Supplemental Figures

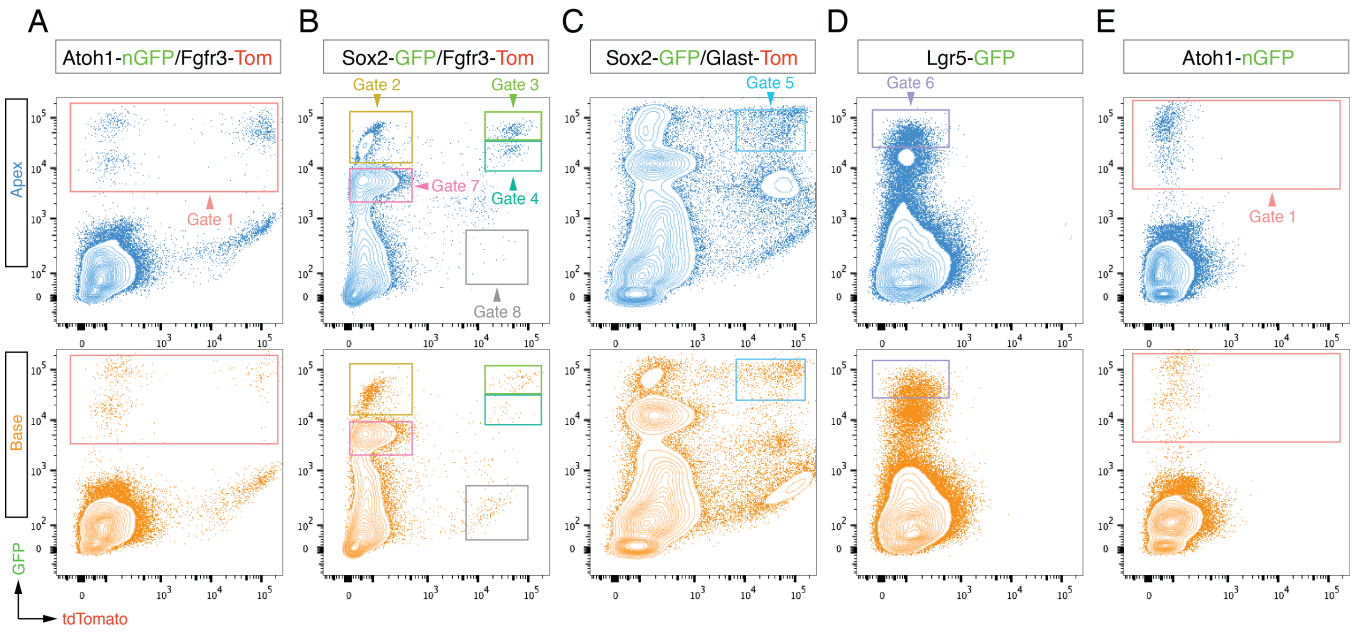


Figure S1. Separate Sorting of Apical versus Basal Originated Organ of Corti Cell Populations. Related to Figure 1.

(**A**) – (**E**) Apical and basal cochlear ducts from all mouse lines investigated were dissected together but then processed separately and sorted independently. Identical gates were applied and equal numbers of cells were collected from apical and basal partitions from each mouse line.

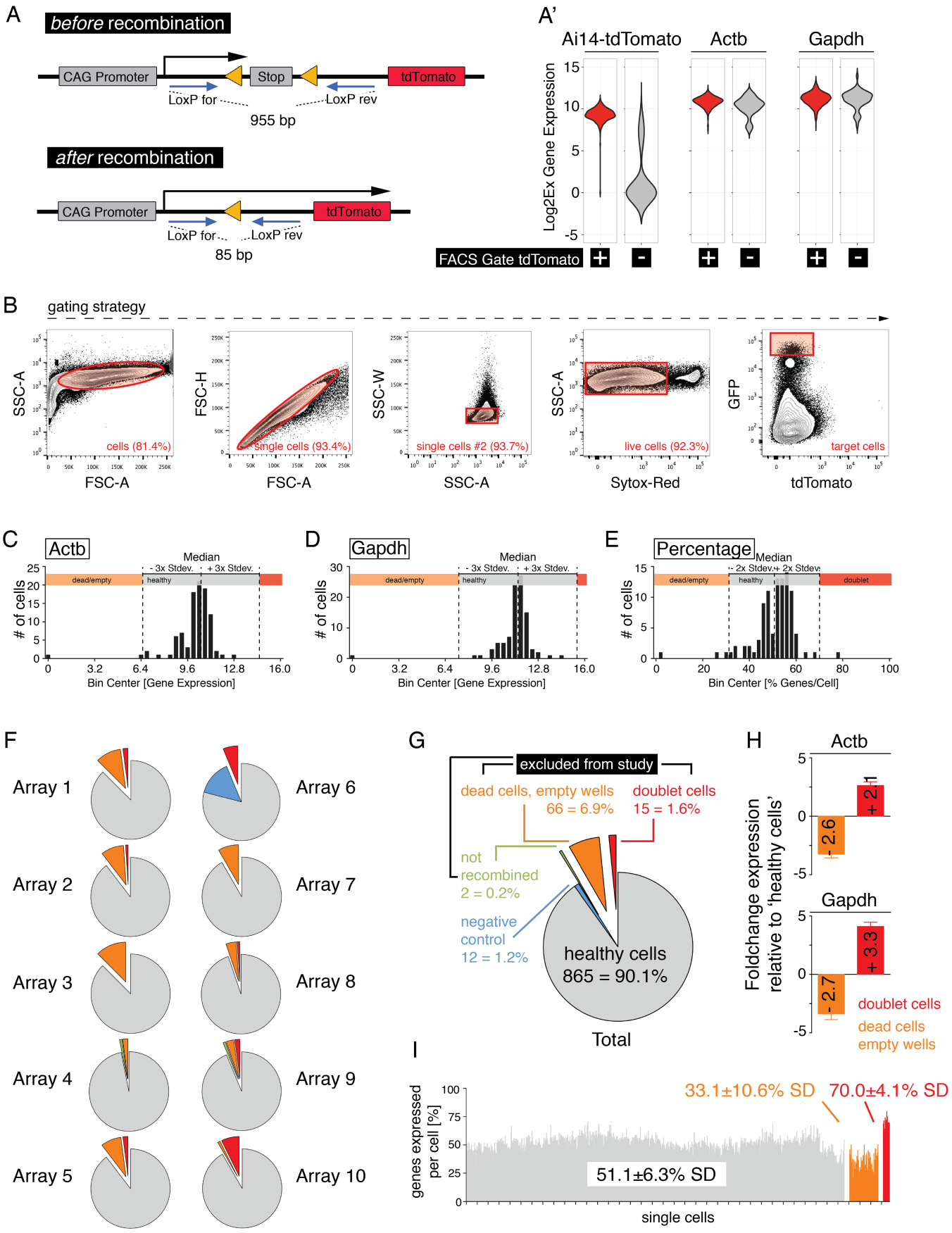

A
before recombination
CAG Promoter
Stop
tdTomato
LoxP for
LoxP rev
955 bp
after recombination
CAG Promoter
tdTomato
LoxP for
LoxP rev
85 bp
A'
Ai14-tdTomato
Actb
Gapdh
Log2Ex Gene Expression
FACS Gate tdTomato
B
gating strategy
SSC-A
cells (81.4%)
FSC-A
FSC-H
single cells (93.4%)
FSC-A
SSC-W
single cells #2 (93.7%)
SSC-A
SSC-A
live cells (92.3%)
Sytox-Red
GFP
target cells
tdTomato
C
Actb
Median
- 3x Stdev.
+ 3x Stdev.
dead/empty
healthy
# of cells
Bin Center [Gene Expression]
D
Gapdh
Median
- 3x Stdev.
+ 3x Stdev.
dead/empty
healthy
# of cells
Bin Center [Gene Expression]
E
Percentage
Median
- 2x Stdev.
+ 2x Stdev.
dead/empty
healthy
doublet
# of cells
Bin Center [% Genes/Cell]
F
Array 1
Array 2
Array 3
Array 4
Array 5
Array 6
Array 7
Array 8
Array 9
Array 10
G
excluded from study
dead cells, empty wells
66 = 6.9%
doublet cells
15 = 1.6%
not
recombined
2 = 0.2%
negative
control
12 = 1.2%
healthy cells
865 = 90.1%
Total
H
Actb
- 2.6
+ 2.1
Gapdh
- 2.7
+ 3.3
Foldchange expression
relative to 'healthy cells'
doublet cells
dead cells
empty wells
I
genes expressed
per cell [%]
33.1±10.6% SD
70.0±4.1% SD
51.1±6.3% SD
single cells

Figure S2. Single Cell qRT-PCR Quality Control. Related to Figures 1 and 2.

(**A**) Schematic overview of the floxed genomic locus of the Ai14-tdTomato reporter mouse. Prior to recombination, primers (LoxP-for and LoxP-rev) cover a 955 bp amplicon. After recombination, the floxed segment with the stop-codon is removed, reducing the amplicon size to 85 bp. The parameters of the single-cell qRT-PCR protocol were such that only the recombined short amplicon was detected.

(**A'**) Violin plots comparing tdTomato recombination, as well as expression level with Actb and Gapdh primers in all tdTomato positive versus all tdTomato negative sorted cells. All tdTomato positive FACS sorted cells carried the recombined locus (minus 2 excluded cells, see below (F)). A few tdTomato negative cells displayed reduced amplification of the presumptively recombined locus, which is possibly the case in cells where recombination might have happened but tdTomato protein expression was not yet detectable.

(**B**) Gating strategy used to exclude debris, doublets and dead cells. Target cells were collected using gates set to distinct tdTomato and GFP fluorescence intensity levels; here a representative example related to Figure 1M is shown. SSC-A: Side Scatter-Area, FSC-A: Forward Scatter-Area, FSH-H: Forward Scatter-Heights, SSC-W: Side Scatter-Width.

(**C**) – (**I**) Illustrates the single cell qRT-PCR quality control criteria used to determine included and excluded cells. Cells expressing internal control genes (**C**) Actb and (**D**) Gapdh within a range of ±3 fold standard deviation were considered as healthy. (**E**) In addition, the percentage of genes expressed per cell exceeding ±2 fold its standard deviation led to exclusion (considered as either dead cells or doublet cells). (**F**) In total, 10 arrays of 96 cells each were quality checked. Colored in gray are healthy and included cells, orange = dead cells/empty wells, red = presumed doublet cells, green cells = failed recombination primer test (e.g. tdTomato+ sorted cells that did not generate a Ct-value indicative for successful recombination), and blue cells represent the 12 intentionally selected negative control cells. (**G**) 865 cells plus 12 additional negative control cells passed quality control. (**H**) Excluded dead cells expressed control genes Actb and Gapdh at lower Log2Ex levels while sorted doublets show elevated Log2Ex values for these genes. (**I**) Healthy cells expressed on average 51.1 ±6.3% (mean ±SD) of the 192 assessed genes, while dead cells and doublets expressed 33.1 ±10.6% and 70.0 ±4.1%, respectively.

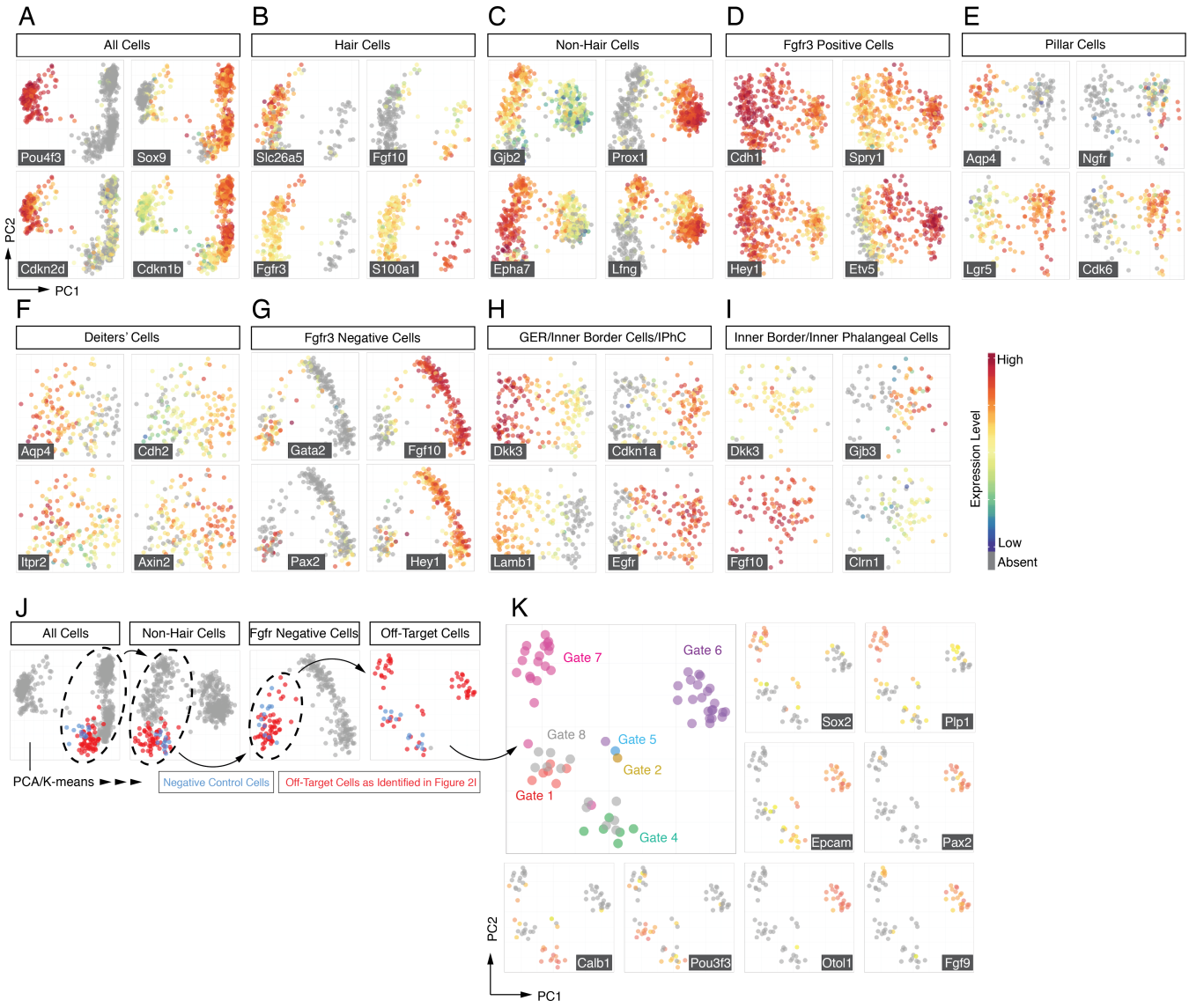


Figure S3. Organ of Corti Subpopulation Specific Gene Expression Patterns. Related to Figure 2.

(**A**) – (**I**) Shown is a sequential series of PCA analyses corresponding to Figure 2. Expression levels of 4 genes per iterative step that are listed in the corresponding variable factor maps shown in Figure 2.

(**J**) Location of off-target cells throughout all clustering levels shows that these excluded cells cluster together.

(**K**) Within the off-target population clustering of unknown cell populations reflects gate identity. Gate 7 sorted cells express Sox2 and Plp1 while Epcam is absent, which is indicative of spiral ganglion glial cell fate. Cells sorted from gate 6 express Epcam but Sox2 is not detected and

they may correspond to a population of Lgr5-GFP expressing cells found in the lateral wall (Chai et al., 2011). The negative control cells cluster with gate 1 and 4 sorted cells. Their cellular identities remain to be determined.

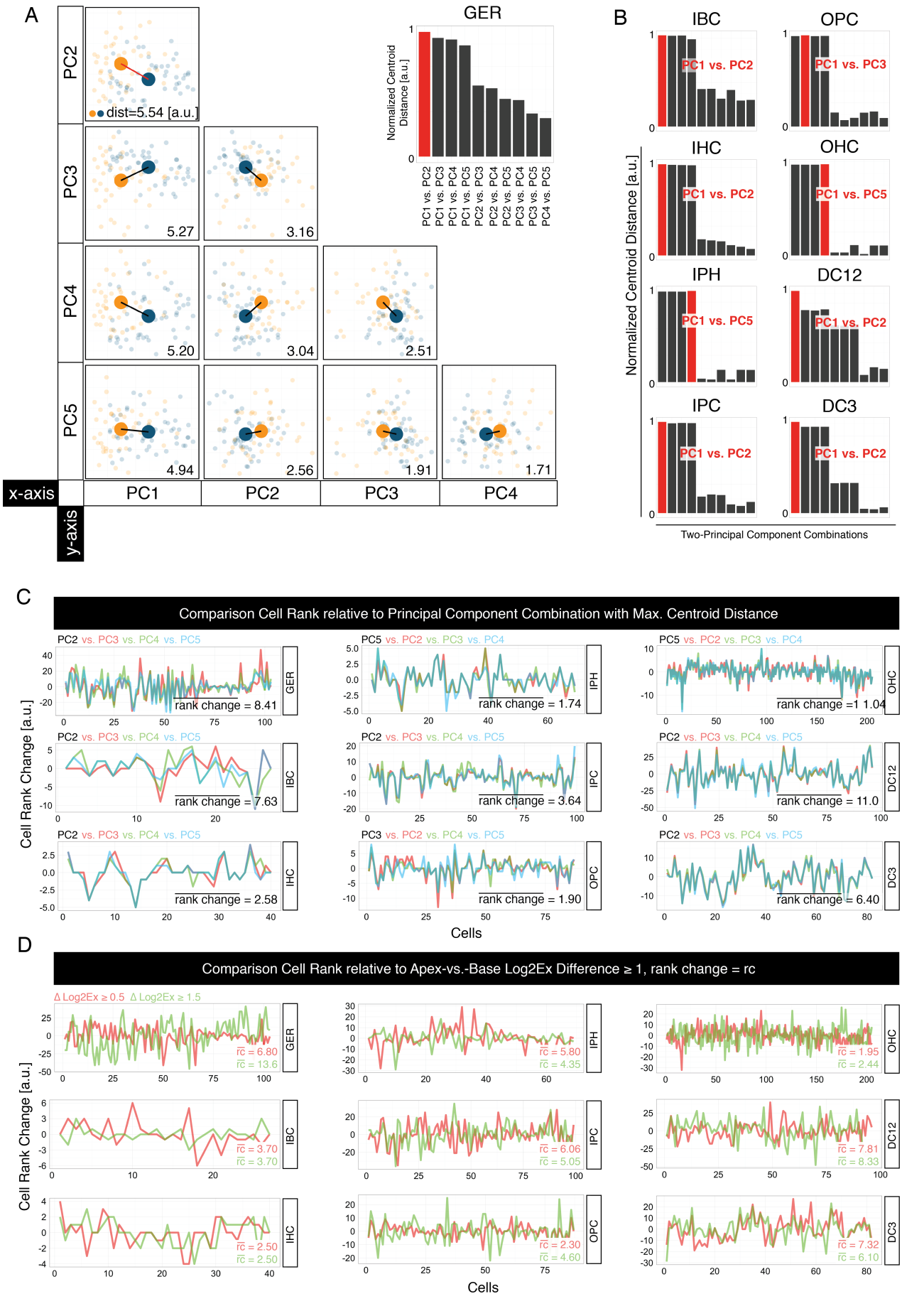

A
PC2
PC3
PC4
PC5
dist=5.54 [a.u.]
5.27
3.16
5.20
3.04
2.51
4.94
2.56
1.91
1.71
x-axis
y-axis
PC1
PC2
PC3
PC4
GER
Normalized Centroid Distance [a.u.]
PC1 vs. PC2
PC1 vs. PC3
PC1 vs. PC4
PC1 vs. PC5
PC2 vs. PC3
PC2 vs. PC4
PC2 vs. PC5
PC3 vs. PC4
PC3 vs. PC5
PC4 vs. PC5
B
IBC
OPC
IHC
OHC
IPH
DC12
IPC
DC3
PC1 vs. PC2
PC1 vs. PC3
PC1 vs. PC2
PC1 vs. PC5
PC1 vs. PC5
PC1 vs. PC2
PC1 vs. PC2
PC1 vs. PC2
Normalized Centroid Distance [a.u.]
Two-Principal Component Combinations
C
Comparison Cell Rank relative to Principal Component Combination with Max. Centroid Distance
PC2 vs. PC3 vs. PC4 vs. PC5
rank change = 8.41
GER
PC5 vs. PC2 vs. PC3 vs. PC4
rank change = 1.74
IPH
PC5 vs. PC2 vs. PC3 vs. PC4
rank change =1 1.04
OHC
rank change = 7.63
IBC
rank change = 3.64
IPC
rank change = 11.0
DC12
PC3 vs. PC2 vs. PC4 vs. PC5
rank change = 2.58
IHC
rank change = 1.90
OPC
rank change = 6.40
DC3
Cell Rank Change [a.u.]
Cells
D
Comparison Cell Rank relative to Apex-vs.-Base Log2Ex Difference ≥ 1, rank change = rc
Δ Log2Ex ≥ 0.5
Δ Log2Ex ≥ 1.5
rc = 6.80
rc = 13.6
rc = 5.80
rc = 4.35
rc = 1.95
rc = 2.44
rc = 3.70
rc = 3.70
rc = 6.06
rc = 5.05
rc = 7.81
rc = 8.33
rc = 2.50
rc = 2.50
rc = 2.30
rc = 4.60
rc = 7.32
rc = 6.10
Cell Rank Change [a.u.]
Cells

Figure S4. Centroid Distance Measurements. Related to Figure 3.

(**A**) Testing of 10 possible principal component combinations iterating PC1 through PC5 for the GER population. Normalized centroid distance is measured [arbitrary unit, a.u.] and plotted as a function of PC combination.

(**B**) Principal component combination reflecting the maximum normalized centroid distance is determined for all nine organ of Corti related cell populations.

(**C**) Rank changes of subpopulations-associated cells along apical-to-basal defined axis. For each of the nine populations the cells' change in position is shown when compared to the PC combination selected in (A), (B). Mean rank change combines all three average rank changes with respect to the subpopulation's total number of cells.

(**D**) Rank changes shown for varying ΔLog2Ex threshold values. Rank changes are calculated in comparison to each cell's rank based on a ΔLog2Ex threshold value ≥ ±1.

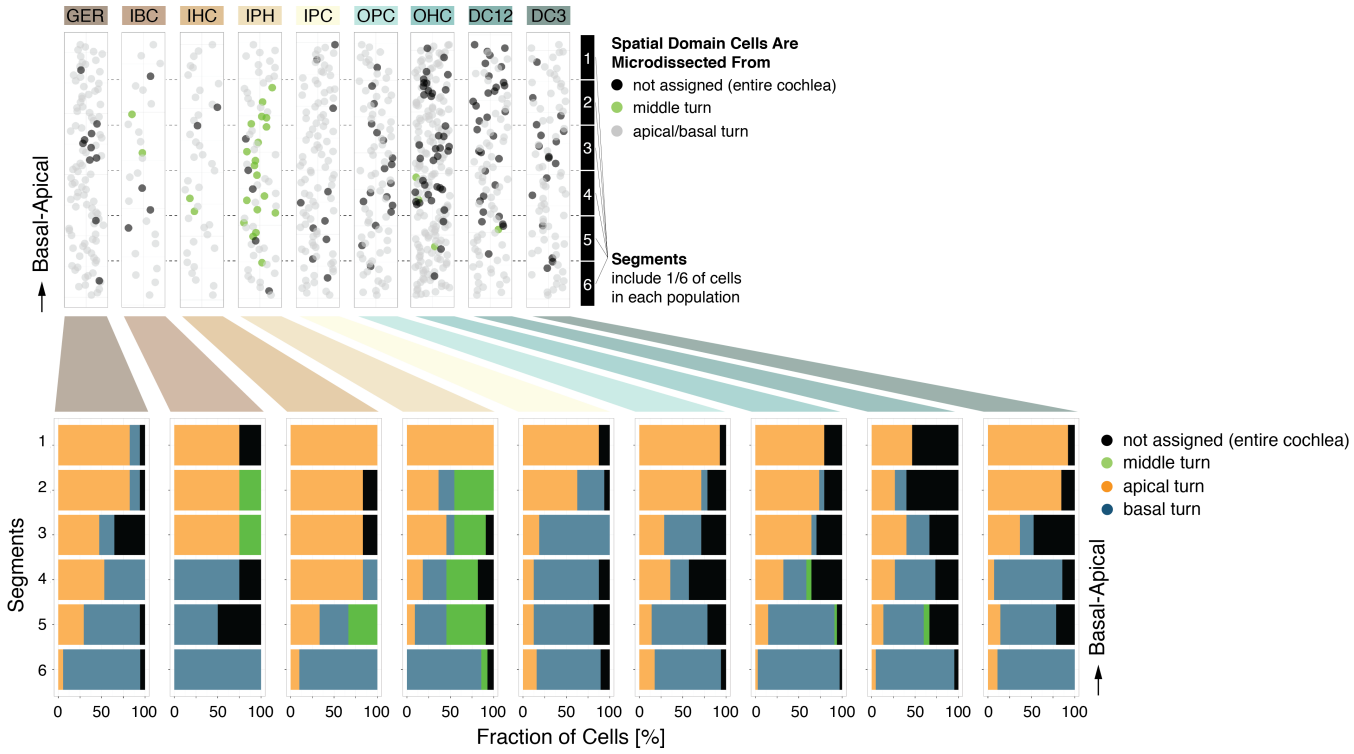
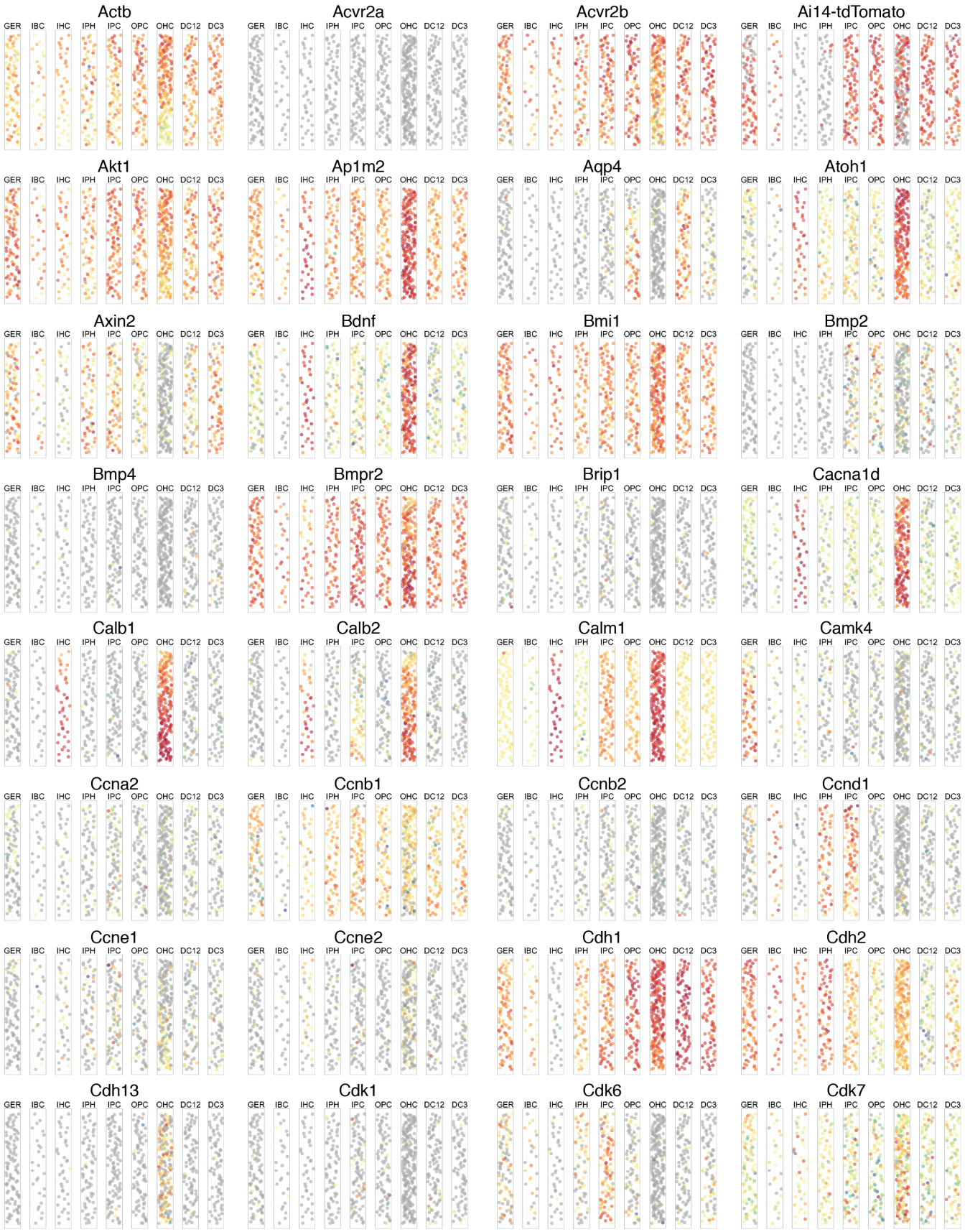
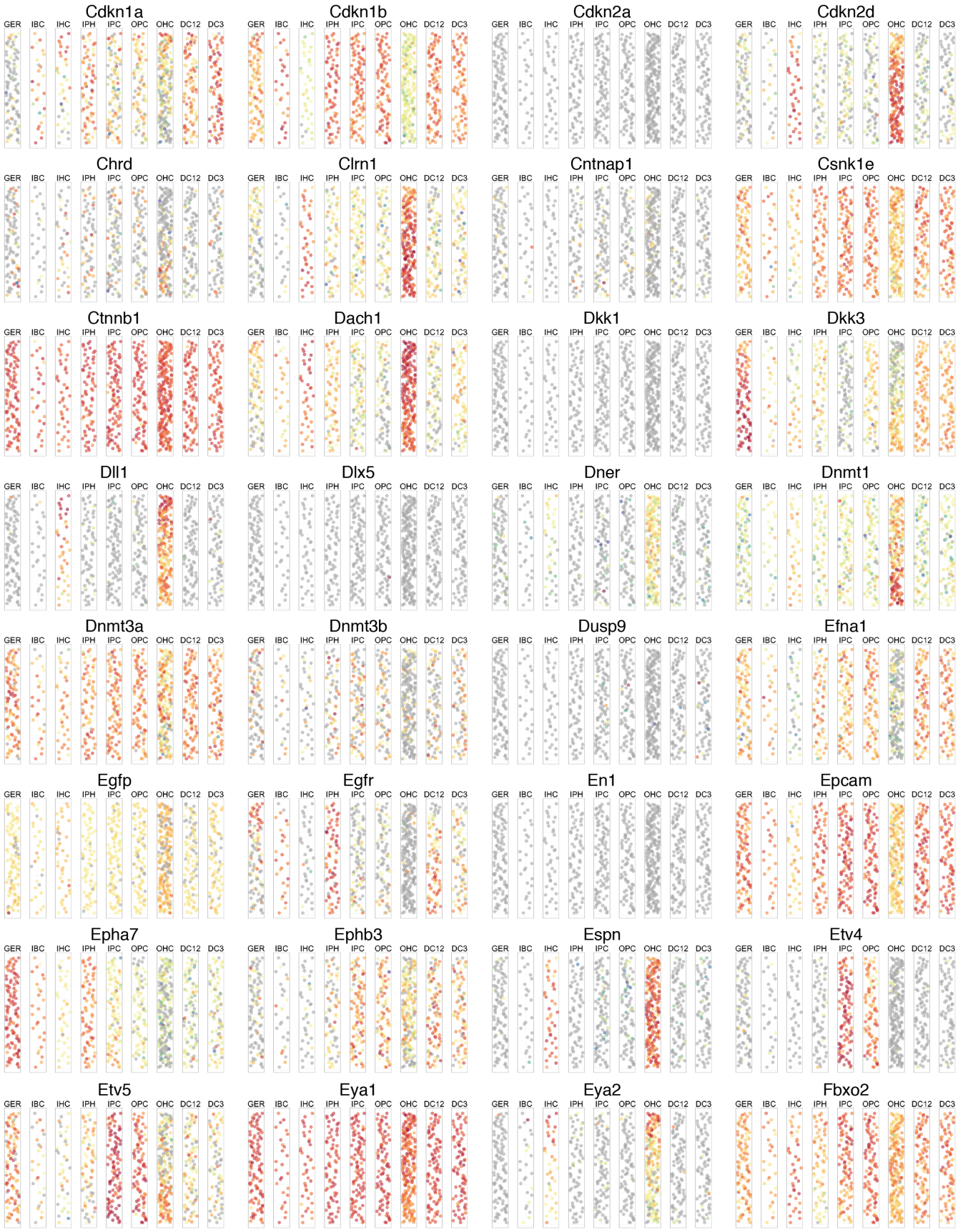
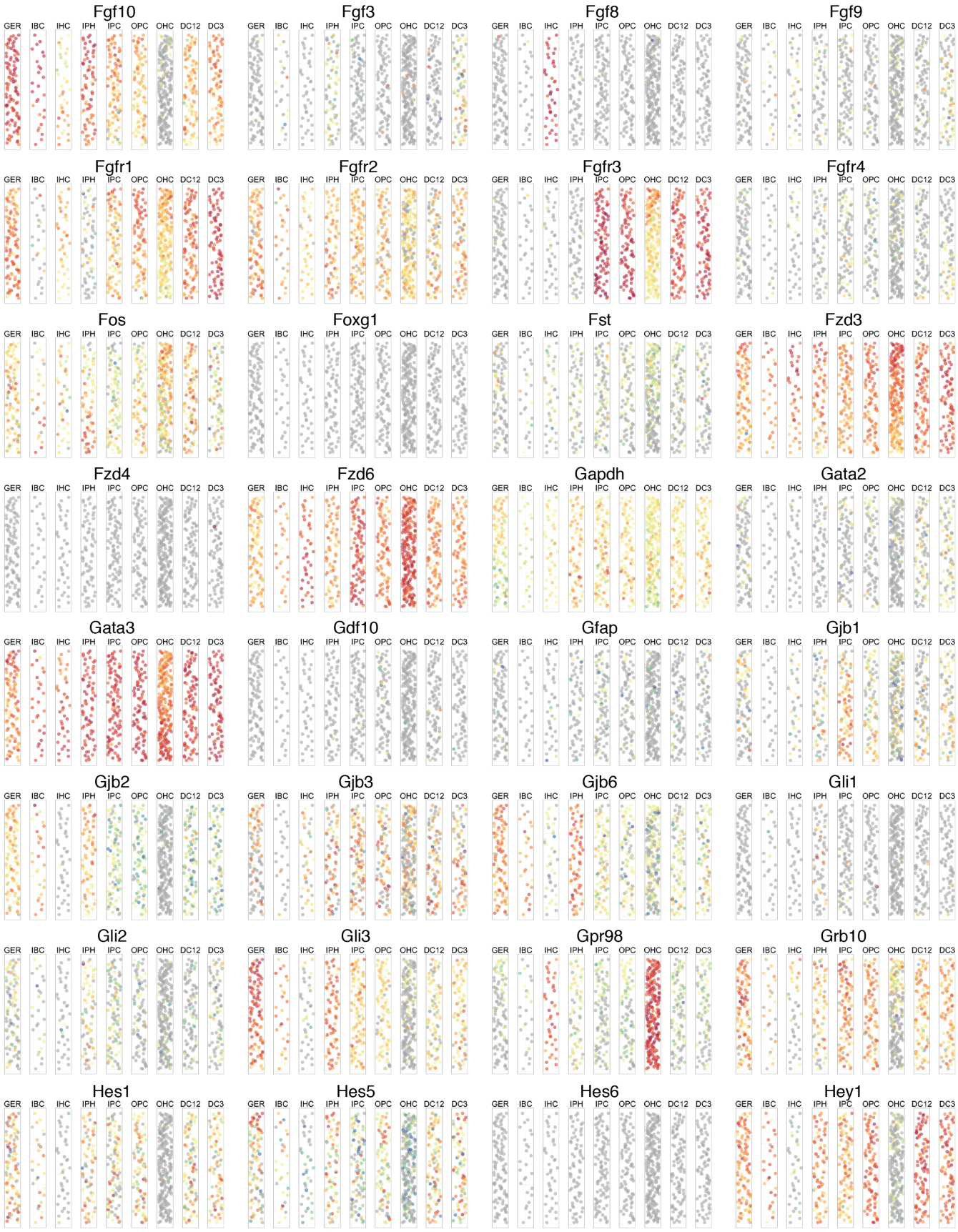


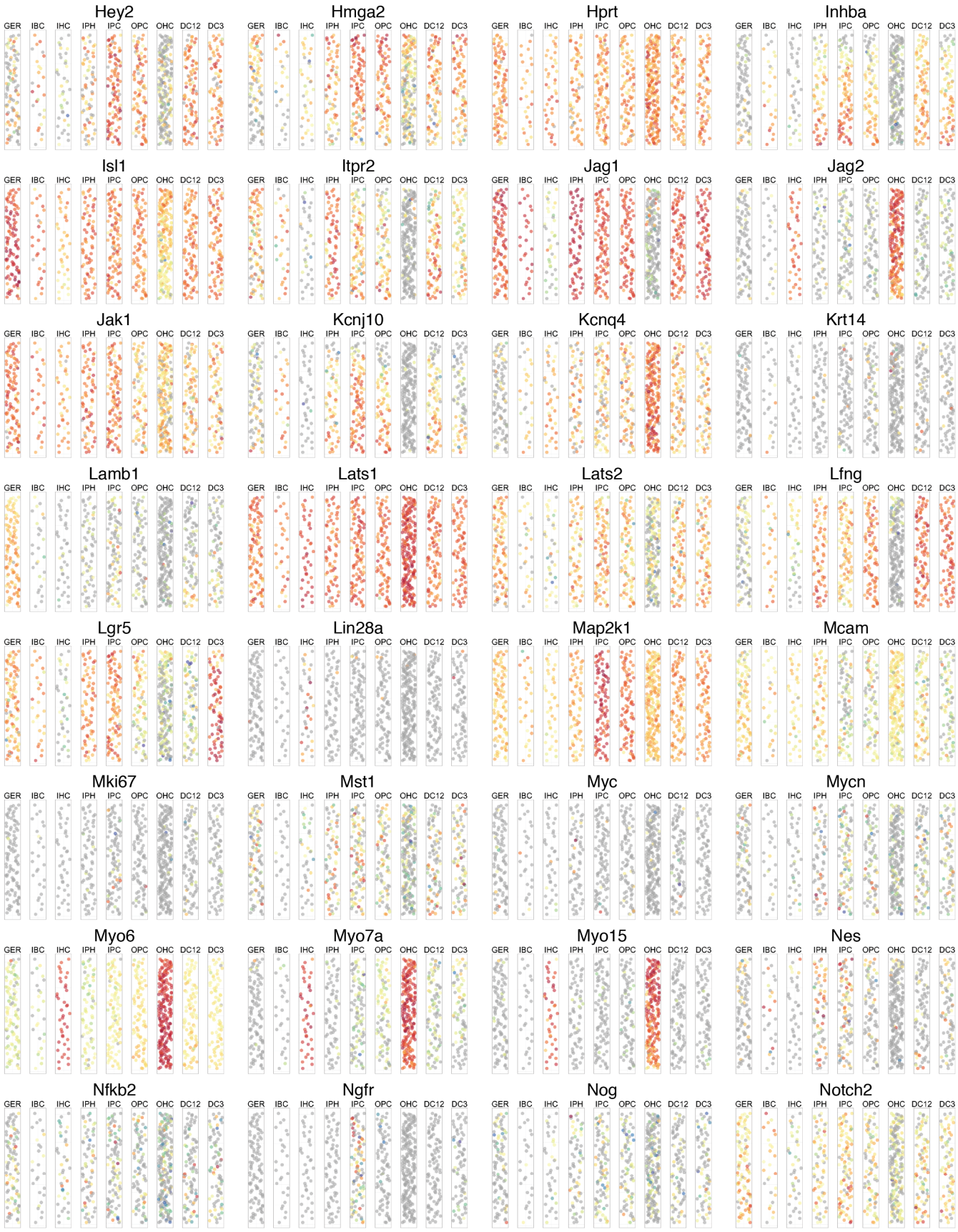
Figure S5. Projection of Cells Isolated from Middle Turn and Whole Cochlear Ducts onto the Organ of Corti Map. Related to Figure 3.

(**Upper**) Shown are the projected locations of control cells purposely isolated from middle turns (27 cells from Atoh1-nGFP/Fgfr3-iCreERT2/Ai14-tdTomato mouse line, corresponding to gate 1; green dots) and from whole cochlear ducts (138 cells from Sox2-GFP/Fgfr3-iCreERT2/Ai14-tdTomato mouse line, corresponding to gates 2, 3 and 4; black dots). Instead of 3 segments, we subdivided each apex-base map into six segments with equal cell numbers, which allowed us to perform a more exact determination of cell location. (**Lower**) The diagram on the lower half of the figure shows the overall distribution of cells isolated from the entire cochlear duct, as well as middle, apical, and basal turns.


Actb
Acvr2a
Acvr2b
Ai14-tdTomato
Akt1
Ap1m2
Aqp4
Atoh1
Axin2
Bdnf
Bmi1
Bmp2
Bmp4
Bmpr2
Brip1
Cacna1d
Calb1
Calb2
Calm1
Camk4
Ccna2
Ccnb1
Ccnb2
Ccnd1
Ccne1
Ccne2
Cdh1
Cdh2
Cdh13
Cdk1
Cdk6
Cdk7
GER IBC IHC IPH IPC OPC OHC DC12 DC3


Cdkn1a
Cdkn1b
Cdkn2a
Cdkn2d
Chrd
Clrn1
Cntnap1
Csnk1e
Ctnnb1
Dach1
Dkk1
Dkk3
Dll1
Dlx5
Dner
Dnmt1
Dnmt3a
Dnmt3b
Dusp9
Efna1
Egfp
Egfr
En1
Epcam
Epha7
Ephb3
Espn
Etv4
Etv5
Eya1
Eya2
Fbxo2
GER IBC IHC IPH IPC OPC OHC DC12 DC3


Fgf10
Fgf3
Fgf8
Fgf9
Fgfr1
Fgfr2
Fgfr3
Fgfr4
Fos
Foxg1
Fst
Fzd3
Fzd4
Fzd6
Gapdh
Gata2
Gata3
Gdf10
Gfap
Gjb1
Gjb2
Gjb3
Gjb6
Gli1
Gli2
Gli3
Gpr98
Grb10
Hes1
Hes5
Hes6
Hey1
GER IBC IHC IPH IPC OPC OHC DC12 DC3


Hey2
Hmga2
Hprt
Inhba
Isl1
Itpr2
Jag1
Jag2
Jak1
Kcnj10
Kcnq4
Krt14
Lamb1
Lats1
Lats2
Lfng
Lgr5
Lin28a
Map2k1
Mcam
Mki67
Mst1
Myc
Mycn
Myo6
Myo7a
Myo15
Nes
Nfkb2
Ngfr
Nog
Notch2
GER IBC IHC IPH IPC OPC OHC DC12 DC3


Ntf3
Oc90
Ocm
Otof
Otol1
Otop1
Otor
Otx1
Otx2
Pax2
Pax8
Pcdh15
Pcgf2
Pcna
Pdfgd
Pdgfra
Plp1
Pou3f3
Pou4f3
Prox1
Ptc1
Ptc2
Ptprq
Pvalb
Rdx
Rpa2
S100a1
Shh
Six1
Six3
Six4
Slc1a3
GER IBC IHC IPH IPC OPC OHC DC12 DC3

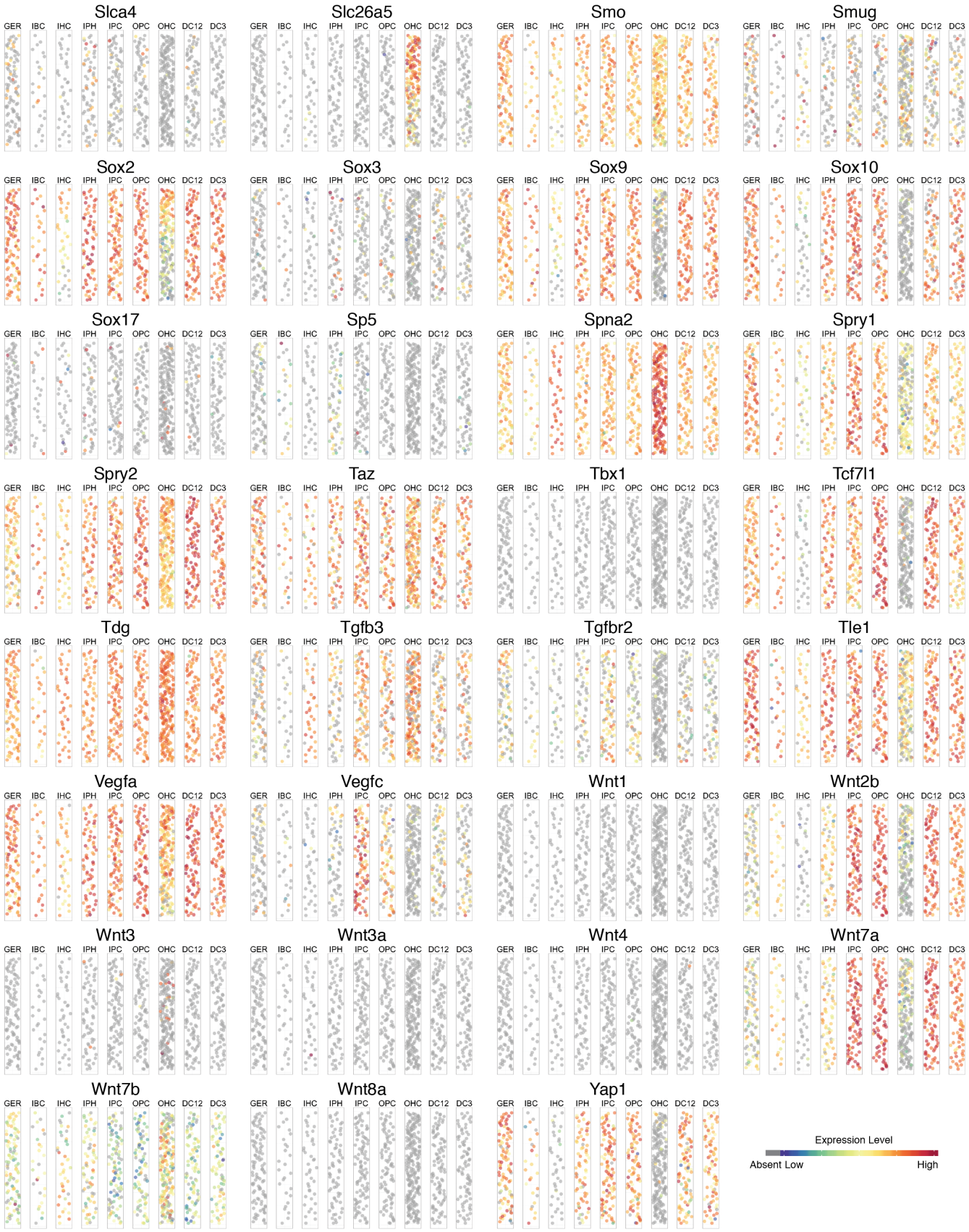

Slca4
Slc26a5
Smo
Smug
Sox2
Sox3
Sox9
Sox10
Sox17
Sp5
Spna2
Spry1
Spry2
Taz
Tbx1
Tcf7l1
Tdg
Tgfb3
Tgfbr2
Tle1
Vegfa
Vegfc
Wnt1
Wnt2b
Wnt3
Wnt3a
Wnt4
Wnt7a
Wnt7b
Wnt8a
Yap1
GER IBC IHC IPH IPC OPC OHC DC12 DC3
Expression Level
Absent Low
High

Figure S6. All Genes Projected onto the Organ of Corti Map. Related to Figure 4

The organ of Corti map for all 192 genes presented in alphabetical order.

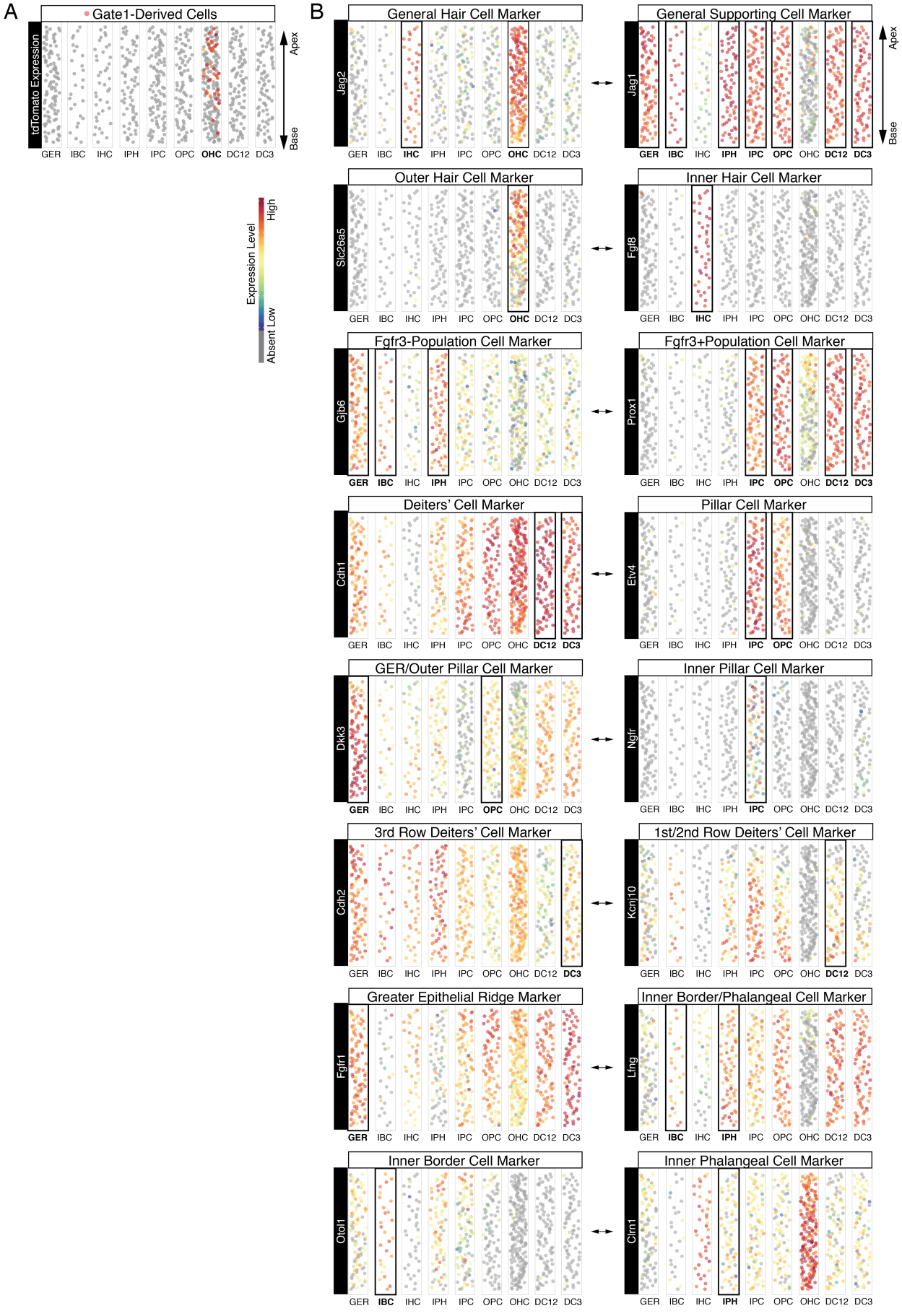

A
Gate1-Derived Cells
tdTomato Expression
Apex
Base
GER IBC IHC IPH IPC OPC OHC DC12 DC3
Expression Level
High
Absent Low
B
General Hair Cell Marker
Jag2
General Supporting Cell Marker
Jag1
Outer Hair Cell Marker
Slc26a5
Inner Hair Cell Marker
Fgf8
Fgfr3-Population Cell Marker
Gjb6
Fgfr3+Population Cell Marker
Prox1
Deiters' Cell Marker
Cdh1
Pillar Cell Marker
Etv4
GER/Outer Pillar Cell Marker
Dkk3
Inner Pillar Cell Marker
Ngfr
3rd Row Deiters' Cell Marker
Cdh2
1st/2nd Row Deiters' Cell Marker
Kcnj10
Greater Epithelial Ridge Marker
Fgfr1
Inner Border/Phalangeal Cell Marker
Lfng
Inner Border Cell Marker
Otol1
Inner Phalangeal Cell Marker
Clrn1

Figure S7. Projections of Selected Candidate Genes onto the Organ of Corti Map. Related to Figures 2 and 4.

(**A**) In the Fgfr3-CreERT2/Ai14-tdTomato/Atoh1-nGFP line only OHCs recombine as visualized by tdTomato expression.

(**B**) Shown maps reveal the conceptual differences between using classic cell type specific marker genes and clustering of populations of cells based on comprehensive analysis of expression of 192 genes. The presented order (from top to bottom) corresponds to sub-cluster levels shown in Figure 2 and illustrates expression of select genes. Most contributing genes are presented as corresponding pairs highlighting expression differences between outlined populations. For example, projection of Jag1 and Jag2 visualizes supporting cell and hair cell populations, but also shows lower level expression of these genes outside of their “assigned” clusters. As shown in Figure 2, the “markers” Jag1 and Jag2 are only two genes of all 192 analyzed genes that ultimately contribute to the segregation of the two populations. Toward the bottom of the figure panel, segregation of inner border cells and inner phalangeal cells is shown. There is a lack of classic marker genes uniquely expressed in these cell types. Nevertheless, analyzing quantitative expression differences of 192 genes in parallel was sufficient to separate both populations. Otol1 and Clrn1 contributed to this separation that ultimately was based on the overall expression analysis of all genes. Although these genes are not classic markers for either of the populations, they were identified as contributors to the separation by variable factor map analysis (Figure 2K) among other genes.

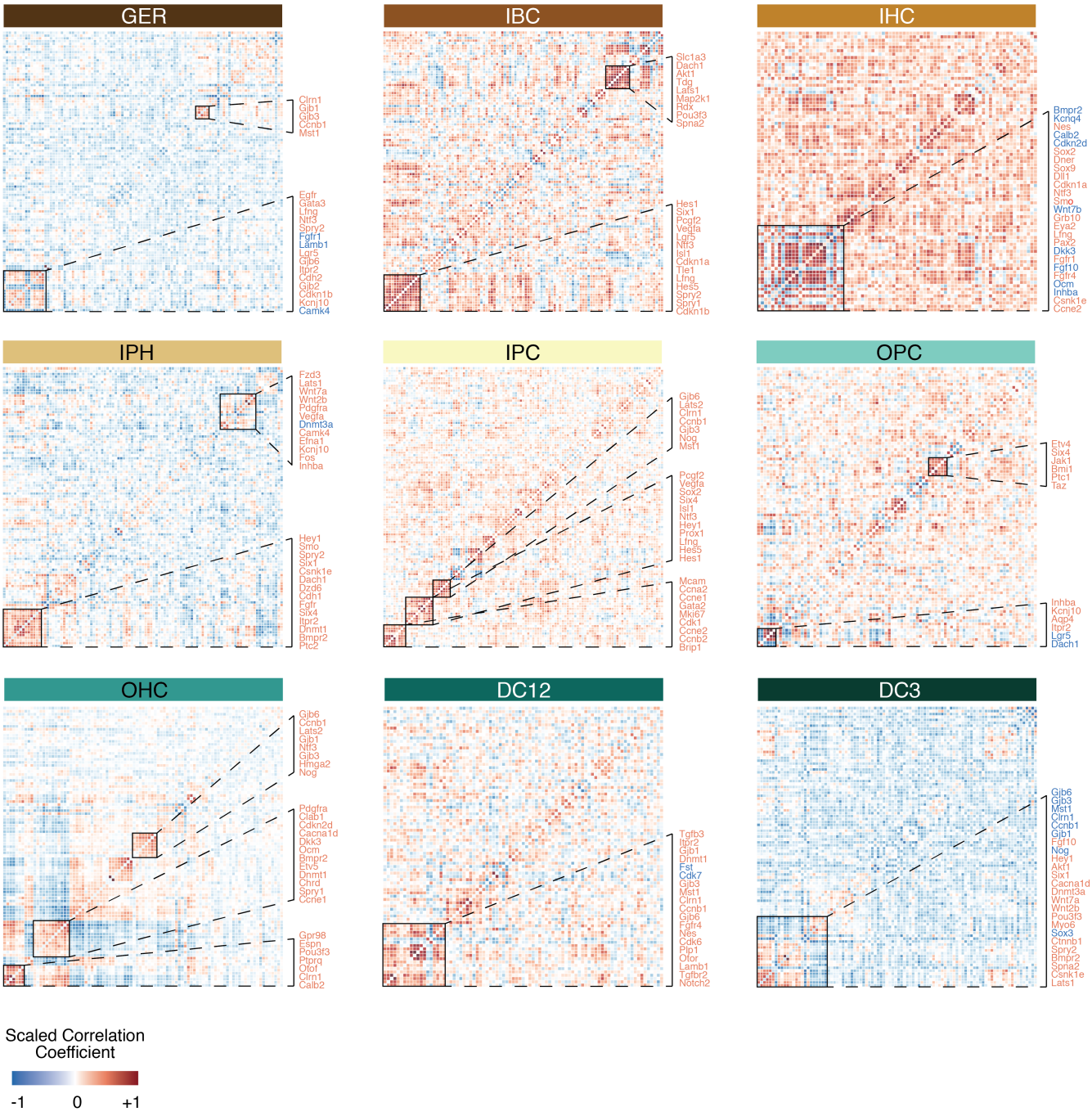


Figure S8. Gene correlation analysis for each of the nine classified organ of Corti populations. Related to Figures 2 and 7.

For each population genes across all cells were correlated with each other (Pearson's method), clustered together, and visualized in heatmap format. Only genes that are expressed in at least 5% of cells and have a variance greater than 0.5, were included in the respective analysis. For each population representative sample gene groups are shown. Color-code is scaled with red indicating high positive correlation and blue indicating high negative correlation.

Supplemental Table Legends

Table S1. Master Table of 877 Cells with Normalized and Quality Controlled Expression Data for all 192 Genes. Related to Figures 2 - 7.

Data are divided into metadata and numerical data. 5 internal control genes were excluded from PCA and *k*-means clustering. 8 additional genes were excluded from cluster calculations for technical reasons such as potential amplification inconsistencies on some plates.

Table S2. Assay Composition and Published Data on Gene Expression. Related to Figures 2 - 7

Listed are all genes used in the study (column A), pathway relations, marker identities and references (PubMed ID (PMID) numbers) for expression patterns and gradients. Correlation of published expression patterns/gradients with data presented in the organ of Corti map format (as shown in Figure S6): green = confirmed; yellow = not contradictory; red = contradictory.

Table S3. Gradually Expressed Genes considering all 9 Organ of Corti Cell Populations. Related to Figures 6 and 7.

Listed are genes with p-values <0.05 in at least one of the three comparisons (Apex versus Middle, Middle versus Base, Apex versus Base).

Table S4. Correlation Coefficients of Gene-to-Gene Pearson's Correlation Analysis for all 9 Organ of Corti Cell Populations. Related to Figures 2 and 7.

Listed are correlation coefficients of all genes that are expressed in at least 10% of each respective cell population.

Supplemental Experimental Procedures

Principal component analysis, *k*-means clustering, hierarchical clustering, statistical analysis and data visualization was performed with R statistical computing software v. 3.1.1. Packages utilized in this study included FactoMineR, kmeans, and heatmap.2. In Figures 2 and S3 principal components were calculated based on feature space that encompassed genes that significantly contributed to the first five components ($p<0.05$). *K*-means algorithm was conducted with one thousand iterations (argument *iter.max* = 1000). Hierarchical clustering was based on Pearson's correlation coefficients with a distance measure method *ward.D2*.

Cells along one-dimensional trajectory PC space were computed based on genes that were differentially expressed between apical-derived and basal-derived cells (ΔLog2Ex values $> \pm 1$). For both axis-associated cell groups the centroids were calculated using the mean coordinates of each group-associated cell in compressed 2D space. Data was rotated by first normalizing each dimensions length (Frobenius norm) and second rotating the data space around the midpoint of each population's 2D image. Successful rotation resulted in the apical-to-basal defined vector (centroid-combining direction) to stay orthogonal to the original direction of principal component 1. Next, PC1 score values for all cells in each population were converted to rank order values (i.e., rank 1 is assigned to the cell with the lowest PC1 score value, rank 2 is assigned to the cell with the second lowest PC1 score value, etc.). Finally for improved visualization purposes the data were randomly spread around a pseudo-axis perpendicular to the abscissa using the *position_jitter* argument in *ggplot2* (avoids overplotting of data). As a result, the absolute position of each cell along the abscissa axis varies randomly, whereas the absolute normalized position of each cell along the ordinate apex-to-base axis stays constant reflecting each cell's rank order position.

Data was fitted by first projecting each cell's position along the defined apical-to-basal direction and then by scaling the cells' relative positions between values of 0 and 1 for comparability between cell populations. Using local polynomial regression fitting with a degree number of polynomials of 2 the expression values (Log2Ex) were fitted along the apical-basal axis.

Statistical analyses in Figure 4 were performed using non-parametric ANOVA test (Kruskal-Wallis) with subsequent post-hoc test for statistical significance using Dunn's method. P-values were corrected for multiple comparisons (Bonferroni approach). Statistical analyses in Figure 2 were performed using Mann-Whitney test. P-values were corrected for multiple comparisons.

Supplemental References

Chai, R., Xia, A., Wang, T., Jan, T.A., Hayashi, T., Bermingham-McDonogh, O., and Cheng, A.G. (2011). Dynamic expression of Lgr5, a Wnt target gene, in the developing and mature mouse cochlea. Journal of the Association for Research in Otolaryngology : JARO *12*, 455-469.