

Current Biology

Epithelial planar bipolarity emerges from Notch-mediated asymmetric inhibition of Emx2

--Manuscript Draft--

Manuscript Number:	CURRENT-BIOLOGY-D-19-01462R2
Full Title:	Epithelial planar bipolarity emerges from Notch-mediated asymmetric inhibition of Emx2
Article Type:	Report
Corresponding Author:	Hernan Lopez-Schier, PhD Helmholtz Zentrum München Neuherberg - Munich, GERMANY
First Author:	Hernan Lopez-Schier, PhD
Order of Authors:	Hernan Lopez-Schier, PhD Eva L Kozak Jeronimo R Miranda-Rodriguez Heiko Lickert Aleksandar Janjic Anika Boettcher Wolfgang Enard Fabian Theis Subarna Palit Anika Böttcher
Abstract:	<p>Most plane-polarized tissues are formed by identically oriented cells[1-2]. A notable exception occurs in the vertebrate vestibular system and lateral-line neuromasts, where mechanosensory hair cells orient along a single axis but in opposite directions to generate bipolar epithelia[3-5]. In zebrafish neuromasts, pairs of hair cells arise from the division of a non-sensory progenitor[6-7], and acquire opposing planar polarity via the asymmetric expression of the polarity-determinant transcription factor Emx2[8-11]. Here, we reveal the initial symmetry-breaking step by decrypting the developmental trajectory of hair cells using single-cell RNA sequencing, diffusion pseudotime analysis, lineage tracing and mutagenesis. We show that Emx2 is absent in non-sensory epithelial cells, begins expression in hair-cell progenitors, and is down-regulated in one of the sibling hair cells via signaling through the Notch1a receptor. Analysis of Emx2-deficient specimens, in which every hair cell adopts an identical direction, indicates that Emx2 asymmetry does not result from auto-regulatory feedback. These data reveal a two-tiered mechanism by which the symmetric monodirectional ground state of the epithelium is inverted by deterministic initiation of Emx2 expression in hair-cell progenitors, and a subsequent stochastic repression of Emx2 in one of the sibling hair cells breaks directional symmetry to establish planar bipolarity.</p>

Epithelial planar bipolarity emerges from Notch-mediated asymmetric inhibition of Emx2

Eva L. Kozak¹, Subarna Palit², Jerónimo R. Miranda-Rodríguez¹, Aleksandar Janjic³, Anika Böttcher^{4,5,6}, Heiko Lickert^{4,5,6,7}, Wolfgang Enard³, Fabian J. Theis^{2,8,9} and Hernán López-Schier^{1,10,*}

¹ Helmholtz Zentrum München, Research Unit of Sensory Biology & Organogenesis, Ingolstädter Landstraße 1, D-85764, Neuherberg, Germany

² Helmholtz Zentrum München, Institute of Computational Biology, Ingolstädter Landstraße 1, D-85764, Neuherberg, Germany

³ Department Biology II, Anthropology and Human Genomics, Ludwig Maximilians University of Munich, Großhaderner Straße 2, D-82152, Martinsried, Germany

⁴ Helmholtz Zentrum München, Institute of Diabetes and Regeneration Research, Ingolstädter Landstraße 1, D-85764, Neuherberg, Germany

⁵ German Center for Diabetes Research, Ingolstädter Landstraße 1, D-85764, Neuherberg, Germany

⁶ Helmholtz Zentrum München, Institute of Stem Cell Research, Ingolstädter Landstraße 1, D-85764, Neuherberg, Germany

⁷ Technical University of Munich, School of Medicine, Ismaninger Straße 22, D-81675, Munich, Germany.

⁸ Technical University of Munich, School of Life Sciences Weihenstephan, Alte Akademie, D-85354, Freising, Germany

⁹ Technical University of Munich, Department of Mathematics, Boltzmannstraße 3, D-85748, Garching, Germany

¹⁰ Lead Contact

* corresponding author: hernan.lopez-schier@helmholtz-muenchen.de

Keywords: Planar polarity / Hair cells / Diffusion pseudotime / Single-cell

RNA sequencing

SUMMARY

Most plane-polarized tissues are formed by identically oriented cells[1-2]. A notable exception occurs in the vertebrate vestibular system and lateral-line neuromasts, where mechanosensory hair cells orient along a single axis but in opposite directions to generate bipolar epithelia[3-5]. In zebrafish neuromasts, pairs of hair cells arise from the division of a non-sensory progenitor[6-7], and acquire opposing planar polarity via the asymmetric expression of the polarity-determinant transcription factor *Emx2*[8-11]. Here, we reveal the initial symmetry-breaking step by decrypting the developmental trajectory of hair cells using single-cell RNA sequencing, diffusion pseudotime analysis, lineage tracing and mutagenesis. We show that *Emx2* is absent in non-sensory epithelial cells, begins expression in hair-cell progenitors, and is down-regulated in one of the sibling hair cells via signaling through the *Notch1a* receptor. Analysis of *Emx2*-deficient specimens, in which every hair cell adopts an identical direction, indicates that *Emx2* asymmetry does not result from auto-regulatory feedback. These data reveal a two-tiered mechanism by which the symmetric monodirectional ground state of the epithelium is inverted by deterministic initiation of *Emx2* expression in hair-cell progenitors, and a subsequent stochastic repression of *Emx2* in one of the sibling hair cells breaks directional symmetry to establish planar bipolarity.

RESULTS AND DISCUSSION

Spatiotemporal analysis of Emx2 expression in neuromasts

Epithelial planar polarity in the vestibular and lateral-line systems is functionally linked to body balance and rheotaxis. In the corresponding sensory epithelia, two populations of mechanoreceptive hair cells orient their apical hair bundles in opposite directions (Figures S1A-E)[3,5]. Despite decades of experimental work, the mechanism controlling such planar bipolarity remains unknown. An opening to this problem first arose from the identification of a homeodomain-containing transcription factor called Emx2[11]. The hair cells that express Emx2 have a direction of planar polarization opposite to those that do not[8-11] (Figure S1F). Therefore, identifying the origin of asymmetric Emx2 expression represents a singular opportunity to understand the emergence of epithelial planar bipolarity. To this end, we conducted single-cell sequencing (scRNASeq) of neuromasts (Figure 1A). Single-cell sequencing is a powerful tool to reveal the extent of cellular heterogeneity within organs[12-14]. It allows unbiased inference of cell-fate trajectories that, when coupled with anatomical references, can define with exceptionally-high accuracy epochs and frequencies of bifurcating cellular differentiation events in developing, regenerating or homeostatic tissues in the natural context[14]. To exclude the inner ear, we used trunks of larval zebrafish to isolate posterior lateral-line neuromast

cells by fluorescence activated cell sorting (FACS) from Tg[myo6b:actb1-EGFP] transgenic specimens expressing the green fluorescent protein EGFP in unipotent hair-cell progenitors (UHCPs) and hair cells[15], and from Tg[Et(krt4:EGFP)sqgw57A] larvae expressing EGFP in supporting cells[16] (Figures S1G-J). We combined the two samples at the level of the barcoded cDNA libraries, which were subsequently sequenced together using the 10X Chromium System. Standard preprocessing and quality control steps were applied to the raw data, including doublet removal, cell filtering based on number of detected genes, count depth and fraction of mitochondrial reads to ensure that only viable cells were used for downstream analyses. We identified neuromast cells based on *bona fide* marker transcripts (Figures S1K-L), and selected 1,167 cells for unbiased graph-based clustering[17]. This analysis generated four major cellular groups (Figures 1B-C), which were annotated *a posteriori* from the expression of pre-selected signature transcripts, including the cell-fate determinant transcription factor Atoh1a (UHCPs and immature hair cells)[18,19], the Notch ligand Dld (UHCPs and immature hair cells)[20], the molecular motor myo6b (immature and mature hair cells)[21,22], the cytosolic calcium buffers pvalb8, s100t and s100s (mature hair cells)[23,24], as well as tekt3[22,26] and tmc2b[27] (mature hair cells), sost[7], klf17, stm and si:dkey-205h13.2[24,28] (non-sensory supporting cells) (Figures 1C-D). A fifth cluster with a lower count depth, higher fraction of mitochondrial genes and fewer transcripts was defined as damaged cells (Figures 1C and S2A-B). Merging our dataset with

a previously-published dataset[25] corroborated our annotations (Figures S2C-F). Because *Emx2* is a transcription factor, we hypothesized that differential gene expression may correlate with hair-cell direction of planar polarization. However, unsupervised clustering using transcriptional profiles did not produce binary separation of opposite-polarized hair cells (Figure 1C), suggesting that a hair-cell's choice of polarity direction occurs under subtle differences in gene expression that were undetectable in our dataset, naturally-high levels of transcriptional noise, or post-transcriptionally. However, we did detect several genes that correlated or anti-correlated with *Emx2* expression (Figure S2G, Table S1). The product of these genes may play a role in planar polarization, or in other known differences between hair cells of opposing polarities, including innervation[9,10] and mechanosensitivity[27,29]. As differentially expressed genes were numerous and may be expressed downstream of *Emx2*, they were not systematically analyzed further because the specific aim of this study is to address the mechanism underlying symmetry-breaking upstream of *Emx2*.

Next, we assessed the spatiotemporal dynamics of *Emx2* expression by tracing the developmental trajectory of hair cells, starting from non-sensory supporting cells to mature hair cells via the intervening UHCPs and immature stages. We applied a method called diffusion pseudotime analysis (DPT), which regards single-cell transcriptomic data as static snapshots of

a dynamic process, defining cellular states and their transitions[13,14,30] (Figure 1E). DPT shows that *Emx2* is not expressed in non-sensory supporting cells, but is co-expressed with *Atoh1a* along the pseudotime trajectory, suggesting that *Emx2* is first transcribed by UHCPs (Figures 1F-G). We corroborated this finding using ultrasensitive whole-mount *in situ* hybridization chain reaction[31] and immunostaining[8] (Figures 1H-I and S2H-I). Assaying *Emx2* expression *in toto* in homozygous *Atoh1a*-mutant specimens, which contain neuromasts with a normal set of non-sensory supporting cells but never produce UHCPs or hair cells[18,19,32], revealed that in contrast to the wild type *Atoh1*-deficient neuromasts completely lack *Emx2*-expressing cells (Figures S2J-K). Gene-expression changes across pseudotime also showed that *Emx2* expression in UHCPs is substantially more uniform than in mature hair cells (Figure 1F), suggesting mosaic down-regulation of *Emx2* with progressive hair-cell maturity. Together with the observations that *Emx2* is sufficient to direct hair-cell orientation[8] and that non-sensory supporting cells are *Emx2*-negative, these results allow us to conclude that the default state of the neuromast epithelium is to coherently orient every hair cell towards to animal's head. As *Emx2* begins expression, UHCPs acquire the potential to develop tail-facing polarity. We found that in wild-type and in *Emx2*-deficient neuromasts, hair cells are born in pairs. Also, we never saw increased hair-cell death or extrusion in the mutants (Video S1), demonstrating that monopolarity in *Emx2*-deficient neuromasts is not caused by a biased

production or maintenance of head-facing hair cells. Therefore, for bipolarity to occur, the intrinsic symmetry of the system must be broken after UHCP division but before hair-cell polarity direction is irreversibly established by the core planar-polarity pathway[8].

Notch1a activity asymmetrically inhibits Emx2 expression

Our pathway enrichment analysis of the top 200-ranked genes for each cell cluster via Kyoto Encyclopedia of Genes and Genomes (KEGG) identified Notch signaling-related transcripts in UHCPs (Figure 2A). Also, several studies have firmly established that Notch signaling in neuromasts controls the production[7,16,20,25,33,34] and polarization[35] of hair cells. Thus, we hypothesized a role of Notch signaling in symmetry breaking, either by controlling Emx2 expression or by affecting Emx2 activity. Thus, we took a genetic approach to nullify Notch signaling using CRISPR/Cas9-mediated genome engineering. We have previously found that the Notch1a and Notch3 receptors are expressed in neuromasts[36]. Notch3 appears in two patches of non-sensory supporting cells on opposite sides of the neuromast, being absent from its center where hair cells are normally present[7,16,20,36]. By contrast, Notch1a is expressed more uniformly[36], including the UHCPs (Figure 2B). Therefore, we chose to mutagenize Notch1a and confirmed that our mutant allele is a loss-of-function based on the irregular boundaries between posterior somites, which is identical to the original loss-of-function allele *deadly seven*[37] (not shown).

Remarkably, Notch1a-mutant zebrafish developed neuromasts exclusively populated by tail-facing hair cells, all of which expressed Emx2 (Figures 2C-G, S3A-B). The Notch1a-mutant phenotype is qualitatively equivalent to that of constitutive expression of Emx2 in every hair cell[8], and double-mutants Emx2;Notch1a produced neuromasts exclusively populated by head-facing hair cells, a phenotype identical to Emx2^{-/-} and opposite to Notch1a^{-/-} (Figure 2H), demonstrating that Emx2 acts downstream of Notch1a. We were surprised by these results because previous work using global pharmacological suppression of Notch signaling by incubating larval zebrafish in the γ -secretase inhibitor DAPT showed neuromasts with a majority of head-facing hair cells[35], a phenotype opposite to the loss of Nocth1a. We confirmed the validity of this result by finding that DAPT treatment biases hair-cell orientation towards the animal's head (Figure 2I), and also showing that it reduces the number of Emx2-expressing hair cells to 32,6% (301 analyzed cells in 14 neuromasts) (Figures 2J and S3C). One explanation for the discrepancy between pharmacological inhibition of global Notch signaling and the loss of Notch1a is that different Notch receptors have opposite effects on Emx2 expression or activity. Alternatively, signaling specifically through Notch1a may control hair-cell planar polarity independently of Emx2. To test these possibilities, we analyzed neuromasts using the double transgenic line Tg[Brn3c:Gal4 ; UAS:NICD][39] to express a constitutively-active Notch1a intracellular domain in hair cells. In these specimens, fewer than normal hair cells

matured, but all were head-facing and, revealingly, none expressed Emx2 (Figures 2K-L and S3D), suggesting that Notch signaling acts upstream of Emx2. We also found that Notch1a mutants treated with DAPT only produced tail-oriented hair cells that were Emx2(+), opposite to DAPT treatment of wild-type specimens (Figures 2M-N and S3D). Altogether, these results reveal that Notch1a breaks the symmetry between sibling hair cells upstream of Emx2, and suggest that it acts downstream of another Notch receptor that is also affected by DAPT. The identity of the additional receptor is unknown. We must thus call it Notch-x (Nx), and hypothesize that it is necessary to initiate Emx2 expression in neuromasts. A corollary of the cumulative data is that the loss of Nx blocks the initiation of Emx2 expression, effectively preventing Notch1a from generating asymmetry via Emx2 inhibition in one of the sibling hair cells. Recently, however, an alternative explanation for the DAPT-generated phenotype was suggested by Jacobo and colleagues[38]. These authors concluded that DAPT acts exclusively through Notch1a, reducing but not eliminating the activity of this receptor. They reason that the incomplete reduction of Notch1a causes a partial inhibition of Emx2 and a polarity bias that is opposite to that resulting from the complete loss of Notch1a. Objectively, because the identity of Nx and partial inhibition of Notch1a by DAPT are hypothetical, both possibilities remain viable.

Anatomical and transcriptional asymmetries are separable

Symmetry-breaking in neuromasts takes two forms. One is genetic and is based on the maintenance of *Emx2* expression in one of the sibling hair cells. The second is anatomical, which is revealed by the opposing direction of hair-cell planar polarity. To ask if both forms of asymmetry always correlate, we closely analyzed the dynamics of *Emx2* expression by comparing scRNA-seq data from wild-type and *Emx2* mutant neuromasts using mcSCRBseq (Figures 3A, S11-J and 4A). After thresholding on outlier peaks, respectively while examining distribution of cells and genes, we measured 168 single cells with a median number of transcripts per cell of approximately 2,500. We detected *Emx2* transcripts in a statistically equivalent fraction of cells in wild-type and *Emx2* mutant samples (37% and 43%) (Figure 3A). We corroborated this finding using whole-mount fluorescent *in situ* hybridization, showing expression of *Emx2* in 49% of the hair cells of *Emx2*-mutant neuromasts (N=194 cells) (Figures 3B-C and S4B-C). Therefore, the asymmetric pattern of *Emx2* expression, and by extension Notch1a-mediated transcriptional symmetry-breaking, are not under auto-regulatory feedback by *Emx2* activity. Importantly, these findings reveal that the loss of anatomical asymmetry is not always coincident with a concurrent loss of transcriptional asymmetry.

In wild-type neuromasts only the anterior-located mature hair cells express *Emx2* and face the tail of the fish[8-10]. However, we found immature hair cells expressing *Emx2* in the anterior or the posterior aspect of the organ

(Figures 3D-F and S4D-E). Therefore, after being born, both the anterior and posterior hair cells are likely to maintain Emx2 expression, suggesting that the initial symmetry-breaking step is a local and stochastic decision between hair-cell siblings. It is important to highlight that this will inevitably lead to the production of pairs of hair cells polarized away from each other (Emx2(-) anterior hair cells oriented towards the animal's head) in 50% of the cases, and towards each other (Emx2(+)) anterior hair cells oriented towards the animal's tail) in the other 50%. Yet, this never occurs because sibling hair cells are eventually always polarized towards each other[5,6,8,15,16,33,35,38,40]. To explain this observation we take into account a cellular rearrangement process that we have previously discovered called planar cell inversion (PCI)[33]. PCI inverts the position of sibling hair cells across the neuromast dorsoventral midline[33,35,40]. PCI occurs after approximately half of UHCP divisions, only once, with an onset between 240 and 290 minutes[33], and is irreversible[33,35]. We suggest that PCI repositions hair cells according to their Emx2-expression status to ensure that every Emx2(+) hair cell is located on the anterior aspect of the neuromast to eventually face the tail of the fish, and *vice versa*. Testing this hypothesis experimentally will require the controlled disruption of the inversion process without affecting Notch signaling or Emx2 expression, which is currently not possible because the molecular mechanism that controls PCI is unknown.

On the basis of the above findings (Figure 4A), we propose a model for the emergence of epithelial planar bipolarity (Figure 4B). Although the direction of planar polarization of mature hair cells is irreversible once it is defined, neuromasts lose and produce hair cells during their entire life of the zebrafish[16], indicating that epithelial planar bipolarity must be constantly re-enacted. The neuromast non-sensory epithelium does not express Emx2, and it is therefore intrinsically monopolar. This conclusion is also supported by Emx2 mutant zebrafish, which develop neuromasts exclusively populated by head-facing hair cells. Emx2 begins expression in hair-cell progenitors (UHCPs), and is likely triggered by a deterministic signal from a Notch receptor, as suggested by DAPT-mediated global inhibition of Notch signaling. Mis-expression studies show that Emx2 is sufficient to impart tail-facing polarity to every cell[8]. Therefore, although planar polarity is not anatomically expressed in UHCPs, Emx2 expression effectively reverses the direction of their planar-polarization potential by 180°. Upon their birth, sibling hair cells progress through a symmetry-breaking process via Notch1a-mediated inhibition of Emx2 expression in one of the siblings. Importantly, the asymmetric expression pattern of Emx2 is not under feedback control through Emx2 activity, indicating that the genetic and anatomical asymmetries of the system are separable. The observation that initially Emx2(+) hair cells can be found at either side of the neuromast dorsoventral midline suggests that symmetry breaking is a stochastic local decision between sibling hair cells (Figure 4C). The model

suggest a role for the enigmatic planar cell inversion (PCI) process in correcting the stochastic Notch1a-mediated lateral inhibition of Emx2 expression, effectively placing all Emx2(+) hair cells on the rostral aspect of the neuromast epithelium, and the Emx2(-) hair cells on the caudal part (Figure 4C). This also explains how mirror symmetry may be established[33]. Finally, as maturing hair cells dissipate their gene-expression profile inherited from UHCPs[41], they stabilize their Emx2(+) or Emx2(-) expression pattern and their direction of planar polarization becomes irreversibly established[8]. Thus, symmetry-breaking must occur within a period of 240-290 minutes after UHCP divisions because it should happen before the onset of PCI[33]. Of note, Jacobo et al. suggest parallel activities of Notch1a signaling during planar polarization. One activity takes place via Emx2 and another independently of it (Jacobo et al. Fig.5C)[38]. Our data, however, is more consistent with a single activity of Notch1a in planar polarization, exclusively canalized through Emx2. Beyond these findings, we suggest that three important questions should be the focus of future work. First, what is the relationship between Emx2 and the core planar-polarity pathway, specifically in reversing the intrinsic monopolar potential of the epithelium[8,11]. Second, what are the intercellular processes that amplify the initial symmetry-breaking signal[42]. Third, which is the ligand for Notch1a[43]. The results presented here provide a framework for testing new hypotheses about epithelial planar polarity.

ACKNOWLEDGMENTS

We thank C.B. Moens (Fred Hutch, Seattle) for providing Atoh1a mutant fish before publication, K. Kindt (NIH, Bethesda) for the Emx2 mutant fish, and A.B. Steiner (Pace, New York) for the DNA clone with the *alpl* promoter. This research was supported by core funding (Helmholtz Association, Germany), and by BRAIN Initiative Grant 1U19NS104653 (NIH, U.S.A.) to H.L-S.

AUTHOR CONTRIBUTIONS

HL-S and ELK conceptualized the project. ELK wrote the first draft of the manuscript. ELK, SP and JRMR conducted experiments, analyzed and interpreted results. AJ and AB conducted scRNA sequencing. HL, WE and FT analyzed data. ELK and HL-S wrote the final paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Figure 1. Single-cell transcriptional analysis of neuromast cells reveals gene-expression dynamics during symmetry breaking

(A) Scheme of the single-cell transcriptional analysis workflow. (B) Heatmap showing the expression of the top 30 ranked genes of each Louvain cell cluster. (C) UMAP plot showing unbiased graph-based Louvain clustering of 1,167 neuromast cells into five clusters: support cells, UHCs, young HCs (including immature HCs), mature HCs and damaged HCs. (D) Violin plots showing expression levels of selected genes in each cluster. (E) (left) Diffusion map showing pseudotime. DPT root cell was fixed as the support cell with the maximum DC1 or minimum DC2 value from prior knowledge.

Dark blue corresponds to the initial start of the differentiation process and yellow indicates end point. (right) Louvain clusters of the same diffusion map as shown in the left panel. The diffusion pseudotime was computed using a random-walk-based cell-to-cell distance metric. (F) Gene expression dynamics of selected markers along the differentiation path in the inferred trajectory from support cells to mature hair cells indicating continuous transition. Cells are ordered using diffusion pseudotime. Expression is shown as a running average over 100 cells along the pseudotime scaled to the maximum observed level per gene. (G) UMAP plot of *Emx2* expression. Log-transformed gene expression used for visualization. (H) Fluorescent in situ hybridization of a horizontal neuromast in the transgenic line *myo6b:EGFP* showing HCs in green, nuclei in blue (DAPI staining) and *Emx2* in red (FISH) in 3 dpf wt larvae. Scale bar is 10 μm . The white arrowhead is highlighting a UHCP, expressing *Emx2*. Insets at the bottom left show the planar polarization of the hair cells ($z=-5\mu\text{m}$), 2x magnification. (I) Immunohistochemical staining of *Emx2* in wt horizontal neuromasts in the transgenic line *myo6b:EGFP* showing HCs in green, nuclei in blue (DAPI staining) and *Emx2* in red (antibody staining) in 2dpf old larvae. The white arrowhead is highlighting a cell (UHCP), expressing *Emx2*. Scale bar is 10 μm . Inset on the bottom left shows the planar polarization of the hair cells ($z=-5\mu\text{m}$), 2x magnification. See also Figures S1, S2 and Table S1.

Figure 2. Notch1a is controls hair-cell planar polarity via *Emx2*

(A) KEGG pathway enrichment analysis of the top 200 ranked genes in UHCPs. The size of the data point represents the number of differentially expressed genes that constitute the given KEGG term. The color indicates the FDR corrected p-value and the x-axis is the ratio of background genes that are contained the given KEGG term. The background genes refer to all

the genes in the dataset, after filtering. **(B)** UMAP plot showing expression of *Notch1a* in wt neuromast cells. Log-transformed gene expression used for visualization. **(C)** Immunohistochemical staining of a horizontal neuromast in the transgenic line *myo6b:EGFP* showing HCs in green, nuclei in blue (DAPI) and *Emx2* in red (antibody) in 2 dpf wt and *Notch1a* mutant larvae. Scale bars are 10 μm . Insets at the bottom left show the planar polarization of the hair cells, with $z=-5 \mu\text{m}$ and 2x magnification. wt n=30 (neuromasts); *Notch1a* $^{-/-}$ n=19 (neuromasts). **(D)** Fluorescent in situ hybridization of a horizontal neuromast in the transgenic line *myo6b:EGFP* showing HCs in green, nuclei in blue (DAPI staining) and *Emx2* in red (FISH) in 3 dpf wt and *Notch1a* mutant larvae. Scale bars are 10 μm . Insets at the bottom left show the planar polarization of the hair cells, with $z=-5 \mu\text{m}$ and 2x magnification. wt n=39 (neuromasts); *Notch1a* $^{-/-}$ n=51 (neuromasts). **(E-I)** Stereociliary hair bundle orientations in the horizontal neuromast of the wt, n=24 neuromasts **(E)**, *Emx2* mutant, n= 18 neuromasts **(F)**, *Notch1a* mutant, n=18 neuromasts **(G)**, double *Emx2*- and *Notch1a* mutant n=51 neuromasts **(H)**, wt treated with DAPT, n=35 neuromasts **(I)**, larvae. **(J)** Immunohistochemical staining of horizontal neuromasts in the wt *myo6b:EGFP* transgenic larvae treated with DAPT. The transgenic line *myo6b:EGFP* is showing HCs in green, nuclei in blue (DAPI staining) and *Emx2* in red (antibody staining) in 4 dpf wt larvae. Insets at the bottom left show the planar polarization of the hair cells, with $z=-5 \mu\text{m}$, n=20 neuromasts. **(K)** Stereociliary hair bundle orientations in the horizontal neuromast of the *brn3c:gal4*, *UAS:NICD* larvae, n=19 neuromasts. The orientation of cells is depicted with circular histograms, 180 corresponding to anterior direction in blue and 0/360 corresponding to the posterior direction in orange. **(L)** Immunohistochemical staining of horizontal neuromasts in *brn3c:gal4*, *UAS:NICD* double transgenic lines, showing nuclei in blue (DAPI staining), *Emx2* in red (antibody staining) and hair bundles in green (phalloidin staining) in 3 dpf wt larvae, n=10 neuromasts.

Insets at the bottom left show the planar polarization of the hair cells, with $z=-5\ \mu\text{m}$. **(M)** Stereociliary hair bundle orientations in the horizontal neuromast of Notch1a mutant treated with DAPT larvae, $n=34$ neuromasts. The orientation of cells is depicted with circular histograms, 180 corresponding to anterior direction in blue and 0/360 corresponding to posterior direction in orange. **(E-I, K, M)** Percentages of anterior (blue) and posterior (orange) oriented cells are written in the bottom of each panel. **(N)** Immunohistochemical staining of horizontal neuromasts in the Notch1a mutant, myo6b:EGFP transgenic larvae treated with DAPT. The transgenic line myo6b:EGFP is showing HCs in green, nuclei in blue (DAPI staining) and Emx2 in red (antibody staining) in 4 dpf wt larvae. Insets at the bottom left show the planar polarization of the hair cells, with $z=-5\ \mu\text{m}$, $n=41$ neuromasts. See also Figure S3 and Video S1.

Figure 3. Emx2 expression in hair cells of wt and mutant larvae

(A) Violin plots showing the expression of selected genes in hair cells from wt and Emx2-mutant larvae and revealing negligible differences in Emx2 expression between the wt and Emx2 mutants. **(B)** Immunohistochemical staining of a horizontal neuromast in the transgenic line myo6b:EGFP showing HCs in green, nuclei in blue (DAPI staining) and Emx2 in red (antibody) in 2 dpf wt and Emx2-mutant larvae. Scale bars are $10\ \mu\text{m}$. Insets at the bottom left show the planar polarization of the hair cells, with $z=-5\ \mu\text{m}$ and 2x magnification. Wt $n=20$ neuromasts, Emx2^{-/-} $n=26$ neuromasts. **(C)** Fluorescent in situ hybridization of a horizontal neuromast in the transgenic line myo6b:EGFP, showing HCs in green, nuclei in blue (DAPI) and Emx2 in red (FISH) in 5 dpf wt and Emx2-mutant larvae. Scale bars are $10\ \mu\text{m}$. Insets at the bottom left show the planar polarization of the hair cells, with $z=-5\ \mu\text{m}$ and 2x magnification. Wt $n=8$ neuromasts, Emx2^{-/-} $n=16$ neuromasts. **(D-E)** Immunohistochemical staining of a horizontal

neuromast in the wt transgenic line *myo6b:EGFP* showing HCs in green, nuclei in blue (DAPI staining) and *Emx2* in red (antibody) in 2 dpf wt larvae. Insets at the bottom left show the planar polarization of the hair cells, with $z = -5 \mu\text{m}$ and 2x magnification. The white arrow highlights a pair of immature sibling hair cells, which have not yet formed the hair bundles. Note that the *Emx2*-expressing immature hair cell is on the anterior side (D) or on the posterior side (E). (F) Quantification of events, when the *Emx2* was detected in the anterior or posterior or none of the sibling immature hair cells. Only the immature hair cells were used for the quantification, for definition of immature hair cells see also Figure S4D. Number of pairs of immature hair cells $n = 55$. See also Figure S4.

Figure 4. The model of epithelial planar bipolarization in neuromasts

(A) Schematic summary of phenotypes in neuromasts showing the wild type and in various defective configurations indicated above each scheme. Magenta hair cells express *Emx2*, whereas blue cell do not. The immature hair cells are depicted in grey. The transversal dashed line depicts the dorsoventral midline. The polarity of the hair cells is indicated with the black arrows below each scheme. (B) A model showing a simple Notch1a-mediated local lateral inhibition of *Emx2* expression between hair-cell siblings, leading to robust symmetry-breaking and planar bipolarity. (C) Representation of the progression of planar polarization of the neuromast. The epithelium formed by non-sensory supporting cells (green) is *Emx2*(-), able to generate head (anterior)-facing coherent planar polarization of hair cells. UHCPs begin to express *Emx2* (pink), effectively reversing the polarization potential of the cell toward the tail (posterior). Upon UHCP division, both daughter cells express *Emx2* initially poised to be tail facing. *Emx2* subsequently remains expressed in one of the sibling hair cells as they mature and polarize, generating bipolarity (confronting black arrows).

STAR METHODS

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hernán López-Schier (hernan.lopez-schier@helmholtz-muenchen.de).

All unique/stable reagents generated in this study are available from the Lead Contact without restriction.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethics statement

Experiments with wild-type, mutant and transgenic zebrafish embryos of undetermined sex were conducted under a protocol number Gz.:55.2-1-54-2532- 202-2014 approved by the “Regierung von Oberbayern”, Germany.

Zebrafish lines and husbandry

Zebrafish eggs and larvae were kept under the standardized conditions at 28.5°C in the 0,3 x Danieau’s solution. ET(krt4:EGFP)SqGw57a[37], Tg[myo6b:actb1-EGFP][15], Tg(pou4f3:GAL4)[53], Tg(UAS:Notch1a-intra)[54], Emx2-mutant (Chr13:fdc5)[8], Atoh1a mutant line[19] have been previously described.

Tg(alpl:nlsEGFP) transgenic line

The Tg(alpl:nlsEGFP) transgenic line was generated by injecting an alpl:nlsEGFP DNA construct and transposase mRNA in the one-cell stage wild-type embryos and selecting carrier fish in the next generation as described previously. The alpl:nlsEGFP DNA construct was generated by Tol2/Gateway zebrafish kit[44]. The p5E-alpl vector was generated after the PCR of alpl promoter was performed using primers with overhangs to add att sites (GGGGACAACCTTTGTATAGAAAAGTTGCGCCCACTTTGGCATAATCTTGC,

GGGGACTGCTTTTTTGTACAAACTTGCAGCGCTGGAAGAACAGACTC) and the AccuPrime Pfx polymerase (Invitrogen). The vectors pDONRP4-P1R, pDestTol2pA and pME-nlsEGFP, p3E-polyA from the Tol2 kit were used.

Notch1a mutant fish line

Notch1a mutant fish line was generated using a CRISPR-Cas9 approach with redundant targeting of single gene[45] in the Tg[myo6b:actb1-EGFP] background. Briefly, a solution of an 1 µg/µl equimolar mixture of 4 sgRNAs, generated using MEGAshortscript™ T7 Transcription Kit (Thermo Fischer), 5 µM Cas9 protein (Sigma), 300 mM KCl, was injected into one-cell stage embryos. The following 4 sgRNAs targeting the Notch1a gene were used: GACTGCAGCATCGCTCGCGACGG, GTGTGTCGGCCGCAGATGCAGGG, GGTGGCATCCCGAAAACCGTCGG, GTGAGGAACCCGTGCACTAATGG. The fish were raised to adulthood and the experiments were performed with the offspring. The selection of Notch1a homozygous mutant larvae was performed by screening 3-4 dpf larvae for abnormal somites, a previously reported phenotype of Notch1a-deficiency[37].

METHOD DETAILS

Somatic mutagenesis of Emx2

Somatic mutagenesis of Emx2 was performed using a CRISPR-Cas9 approach with redundant sgRNAs targeting of single gene[45]. The following sgRNAs, generated using MEGAshortscript™ T7 Transcription Kit (Thermo Fischer), were used to target Emx2: GGTA AACACCTCTTCGGTG, GGA CTGTGCGAAGACGACAG, CTCTTTTCGCAAGCCAGCAA, CCTGAGTTTCTGTGAGGCTA. Briefly, a solution of an 1 µg/µl equimolar mixture of 4 sgRNAs, 5 µM Cas9 protein, 300 mM KCl, was injected into one-cell stage embryos (wt or Notch1a mutant). The efficiency of somatic mutagenesis was assessed by observing the Emx2-deficient hair cell polarity phenotype in 2-3 dpf larvae.

Pharmacology

N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine-t-butyl ester (DAPT) (Sigma) was used for global inhibition of Notch signalling. 2 dpf-old larvae were treated with 100 μ M DAPT in 0.3x Danieau's solution for 24h and left to recover in 0.3x Danieau's solution for 24h. The pharmacological experiments were performed at 28.5°C and equal amounts of DMSO (Sigma) were diluted in 0.3x Danieau's solution for control specimens.

Analysis of hair cells orientation

The orientation of hair cell bundles was quantified in wt, or mutant, *myo6b:GFP* transgenic larvae and wt or transgenic *brn3c:gal4*, *UAS:NICD* larvae stained with phalloidin by analyzing the images of horizontal neuromasts and manually drawing a line corresponding to the orientation of each cell. The angles of all lines were plotted in a circular histogram. The hair cells, in which the polarity was not well visible as well as neuromasts containing only one hair cell with visible polarity, were excluded from the analysis. Additionally, the orientation of hair cells in *Emx2*- and *Notch1a*-double mutant larvae was analyzed in larvae of the *Notch1a*-mutant line with somatically mutated *Emx2* as described in the previous section. To this end, was the efficiency of the somatic mutagenesis of *Emx2* first assessed by injecting the wt embryos and quantifying the hair cells orientation. Subsequently, approximately 10% of background was expected in the *Emx2*-, *Notch1a*-double mutant, corresponding to cells, which did not carry mutation in *Emx2* gene.

Sample preparation for scRNA-seq

Embryo dissociation

The trunks of sacrificed 6 and 7 dpf zebrafish larvae were cut with a sharp blade after the yolk to collect only the cells from the posterior lateral line neuromasts and simultaneously discard the GFP-positive cells in the heart of *myo6b:EGFP* line. The larval trunks were collected in ice-cold Ringer's solution (116 mM NaCl, 2.6 mM KCl, 1.8 mM CaCl₂, 5 mM HEPES, pH=7).

The trunks were dissociated in a 1:1 mixture of Accutase solution (Sigma) and Accumax solution (Sigma) for 25 min and gently triturated with 1 ml pipette tip every 3 min at room temperature. The undissociated parts of the tissues were let to sink on the bottom of the tube and the supernatant containing dissociated cells was collected and washed with cold Ringer's solution without CaCl₂ and resedimented by centrifugation at 17000rpm. Following, the cells were kept on ice and resuspended in sort buffer (Ringer's solution without CaCl₂, 2 mM EDTA, 10% FBS), filtered twice through 40 µm strainer (Falcon) and stained with 7-AAD (5 µg/mL, BD).

Fluorescence activated cell sorting

The single-cell solution was FAC-sorted immediately after dissociation using the following gating strategy: main population > single cells > live cells > GFP-positive cells, the latter collected in the sort collection buffer (Ringer's solution without CaCl₂, 2% FBS, 10 µM EDTA). Only the cells sorted in the first 25 min were used for the transcriptional analysis to avoid sequencing of stressed cells. For the scRNAseq with 10x Chromium were the GFP-positive cells from myo6b:GFP and Gw57a transgenic lines FAC-sorted and collected separately. GFP-negative cells were added to both samples for control of gating specificity. For the scRNAseq with mcSCRSeq, the EGFP-positive cells (myo6b:GFP and myo6b:GFP; Emx2-mutant) were collected one cell per well into two 96-well plates containing lysis buffer, consisting of 5 M Guanidine hydrochloride (Sigma-Aldrich), 1% 2-mercaptoethanol (Sigma-Aldrich) and a 1:500 dilution of Phusion HF buffer (New England Biolabs)[46] (updated in [dx.doi.org/10.17504/protocols.io.2aegabe](https://doi.org/10.17504/protocols.io.2aegabe)), and immediately frozen on dry ice.

Ultrasensitive chain-reaction whole-mount *in situ* hybridization

For Emx2 whole-mount fluorescent in-situ hybridization[31], a set of 26 probe pairs was ordered from Molecular Instruments, Inc. and used following the protocol as described by the manufacturer. Briefly, the larvae

were fixed in 4% paraformaldehyde (PFA) for 24 h at 4 °C, permeabilized with methanol and cooled to -20°C. Rehydration, treatment with proteinase K and postfixation followed. The samples were washed with PBST between the steps. Probe hybridization buffer was used for the prehybridization for 30 minutes at 37°C and the samples were incubated in the probe solution, prepared following the manufacturer's instructions, overnight at 37°C. After removing the probe solution, washing the samples and incubating them in the pre-amplification buffer, the samples were incubated in the hairpin mixture overnight in the dark at room temperature. Finally, after several washes with SSCT, the cell nuclei were stained with DAPI (4',6-diamidino-2-phenylindole, Sigma) at room temperature and the samples were mounted for imaging. The larvae from mutant or transgenic lines were stained simultaneously as wt control.

Immunohistochemistry

First, zebrafish embryos were dechorionated manually and fixed in 4% PFA solution in PBS medium for 4h at room temperature. Second, the samples were washed 4 times with PBST (0,05% tween-20) and permeabilized with pre chilled acetone at -20°C for 5 minutes and rinsed with PBST. An incubation for 2h at room temperature in blocking solution (1,5% BSA, 1,5% sheep serum in PBST) followed. Third, the samples were incubated with primary Ab (Emx2 1:200) for 40 hours at 4 °C. After 3 washing steps the secondary antibodies were added (GaRb 594) and incubated overnight at 4 °C. Finally, the samples were washed with PBST and mounted for imaging. The Emx2 primary antibodies were obtained from TransGenic (Fukuoka, Japan). Following the immunostaining, the cell nuclei were stained with DAPI (4',6-diamidino-2-phenylindole, Sigma) at room temperature. The larvae from mutant or transgenic lines were stained simultaneously as wt control.

Imaging and video-microscopy

For imaging of Emx2 immunohistochemically stained samples, the samples after fluorescence in situ hybridization (FISH) and screening of live Emx2-mutant embryos at 2-3 dpf, a custom-built inverted spinning disc microscope (Zeiss AxioScope) with 63× water immersion objective was used. Larvae used for in vivo imaging were anesthetized in MS-222 (tricaine) 0.16 g/L and mounted in 1% low-melting-point agarose with diluted anesthetic on the cover slip of a glass-bottom dish (MatTek, Ashland, MA). The dishes containing live larvae were bathed in 0.3x Danieau's solution with MS-222 0.16 g/L, except during the live video-microscopy, where the MS-222 concentration was reduced to 0.08 g/L. Acquisition of live larvae was performed at 28.5 °C. The video-microscopy included time-lapse image acquisition every 150 - 240 seconds. All images were processed with Fiji software[47], including the contrast correction, which was performed for separate channels. Additionally, were the movies registered for image stabilization and the UHCPs and newly arisen immature hair cells were tracked with Fiji plugin MTrackJ[48].

QUANTIFICATION AND STATISTICAL ANALYSIS

Single cell sequencing by 10X Chromium and data preprocessing

Single cell libraries were generated using the Chromium Single Cell 3' library and gel bead kit v2 (PN #120237) from 10x Genomics. The FAC-sorted cells were loaded onto the 10x Chromium™ Chip to produce Gelbeads-in-Emulsions. After the reverse transcription, including the barcoding of RNA, the cDNA was amplified and 3' gene expression libraries were generated from both samples (cells isolated from Gw57a line, named support cells sample, and cells isolated from myo6b:EGFP line, named hair cells sample). Libraries were pooled and sequenced together to minimize the batch effect. The libraries were sequenced to a depth of approximately 100,000 reads per cell on the HiSeq4000 (Illumina) with 150 bp paired-end

sequencing. Following, the data was demultiplexed by sample indices and mapped to the zebrafish genome GRCz11.

Single cell sequencing by mcSCRBseq and data preprocessing

Single cell libraries were generated using an adapted version of mcSCRB-seq[46]. Following the FACS were the lysed cells cleaned using SPRI beads and resuspended with 9 μ L of reverse transcription mastermix and 1 μ L 2 μ M barcoded oligo-dT primer. The remaining protocol was carried out as outlined in mcSCRB-seq. Libraries were paired-end sequenced on a high output flow cell of a HiSeq 1500 (Illumina) to an average depth of about 811,000 reads per cell. The following sequencing setup was used: 16 bases for the cellular barcode and UMI, 8 bases for the i7 barcode, and 50 bases for the cDNA read. Fastq files were assessed using FastQC (v0.10.1) and preprocessed using zUMIs (v2.0) using the default settings. The data were mapped to the zebrafish genome (danRer11) and gene annotations were obtained from Ensembl (GRCz11.94).

Quality Control and Processing of 10x Data

Standard quality control steps were used to filter noisy and unreliable cells and genes including assessing the number of molecule counts or UMIs (Unique Molecular Identifiers), the number of detected genes and the fraction of counts from mitochondrial genes with a high fraction indicating potential cellular stress. By looking at the distribution of these QC covariates jointly, the cells were investigated for anomalous summary statistics, which would then be subject to filtering using thresholding decisions. Specifically, by visually tracing a Gaussian kernel density around the number of counts distribution, we estimate the population to range until 30,000 and 40,000 for hair and support cells respectively. Moreover, cells with less than 800 genes were filtered out arguing that below this threshold, cells could potentially be dying or be empty droplets driven by ambient RNA contamination. After filtering cells, we follow the basic library size normalization to counts per million and log-transform the gene expression

matrix. Additionally, we used functions from the Python package Scanpy (v1.4.3)[12] with their default parameters to identify the top 4000 highly variable genes, which is used to perform principal component analysis (PCA). Also, technical effects due to count depths were regressed out prior to PCA with a view to improve the inference results. Following dimensionality reduction, we clustered the data by Louvain clustering with resolution = 0.4. Prior to Louvain clustering, we computed the neighbourhood graph of cells with the size of local neighbourhood set to 25 using the *scanpy.api.pp.neighbors* function. Eight clusters were identified and two of them were subsequently annotated as neuromasts on the basis of cluster-specific markers. The neuromast cells were followed up and sub-clustered using Lovain clustering at resolution = 0.55 and a local neighborhood of 15 to re-estimate the manifold. Here an additional cluster with relatively low counts, high mitochondrial reads and few detected genes as well as no strongly cluster-specific genes was observed, which suggested that it could be an artifact due to cellular stress.

Quality Control and Processing of mcSCRBseq Data

Similar to the previous section, the mcSCRBseq data was subject to QC measures. The distribution of count depth was estimated to range until 20,000 reads. Additionally, we kept cells with at least 1,250 genes measured. The cells were normalised by count depth scaling and log-transformed. Since plate-based scRNA-seq data tend to exhibit batch effects between plates we used Combat to fit batch parameters[49]. PCA and Louvain clustering were applied similar to the previous cluster analysis.

Selection of neuromast cell clusters in 10x Data

We next set to identify cell types using the log-transformed expression data. We used the UMAP algorithm for visualizations as it has been shown to accurately represent the topology of the data. Particularly for our scRNA-seq analysis of wild-type neuromast cells, we used graph-based louvain clustering with the resolution parameter set to 0.4. The merged data

included neuromast hair cells, neuromast support cells, the EGFP-negative cells (control of sorting specificity) and the cells, which were sorted as EGFP-positive cells from the Gw57 line, but belonged to other non-neuromast cell types. The neuromast cell clusters were carefully selected based on the expression of neuromast marker genes and genes predominantly expressed in neuromasts in each cluster. 2 out of 7 clusters were identified as neuromast cells and selected for second-level clustering. The classification was performed on these cells using the default parameters. The UMAP plots, shown in this study as part of second-level clustering results have been obtained with Scanpy as described above[12].

Doublet Removal

In order to identify technical artifacts of cell doublets, we applied the Scrublet v0.2.1 pipeline[50] to the support cells and hair cells sample individually with default parameters for doublet score calculation. Cells with doublet score of 0.5 and 0.42 for hair and support cells respectively were annotated as potential doublets and subsequently removed.

Pseudotime Analysis

We performed pseudotime analysis on 1,032 and 135 cells selected from the subclusters corresponding to neuromast hair and support cells respectively based on second-level clustering results. Pseudotime analysis was performed using the Scanpy toolkit (v1.4.3) that already has DPT integrated into it and can be directly used in the Scanpy workflow using the function *scanpy.api.tl.dpt*. Prior to analysis DPT requires a root cell to be assigned in order to infer pseudotime. Looking at the first two diffusion components (DCs) we posit that the most appropriate root cell will be the support cell with the maximum DC1 or minimum DC2 value. Furthermore, we use the *scanpy.api.pl.paga_path* plotting function of Scanpy to investigate how the expression of selected genes varies across the pseudotime to better describe the differentiation process[12]. Briefly we

used 100 data points to compute the running average expression and scaled it from 0 to 1 per gene.

Data Integration

Integration of our dataset with Lush et al. (2019)[25] was performed using the batch alignment tool BBKNN (v1.3.6)[51], which works by identifying the k nearest neighbors for each cell across technical batches. This way it relates cells belonging to the same cell-type or cell state from different batches without the need to modify counts. Briefly, we downloaded the count matrices for Lush et al study from the Gene Expression Omnibus (GEO) database (accession no. GSE123241) and followed their pre-processing steps for compatibility and filtered out genes that were expressed in less than 3 cells from their WT dataset. This was followed by merging the two datasets, log normalisation of the remaining UMI counts and calculating highly variable genes following default Scanpy workflow. We next applied BBKNN to the data in principal component space and observed that it successfully corrects for technical batch effects and aligns the two hair cell progenitor populations across studies.

Differential Expression Analysis

Differential expression analysis comparing hair cells from wt and *emx2*-mutant larvae was performed using the R package limma (v3.38.3) and edgeR (v3.24.1). For this purpose, log-normalized counts are used that accounts for library size differences. As we are interested to see which genes are expressed at differential levels between wild-type and loss-of-function phenotype a design matrix is first constructed with the cell population information. Linear models are fit to the expression of each gene using the *lmFit* function. Next, an empirical Bayes method, *eBayes*, is used to rank genes with enhanced statistical power to detect significant differential expression. Finally the top ranked genes are extracted from the linear model fit after correcting the p-values for multiple testing using the default adjustment method, Benjamini–Hochberg or BH.

Young hair cell population from the 10X study was manually split into *emx2*-expressing cells and cells where *emx2* could not be detected and subsequently analyzed for differential genes using the *scanpy.api.tl.rank_genes_groups* function and default parameters.

KEGG Pathway Analysis

KEGG pathway analysis was performed using the Python toolkit GProfiler (v1.0.0)[52] to identify top enriched biological pathways using differentially expressed genes between two clusters of cells. Briefly, the top 200 genes characterizing each cluster were computed using the *scanpy.api.tl.rank_genes_groups* function using default parameters and subsequently used as an input to the GProfiler function to perform statistical enrichment analysis from KEGG. The results were corrected for multiple testing using *fdr* and visualized using a custom plotting function.

DATA AND CODE AVAILABILITY

The single-cell RNA sequencing datasets reported in this study is available in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) under the accession ID <to be determined>. All software resources that were used to support the findings of this study have been indicated above, and are available at GitHub.

Table S1

List of differentially expressed genes in young hair cells expressing *Emx2* vs young hair cell not expressing *Emx2*. Related to Figure 1.

A list of top 100 differentially expressed genes, identified in comparison of transcriptional profiles of young hair cells with *Emx2* expression > 0 vs young hair cells, in which *Emx2* was not detected.

Video S1

Development of hair cells in wild type, and *Emx2*- and *Notch1a*-mutants. Related to Figure 2.

A 15:59h time-lapse video showing an example of HCs development in a horizontal neuromast of wild-type, *myo6b:EGFP* and *Emx2*-mutant, *myo6b:GFP* larvae on the left and wild-type, *myo6b:EGFP*, *alpl:nlsEGFP* and *myo6b:GFP*, *alpl:nlsEGFP*, *Notch1a*-mutant larvae on the right. In all panels the UHCs and the hair cells are visible, and additionally are in the right two panels nuclei of the neuromast cells labelled as well (*alpl:nlsEGFP* line). For clarity, are the UHCs and pairs of newly arisen immature HCs labeled with dots in random colors with two sibling hair cells always shown in the same color. The video showing exemplary neuromast demonstrates that the hair cells in wt and *Emx2* - and *Notch1a*-mutant larvae arise in pairs and that there is no abnormal hair-cell death in the mutant lines.

REFERENCES

1. Butler, M.T., and Wallingford, J.B. (2017). Planar cell polarity in development and disease. *Nat. Rev. Mol. Cell Biol.* *18*, 375-388.
2. Zallen, J.A. (2007). Planar polarity and tissue morphogenesis. *Cell* *129*, 1051-1063.
3. Tarchini, B., and Lu, X. (2019). New insights into regulation and function of planar polarity in the inner ear. *Neurosci. Lett.* *709*, 134373.
4. Duncan, J.S., Stoller, M.L., Francl, A.F., Tissir, F., Devenport, D., and Deans, M.R. (2017). *Celsr1* coordinates the planar polarity of vestibular hair cells during inner ear development. *Dev. Biol.* *423*, 126-137.
5. López-Schier, H., Starr, C.J., Kappler, J.A., Kollmar, R., and Hudspeth, A.J. (2004). Directional cell migration establishes the axes of planar polarity in the posterior lateral-line organ of the zebrafish. *Dev. Cell* *7*, 401-412.

6. López-Schier, H., and Hudspeth, A.J. (2006). A two-step mechanism underlies the planar polarization of regenerating sensory hair cells. *Proc Natl Acad Sci U S A* *103*, 18615-18620.
7. Thomas, E.D., Raible, D.W. (2019). Distinct progenitor populations mediate regeneration in the zebrafish lateral line. *Elife*. *8*, e43736.
8. Jiang, T., Kindt, K., and Wu, D.K. (2017). Transcription factor Emx2 controls stereociliary bundle orientation of sensory hair cells. *Elife* *6*.
9. Ji, Y.R., Warriar, S., Jiang, T., Wu, D.K., and Kindt, K.S. (2018). Directional selectivity of afferent neurons in zebrafish neuromasts is regulated by Emx2 in presynaptic hair cells. *Elife* *7*.
10. Lozano-Ortega, M., Valera, G., Xiao, Y., Faucherre, A., and López-Schier, H. (2018). Hair cell identity establishes labeled lines of directional mechanosensation. *PLoS Biol.* *16*, e2004404.
11. Holley, M., Rhodes, C., Kneebone, A., Herde, M.K., Fleming, M., and Steel, K.P. (2010). Emx2 and early hair cell development in the mouse inner ear. *Dev. Biol* *340*, 547-556.
12. Wolf, F.A., Angerer, P., Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* *19*, 15.
13. Francesconi, M., Lehner, B. (2015). Reconstructing and analysing cellular states, space and time from gene expression profiles of many cells and single cells. *Mol. Biosyst.* *11*, 2690-2698.
14. Tritschler, S., Buttner, M., Fischer, D.S., Lange, M., Bergen, V., Lickert, H., and Theis, F.J. (2019). Concepts and limitations for learning developmental trajectories from single cell genomics. *Development* *146*: dev170506.
15. Kindt, K.S., Finch, G., and Nicolson, T. (2012). Kinocilia mediate mechanosensitivity in developing zebrafish hair cells. *Dev. Cell* *23*, 329-341.
16. Pinto-Teixeira, F., Viader-Llargués, O., Torres-Mejía, E., Turan, M., González-Gualda, Pola-Morell, L., and López-Schier, H. (2015).

Inexhaustible hair-cell regeneration in young and aged zebrafish. *Biol Open* 4, 903-909.

17. Blondel, V.D., Guillaume, J-L, Lambiotte, R., Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory and Experiment.* 10, P10008. doi:10.1088/1742-5468/2008/10/P10008

18. Sarrazín, A.F., Villablanca, E.J., Nuñez, V.A., Sandoval, P.C., Ghysen, A., and Allende, M.L. (2006). Proneural gene requirement for hair cell differentiation in the zebrafish lateral line. *Dev Biol* 295, 534-545.

19. Kidwell, C.U., Su, C.Y., Hibi, M., Moens, C.B. (2018). Multiple zebrafish *atoh1* genes specify a diversity of neuronal types in the zebrafish cerebellum.

Dev Biol. 438, 44-56.

20. Ma, E.Y., Rubel, E.W., Raible, D.W. (2008). Notch signaling regulates the extent of hair cell regeneration in the zebrafish lateral line. *J. Neurosci.* 28,2261-73.

21. Kappler, J.A., Starr, C.J., Chan, D.K., Kollmar, R., and Hudspeth, A.J. (2004). A nonsense mutation in the gene encoding a zebrafish myosin VI isoform causes defects in hair-cell mechanotransduction. *P.N.A.S U.S.A.* 101:13056-13061

22. Seiler, C., Ben-David, O., Sidi, S., Hendrich, O., Rusch, A., Burnside, B., Avraham, K.B., and Nicolson, T. (2004). Myosin VI is required for structural integrity of the apical surface of sensory hair cells in zebrafish. *Dev. Biol.* 272:328-338

23. Hsiao, C.D., Tsai, W.Y., and Tsai, H.J. (2002). Isolation and expression of two zebrafish homologues of parvalbumin genes related to chicken CPV3 and mammalian oncomodulin. *Mech. Dev.* 119 Suppl 1, S161-166.

24. Thisse, B., Heyer, V., Lux, A., Alunni, V., Degraeve, A., Seiliez, I., Kirchner, J., Parkhill, J.P., Thisse, C. (2004). Spatial and temporal expression of the zebrafish genome by large-scale *in situ* hybridization screening. *Methods Cell Biol.* 77, 505-519.

25. Lush, M.E., Diaz, D.C., Koenecke, N., Baek, S., Boldt, H., St Peter, M.K., Gaitán-Escudero, T., Romero-Carvajal, A., Busch-Nentwich, E.M., Perera, A.G., et al. (2019). scRNA-Seq reveals distinct stem cell populations that drive hair cell regeneration after loss of Fgf and Notch signaling. *Elife* *8*.
26. Erickson, T., and Nicolson, T. (2015). Identification of sensory hair-cell transcripts by thiouracil-tagging in zebrafish. *BMC Genomics* *16*, 842.
27. Chou, S.W., Chen, Z., Zhu, S., Davis, R.W., Hu, J., Liu, L., Fernando, C.A., Kindig, K., Brown, W.C., Stepanyan, R., McDermott, B.M. Jr. (2017). A molecular basis for water motion detection by the mechanosensory lateral line of zebrafish. *Nat. Commun.* *8*, 2234.
28. Elkon, R., Milon, B., Morrison, L., et al. (2015). RFX transcription factors are essential for hearing in mice. *Nat. Commun.* *6*:8549
29. Pichler, P., Lagnado, L. (2019). The Transfer Characteristics of Hair Cells Encoding Mechanical Stimuli in the Lateral Line of Zebrafish. *J. Neurosci.* *2*;39:112-124.
30. Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., Theis F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* *13*, 845-8.
31. Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A., Pierce, N.A. (2018). Third-generation *in situ* hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development.* *145*.
32. Pujol-Martí, J., Zecca, A., Baudoin, J.P., Faucherre, A., Asakawa, K., Kawakami, K., López-Schier, H. (2012). Neuronal birth order identifies a dimorphic sensorineural map. *J. Neurosci.* *29*;32(9):2976-87.
33. Wibowo, I., Pinto-Teixeira, F., Satou, C., Higashijima, S., and López-Schier, H. (2011). Compartmentalized Notch signaling sustains epithelial mirror symmetry. *Development* *138*, 1143-1152.

34. Viader-Llargués, O., Lupperger, V., Pola-Morell, L., Marr, C., and López-Schier, H. (2018). Live cell-lineage tracing and machine learning reveal patterns of organ regeneration. *Elife* 7.
35. Mirkovic, I., Pylawka, S., and Hudspeth, A.J. (2012). Rearrangements between differentiating hair cells coordinate planar polarity and the establishment of mirror symmetry in lateral-line neuromasts. *Biol. Open* 1, 498-505.
36. Fröhlicher, M., Liedtke, A., Groh, K., López-Schier, H., Neuhaus, S.C., Segner, H., Eggen, R.I. (2009). Estrogen receptor subtype beta2 is involved in neuromast development in zebrafish (*Danio rerio*) larvae. *Dev. Biol.* 330, 32-37.
37. Sieger, D., Tautz, D., Gajewski, M. (2003). The role of Suppressor of Hairless in Notch mediated signalling during zebrafish somitogenesis. *Mech. Dev.* 120(9):1083-1094.
38. Jacobo, A., Dasgupta, A., Erzberger, A., Siletti, K. Hudspeth, A.J. (2019). Notch-Mediated Determination of Hair-Bundle Polarity in Mechanosensory Hair Cells of the Zebrafish Lateral Line. *Curr. Biol.* 29(21):3579-3587.
39. Faucherre, A., López-Schier, H. (2011). Delaying Gal4-driven gene expression in the zebrafish with morpholinos and Gal80. *PLoS One* 6(1):e16587.
40. Navajas Acedo, J., Voas, M.G., Alexander, R., Woolley, T., et al. (2019). PCP and Wnt pathway components act in parallel during zebrafish mechanosensory hair cell orientation. *Nat. Commun.* 10(1):3993.
41. Richard, A., Boullu, L., Herbach, U., et al. (2016). Single-Cell-Based Analysis highlights a surge in cell-to-cell molecular variability preceding irreversible commitment in a differentiation process. *PLoS Biol.* 14(12):e1002585
42. Warrington, S.J., Strutt, H., Fisher, K.H., and Strutt, D. (2017). A Dual Function for Prickle in Regulating Frizzled Stability during Feedback-Dependent Amplification of Planar Polarity. *Curr. Biol* 27, 2784-2797 .

43. Henrique, D., Schweisguth, F. (2019). Mechanisms of Notch signaling: a simple logic deployed in time and space. *Development*. 146(3).
44. Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant J.M., Yost, H.J., Kanki, J.P., Chien, C.B. (2007). The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* 236, 3088-99.
45. Wu, R.S., Lam, I.I., Clay, H., Duong, D.N., Deo, R.C., & Coughlin, S.R. (2018). A rapid method for directed gene knockout for screening in G0 zebrafish. *Dev. Cell*, 46, 112-125.
46. Bagnoli, J.W., Ziegenhain, C., Janjic, A., Wange, L.E., Vieth, B., Parekh, S., Geuder, J., Hellmann, I., Enard, W. (2018). Sensitive and powerful single-cell RNA sequencing using mcSCR-seq. *Nat. Commun.* 9, 2937
47. Schindelin, J., Arganda-Carreras, I., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. 9, 676-82.
48. Meijering, E., Dzyubachyk, O., Smal, I. (2012) Methods for cell and particle tracking. *Methods in enzymology* 504:183–200.
49. Johnson, W.E., Li, C., Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, 8, 118–127.
50. Wolock, S.L., López, R., Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* 8, 281-291.e9.
51. Polański, K., Young, M.D., Miao, Z., Meyer, K.B., Teichmann, S.A., Park, J.E. (2019). BBKNN: Fast Batch Alignment of Single Cell Transcriptomes. *Bioinformatics*, 1-2.
52. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47(W1), 191-198.

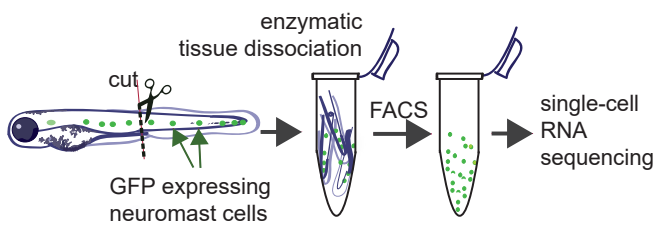
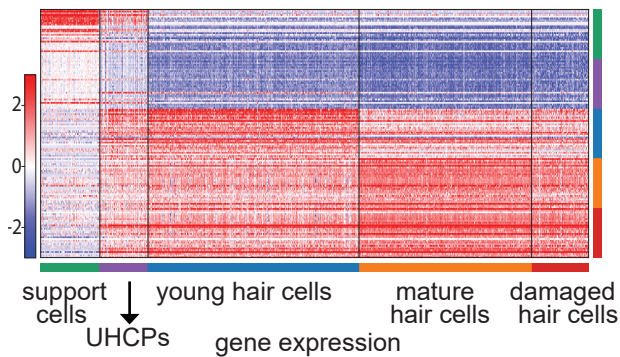
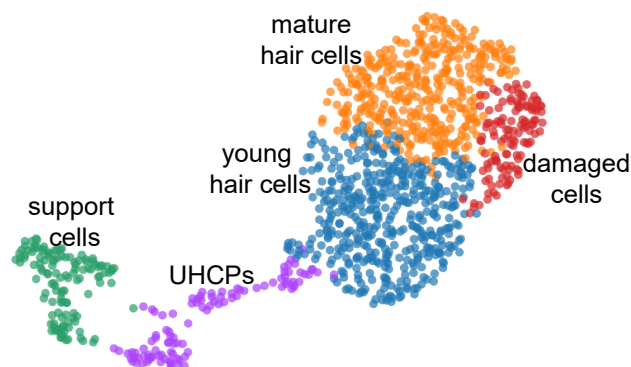
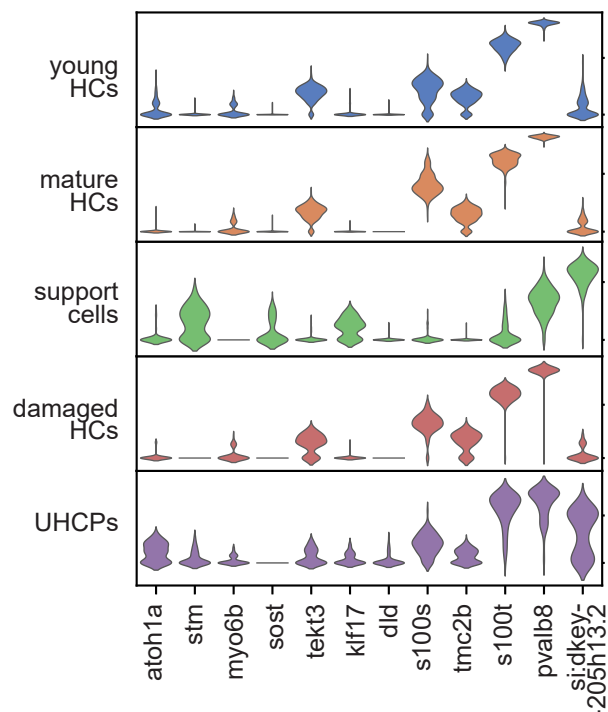
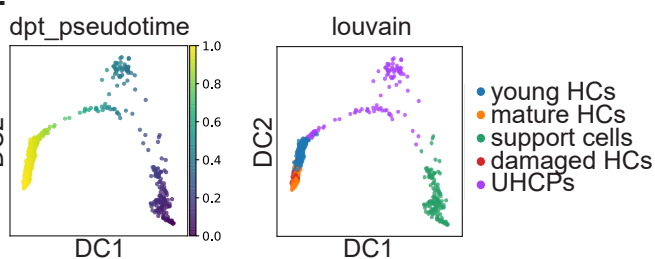
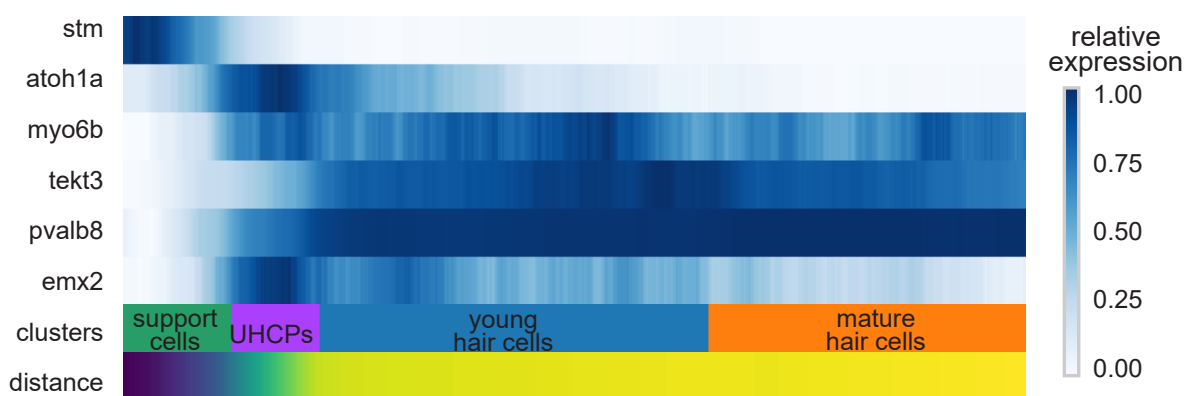
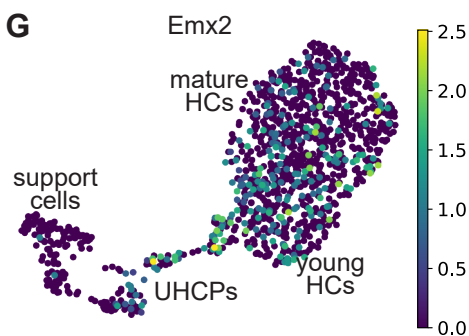
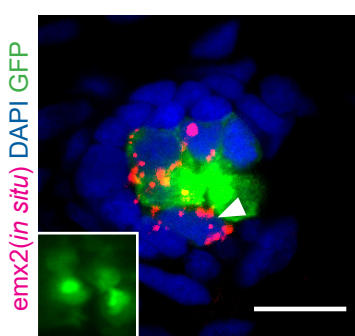
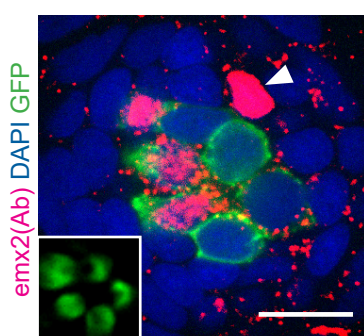
53. Xiao, T., and Baier, H. (2007). Lamina-specific axonal projections in the zebrafish tectum require the type IV collagen Dragnet. *Nature Neuroscience*. 10(12):1529-1537.
54. Scheer, N. and Campos-Ortega, J.A. (1999). Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mechanisms of Development*. 80(2):153-158.
55. R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
56. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. 2015;43:e47.
57. Robinson, M.D., McCarthy, D.J., Smyth, G.K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–140.

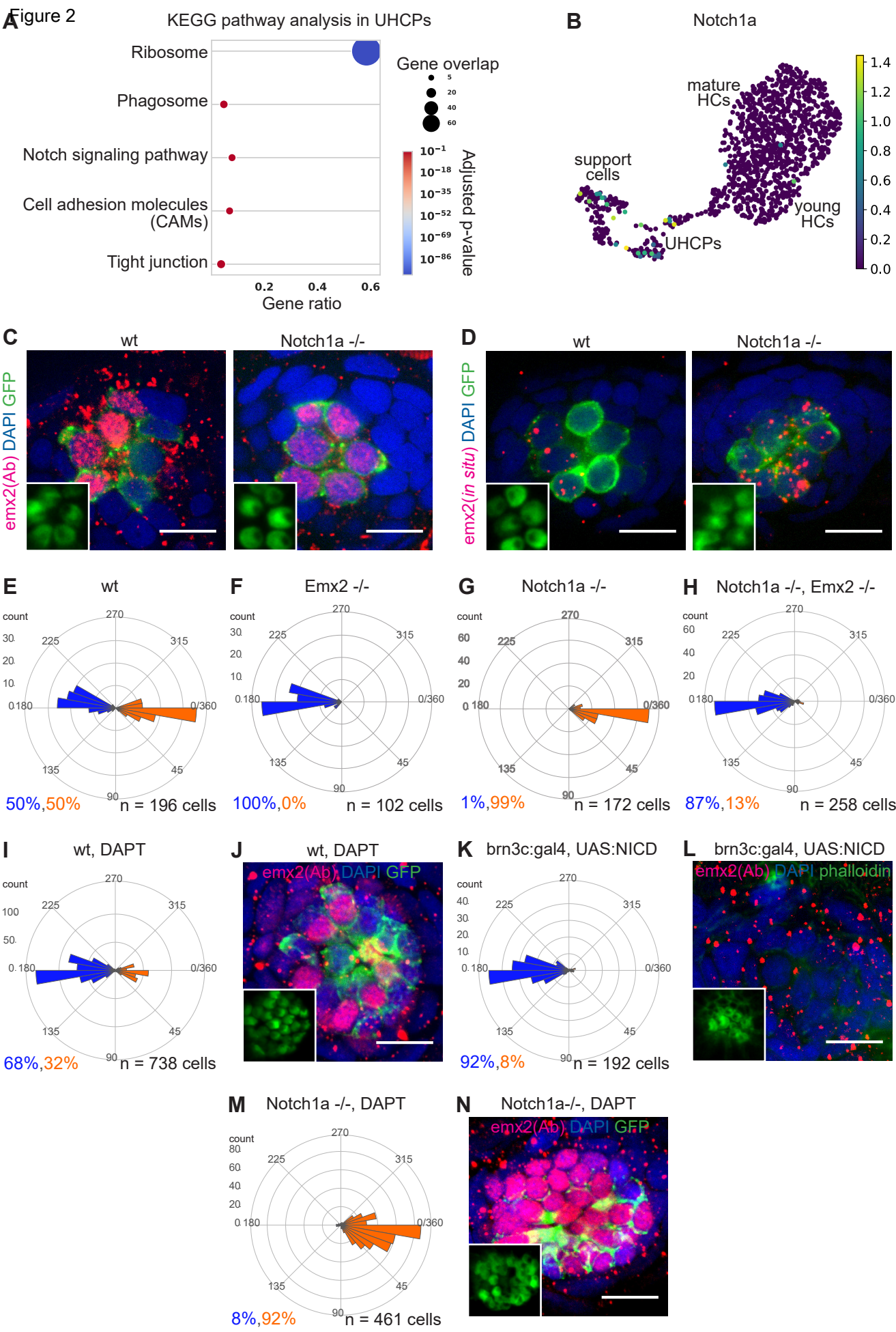
KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti Emx2	Trans Genic (Fukuoka, Japan)	Cat#KO609
Goat anti rabbit Alexa Fluor-594	Abcam	Cat#ab150080
Chemicals, Peptides, and Recombinant Proteins		
4',6-diamidino-2-phenylindole (DAPI)	Sigma-Aldrich	Cat#D9542
N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine-t-butyl ester (DAPT)	Sigma-Aldrich	Cat# D5942
Cas9 protein	Sigma-Aldrich	Cat#CAS9PROT
Alexa Fluor™ 488 Phalloidin	Thermo Fischer	Cat#A12379
MS-222	PHARMAQ	Cat#QN01A X93
Guanidine hydrochloride	Sigma-Aldrich	Cat#G3272
2-mercaptoethanol	Sigma-Aldrich	Cat#M6250
Phusion HF buffer	New England Biolabs	Cat#M0530L
7-Amino-Actinomycin D (7-AAD)	BD	Cat#BDB559925
Accutase solution	Sigma-Aldrich	Cat#A6964
Accumax solution	Sigma-Aldrich	Cat#A7089
AccuPrime Pfx polymerase	Invitrogen	Cat#12344024
Critical Commercial Assays		
Gateway™ LR Clonase™ II Enzyme mix	Thermo Fischer	Cat#11791100
Gateway™ BP Clonase™ II Enzyme mix	Thermo Fischer	Cat#11789100
MEGAscript™ T7 Transcription Kit	Thermo Fischer	Cat# AM1354
Emx2-set of probes	Molecular Instruments, Inc	N/A
Single Cell 3' Library and Gel Bead Kit v2	10x Genomics	Cat#PN-120237
Deposited Data		
Single cell RNAseq data (10x Genomics)	This paper	<to be determined>
Single cell RNAseq data (mcSCRbseq)	This paper	<to be determined>
Single cell RNAseq data (10x Genomics)	[25]	GSE123241
Experimental Models: Organisms/Strains		
ET(krt4:EGFP)SqGw57a	Korzh lab	[16]
Tg[myo6b:actb1-EGFP]	Kindt lab	[15]
Tg(pou4f3:GAL4)	Baier lab	[53]
Tg(5xUAS-E1b:6xMYC-notch1a-intra)	[54]	[54]

Emx2-mutant (Chr13:cdc5)	Wu and Kindt lab	[8]
Atoh1a mutant line[S5]	Moens lab	[16]
Tg(alpl:nlsEGFP)	This paper	N/A
Tg(myo6b:Gal4VP16-pA)	This paper	N/A
Notch1a mutant line	This paper	N/A
Oligonucleotides		
5E_p-alpl_F: GGGGCAACTTTGTATAGAAAAGTTGCGCC CACTTTGGCATAATCTTGC	This paper	N/A
5E_p-alpl_R: GGGGACTGCTTTTTTTGTACAACTTGCAGC GCTGGAAGAACAGACTC	This paper	N/A
sgRNA_Notch1a_1: GACTGCAGCATCGCTCGCGACGG	This paper	N/A
sgRNA_Notch1a_2: GTGTGTCGGCCGCAGATGCAGGG	This paper	N/A
sgRNA_Notch1a_3: GGTGGCATCCCGAAAACCGTCGG	This paper	N/A
sgRNA_Notch1a_4: GTGAGGAACCCGTGCACTAATGG	This paper	N/A
sgRNA_Emx2_1: GGTAAAACACCTCTTCGGTG	This paper	N/A
sgRNA_Emx2_2: GGACTGTGCGAAGACGACAG	This paper	N/A
sgRNA_Emx2_3: CTCTTTTCGCAAGCCAGCAA	This paper	N/A
sgRNA_Emx2_4: CCTGAGTTTCTGTGAGGCTA	This paper	N/A
Recombinant DNA		
alpl:nlsEGFP	This paper	N/A
Software and Algorithms		
R (3.5.0)	https://www.r-project.org/	[16]
ImageJ	https://imagej.nih.gov	[47]
Python (3.7)	https://www.python.org/	N/A
Scanpy (v1.4.3)	https://scanpy.readthedocs.io/en/stable/	[12]
Scrublet (v0.2.1)	https://github.com/AllonKleinLab/scrublet	[50]
BBKNN (v1.3.6)	https://github.com/Teichlab/bbknn	[51]
limma (v3.38.3)	https://bioconductor.org/packages/rel	[56]

	ease/bioc/html/limma.html	
edgeR (v3.24.1)	https://bioconductor.org/packages/release/bioc/html/edgeR.html	[57]
GProfiler (v1.0.0)	https://pypi.org/project/gprofiler-official/	[52]
Code for clustering and analyses of single cell RNAseq data (10x Genomics)	This paper	<to be determined>
Code for clustering and analyses of single cell RNAseq data (mcSCRBseq)	This paper	<to be determined>

A Figure 1**B****C****D****E****F****G****H****I**



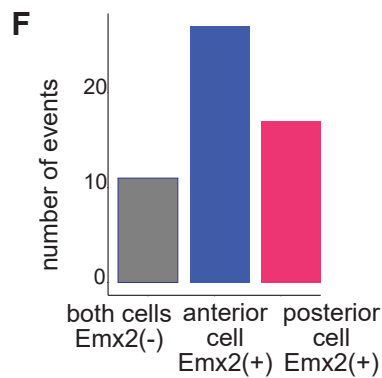
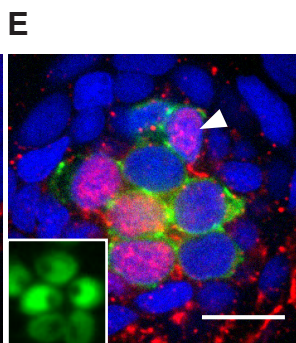
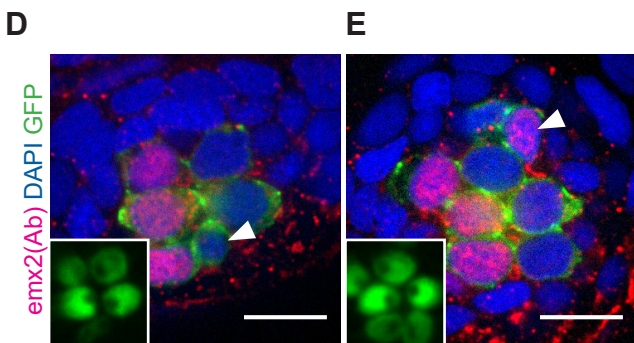
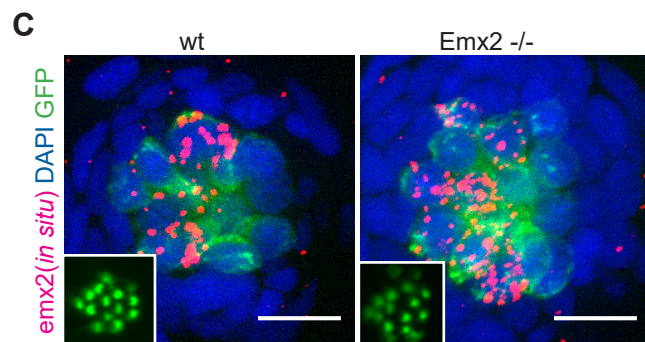
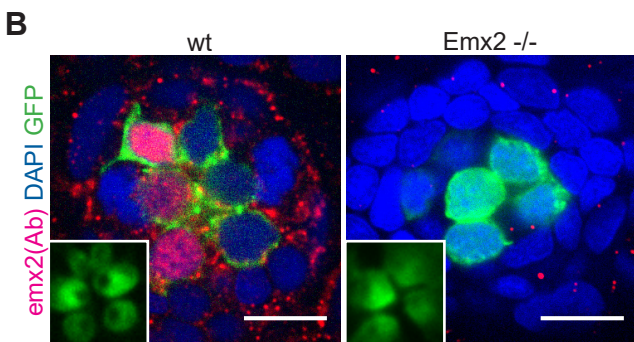
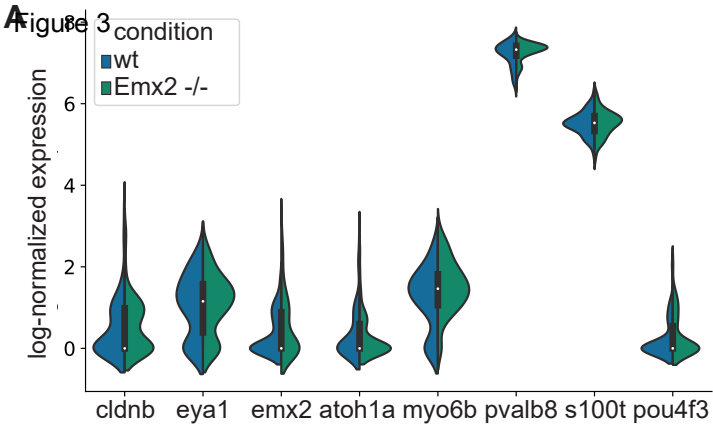
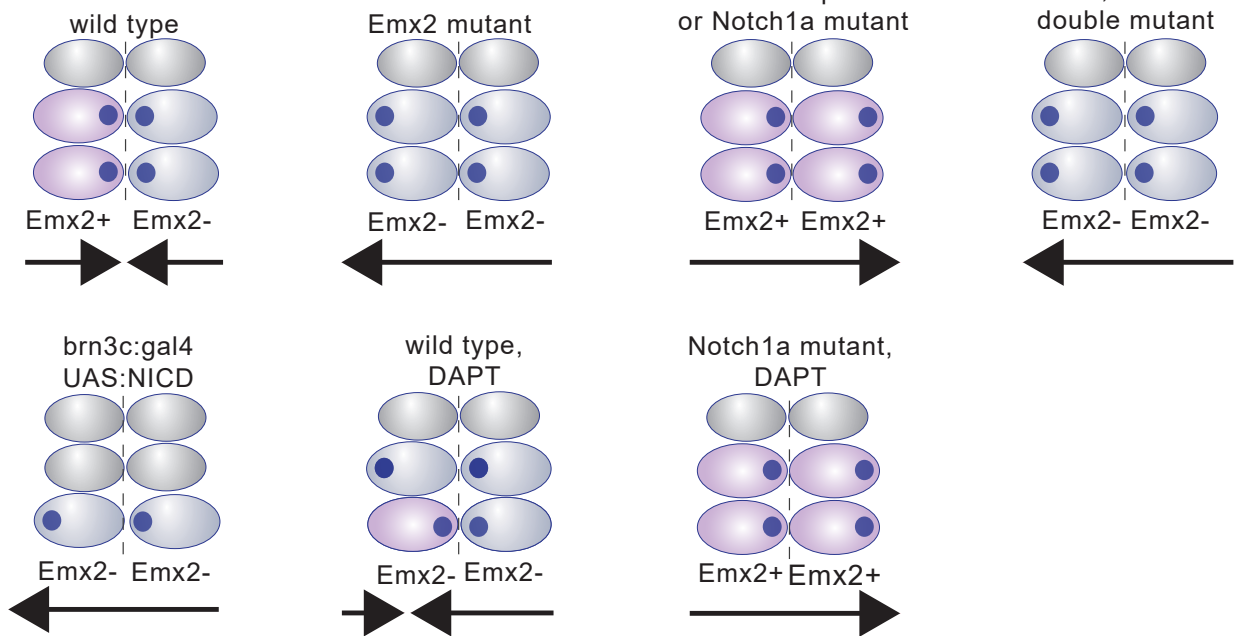
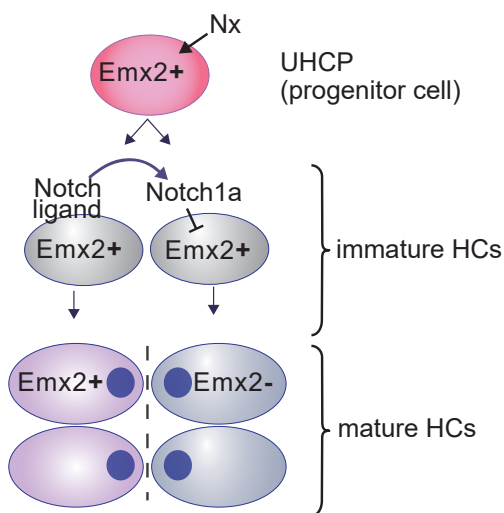


Figure 4**B****C**