clustering on the expression panel following dimension reduction into three independent components. We quantified the concordance between the observed clustering C and the otic / non-otic cell classification $C^{'}$ proposed in the original work by means of the Jaccard index, i.e. $J = \left| C \cap C^{'} \right| / \left| C \cup C^{'} \right|$ with $J \in [0, 1]$. To score the otic resemblance of an individual cell, we computed the reciprocal for the ratio between the average Euclidean distance to its 10-nearest neighbors in the otic reference population d_i and the expected distance to by chance $d_i^{'}$ to a null reference population, i.e. $s_i = (d_i / d_i^{'})^{-1}$. The 272 reference otic cells were selected according to Durruthy-

Durruthy *et al.* (2014) (16). We generated 500 null reference populations by permuting the expression matrix of the otic population, $\boldsymbol{X}_{\mathbf{otic}} \rightarrow \mathbb{R}^{272 \times 23}$.

Immunocytochemistry

Cultures were fixed in 4% PFA for 15 minutes at room temperature, blocked for 1 hour in 1% BSA, 0.2% Triton-X. Primary antibodies were incubated overnight at 4°C. The following primary antibodies were used: mouse anti-AP-2α 1:50 (DSHB), rabbit anti-Gata3 1:400 (Cell Signaling), rabbit anti-Six1 1:200 (Sigma), rabbit anti-Pax2 1:200 (Covance), mouse anti-Pax6 1:100 (DSHB), mouse anti-Pax3 1:100 (DSHB), rabbit anti-GATA6 1:100 (Santa Cruz), and rabbit anti-Brachyury 1:100 (Santa Cruz). Species-specific Alexa-conjugated secondary antibodies diluted 1:500 in 0.5% BSA, 0.2% Triton-X were incubated for 1 hour at room temperature. Nuclei were stained with DAPI. Confocal microscopic imaging was performed with a Zeiss LSM 700 microscope. Cell counts were performed using single channel data in ImageJ. Data are presented as mean ± standard deviation (SD); *n* represents number of independent experiments.

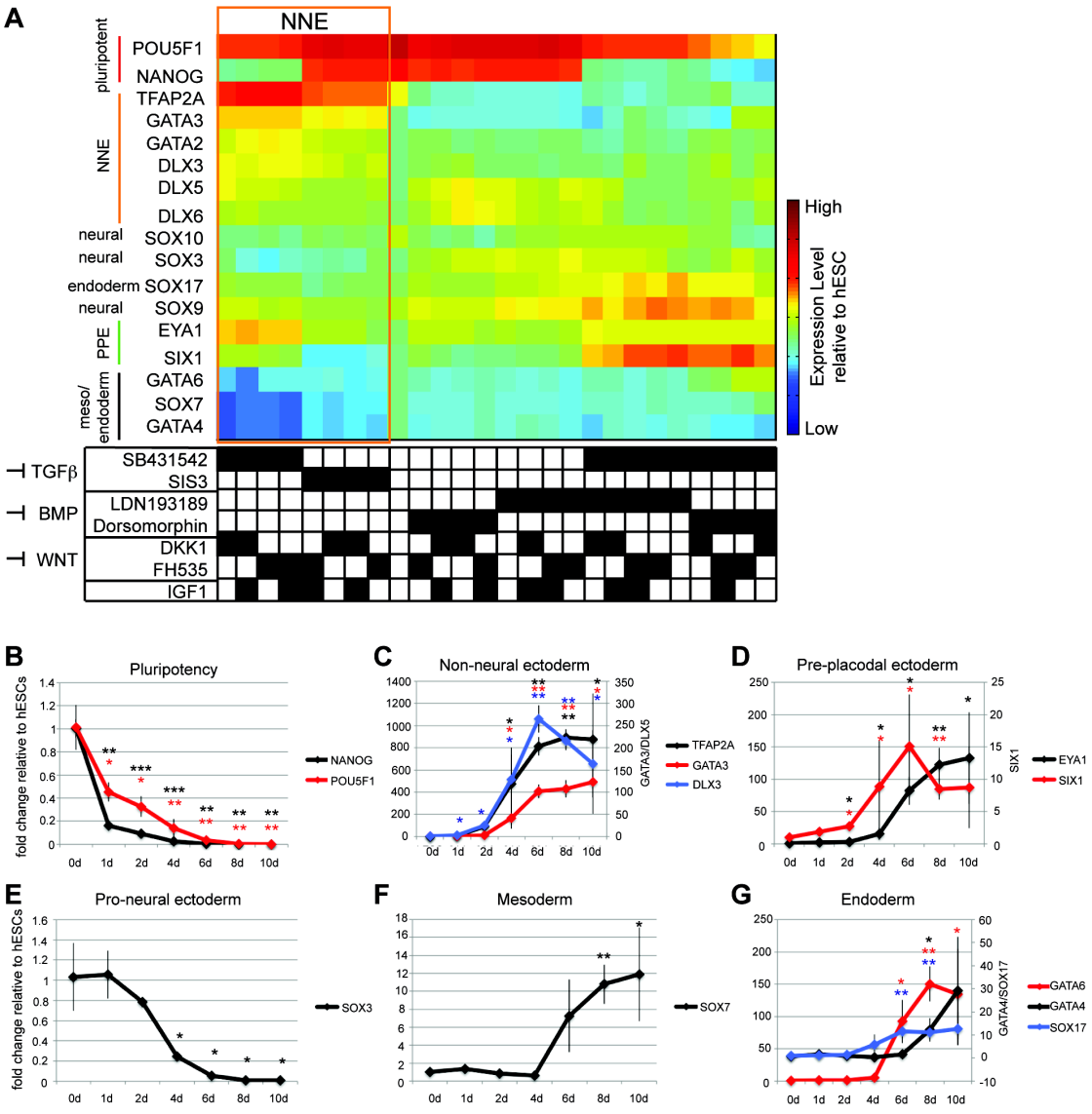


Fig. S1. Identification of conditions leading to NNE marker upregulation. (*A*) Screen for conditions that led to up-regulation of non-neural ectoderm markers in monolayer-grown hESCs maintained for 6 days in culture. TGFß, BMP, and WNT inhibitors, as well as IGF1 were tested in different combinations as indicated. Mean gene expression values relative to hESCs are represented in the heat map for 26 different combinations tested in triplicate. Treatment with SB431452

showed the most prominent down-regulation of pluripotency markers *NANOG* and *POU5F1.* Up-regulation of NNE markers *TFAP2A*, *GATA3*, *GATA2*, and *DLX3* was most consistently observed in cultures that received WNT inhibitor such as FH535 or DKK1. BMP inhibition conducted either isochronally with WNT and TGFß inhibition or simultaneously with inhibition of either WNT or TGFß signaling did not lead to NNE marker gene expression. Black boxes indicate presence of factor in each culture condition. Presumptive NNE-promoting conditions requiring TGFß and WNT inhibition and absence of BMP blockers are marked with a box. (*B*-*G*) Time course of changes in marker gene expression determined with quantitative RT-PCR after 1d, 2d, 4d, 6d, 8d, and 10d in culture in comparison to hESCs at day 0, normalized to *GAPDH*. Cultures were treated with SB431452 and FH535. Secondary ordinates are shown in (*C*, *D*, and *G*) and are indicated on the right side of the corresponding plots. (*B*) Pluripotency markers progressively became downregulated with increasing culture times. (*C*) Significant upregulation of multiple NNE markers was observed already after 4d and stabilized at 6d. Longer culture periods (8d and 10d) did not result in higher expression of NNE markers and expression of some markers such as *DLX3* declined. (*D*) PPE markers are detectable from 4d onward. (*E*) Proneural ectoderm markers were initially expressed but significantly declined from 4d onward. (*F*) Mesoderm and endoderm (*G*) marker genes became detectable first at d6 and increased further at d8 and d10.

Shown are mean fold changes ± SD; n = 3. Unpaired Student's t-test was conducted to indicate * $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$ in comparison to hESC/day 0.

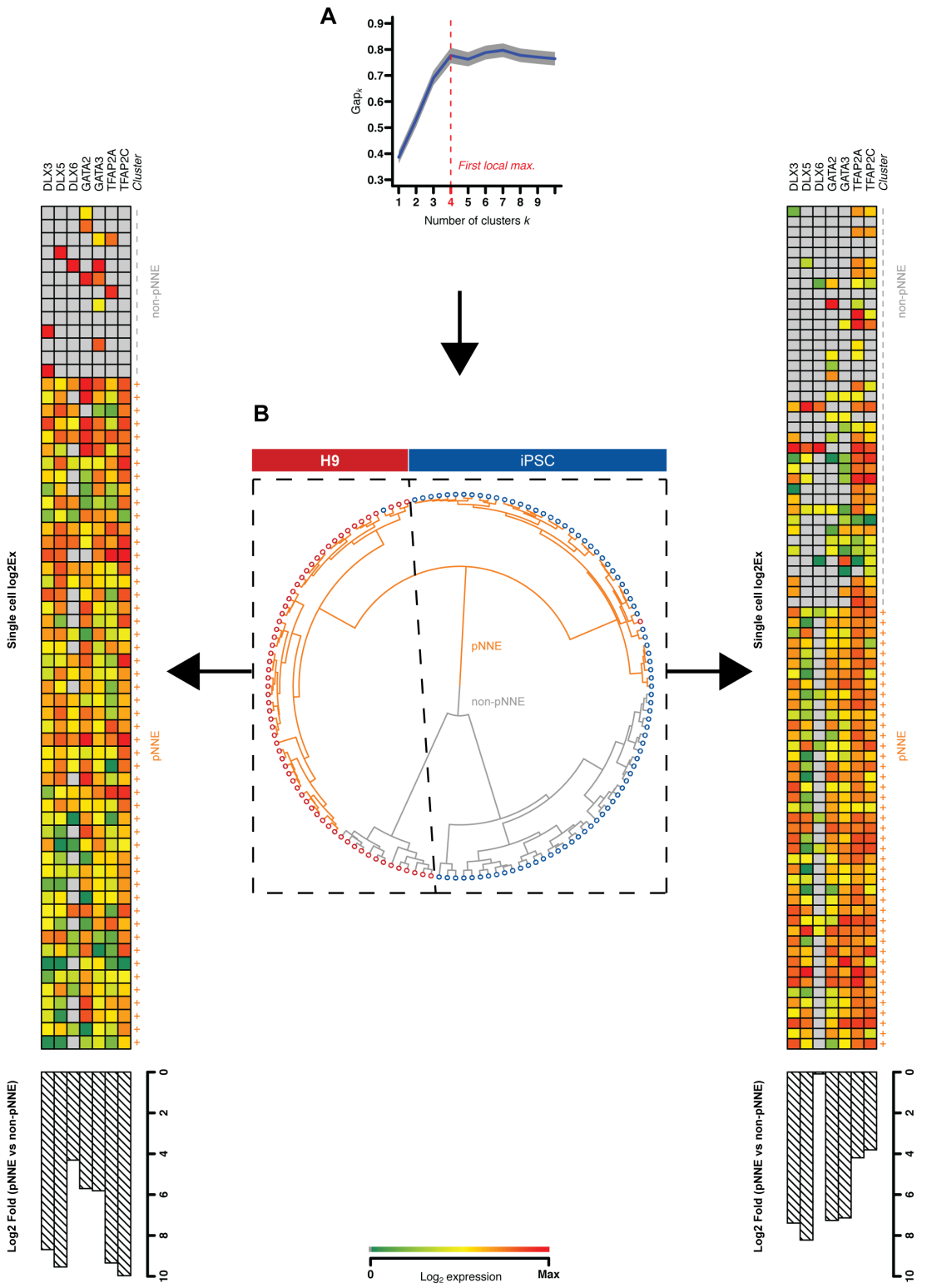


Fig. S2. Cluster analysis of cells 6 days post-induction. *(A)* The gap statistic indicates 4 clusters in this population, i.e. the least *k* for which the within-cluster dispersion is lowest compared to a reference null distribution. *(B)* Hierarchical

clustering result. The clustering dendrogram implies that the cell population intrinsically splits in two major clusters: presumptive non-neural ectoderm (pNNE) cells and non-pNNE cells. Both groups can be further separated by cell line (H9 hESC and iPSC) forming *k* = 4 cohesive clusters. Differential gene expression analysis (left and right panel; for H9 hESCs and iPSCs, respectively) exhibits up-regulation of 6-7 NNE marker genes in the designated pNNE population within both cell lines.

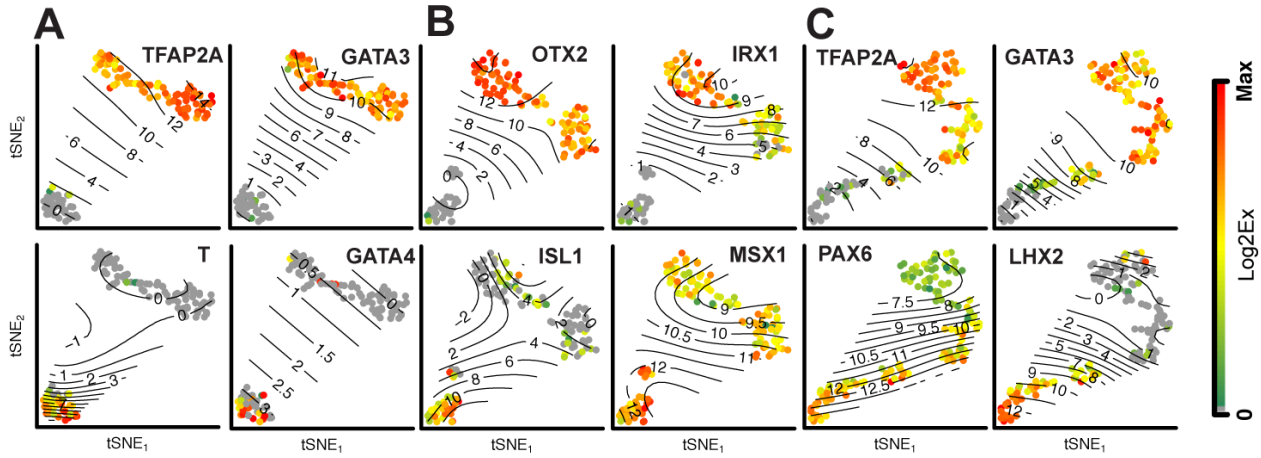


Fig. S3. Surface plots of single cell expression values of example marker genes for pNNE and mesendoderm *(A)*, intermediate mesoderm *(B)*, and neural ectoderm *(C)*. Colorbar values represent dynamic range of Log2Ex values for each gene.

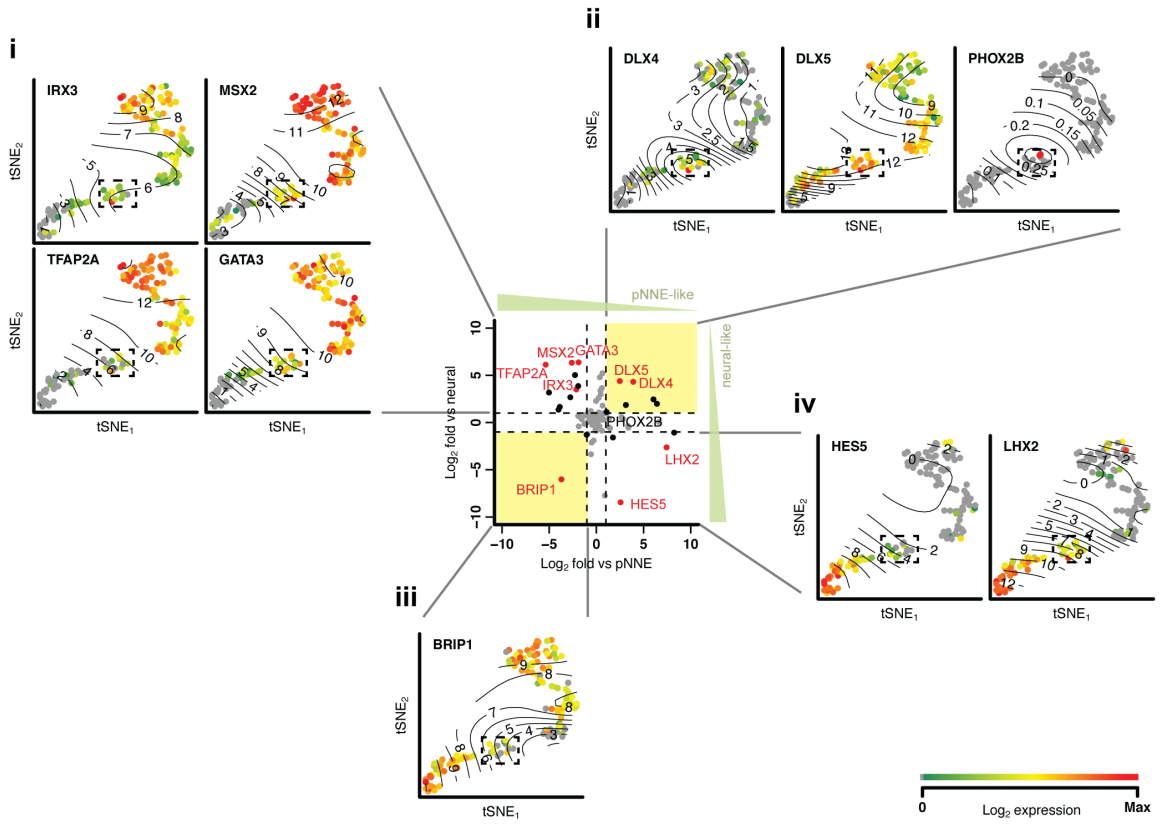


Fig. S4 Identification of putative neural crest cells. The tSNE visualization indicates a minor cell population located between neural and pNNE cells in terms of transcriptional similarity (the corresponding population is tagged with a dotted box in the surface plots). Shown in the center is the differential gene expression analysis compared to neural (Y axis) and pNNE cells (X axis). Genes exhibiting significant differences between cell ensembles (Holm corrected Mann-Whitney test $P < 0.01$) are highlighted in red and their cell-specific expression profile is shown, respectively (i – iv). Besides their expression of neural cell-like and pNNE cell-like genes (bottom right and top left panel), those cells have a higher expression of the known neural crest genes *DLX4* (17), *DLX5* (18), and *PHOX2B* (top right). The latter gene ($P < 0.05$) was described inevitable for neural crest development (19) and is exclusively found in this subpopulation ($n = 2$, Log2Ex = 4.02).

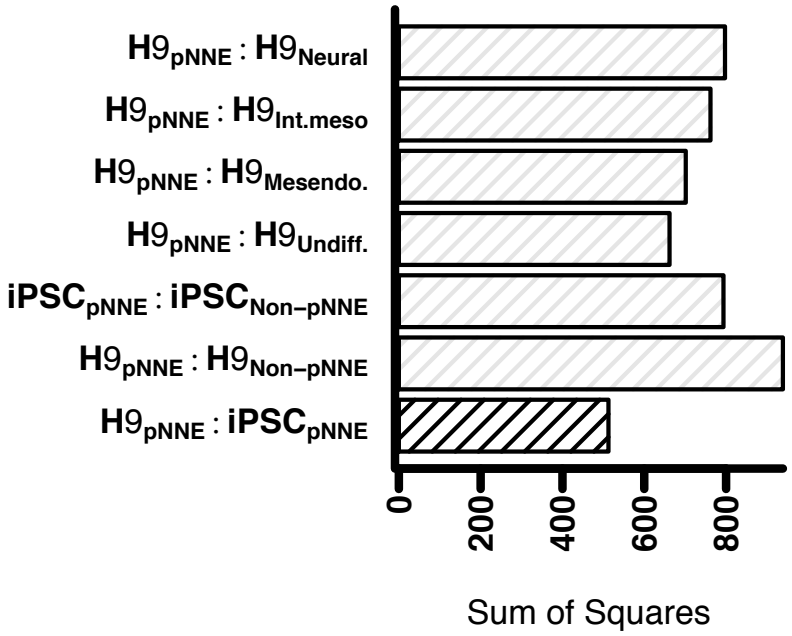


Fig. S5. Sum of squares comparison of variability between culture conditions shown in Fig. 2 and Fig. 3. Comparison of pNNE generated from H9 ESCs and iPSCs ($H9_{pNNE}$:$iPSC_{pNNE}$) showed lower heterogeneity (average distance from the population mean) than any other cell type compared, including the non-pNNE cells, undifferentiated H9 hESCs, mesendoderm, intermediate mesoderm, and neural cells.

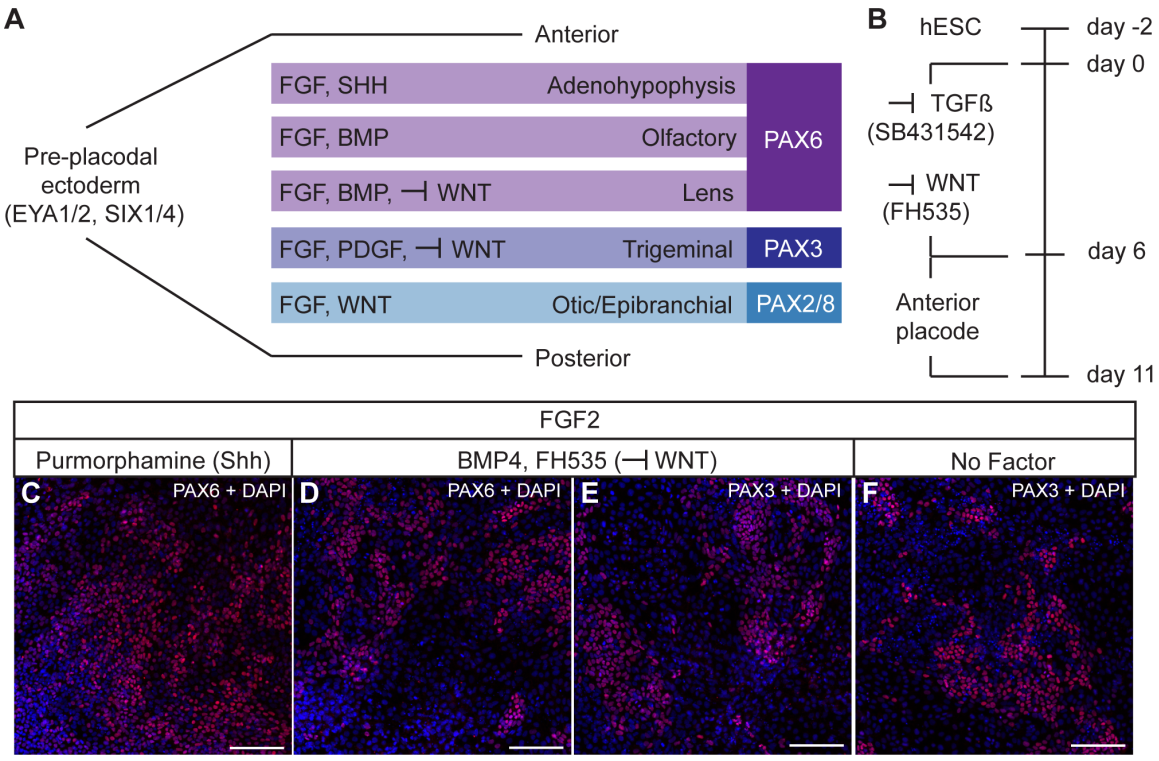


Fig. S6. pNNE is competent to generate anterior placode fates. *(A)* Signaling events required for cranial placode generation from pre-placodal ectoderm. *(B)* Schematic for generation of anterior placodes. *(C)* PAX6 expression in pNNE cells treated for 5 days with FGF2 and Shh agonist purmorphamine. *(D-E)* PAX6 *(D)* and PAX3 *(E)* expression in pNNE cells treated for 5 days with FGF2, BMP4 and WNT antagonist FH535. *(F)* PAX3 expression in pNNE treated for 5 days with FGF2. Scale bars: 100 µm

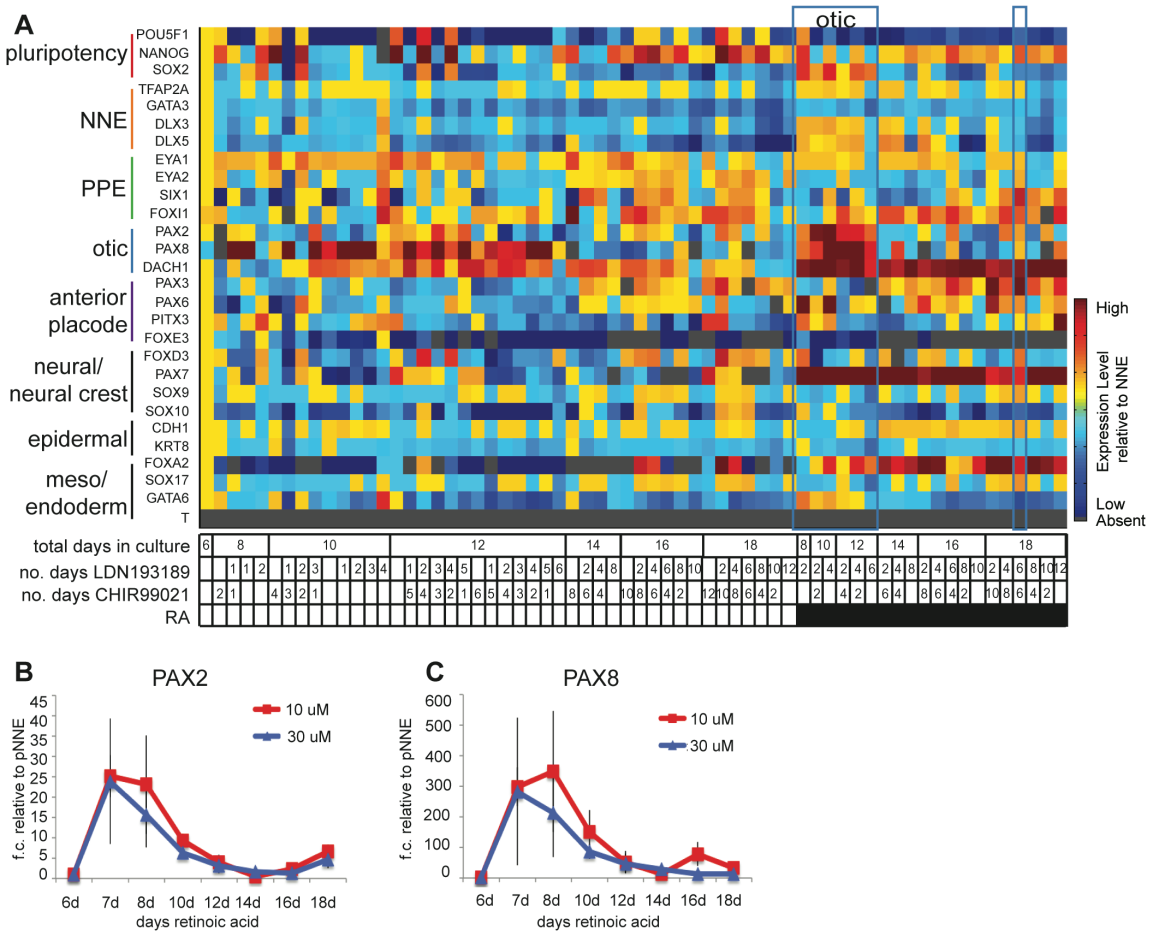
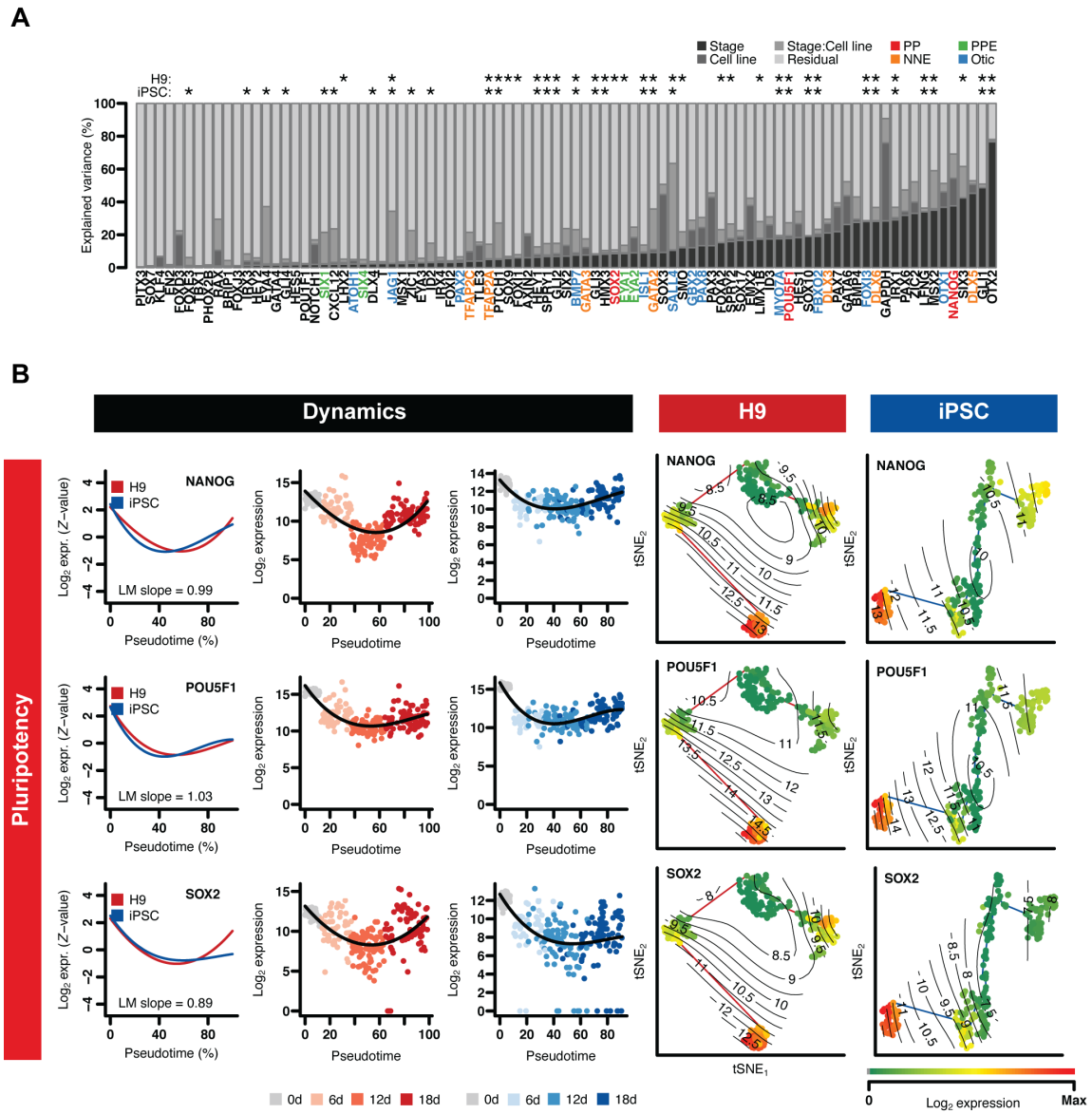
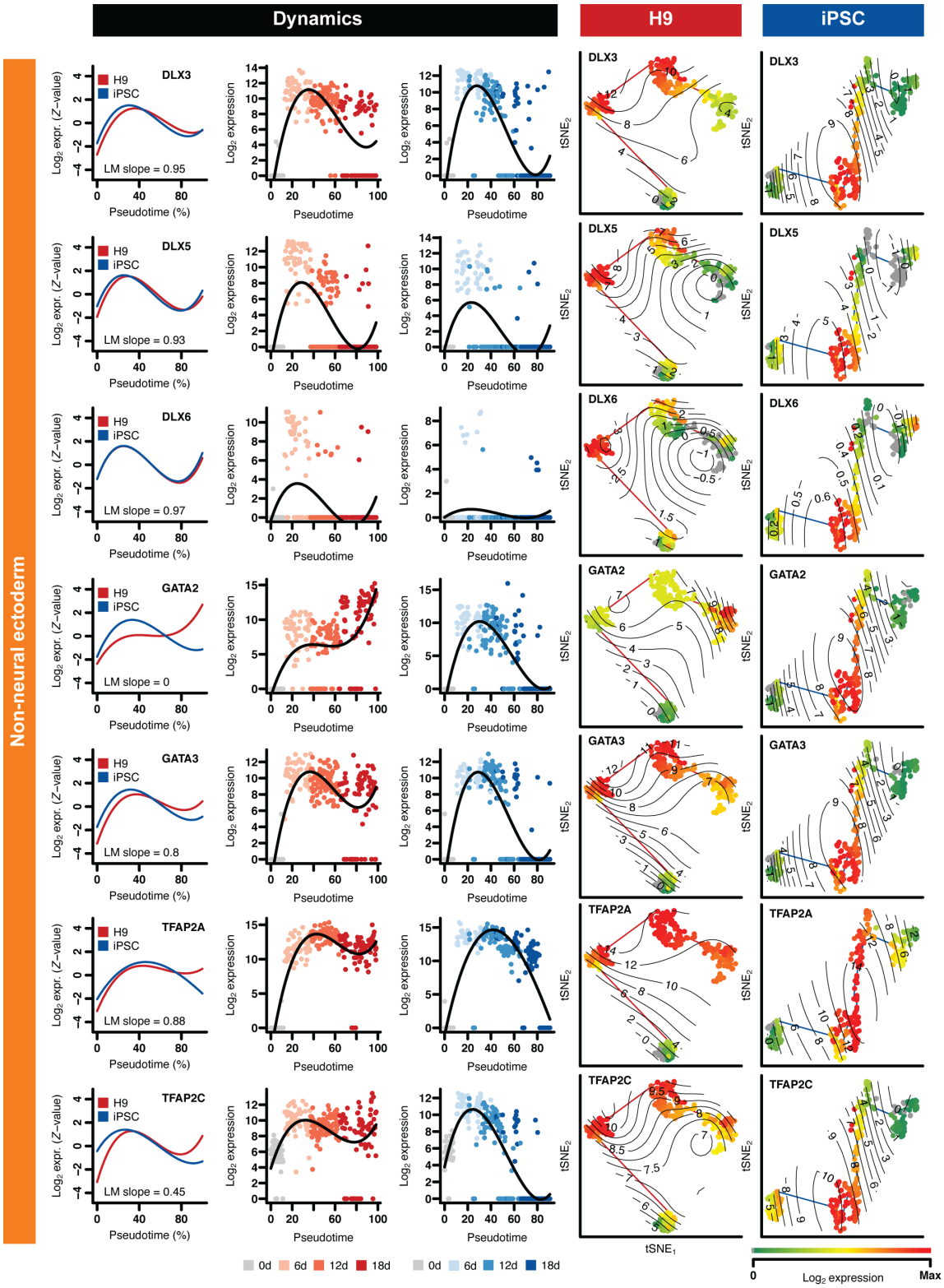
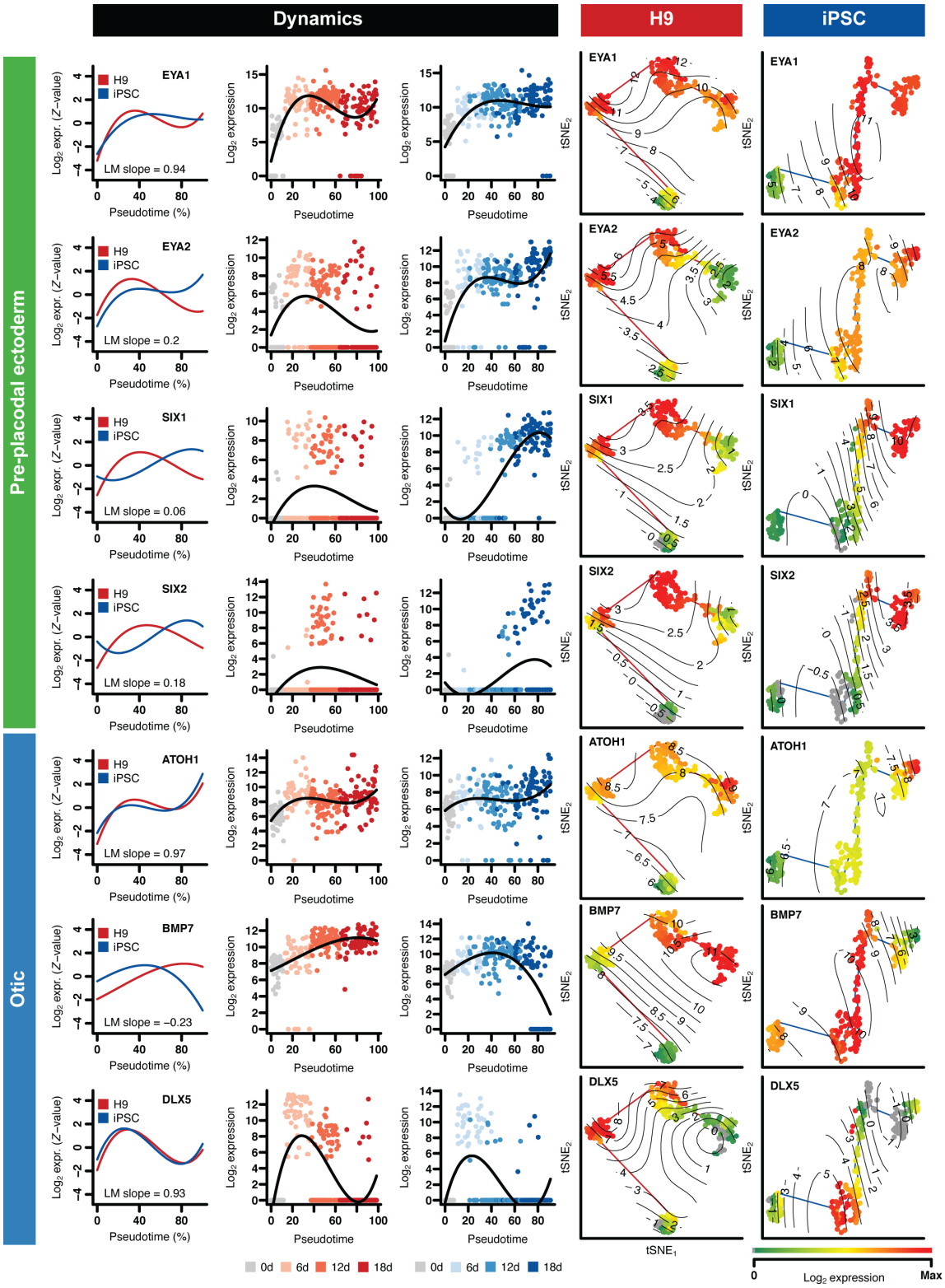
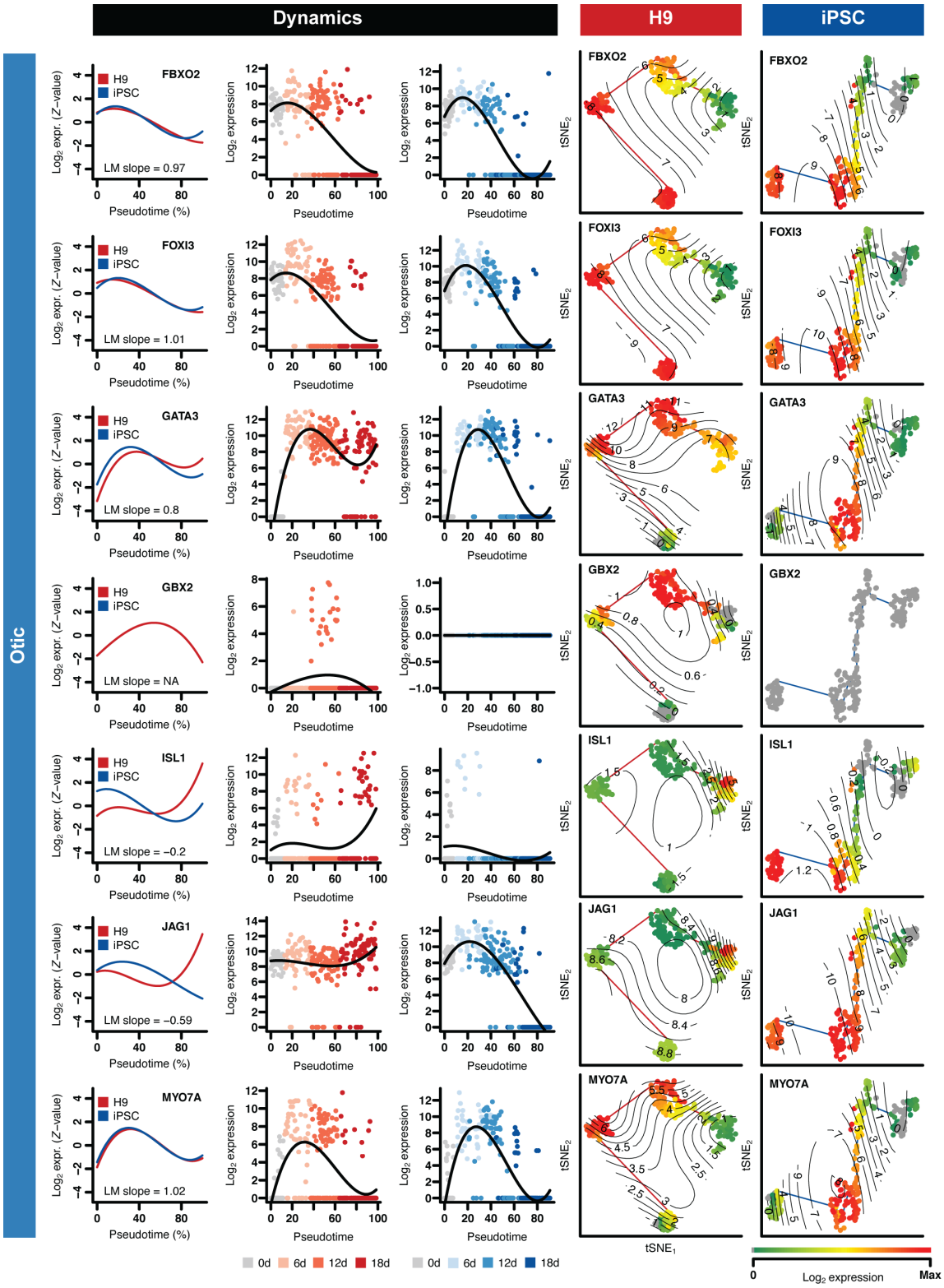


Fig. S7. Otic induction from NNE. *(A)* Screen for conditions to generate otic placode from pNNE. Mean expression values for each combination relative to NNE expression are represented in the heat map. Total days in culture and number of days treated with BMP inhibition, WNT agonist and whether retinoic acid was added (black) are indicated for each experiment. FGF2 was present in all conditions tested. Conditions yielding the highest *PAX8* and/or *PAX2* expression were used for further analyses (indicated by blue boxes). Heat map values represent the average of three technical replicates of triplicate biological replicates for each condition. *(B-C)* Time course for retinoic acid exposure to obtain posterior placode marker expression. Error bars: SD


A
Stage
Cell line
Stage:Cell line
Residual
PP
NNE
PPE
Otic
H9:
iPSC:
Explained variance (%)
B
Dynamics
H9
iPSC
Pluripotency
NANOG
POU5F1
SOX2
H9
iPSC
LM slope = 0.99
LM slope = 1.03
LM slope = 0.89
Log2 expr. (Z-value)
Pseudotime (%)
Log2 expression
Pseudotime
tSNE2
tSNE1
0d 6d 12d 18d 0d 6d 12d 18d
0
Log2 expression
Max


Dynamics
H9
iPSC
Non-neural ectoderm
DLX3
DLX5
DLX6
GATA2
GATA3
TFAP2A
TFAP2C
LM slope = 0.95
LM slope = 0.93
LM slope = 0.97
LM slope = 0
LM slope = 0.8
LM slope = 0.88
LM slope = 0.45
Log2 expr. (Z-value)
Pseudotime (%)
Log2 expression
Pseudotime
tSNE1
tSNE2
0d 6d 12d 18d
0 Log2 expression Max


Dynamics
H9
iPSC
Pre-placodal ectoderm
Otic
EYA1
EYA2
SIX1
SIX2
ATOH1
BMP7
DLX5
LM slope = 0.94
LM slope = 0.2
LM slope = 0.06
LM slope = 0.18
LM slope = 0.97
LM slope = −0.23
LM slope = 0.93
Log2 expr. (Z-value)
Pseudotime (%)
Log2 expression
Pseudotime
tSNE1
tSNE2
0d 6d 12d 18d
0
Log2 expression
Max


Dynamics
H9
iPSC
Otic
FBXO2
FOXI3
GATA3
GBX2
ISL1
JAG1
MYO7A
LM slope = 0.97
LM slope = 1.01
LM slope = 0.8
LM slope = NA
LM slope = −0.2
LM slope = −0.59
LM slope = 1.02
Log2 expr. (Z-value)
Pseudotime (%)
Log2 expression
Pseudotime
tSNE1
tSNE2
0d 6d 12d 18d
0
Log2 expression
Max

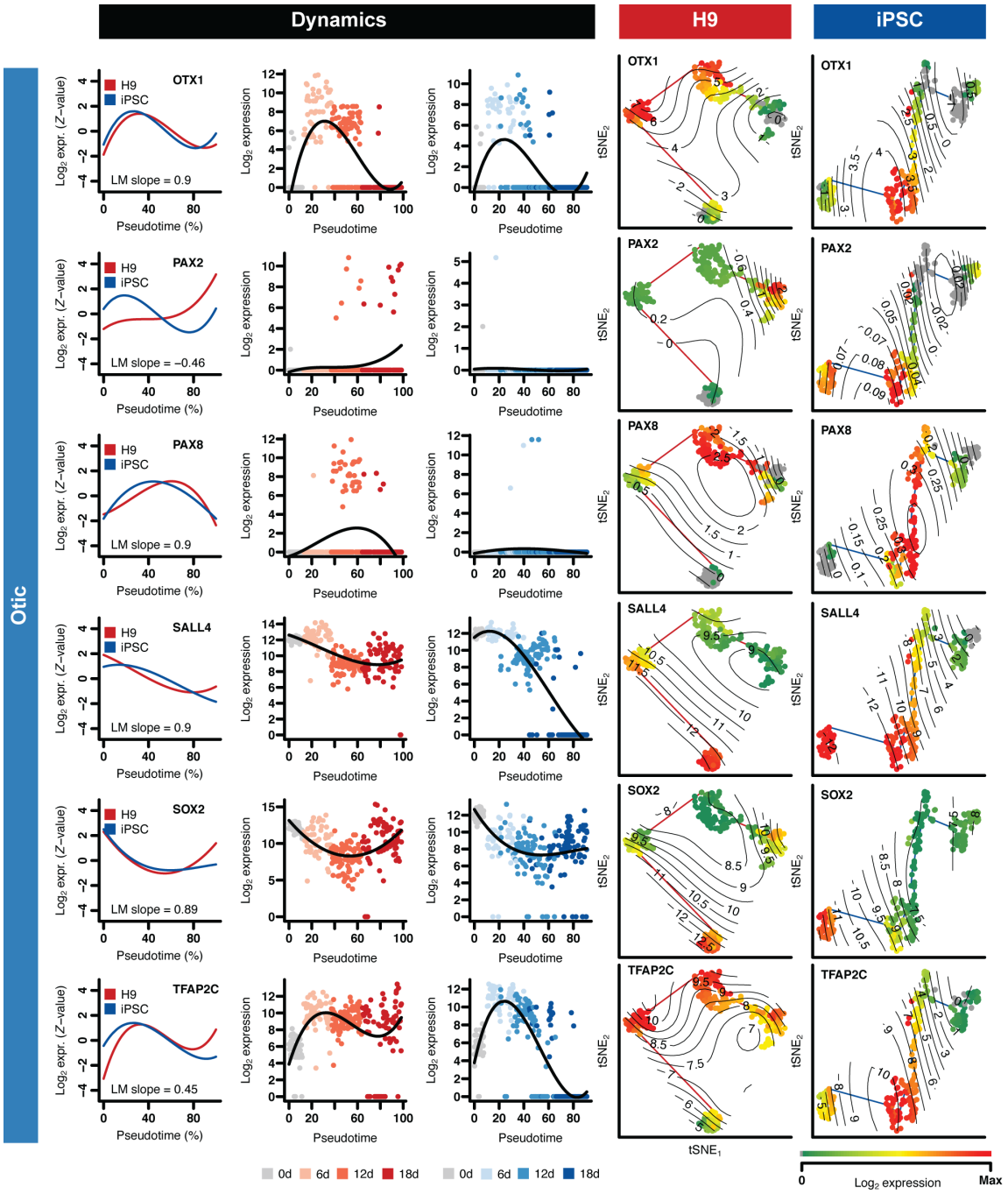


Fig. S8. Pseudotemporal ordering of H9 hESCs and iPSCs. (*A*) Two-factorial ANOVA unveils the effect of the independent variables *stage* and *cell line* on the gene expression profile (% of explained variance is indicated in black: *cell line*, dark grey: *stage*, medium grey: *stage* and *cell line* interaction, light grey: residual variance not explained by *stage* or *cell line*). Genes for which the factor *cell line*

masked the temporal variance, i.e. cell line specific genes, were not used to compute the temporal trajectories (e.g. *SOX3* and *PAX3* have more variance explained by their differential expression between cell lines than the stage of otic induction, whereas OXT2 shows stage specific expression within each *cell line* linage). Genes marked by asterisks exhibit a significant fraction of variance explained by the manifest variable *stage* and were used to compute the temporal trajectory for the indicated cell line. Gene color code indicates stage of induction. (*B*) Gene expression dynamics across the temporal trajectories indicated in H9 hESCs and iPSCs. The first column shows the comparison of fitted gene expression dynamics between H9- and iPSC-derived cells. The second and third columns show the fitting of gene expression over time, for each cell line, respectively. The fourth and fifth column visualize the tSNE-based surface plots projecting the gene expression on each cell. LM slope: slope of the linear model regression fit.

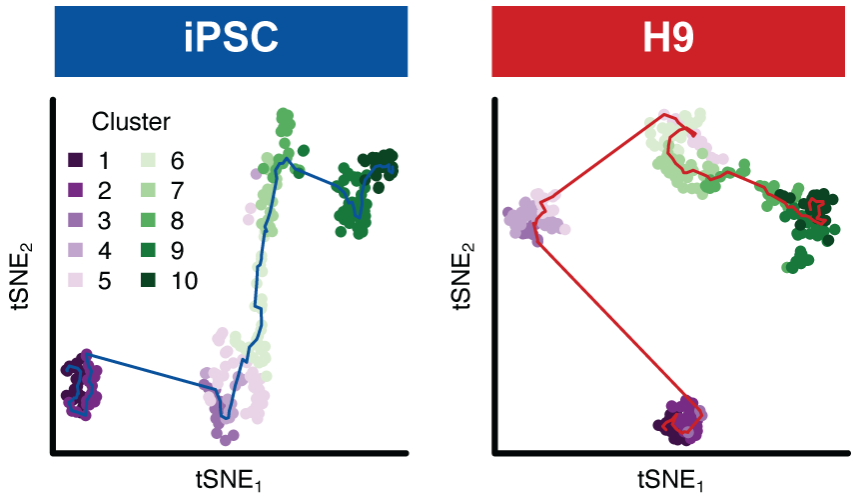


Fig. S9. Temporal clustering. Since the cell lines were of different size, cells were clustered along the trajectory around 10 medoids using the PAM algorithm.

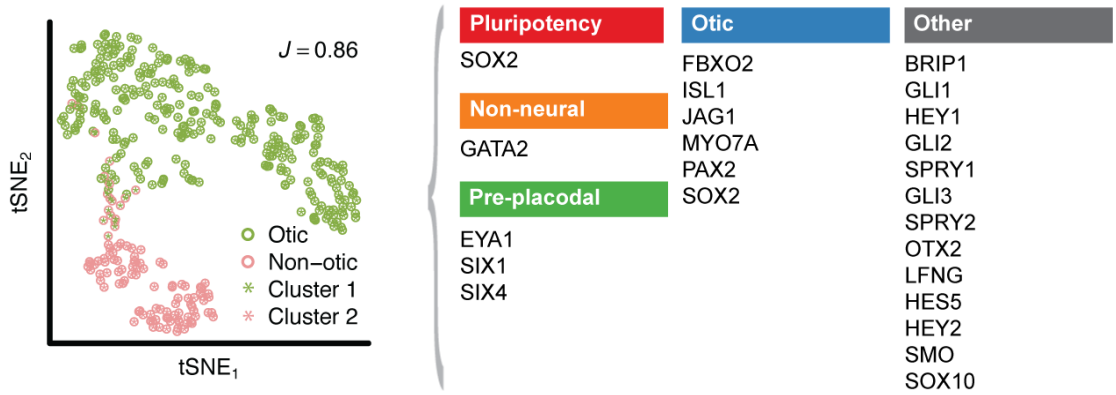


Fig. S10 Clustering of murine cells from otocyst and proximal tissue taken from E10.5 mice. Cells were classified otic / non-otic (neuroblast) according to Durruthy-Durruthy *et al.* (2014) (16). Hierarchical clustering (colored asterisks) based on a subset of 23 homologous genes was in agreement with the original annotation using the whole assay indicated by colored circles (Jaccard index = 0.86).

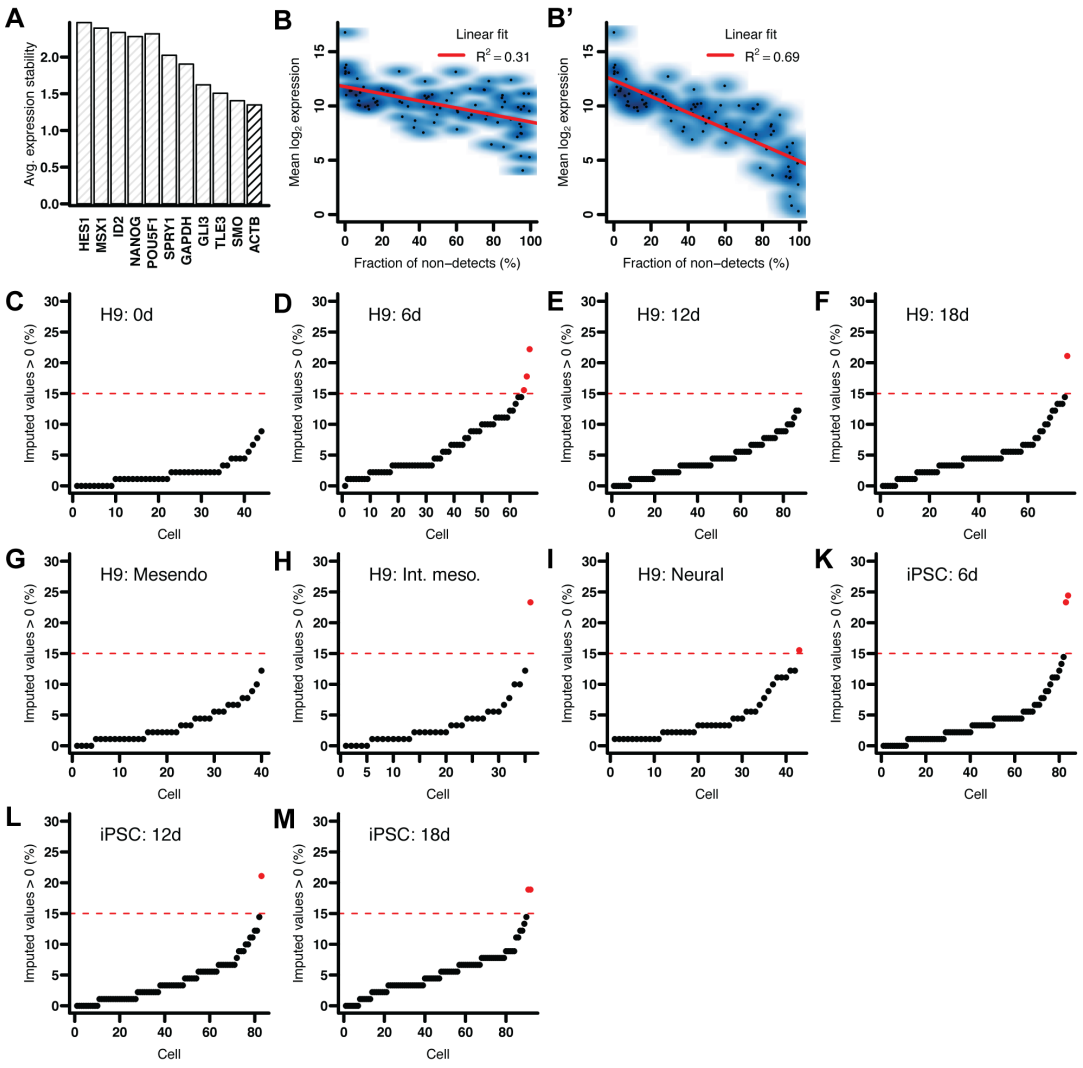


Fig. S11. Single cell expression data normalization and quality control. (*A*) Determination of reference genes for normalization based on average expression stability of 11 genes expressed in at least 95% of cells; *ACTB* was most stably expressed and was used for normalization of Log2Ex values. (*B*) Imputation of non-random non-detect values using a *k*-nearest neighbor approach. Gene expression was fitted as a cubic smoothed function of cell ordering of the *k*-neighborhood in the first independent component, before (*B*) and after (*B'*) imputation of drop-outs. (*C-M*) Percent imputed values per cell for each culture condition and time point. Cells with >15% drop-outs (red) were removed from the data frame.

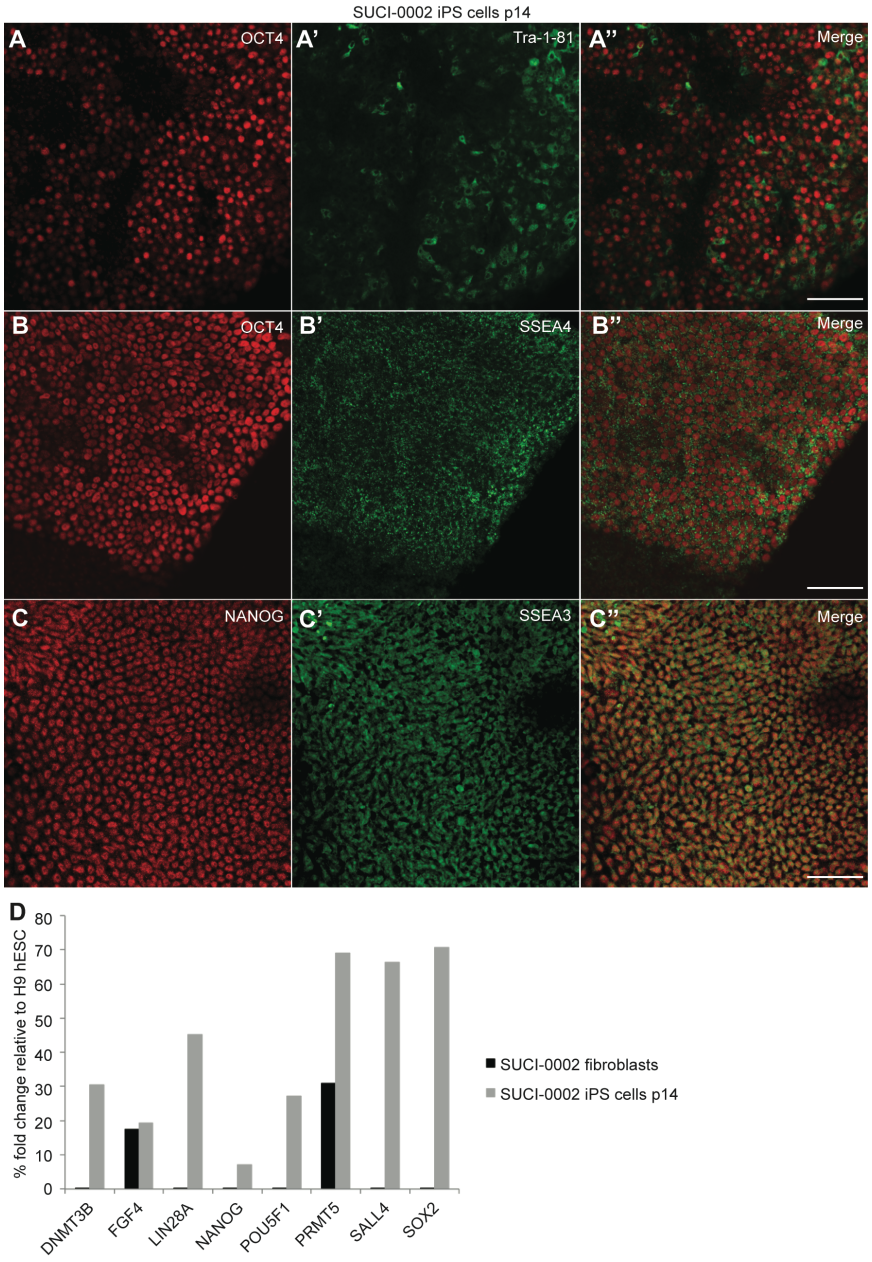


Fig. S12. iPS cell characterization. *(A-C)* Immunocytochemistry of SUCI-0002 pluripotent cells at passage 14. Cells stained positive for OCT4, Tra-1-81, SSEA4, NANOG, and SSEA3. Scale bars: 50 µm. *(D)* qRT-PCR of pluripotency genes up-regulated in SUCI-0002 iPSCs. Percent expression relative to H9 hESCs are shown and SUCI-0002 fibroblasts are shown for comparison.

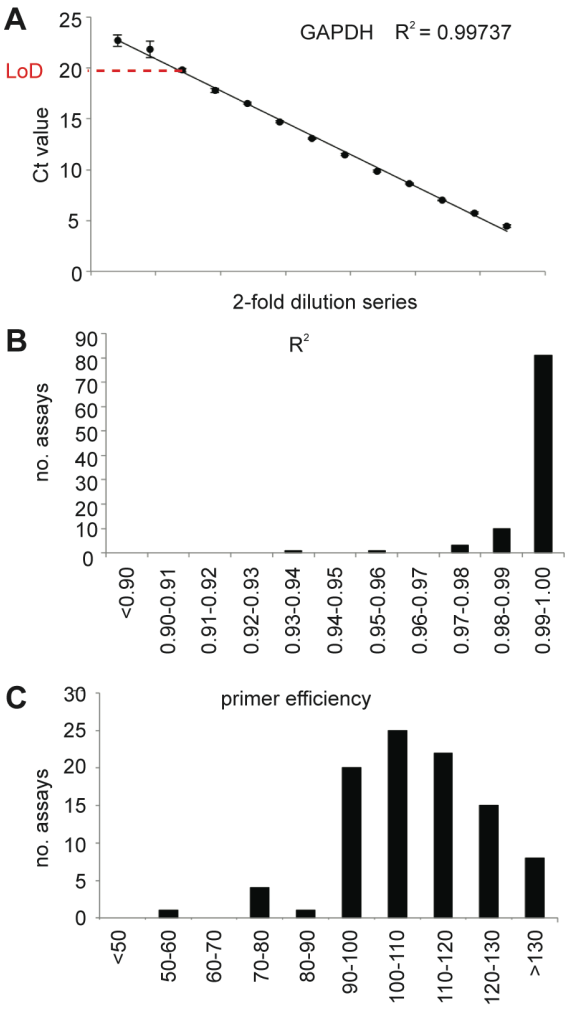


Fig. S13. Primer validation for 96-gene assay. *(A)* Example standard curve for *GAPDH*, error bars: S.D. Slope of the fitted linear regression line was used to determine the primer efficiency. R2 value was determined for the fit of the linear regression line to the standard curve data points. Limit of detection (LoD) was determined by the presence of amplification for all 6 replicates and standard deviation less than 1. *(B)* R^2 values for all assays. *(C)* Primer efficiencies of all assays.

Table S1. 96 genes for human stem cell otic induction single cell analysis.

Gene name	Rationale for inclusion	Reference	Forward Primer	Reverse Primer
ACTB	reference gene		CCAACCGCGAGAAGATGAC	TAGCACAGCCTGGATAGCAA
ATOH1	pro-hair cell development	Bermingham et al., 1999 (20)	GCAATGTTATCCCGTCGTTCAA	TCGGACAAGGCGTTGATGTA
AXIN2	activated by Wnt signaling	Behrens et al., 1998 (21)	CGAAATGCAAAAGCCACTCC	TGCGTTTGGGCAAGGTAC
BMP4	expressed in pre-placodal region at gastrula stages	Saint-Jeannet and Moody, 2014 (22)	CCACAGCACTGGTCTTGAGTA	GGTCCCTGGGATGTTCTCC
BMP7	otic pit invagination	Solloway and Robertson, 1999 (23)	ACGTTCCGGATCAGCGTTTA	CTGTCGAGCAGGAAGAGATCC
BRIP1	otic	Durruthy-Durruthy et al., 2014, Hartman et al. 2015 (16, 24)	TCACCACTGCTACTTTTCCC	CTCTTGCCTCCTCTTTACCATAA
CXCL12	olfactory	Miyasaka et al., 2007 (25)	GCTGGTCCTCGTGCTGAC	GAATCGGCATGGGCATCTGTA
*CXCR4**	definitive endoderm; olfactory	Miyasaka et al., 2007 (25)	CCCGACTTCATCTTTGCCAAC	ACACAACCACCCACAAGTCA
DLX3	otic-epibranchial; DLX+SIX+FOXI1 overlap in pre-placodal region	Saint-Jeannet and Moody, 2014, Brown et al. 2005 (22, 26)	AGCCTCCTACCGGCAATAC	TTCCGGCTCCTCCTTCAC

DLX4	DLX family ortholog	Quint et al., 2000 (27)	GTTTCCAGCACACGCAGTAC	GCGTTTGTTCTGAAACCAGATCTT
DLX5	non-neural ectoderm; pre-placodal region	Lleras-Forero and Streit, 2012, Brown et al. 2005 (26, 28)	GCTAGCTCCTACCACCAGTAC	GGTTTGCCATTCACCATTCTCA
DLX6	non-neural ectoderm; pre-placodal region	Lleras-Forero and Streit, 2012, Brown et al. 2005 (26, 28)	CAGCTTCCTTAGGACTGACACA	GAGGATTACTGCCCTGCTTCA
*DUSP6**	otic vesicle and branchial arches	Urness et al., 2008 (29)	AGCAGCCCATGTGACAACA	TGCAGAGAGTCCACCTGGTA
EMX2	neural	Surmacz et al., 2012, Holley et al. 2010 (30, 31)	GCCCCATAAATCCGTTCCTCA	CAAGTCCGGGTTGGAGTAGAC
EN2	neural	Hidalgo-Sanchez et al., 2005 (32)	CGGCGTGGGTCTACTGTAC	AGCTGCTCGGCGGTAAA
*ETV4**	activated by FGF signaling	Mahoney Rogers et al., 2011 (33)	TGCGCATCAATGTACCTCCA	TCGCAGAGGTTTCTCATAGCC
EYA1	pre-placodal region	Lleras-Forero and Streit, 2012, Zou et al. 2004, Zou et al. 2006 (28, 34, 35)	ACAGCCGACGGGTCTTTAA	TTGGTCGTGGGCTGAAACTA
EYA2	pre-placodal region	Lleras-Forero and Streit, 2012 (28)	ACCAGATCCACGTTGATGAC	CGTCAGCGGAGAAGTTGTA

EYA3	eye field development; not in otic placode	Soker et al., 2008 (36)	AGCAGTAGCCAGCATCTCAA	AGGCCTGGTACTGATTCTGAC
EYA4	otic	Modrell and Baker, 2012 (37)	CCCTTGAACAGCAGTGAAACC	GGGCTGTAGCCACTACTTGTAA
FBXO2	otic	Nelson et al., 2007; Hartman et al. 2015 (24, 38)	GCTGCTGGACACGACTCA	GAGCTCGTAGAGGCAACCA
FOXA2	definitive endoderm	Kroon et al., 2008 (2)	ACTGGAGCAGCTACTATGCA	TGTTCATGCCGTTCATCCC
FOXD3	neural crest	Dottori et al., 2001 (39)	CTCATGGCCACCCACCAA	GGAGAGTGGCACGCTAAGAA
FOXE3	lens	Dimanlig et al., 2001 (40)	GAAGCCGCCCTACTCGTAC	GCAAAGCGTTCGGTGATGAA
FOXI1	NNE; DLX+SIX+FOXI1 overlap in pre-placodal region	Lleras-Forero and Streit, 2012 (28)	GACAAGCGCCTCACTCTCA	CCGGCCTTGCTCTTGTTGTA
FOXI2	cranial ectoderm excluded from otic	Ohyama and Groves, 2004 (41)	CTTTCTACAAGCGCAGCAA	CGGGGCACCTTCTTGAA
FOXI3	pre-placodal region, otic-epibranchial	Saint-Jeannet and Moody, 2014 (22)	GAAGCCGCCCTACTCGTAC	GCAGTTCGGATCAAGAGTCCAA
FOXJ3	neural crest	Landgren and Carlsson, 2004 (42)	GGTGCCATGCATCCAACAAA	CTGAAGGAGGGAGATTTTGCCTA
GAPDH	reference gene		ACACCATGGGGAAGGTGAAG	GTGACCAGGCGCCCAATA

GATA2	pre-placodal region	Lleras-Forero and Streit, 2012, Lillevali et al. 2004 (28, 43)	GCCTGTGGCCTCTACTACAA	GTCTGGATCCCTTCCTTCTTCA
GATA3	pre-placodal region	Lleras-Forero and Streit, 2012, Lillevali et al. 2004 (28, 43)	CACGGTGCAGAGGTACCC	AGGGTAGGGATCCATGAAGCA
GATA4	endoderm	Fletcher et al., 2006 (44)	AAAACGGAAGCCCAAGAACC	AAGGCTCTCACTGCCTGAA
GATA6	endoderm	Fletcher et al., 2006 (44)	GGGCTCTACAGCAAGATGAAC	GTTGGCACAGGACAATCCAA
GBX2	posterior pre-placodal region; otic; activated by Wnt signaling	Hidalgo-Sanchez et al., 2005, Hidalgo-Sanchez et al. 2000 (32, 45)	CAAAGGCTTCCTGGCCAAA	CACCTTTGACTCGTCTTTCCC
GLI1	activated by Shh signaling	Villavicencio et al., 2000 (46)	CCTTCAGCAATGCCAGTGAC	GCAGCCAGGGAGCTTACATA
GLI2	activated by Shh signaling	Sasaki et al., 1999 (47)	GCAGCAACTGTCTGAGTGACA	TGAATGGCGACAGGGTTGAC
GLI3	activated by Shh signaling; may also be induced by WNT signaling	Sasaki et al., 1999 (47)	CATGGACCCCAGGAATGGTTA	AGGATGGAAGGCAGGGAAAA
GLI4	activated by Shh signaling	Villavicencio et al., 2000 (46)	TTCACCTCCATGGGCATCAA	TCCTCTACGTCTTGGAGATCC
HES1	activated by Wnt signaling	Saint-Jeannet and Moody, 2014 (22)	CAACACGACACCGGATAAAC	TGCTCTTCGTCTTTTCTCCA

HES5	activated by Notch signaling	Kiernan, 2013 (48)	AGCTGCTCAGCCCCAAAGA	TGCTCGATGCTGCTGTTGATG
*HES7**	activated by Notch signaling	Kiernan, 2013 (48)	ATCAACCGCAGCCTGGAA	TTCTCCAGCTTCGGGTTCC
HEY1	activated by Notch signaling	Kiernan, 2013 (48)	GAGACCATCGAGGTGGAGAA	GTAGTTGGGGACATGGAACCTA
HEY2	activated by Notch signaling	Kiernan, 2013 (48)	GTGGGGAGCGAGAACAATTA	TGTTGTTGGAGAATTCAATCTAATCAC
HMX3	otic placode	Rinkwitz-Brandt et al., 1995 (49)	TGGTACCCCTACACCCTGAC	CTCTCAGCAAGGCCTTCTCC
ID2	activated by TGFβ1 and BMP signaling	Miyazono and Miyazawa, 2002 (50)	CTCAACACGGATATCAGCATCC	CACACAGTGCTTTGCTGTCA
ID3	activated by TGFβ1 and BMP signaling	Miyazono and Miyazawa, 2002 (50)	AAAAGGAGCTTTTGCCACTGAC	TTCCGGCAGGAGAGGTTCC
IRX1	non-neural ectoderm; pre-placodal region	Saint-Jeannet and Moody, 2014 (22)	CTACGCCGCGGATCTCA	CCAGGGTTGTCCTTCAGTTCA
*IRX2**	non-neural ectoderm; pre-placodal region	Saint-Jeannet and Moody, 2014 (22)	ACGCTACCAGAAGCAAGGAC	GTGATCCGTGAGCGAGTCC
IRX3	non-neural ectoderm; pre-placodal region	Saint-Jeannet and Moody, 2014 (22)	GCGCCTTCCTGCCCTA	GGGCTGTCCTTCAGCTCATA
IRX4	pre-placodal region	Feijoo et al., 2009 (51)	CGCAGGGCTATGGCAACTA	CCCGAACCATCCTTGGAATCA
ISL1	mesoderm, otic development	Kang et al., 2009; Li et al., 2004 (52, 53)	TCGCCTTGCAGAGTGACATA	CCCGGTCCTCCTTCTGAAAA

JAG1	activated by Wnt signaling; otic prosensory domain	Lewis et al. 1998 (54)	AACAAAGGCTTCACGGGAAC	CAAGTGCCACCGTTTCTACAA
KLF4	pluripotent	Takahashi and Yamanaka, 2006 (55)	CTGCGGCAAAACCTACACAA	CGTCCCAGTCACAGTGGTAA
LFNG	otic prosensory domain	Bok et al., 2011 (56)	ACTCCCACCTGGAGAACC	GCGTTCCGCTTGTTTTCAA
LHX2	neural	Surmacz et al., 2012 (30)	CAAAAGACGGGCCTCACCAA	CGTAAGAGGTTGCGCCTGAA
LMX1B	activated by BMP signaling	Groves and Fekete, 2012 (57)	GCTGTGCAAGGGTGACTAC	TCATGTCCCCATCTTCATCC
MSX1	NNE	Saint-Jeannet and Moody, 2014; Mackenzie et al. 1991 (22, 58)	CGCAGGTGAAGATATGGTTCC	CTCCAGCTCTGCCTCTTGTA
MSX2	neural crest	Kwang et al., 2002 (59)	GCCTCGGTCAAGTCGGAAAA	CCTCAGGGTGCAGGTGGTA
MYO7A	hair cell marker, otic vesicle	Durruthy-Durruthy et al., 2014 (16)	TGAGACCCAGTTTGGCATCA	GGTGTCTCGGTTCTTCTCCA
NANOG	pluripotency marker	Takahashi and Yamanaka, 2006 (55)	TGCAGAGAAGAGTGTCGCAAA	GCTGGGTGGAAGAGAACACA
NOTCH1	activated by Wnt signaling	Lewis et al. 1998 (54)	ACGGCGTGAACACCTACAA	TGGCACTCGTCCACATCC

OTX1	otic	Hammond and Whitfield, 2006 (60)	GACCTCCTGCACCCATCC	CAGCTGTGAACGCGTGAA
OTX2	anterior NNE	Hidalgo-Sanchez et al., 2005; Hidalgo-Sanchez et al., 2000 (32, 45)	AGGAGGTGGCACTGAAAATCA	CTGTTGTTGGCGGCACTTA
PAX2	otic-epibranchial placode	Hans et al., 2004; Hidalgo-Sanchez et al., 2000 (45, 61)	CGGCTGTGTCAGCAAAATCC	GCTTGGAGCCACCGATCA
PAX3	trigeminal placode	Baker et al., 1999 (62)	GCGGTCTGTGATCGAAACA	TCCTCCTCTTCACCTTTCCC
PAX6	anterior with SIX3 and SIX6; neural progenitors	Saint-Jeannet and Moody, 2014; Surmacz et al., 2012 (22, 30)	TTGCCCGAGAAAGACTAGCA	TCTCCATTTGGCCCTTCGATTA
PAX7	neural crest	Basch et al., 2006 (63)	ACAGCATCGACGGCATCC	CAGGTTCCGACTCCACATCC
PAX8	otic-epibranchial placode	Hans et al., 2004 (61)	GCCCAGTGTCAGCTCCATTA	GCTGTCCATAGGGAGGTTGAA
PHOX2B	epibranchial neuroblasts	Saint-Jeannet and Moody, 2014 (22)	GCCCTGAAGATCGACCTCAC	CGCTCCTGCTTGCGAAA
PITX3	pituitary/lens placodes; intermediate mesoderm	Pommereit et al., 2001 (64)	TTCCAGAGGAACCGCTACC	CGCGCCGGTTCTTGAA

POU1F1	activated by Wnt signaling	Potok et al., 2008 (65)	AGGAACTCAGGCGGAAAAGTA	CAGGGCCTCCCCAACA
POU5F1	pluripotency marker	Takahashi and Yamanaka, 2006 (55)	TGGGATATACACAGGCCGATG	GATGGTCGTTTGGCTGAATACC
PTCH1	activated by Shh signaling	Villavicencio et al., 2000 (46)	AAGCCGACTACATGCCTGAA	GGTAGAAAGGGAACTGGGCATA
RAX	developing eye field marker	Surmacz et al., 2012 (30)	CCCACTACCCGGACGTGTAC	TTAGCCCGTCGGTTCTGGAA
*RBPJ**	otic	Kiernan, 2013 (48)	GAGTGTGGTTTGGGGATGTA	AGAAATGTCTGGGACGACAC
SALL4	otic	Barembaum and Bronner-Fraser, 2010 (66)	CACTGGAGAGAAGCCTTTTGTG	CCCCGTGTGTCATGTAGTGA
SIX1	DLX+SIX+FOXI1 overlap in pre-placodal region	Zou et al., 2004; Ozaki et al., 2004 (34, 67)	GGTTTAAGAACCGGAGGCAAA	TGCTTGTTGGAGGAGGAGTTA
SIX2	otic placode	Ghanbari et al., 2001 (68)	ACAGGTCAGCAACTGGTTCA	AGCGGGTTGTGGCTGTTA
SIX3	anterior placode with PAX6 and SIX6	Saint-Jeannet and Moody, 2014 (22)	GGCCTCACTCCCACACAA	ATGCCGCTCGGTCCAA
SIX4	pre-placodal region; anterior placode	Lleras-Forero and Streit, 2012 (28)	GGGAGCATTGGATTCTCTCCA	CCATCTGAAGTGCTTGAGCTTAC
SMO	activated by Shh signaling	Villavicencio et al., 2000 (46)	CCAGCAAGATCAACGAGACC	CAGCTGAAGGTAATGAGCACAA
SOX10	neural crest	Kelsh, 2006 (69)	TCGCGGACCAGTACCC	GCGCTTGTCACTTTCGTTCA

SOX17	definitive endoderm	Kroon et al., 2008 (2)	CACAACGCCGAGTTGAGCAA	GCTCTGCCTCCTCCACGAA
SOX2	pluripotency; neural plate; otic pro-sensory	Takahashi and Yamanaka, 2006; Lleras-Forero and Streit, 2012; Kiernan et al., 2005 (28, 55, 70)	AGCTCGCAGACCTACATGAA	GGAGTGGGAGGAAGAGGTAAC
SOX3	pro-neural ectoderm marker	Lleras-Forero and Streit, 2012 (28)	GTGTGAAACGGCCCATGAAC	GTGCATCTTGGGGTTCTCCA
SOX7	mesoderm	Pendeville et al., 2008 (71)	GGCCAAGGACGAGAGGAAAC	TCCGCCTCGTCCACGTA
SOX9	otic-epibranchial/neural crest	Barrionuevo et al., 2008 (72)	AACGCCGAGCTCAGCAA	CGCTTCTCGCTCTCGTTCA
SPRY1	otic	Mahoney Rogers et al., 2011 (33)	ACCATCCTGTTTGGCCTGTA	GACTAAGCACATGCAGGTTCC
SPRY2	otic	Mahoney Rogers et al., 2011 (33)	TTTGCACATCGCAGAAAGAA	TCCAGCAGGCTTAGAACACA
T	mesoderm	Mae et al., 2013 (3)	CGCTTCAAGGAGCTCACCAA	GCCAGACACGTTCACCTTCA

TBX1	posterior pre-placodal region; may be regulated by RA; activated by WNT	Freyer and Morrow, 2010 (73)	TCGACAAGCTCAAGCTGAC	GCTGGTATCTGTGCATGGAA
TFAP2A	pre-placodal region	Lleras-Forero and Streit, 2012 (28)	TAAAGCTGCCAACGTTACCC	GCACACGTACCCAAAGTCC
TFAP2C	pre-placodal region	Lleras-Forero and Streit, 2012 (28)	TGTTGCTGCACGATCAGAC	CTTCTGACAGGGGAGGTTCA
TLE3	pre-placodal region	Lleras-Forero and Streit, 2012 (28)	CTGGTGGTGGATGTTTCCAA	CCTTGTCCAGCCCATTTTCA
ZIC1	neural plate border	Fishwick et al., 2012 (74)	GCGCGCTCCGAGAATTTA	CCCTCAAACTCGCACTTGAA
ZIC2	neural plate, neural crest	Brewster et al., 1998 (75)	GCTCCGAGAACCTCAAGATCC	CGGTCGCAGCCCTCAAA

* Failed quality control and gene was excluded from single cell data analysis.

Table S2. Comparison of single cell populations. Listed are the $\log_2$ fold-changes and the *q*-Values (multiple-testing corrected Mann-Whitney test *P*-values) for each single cell population contrast for all assayed genes. Fold induction or reduction is relative to the presumptive non-neural ectoderm (pNNE) population.

	pNNE / non-pNNE		pNNE / hESC		pNNE / Mesendo		pNNE / Int. meso.		pNNE / Neural	
Gene	**Fold**	***q*-Value**	**Fold**	***q*-Value**	**Fold**	***q*-Value**	**Fold**	***q*-Value**	**Fold**	***q*-Value**
ACTB	0.00	--	0.00	--	0.00	--	0.00	--	0.00	--
ATOH1	-2.76	1.28E-04	2.10	1.52E-10	1.10	8.87E-03	0.50	1.00E+00	0.67	1.00E+00
AXIN2	-0.79	5.76E-01	1.32	1.91E-04	-0.52	1.00E+00	1.15	2.59E-01	1.44	2.09E-02
BMP4	5.16	1.07E-04	3.83	8.55E-17	2.77	8.66E-12	1.66	1.00E+00	2.37	3.18E-02
BMP7	4.02	7.14E-02	1.10	2.51E-06	0.07	1.00E+00	-0.42	1.00E+00	-0.43	1.00E+00
BRIP1	2.49	1.00E+00	-1.77	6.94E-01	-1.10	1.00E+00	3.26	4.69E-02	-1.18	9.62E-01
CXCL12	1.46	8.29E-01	-5.96	4.26E-16	-4.14	6.48E-06	2.58	1.59E-02	2.70	1.97E-03
DLX3	7.59	3.99E-11	10.93	8.20E-20	9.66	4.43E-18	4.21	4.13E-08	2.33	4.57E-01
DLX4	1.66	1.48E-03	1.79	8.27E-04	1.82	2.00E-03	1.81	8.45E-03	-0.55	1.00E+00
DLX5	8.77	2.48E-15	10.31	6.93E-20	9.02	5.54E-16	2.23	1.33E-01	1.03	1.00E+00
DLX6	2.34	3.36E-02	3.35	1.70E-04	2.98	4.43E-03	-0.87	1.00E+00	-3.91	4.83E-04
EMX2	-2.31	6.82E-05	2.37	1.89E-11	1.18	1.95E-02	0.11	1.00E+00	0.29	1.00E+00
EN2	-0.09	1.00E+00	0.08	1.00E+00	0.08	1.00E+00	0.08	1.00E+00	0.08	1.00E+00
EYA1	1.32	1.00E+00	5.37	7.31E-16	8.06	1.01E-16	0.74	1.00E+00	1.87	5.12E-02
EYA2	6.28	7.68E-11	5.35	2.04E-11	6.64	1.41E-12	0.70	1.00E+00	5.27	5.80E-08
EYA3	1.40	1.00E+00	-0.15	1.00E+00	0.39	2.21E-01	2.30	1.28E-01	-0.75	1.92E-01
EYA4	0.25	1.00E+00	0.72	1.00E+00	0.87	5.24E-01	0.87	1.00E+00	-0.28	1.00E+00
FBXO2	3.44	7.81E-02	0.42	5.03E-03	0.94	1.20E-03	-0.17	1.00E+00	1.53	4.02E-02
FOXA2	-0.26	1.00E+00	-0.31	1.00E+00	-0.14	1.00E+00	0.08	1.00E+00	0.08	1.00E+00
FOXD3	-2.39	2.59E-03	-0.59	1.00E+00	1.10	1.88E-02	1.12	1.00E+00	2.22	9.63E-03

	pNNE / non-pNNE		pNNE / hESC		pNNE / Mesendo		pNNE / Int. meso.		pNNE / Neural	
FOXE3	0.10	1.00E+00	0.38	1.00E+00	0.27	1.00E+00	0.38	1.00E+00	0.38	1.00E+00
FOXI1	-2.45	2.97E-05	1.94	1.88E-10	0.89	1.78E-03	-0.21	1.00E+00	0.11	1.00E+00
FOXI2	-1.12	1.00E+00	0.86	1.36E-01	-0.24	1.00E+00	0.31	1.00E+00	0.62	4.57E-01
FOXI3	7.14	3.04E-12	1.68	3.06E-08	1.13	1.60E-03	4.30	1.03E-08	8.01	7.34E-16
FOXJ3	2.20	1.00E+00	1.01	2.03E-10	0.97	2.94E-06	3.21	2.05E-08	-0.18	1.00E+00
GAPDH	-0.16	1.00E+00	-0.19	6.94E-01	0.97	6.30E-08	-0.39	4.25E-01	0.50	1.23E-02
GATA2	6.19	1.24E-10	8.93	3.76E-19	7.83	4.72E-16	1.29	1.00E+00	6.16	3.35E-10
GATA3	6.85	2.09E-12	10.18	7.87E-20	10.03	2.25E-18	0.64	4.55E-01	6.89	2.00E-16
GATA4	-0.09	1.00E+00	0.29	1.00E+00	-2.86	2.98E-08	0.02	1.00E+00	0.05	1.00E+00
GATA6	0.43	1.00E+00	0.75	6.02E-01	-0.32	1.00E+00	-1.05	8.98E-01	-0.72	1.00E+00
GBX2	-0.01	1.00E+00	0.06	1.00E+00	-0.58	1.13E-03	0.06	1.00E+00	0.06	1.00E+00
GLI1	-1.32	1.00E+00	-4.36	1.21E-13	-2.99	1.88E-05	-2.18	5.57E-02	1.20	9.62E-01
GLI2	1.61	1.00E+00	-1.15	2.08E-02	0.91	4.52E-02	0.12	1.00E+00	-1.04	1.00E+00
GLI3	0.98	1.00E+00	2.91	6.08E-18	3.23	2.82E-16	2.96	1.95E-09	-0.62	2.83E-01
GLI4	1.11	1.00E+00	0.56	4.01E-03	0.26	1.80E-01	1.24	1.55E-01	0.58	1.00E+00
HES1	-1.37	1.33E-02	2.94	4.88E-13	0.27	1.00E+00	0.10	1.00E+00	-0.08	1.00E+00
HES5	-0.38	1.00E+00	0.47	1.00E+00	0.54	1.00E+00	0.26	1.00E+00	-9.60	4.46E-21
HEY1	-0.77	1.00E+00	0.72	1.00E+00	2.01	3.09E-01	-3.06	2.54E-02	-2.30	1.02E-01
HEY2	0.30	1.00E+00	-6.20	1.15E-15	-5.68	1.48E-11	0.60	1.00E+00	0.85	1.00E+00
HMX3	0.47	1.00E+00	1.08	5.22E-02	0.76	1.00E+00	1.02	5.57E-01	1.08	9.41E-02
ID2	-1.02	5.65E-03	3.84	2.40E-16	0.99	7.14E-03	0.77	1.87E-01	1.67	5.49E-04
ID3	-0.37	2.16E-01	0.89	2.77E-02	0.96	6.13E-03	-1.01	2.71E-01	0.82	2.83E-01
IRX1	4.19	2.52E-04	7.80	4.26E-16	5.50	9.23E-10	7.33	1.05E-12	5.31	7.22E-10
IRX3	5.33	1.60E-09	7.02	1.52E-18	7.22	1.37E-17	7.40	6.19E-16	4.48	7.52E-09
IRX4	1.40	6.71E-02	-0.23	1.00E+00	0.71	1.00E+00	0.95	1.00E+00	-0.28	1.00E+00

	pNNE / non-pNNE		pNNE / hESC		pNNE / Mesendo		pNNE / Int. meso.		pNNE / Neural	
ISL1	2.10	1.23E-02	1.46	1.00E+00	2.32	1.04E-02	-7.27	1.45E-11	2.32	1.05E-02
JAG1	1.36	1.00E+00	1.21	2.50E-06	0.42	1.50E-01	-1.31	8.38E-02	-1.41	1.51E-02
KLF4	-0.71	1.00E+00	1.92	2.08E-02	-0.61	1.00E+00	-0.14	1.00E+00	2.49	1.45E-01
LFNG	1.66	6.73E-01	0.19	1.00E+00	1.65	4.27E-04	2.34	1.32E-03	-3.76	8.04E-12
LHX2	0.21	1.00E+00	1.36	1.55E-01	1.24	5.44E-01	1.22	7.02E-01	-9.33	3.25E-20
LMX1B	1.51	1.00E+00	6.32	6.14E-10	6.28	5.03E-09	6.30	5.36E-08	4.27	6.99E-05
MSX1	-2.51	1.17E-06	3.11	3.14E-13	1.16	1.07E-02	-2.42	3.25E-07	1.63	1.10E-03
MSX2	1.27	1.00E+00	10.09	1.88E-19	4.97	3.75E-16	-0.86	5.03E-01	7.66	1.06E-16
MYO7A	6.65	4.23E-13	6.68	1.19E-14	7.42	6.27E-15	7.44	7.16E-13	3.66	6.95E-03
NANOG	-2.10	1.24E-03	-1.87	2.97E-13	-2.50	2.42E-14	0.43	1.00E+00	0.27	1.00E+00
NOTCH1	0.60	1.00E+00	1.73	1.84E-08	0.53	5.33E-01	2.84	3.89E-06	-0.84	7.74E-01
OTX1	6.07	2.47E-12	7.03	5.35E-17	7.38	2.04E-16	7.38	9.93E-15	4.79	1.65E-08
OTX2	5.78	4.31E-04	1.26	3.65E-03	5.98	1.41E-13	11.95	1.48E-16	-0.09	1.00E+00
PAX2	-0.58	1.00E+00	0.01	1.00E+00	-0.09	1.00E+00	0.06	1.00E+00	-0.11	1.00E+00
PAX3	-1.80	1.91E-03	6.54	9.55E-15	6.15	2.04E-12	6.54	3.76E-12	3.41	1.60E-04
PAX6	-1.43	2.28E-01	3.60	3.14E-13	2.71	1.10E-06	1.51	5.03E-01	-3.47	3.17E-08
PAX7	-0.50	1.00E+00	3.54	1.21E-04	2.58	2.76E-01	3.62	7.08E-04	-4.43	2.22E-06
PAX8	-0.39	1.00E+00	0.27	1.00E+00	0.27	1.00E+00	0.27	1.00E+00	-1.27	9.57E-02
PHOX2B	0.09	1.00E+00	0.09	1.00E+00	0.09	1.00E+00	0.09	1.00E+00	0.09	1.00E+00
PITX3	-0.39	1.00E+00	0.02	1.00E+00	-0.05	1.00E+00	-1.56	3.16E-04	-0.22	1.00E+00
POU1F1	0.17	1.00E+00	0.17	1.00E+00	0.17	1.00E+00	0.17	1.00E+00	0.17	1.00E+00
POU5F1	-2.01	1.15E-03	-3.62	1.02E-17	-3.04	3.80E-15	0.17	1.00E+00	0.21	1.00E+00
PTCH1	0.52	1.00E+00	0.84	3.99E-03	0.37	1.00E+00	2.60	8.23E-02	-1.08	4.60E-01
RAX	-0.11	1.00E+00	0.53	1.00E+00	0.63	1.00E+00	0.63	1.00E+00	-8.24	5.29E-21
SALL4	2.26	2.83E-02	0.07	1.00E+00	1.52	2.51E-11	4.49	6.19E-13	0.17	1.00E+00

	pNNE / non-pNNE		pNNE / hESC		pNNE / Mesendo		pNNE / Int. meso.		pNNE / Neural	
SIX1	0.85	1.00E+00	2.76	2.19E-03	2.96	8.20E-04	2.96	3.15E-03	1.16	1.00E+00
SIX2	0.25	1.00E+00	0.15	1.00E+00	0.01	1.00E+00	0.25	1.00E+00	-0.30	1.00E+00
SIX3	1.83	3.47E-03	2.12	4.40E-04	2.12	1.14E-03	2.12	4.30E-03	-6.48	2.68E-16
SIX4	1.77	1.00E+00	0.56	3.03E-02	1.23	7.86E-05	2.62	1.46E-05	0.07	1.00E+00
SMO	-1.25	6.22E-06	0.09	1.00E+00	1.24	6.34E-06	1.78	8.42E-08	-0.34	1.00E+00
SOX10	-1.52	1.00E+00	1.54	1.00E+00	2.39	1.26E-01	3.79	2.45E-02	-1.54	1.00E+00
SOX17	3.54	1.62E-02	3.96	6.50E-16	4.97	2.49E-12	-0.63	1.00E+00	-0.19	1.00E+00
SOX2	-2.35	3.10E-06	-2.51	2.10E-11	0.49	1.00E+00	1.72	5.15E-02	-3.01	2.61E-10
SOX3	-1.10	1.00E+00	-0.43	1.00E+00	-1.65	2.97E-01	0.06	1.00E+00	-2.12	6.48E-03
SOX7	0.31	1.00E+00	-0.08	1.00E+00	-0.11	1.00E+00	-1.32	5.15E-02	0.14	1.00E+00
SOX9	4.49	3.08E-04	4.08	2.33E-16	5.34	8.33E-13	-0.87	3.60E-01	-0.19	1.00E+00
SPRY1	-1.49	1.00E+00	0.31	1.00E+00	0.51	1.64E-01	0.27	1.00E+00	-2.03	4.62E-08
SPRY2	3.30	1.87E-01	-0.69	1.74E-01	-1.48	1.80E-05	1.60	1.00E+00	-1.84	4.43E-07
T	0.05	1.00E+00	0.05	1.00E+00	-7.30	4.08E-22	0.05	1.00E+00	0.05	1.00E+00
TBX1	0.68	1.00E+00	0.75	1.00E+00	-0.17	1.00E+00	0.40	1.00E+00	-0.65	1.00E+00
TFAP2A	4.84	4.23E-02	12.39	7.61E-20	12.05	2.25E-18	2.73	1.29E-03	10.58	1.67E-17
TFAP2C	5.09	1.22E-05	4.94	6.80E-18	2.33	2.44E-10	3.84	1.20E-08	1.87	1.00E+00
TLE3	-0.60	2.81E-01	1.09	5.09E-07	1.12	7.62E-05	1.44	1.30E-03	1.08	5.57E-03
ZIC1	-0.55	1.00E+00	0.40	1.00E+00	-0.44	1.00E+00	-0.59	1.00E+00	-7.03	3.30E-15
ZIC2	2.00	1.00E+00	-1.90	3.44E-04	-1.91	1.16E-03	9.16	1.34E-14	-2.35	3.10E-05

Supporting Information References

1. Oshima K*, et al.* (2010) Mechanosensitive hair cell-like cells from embryonic and induced pluripotent stem cells. *Cell* 141(4):704-716.
2. Kroon E*, et al.* (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nature biotechnology* 26(4):443-452.
3. Mae S*, et al.* (2013) Monitoring and robust induction of nephrogenic intermediate mesoderm from human pluripotent stem cells. *Nature communications* 4:1367.
4. Ronaghi M*, et al.* (2014) Inner ear hair cell-like cells from human embryonic stem cells. *Stem cells and development* 23(11):1275-1284.
5. Vandesompele J*, et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* 3(7):RESEARCH0034.
6. Hastie T & Tibshirani R (1986) Generalized additive models. *Statistical Science* 1(3):297-318.
7. Hyvarinen A & Oja E (2000) Independent component analysis: Algorithms and applications. *Neural Networks* 13(4-5):411-430.
8. van der Maaten LJP (2013) Barnes-Hut-SNE. *In Proceedings of the Intenational Conference on Learning Representations*.
9. Wood SN (2003) Thin plate regression splines. *J.R. Statist. Soc. B* 65:95-114.
10. Ward JH, Jr. (1963) Hierarchical grouping to optimize and objective function. *Journal of the American Statistical Association* 58:236-244.
11. TIbshirani RW, G; Hastie, T (2001) Estimating the number of clusters in a data set via the gap statistic. *Journal for the Royal Statistical Society: Series B (Statistical Methodology)* 63(2):411-423.
12. Prim RC (1957) Shortest connection networks and some generalizations. *Bell System Technical Journal* 37:1389-1401.
13. Trapnell C*, et al.* (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 32(4):381-386.
14. Yee TW & Wild CJ (1996) Vector generalized additive models. *J Roy Stat Soc B Met* 58(3):481-493.
15. Reynolds A, Richards G, de la Iglesia B, & Rayward-Smith V (1992) Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms* 5:475-504.
16. Durruthy-Durruthy R*, et al.* (2014) Reconstruction of the mouse otocyst and early neuroblast lineage at single-cell resolution. *Cell* 157(4):964-978.
17. Ellies DL, Langille RM, Martin CC, Akimenko MA, & Ekker M (1997) Specific craniofacial cartilage dysmorphogenesis coincides with a loss of dlx gene expression in retinoic acid-treated zebrafish embryos. *Mechanisms of development* 61(1-2):23-36.

18. Yang L*, et al.* (1998) An early phase of embryonic Dlx5 expression defines the rostral boundary of the neural plate. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 18(20):8322-8330.
19. Pattyn A, Morin X, Cremer H, Goridis C, & Brunet JF (1999) The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives. *Nature* 399(6734):366-370.
20. Bermingham NA*, et al.* (1999) Math1: an essential gene for the generation of inner ear hair cells. *Science* 284(5421):1837-1841.
21. Behrens J*, et al.* (1998) Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. *Science* 280(5363):596-599.
22. Saint-Jeannet JP & Moody SA (2014) Establishing the pre-placodal region and breaking it into placodes with distinct identities. *Developmental biology* 389(1):13-27.
23. Solloway MJ & Robertson EJ (1999) Early embryonic lethality in Bmp5;Bmp7 double mutant mice suggests functional redundancy within the 60A subgroup. *Development* 126(8):1753-1768.
24. Hartman BH, Durruthy-Durruthy R, Laske RD, Losorelli S, & Heller S (2015) Identification and characterization of mouse otic sensory lineage genes. *Front Cell Neurosci* 9:79.
25. Miyasaka N, Knaut H, & Yoshihara Y (2007) Cxcl12/Cxcr4 chemokine signaling is required for placode assembly and sensory axon pathfinding in the zebrafish olfactory system. *Development* 134(13):2459-2468.
26. Brown ST, Wang J, & Groves AK (2005) Dlx gene expression during chick inner ear development. *The Journal of comparative neurology* 483(1):48-65.
27. Quint E, Zerucha T, & Ekker M (2000) Differential expression of orthologous Dlx genes in zebrafish and mice: implications for the evolution of the Dlx homeobox gene family. *J Exp Zool* 288(3):235-241.
28. Lleras-Forero L & Streit A (2012) Development of the sensory nervous system in the vertebrate head: the importance of being on time. *Curr Opin Genet Dev* 22(4):315-322.
29. Urness LD, Li C, Wang X, & Mansour SL (2008) Expression of ERK signaling inhibitors Dusp6, Dusp7, and Dusp9 during mouse ear development. *Dev Dyn* 237(1):163-169.
30. Surmacz B*, et al.* (2012) Directing differentiation of human embryonic stem cells toward anterior neural ectoderm using small molecules. *Stem cells* 30(9):1875-1884.
31. Holley M*, et al.* (2010) Emx2 and early hair cell development in the mouse inner ear. *Dev Biol* 340(2):547-556.
32. Hidalgo-Sanchez M, Millet S, Bloch-Gallego E, & Alvarado-Mallart RM (2005) Specification of the meso-isthmo-cerebellar region: the Otx2/Gbx2 boundary. *Brain Res Brain Res Rev* 49(2):134-149.
33. Mahoney Rogers AA, Zhang J, & Shim K (2011) Sprouty1 and Sprouty2 limit both the size of the otic placode and hindbrain Wnt8a by antagonizing FGF signaling. *Developmental biology* 353(1):94-104.

34. Zou D, Silvius D, Fritzsch B, & Xu PX (2004) Eya1 and Six1 are essential for early steps of sensory neurogenesis in mammalian cranial placodes. *Development* 131(22):5561-5572.
35. Zou D, Silvius D, Rodrigo-Blomqvist S, Enerback S, & Xu PX (2006) Eya1 regulates the growth of otic epithelium and interacts with Pax2 during the development of all sensory areas in the inner ear. *Dev Biol* 298(2):430-441.
36. Soker T*, et al.* (2008) Pleiotropic effects in Eya3 knockout mice. *BMC developmental biology* 8:118.
37. Modrell MS & Baker CV (2012) Evolution of electrosensory ampullary organs: conservation of Eya4 expression during lateral line development in jawed vertebrates. *Evolution & development* 14(3):277-285.
38. Nelson RF*, et al.* (2007) Selective cochlear degeneration in mice lacking the F-box protein, Fbx2, a glycoprotein-specific ubiquitin ligase subunit. *J Neurosci* 27(19):5163-5171.
39. Dottori M, Gross MK, Labosky P, & Goulding M (2001) The winged-helix transcription factor Foxd3 suppresses interneuron differentiation and promotes neural crest cell fate. *Development* 128(21):4127-4138.
40. Dimanlig PV, Faber SC, Auerbach W, Makarenkova HP, & Lang RA (2001) The upstream ectoderm enhancer in Pax6 has an important role in lens induction. *Development* 128(22):4415-4424.
41. Ohyama T & Groves AK (2004) Expression of mouse Foxi class genes in early craniofacial development. *Developmental dynamics : an official publication of the American Association of Anatomists* 231(3):640-646.
42. Landgren H & Carlsson P (2004) FoxJ3, a novel mammalian forkhead gene expressed in neuroectoderm, neural crest, and myotome. *Developmental dynamics : an official publication of the American Association of Anatomists* 231(2):396-401.
43. Lillevali K, Matilainen T, Karis A, & Salminen M (2004) Partially overlapping expression of Gata2 and Gata3 during inner ear development. *Dev Dyn* 231(4):775-781.
44. Fletcher G, Jones GE, Patient R, & Snape A (2006) A role for GATA factors in Xenopus gastrulation movements. *Mechanisms of development* 123(10):730-745.
45. Hidalgo-Sanchez M, Alvarado-Mallart R, & Alvarez IS (2000) Pax2, Otx2, Gbx2 and Fgf8 expression in early otic vesicle development. *Mech Dev* 95(1-2):225-229.
46. Villavicencio EH, Walterhouse DO, & Iannaccone PM (2000) The sonic hedgehog-patched-gli pathway in human development and disease. *American journal of human genetics* 67(5):1047-1054.
47. Sasaki H, Nishizaki Y, Hui C, Nakafuku M, & Kondoh H (1999) Regulation of Gli2 and Gli3 activities by an amino-terminal repression domain: implication of Gli2 and Gli3 as primary mediators of Shh signaling. *Development* 126(17):3915-3924.
48. Kiernan AE (2013) Notch signaling during cell fate determination in the inner ear. *Semin Cell Dev Biol* 24(5):470-479.

49. Rinkwitz-Brandt S, Justus M, Oldenettel I, Arnold HH, & Bober E (1995) Distinct temporal expression of mouse Nkx-5.1 and Nkx-5.2 homeobox genes during brain and ear development. *Mech Dev* 52(2-3):371-381.
50. Miyazono K & Miyazawa K (2002) Id: a target of BMP signaling. *Science's STKE : signal transduction knowledge environment* 2002(151):pe40.
51. Feijoo CG, Saldias MP, De la Paz JF, Gomez-Skarmeta JL, & Allende ML (2009) Formation of posterior cranial placode derivatives requires the Iroquois transcription factor irx4a. *Molecular and cellular neurosciences* 40(3):328-337.
52. Kang J, Nathan E, Xu SM, Tzahor E, & Black BL (2009) Isl1 is a direct transcriptional target of Forkhead transcription factors in second-heart-field-derived mesoderm. *Developmental biology* 334(2):513-522.
53. Li H*, et al.* (2004) Islet-1 expression in the developing chicken inner ear. *The Journal of comparative neurology* 477(1):1-10.
54. Lewis AK, Frantz GD, Carpenter DA, de Sauvage FJ, & Gao WQ (1998) Distinct expression patterns of notch family receptors and ligands during development of the mammalian inner ear. *Mechanisms of development* 78(1-2):159-163.
55. Takahashi K & Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4):663-676.
56. Bok J*, et al.* (2011) Transient retinoic acid signaling confers anterior-posterior polarity to the inner ear. *Proceedings of the National Academy of Sciences of the United States of America* 108(1):161-166.
57. Groves AK & Fekete DM (2012) Shaping sound in space: the regulation of inner ear patterning. *Development* 139(2):245-257.
58. Mackenzie A, Leeming GL, Jowett AK, Ferguson MW, & Sharpe PT (1991) The homeobox gene Hox 7.1 has specific regional and temporal expression patterns during early murine craniofacial embryogenesis, especially tooth development in vivo and in vitro. *Development* 111(2):269-285.
59. Kwang SJ*, et al.* (2002) Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest. *Development* 129(2):527-538.
60. Hammond KL & Whitfield TT (2006) The developing lamprey ear closely resembles the zebrafish otic vesicle: otx1 expression can account for all major patterning differences. *Development* 133(7):1347-1357.
61. Hans S, Liu D, & Westerfield M (2004) Pax8 and Pax2a function synergistically in otic specification, downstream of the Foxi1 and Dlx3b transcription factors. *Development* 131(20):5091-5102.
62. Baker CV, Stark MR, Marcelle C, & Bronner-Fraser M (1999) Competence, specification and induction of Pax-3 in the trigeminal placode. *Development* 126(1):147-156.
63. Basch ML, Bronner-Fraser M, & Garcia-Castro MI (2006) Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature* 441(7090):218-222.

64. Pommereit D, Pieler T, & Hollemann T (2001) Xpitx3: a member of the Rieg/Pitx gene family expressed during pituitary and lens formation in Xenopus laevis. *Mechanisms of development* 102(1-2):255-257.
65. Potok MA*, et al.* (2008) WNT signaling affects gene expression in the ventral diencephalon and pituitary gland growth. *Developmental dynamics : an official publication of the American Association of Anatomists* 237(4):1006-1020.
66. Barembaum M & Bronner-Fraser M (2010) Pax2 and Pea3 synergize to activate a novel regulatory enhancer for spalt4 in the developing ear. *Developmental biology* 340(2):222-231.
67. Ozaki H*, et al.* (2004) Six1 controls patterning of the mouse otic vesicle. *Development* 131(3):551-562.
68. Ghanbari H, Seo HC, Fjose A, & Brandli AW (2001) Molecular cloning and embryonic expression of Xenopus Six homeobox genes. *Mechanisms of development* 101(1-2):271-277.
69. Kelsh RN (2006) Sorting out Sox10 functions in neural crest development. *BioEssays : news and reviews in molecular, cellular and developmental biology* 28(8):788-798.
70. Kiernan AE*, et al.* (2005) Sox2 is required for sensory organ development in the mammalian inner ear. *Nature* 434(7036):1031-1035.
71. Pendeville H*, et al.* (2008) Zebrafish Sox7 and Sox18 function together to control arterial-venous identity. *Developmental biology* 317(2):405-416.
72. Barrionuevo F*, et al.* (2008) Sox9 is required for invagination of the otic placode in mice. *Dev Biol* 317(1):213-224.
73. Freyer L & Morrow BE (2010) Canonical Wnt signaling modulates Tbx1, Eya1, and Six1 expression, restricting neurogenesis in the otic vesicle. *Developmental dynamics : an official publication of the American Association of Anatomists* 239(6):1708-1722.
74. Fishwick KJ, Kim E, & Bronner ME (2012) ILF-3 is a regulator of the neural plate border marker Zic1 in chick embryos. *Developmental dynamics : an official publication of the American Association of Anatomists* 241(8):1325-1332.
75. Brewster R, Lee J, & Ruiz i Altaba A (1998) Gli/Zic factors pattern the neural plate by defining domains of cell differentiation. *Nature* 393(6685):579-583.