PNAS

# Supporting Information for

SoxC transcription factors shape the epigenetic landscape to establish competence for sensory differentiation in the mammalian organ of Corti.

Xizi Wang[1,2], Juan Llamas[1,2], Talon Trecek[1,2], Tuo Shi[2], Litao Tao[1,2], Welly Makmura[1,2], J. Gage Crump[2], Neil Segil[1,2,3], and Ksenia Gnedeva[1,2,*]

[1] USC Caruso Department of Otolaryngology – Head and Neck Surgery, and
[2] Department of Stem Cell Biology and Regenerative Medicine, Keck School of Medicine of University of Southern California
[3] Deceased

* to whom correspondence may be addressed
Email: gnedeva@usc.edu

**This PDF file includes:**

- Supporting text
- Figures S1 to S7
- Legends for Datasets S1 to S3
- SI References

**Other supporting materials for this manuscript include the following:**

- Datasets S1 to S3

## Materials and Methods

### Primary cochlear progenitor and supporting cell isolation and cultures

Organotypic cochlear cultures were described previously (1). Briefly, cochleae were dissected at E12.0 and E13.5 from *Atoh1*-GFP mice. Isolated organs were placed in 400 ul collagen matrix containing 100ul collagen I gel solution (Sigma), 70 ul 0.9M sodium bicarbonate (Sigma), 133 ul 0.34M sodium hydroxide (Sigma), 40 ul 1M 4(2hydroxyethyl)-1 piperazineethanesulfonic acid (HEPES), and 50 ul 10x PBS. After embedding the organs and gel solidification, the wells were flooded with 500 ul of hair cell medium containing DMEM-F12, N2, B27, 20mg/L epidermal growth factor, 20mg/L fibroblast growth factor,15mM HEPES, and Insulin Transferin-sodium Selenite.

For E12.0-14.5 cultures, single-cell suspension was acquired from micro-dissected cochlear ducts. Inner ears were first dissected in sterile PBS and enzymatically digested for 10 min at RT in 0.25% dispase (Gibco)/0.25% collagenase I (Gibco) diluted in DMEM. Ears were then transferred to a new dish field with sterile PBS and epithelial cochlear ducts were dissected and further enzymatically digested for 10 min at 37°C with 0.125% Trypsin-EDTA (Sigma). Enzymatic reaction was stopped by adding 10% FBS (Signa) and cells were spun down for 5 min at 300 RCF at 4 °C. Cells were plated onto 10% Matrigel-coated (Sigma) 96-well tissue culture plates and cultured in the hair cell medium for 2-3 days.

For P6 cultures, cochlear ducts were dissected and enzymatically digested for 10 min at 37°C with 0.125% Trypsin-EDTA (Sigma). FBS was added to stop the reaction and cells were dissociated into single-cell suspension by trituration for 5 minutes. Cells were then filtered through 40μm cell strainer (BD Biosciences) to remove residual undissociated ECM and Lfng-GFP positive supporting cells were isolated via FACS. To improve cell survival, cells were collected directly into hair cell medium containing ROCK Inhibitor (Y-27632; Sigma). Supporting cells were mixed with peri-otic mesenchyme (2) and plated onto poly-D-lysine- (0.5 mg/ml; Sigma) and fibronectin- (25 ug/ml; Gibco) coated cover slips.

For both embryonic and postnatal primary cell cultures, cells were infected with Adenovirus Type 5 (Ad5) containing the full-length coding sequence of *RFP*, *Sox4-IRES-RFP*, *Sox11-IRES-RFP*, *Sox2-IRES-RFP*, or *Atoh1-IRES-RFP*. The medium containing virus was replaced after 24hr with fresh hair cell medium.

### Immunohistochemistry and RNA *in situ* hybridization.

For whole-mount organ staining, cochlea ducts were obtained as described above and fixed for 30 min in 4% paraformaldehyde. Organs were blocked for 4 hrs at room temperature using the normal donkey serum blocking buffer containing 20 mM Tris-Buffered Saline (10x TBS; Bio-Rad), 0.1% Tween-20 (Sigma-Aldrich), and 5% normal donkey serum (Sigma-Aldrich). The primary antibodies were diluted in the same blocking buffer and incubated overnight at 4 ˚C. Primary antibodies used include goat anti-Sox2 (Invitrogen), rabbit anti-Myo7A (Proteus Bioscience), chicken anti-GFP (Abcam), rabbit anti-GFP (Torrey Pines Biolabs), guinea pig anti-Sox4 (3). Samples were washed with 20 mM TBS supplemented with 0.1% Tween 20 (Sigma-Aldrich) (TBST) 3 times, 5 min each. Secondary antibodies conjugated to Alexa Fluor dyes (Life Technologies) were diluted in TBST for 2hr at room temperature or at 4˚C overnight. 3 μM DAPI (Sigma-Aldrich) were used for nuclei labeling.

For cryo-sections, the whole inner ears were dissected and fixed using 4% paraformaldehyde for 45 min at RT or overnight at 4˚C. Tissue was then transferred to 30% sucrose in PBS overnight at 4˚C, embedded in Tissue-Tek O.C.T. (Sakura), and snap frozen on dry ice. Cryo-sections at 12 μm thickness were obtained using cryostats and kept at room temperature for 24hr. For immunohistochemistry, if necessary, antigen retrieval was performed in the citrate acid buffer for 5 min at 95˚C. Sections were then blocked using the previously described blocking buffer at room temperature for 1 hr to overnight. Primary antibodies including mouse anti-Isl1 (Invitrogen), rabbit anti-Ebf1 (Millipore), mouse anti-Bcl11a (Abcam), goat anti-Sox2 (Invitrogen) were applied in the blocking solution overnight at 4 ˚C. Secondary antibodies (Life Technologies) were diluted in TBST for 2hr at room temperature or 4 ˚C overnight. 3 μM DAPI (Sigma-Aldrich) were used for nuclei labeling. For RNA *in situ* hybridization, we utilized RNAScope Assay per manufacturer instructions with following modifications (ACDBio). Pretreatment steps were performed using RNAScope H2O2 and Protease Reagents (ACDBio 322831). Inner ear sections were washed in 1x phosphate-buffered saline (PBS) for 5 min at room temperature to remove OCT. Slides were then baked for 30 min at 60 °C and post-fixed in pre-chilled 4% paraformaldehyde for 15 min at 4 °C. This was followed by dehydration in a series of ethanol (EtOH)/PBS (50%, 70%, and 100% EtOH) immersion for 5 min at room temperature. Slides were then air dried for 5 min at room temperature and hydrophobic

barriers were created surrounding sections using ImmEdge hydrophobic barrier pen (Vector Laboratories H-4000). RNAScope Hydrogen Peroxide was then applied to each slide, incubated for 10 min at room temperature, and washed twice in distilled water at room temperature. Without performing the target retrieval step, RNAScope Protease III was applied to each slide and incubated for 4 min at 40 °C and then washed twice in distilled water at room temperature. RNAScope probes for *Sox4* (Mm-Sox4-C2; ACDBio 471381-C2) and *Sox11* (Mm-Sox11-O6; ACDBio 852081) were prewarmed to 40 °C for 10 min and then cooled to room temperature to ensure there was no precipitation in the solution. Immediately following pretreatment, enough probe mixture was applied to each slide to cover tissue sections. They were then incubated for 2 hours at 40 °C. Following probe hybridization, slides were washed in 1x RNAScope Wash Buffer (ACDBio 310091) twice for 3 min each at room temperature. Slides were then submerged in 5x saline sodium citrate (SSC) (diluted from 20x stock of 3M NaCl and 0.3M Sodium Citrate) overnight at room temperature for storage. Amplification and signal development steps were performed using RNAScope Multiplex Fluorescent Detection Kit (ACDBio 323110) and Opal fluorophore from Akoya Biosciences (Opal 570, PN FP1488001KT; Opal 690, PN FP1497001KT). Slides were removed from overnight 5x SSC and washed twice in 1x wash buffer for 3 min each at room temperature, followed by sequential incubation at 40 °C in RNAScope AMP1 (30 min), AMP2 (30 min), and AMP3 (15min) reagents. After each incubation, slides were washed twice in 1x wash buffer for 3 min each at room temperature. After AMP3 incubation and washes, slides were then incubated at 40 °C with RNAScope C1-HRP for 15 min, Opal 570 fluorophore (1:1500 dilution in TSA buffer) for 30 min, and HRP blocker for 15 min to develop signal for *Sox11*. Following each incubation, slides were washed twice in 1x wash buffer for 3 min each at room temperature. This was then repeated with C2-HRP, Opal 690 (1:1500 dilution in TSA buffer), and HRP blocker to develop signal for *Sox4*. Following the last wash, slides were then blocked and immunostained with goat-anti-Sox2 (Invitrogen) and donkey-anti-goat (Abcam) primary and secondary antibodies as described above.

**Molecular cloning of viral plasmids and virus production**

Adenovirus for *Sox4* and *Sox11* overexpression were described previously (4). Full-length coding sequences for *Sox4* and *Sox11* genes also cloned in frame with T2A-GFP to be expressed under the control of a cytomegalovirus promoter in the AAV vectors using Gibson assembly (NEB). The pAnc80L65AAP viral capsid ((5); Addgene plasmid 92307) was used to package adeno-associated virus in HEK293T cells. Virus was then purified by CsCl-gradient centrifugation followed by dialysis (Viral Vector Core Facility, Sanford Burnham Prebys Medical Discovery Institute). To achieve transduction of the inner ear sensory organs, each animal was injected at P5 into the lateral ventricle with 5 ul of the virus at a titer of $10^{12}$ PFU/mL (6).

**Single-cell RNA-sequencing and analysis**

Cochlear duct epithelia from E13.5 wildtype and *SoxC* conditional double knockout animals were collected as described above. Single-cell suspensions were obtained by trypsin treatment for 10min at 37 ˚C. The standard protocol for the 10X single cell kit (V3.0) was followed and each sample was loaded onto a separate lane on the same chip. Samples were then sequenced using Illumina NextSeq kit.

Sequencing data from 10x runs was aligned and quantified using the CellRanger software using default parameters. Sequencing results for both reached 50% saturation. Quality controls and data analysis were done using Seurat (7, 8) and Scenic packages (9).

**ATAC sequencing**

Cochlear duct epithelia from E13.5 wildtype and *SoxC* conditional double knockout animals were collected as described above. Whole epithelia were lysed in the buffer containing 10mM TrisHCl (pH=8.0), 5mM $MgCl_2$, 10% DMF, 0.2% NP40 in $ddH_2O$. Transposition buffer containing 10 mM Tris-HCL (pH=8.0), 5 mM $MgCl_2$, 10% DMF, 0.2% NP40, and Tn5 transposes was then added and incubate at 37°C for 20 min. After DNA purification, amplification and library construction was performed through PCR. At least 50 million paired-end reads were sequenced for each sample. Encode pipelines were adapted for analysis for ATAC-seq data.

**CUT&RUN**

The CUT&RUN method for *in situ* chromatin immunoprecipitation was described previously (10) and was used to profile *Sox4* occupancy of the chromatin in E13.5 progenitors. Protein A-MNase fusion

protein was obtained from Dr. Henikoff's laboratory at Fred Hutchinson Cancer Research Center. Guinea pig anti-Sox4 antibody was used (3). CUT&RUN libraries were constructed using Accel-NGS 2S plus DNA prep kits with single index and MIDs (Swift Bioscience). At least 20 million paired-end reads were sequenced for each sample.

Encode pipelines were adapted for analysis of CUT&RUN data. Reads from the raw fastq files were aligned to GRCm38/mm10 genome assembly using STAR package (11). PCR duplicates were detected, and peaks were called by Model-based analysis of ChIP-Seq (MACS2) with FDR< 0.01 and the dynamic lambda (--nolambda) option for individual replicates. For each sample, IDR peaks, overlap peaks, and pooled peaks were identified between the biological replicates. BamCoverage files were generated directly from BAM files after sorting. Bigwig files were generated with deepTools (12). IGV (13) was used to visualize the genomic loci from Bigwig files. Heatmaps were generated with deepTools based on normalized Bigwig signal files. HOMER (14) suite was used for identifying transcription factor motif enrichment of subsets of the picked genomic regions. Closest genes and overlapping genes were identified using BEDOPS package with "closest-features" and "bedmap" functions respectively.

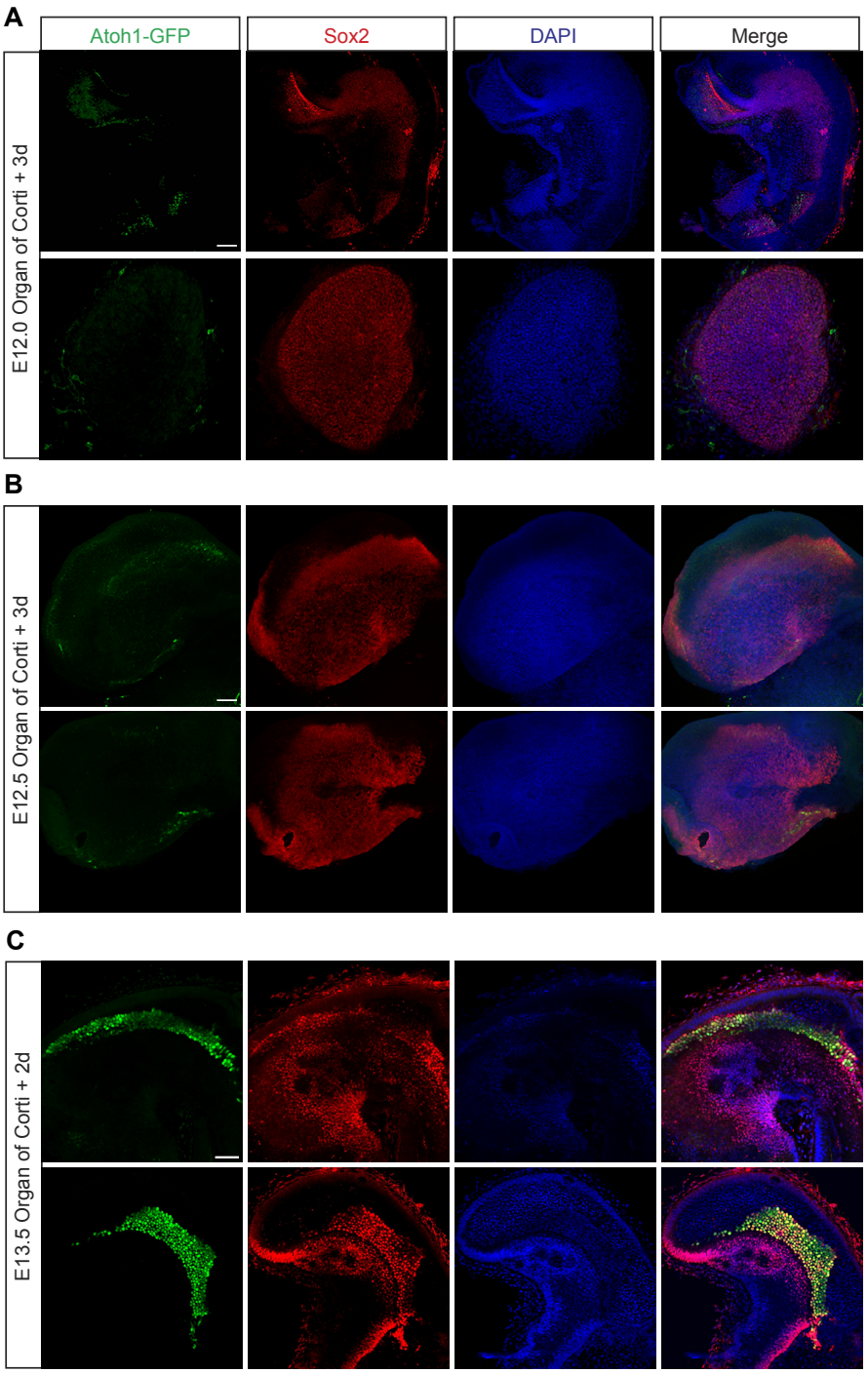


**Figure S1.** Prosensory progenitors in the cochlear duct acquire competence to differentiate as hair cells and supporting cells between E12.0 and E13.5. Representative immunofluorescent images of the whole cochleae isolated from (A) E12.0, (B) E12.5, and (C) E13.5 *Atoh1-GFP* transgenic reporter mice and harvested for characterization after 3-day or 2-day in culture. Atoh1-GFP-positive hair cells (green) and Sox2-positive progenitor cells/supporting cells (red) are labeled. The nuclei are stained with DAPI (blue). Scale bar = 100 μm.

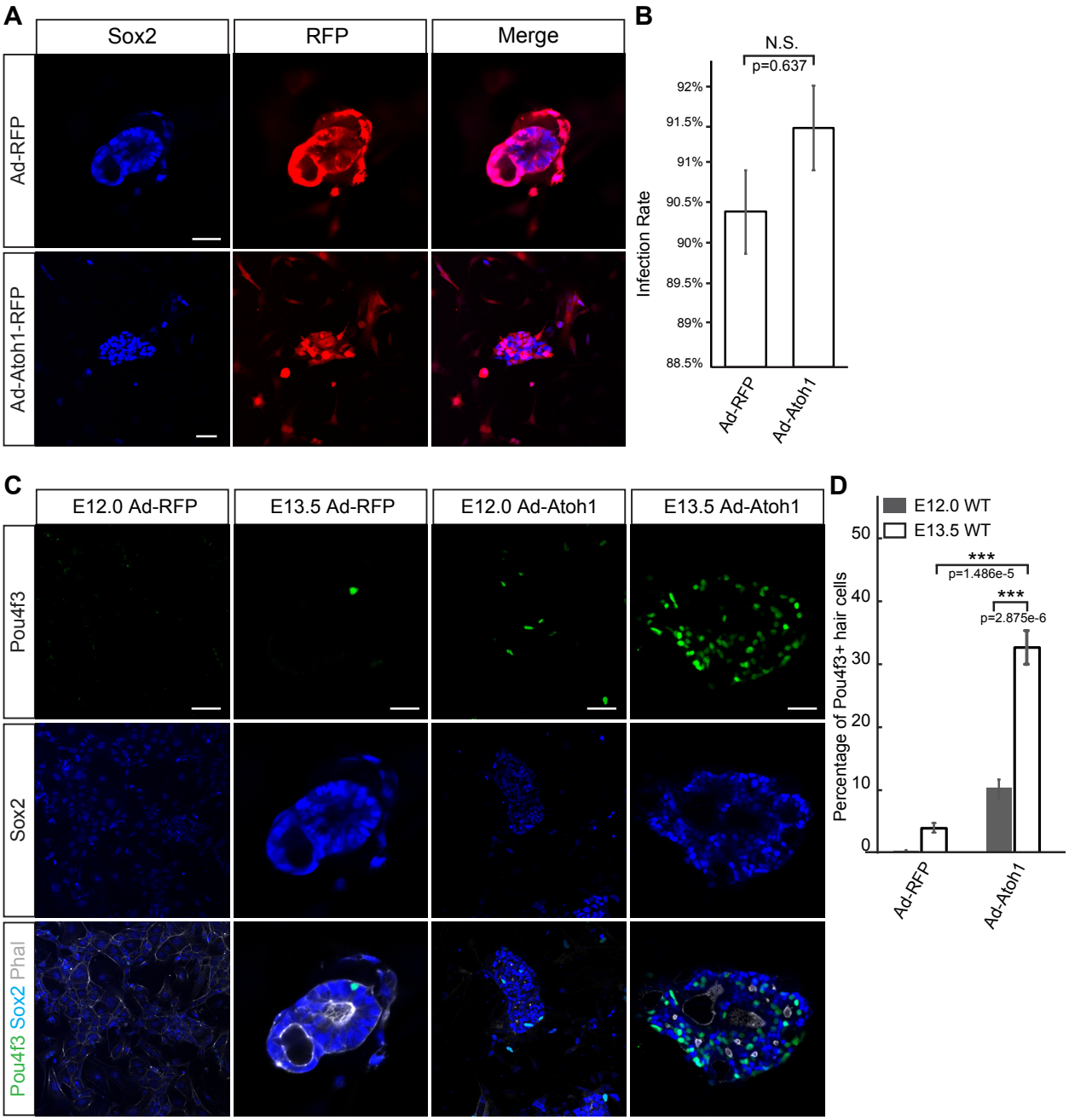


**Figure S2.** The competent chromatin state established in the postmitotic progenitor cells is required for Atoh1-mediated hair cell differentiation. (A) Immunofluorescence analysis of the dissociated cochlear cells isolated from E12.0 animals and infected with *Ad-RFP* or *Ad-Atoh1-RFP* viral vectors demonstrates the viral transduction efficiency. Widespread RFP expression (red) is observed in both conditions. Scale bar = 50 μm. (B) Bar graphs show the quantitative analysis of viral infection rate in A. (C) Representative immunofluorescent images show the dissociated progenitor cells isolated at E12.0 and E13.5 cochlea from the wildtype animals, infected with *Ad-RFP* control (left two columns) or *Ad-Atoh1-RFP* (right two columns) virus, and maintained in culture for 3 days. Note that only in E13.5 cultures *Atoh1* overexpression results in formation of the sensory rosettes with a small lumen formed by the actin-rich (Phalloidin, white) apical surfaces of the polarized Pou4f3-positive hair cells (green) and surrounding Sox2-positive supporting cells (blue). Scale bar = 50 μm. (D) Bar graphs show the quantitative analysis of

cultures in C. Compared to other conditions, a significant increase of the percentage of Pou4f3-positive hair cells is observed in the E13.5 cultures where *Atoh1* is overexpressed (n=3).

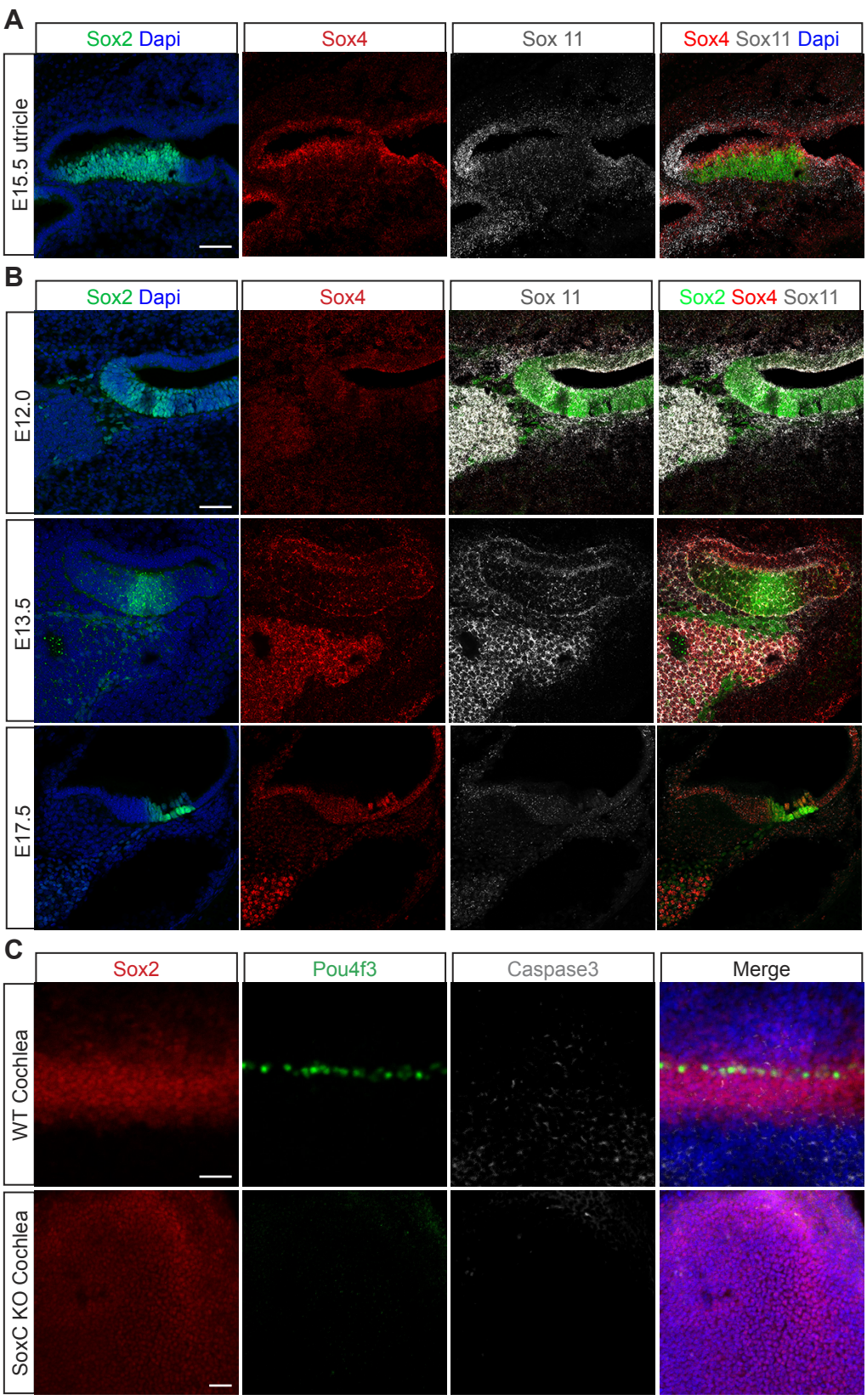


**Figure S3.** SoxC transcription factors are essential for initiation of sensory differentiation during the development of organ of Corti. (A) Representative images show RNAscope *in situ* hybridization for *Sox4*

(red) and *Sox11* (white) in the utricles from WT embryos at E15.5. Sox2-positive supporting cells (green) are also immunolabeled. Scale bar = 50 μm. (B) Representative images show RNAscope *in situ* hybridization for *Sox4* (red) and *Sox11* (white) in the developing organ of Corti (E12.0-E17.5). While *Sox11* is highly expressed in the cochlear duct and the spiral ganglion with the signal strongest at E12.0 and decreasing by E17.5. *Sox4* expression is increased from E12.0 to E13.5 and is strongest in the developing hair cells and spiral ganglion neurons at E17.5. Sox2-positive supporting cells (green) are also immunolabeled. Scale bar = 50 μm. (C) Representative immunofluorescent images show the whole cochlea isolated from the WT and *SoxC* KO littermate embryos at E14.5. Sox2-positive supporting cells (red) and one row of Pou4f3-positive inner hair cells (green) are detected in the WT but not in the *SoxC* KO organs. Note that no apoptotic cells are detected by activated caspase 3 staining (white) in either condition. Cell nuclei are labeled with DAPI (blue). Scale bar = 50 μm.

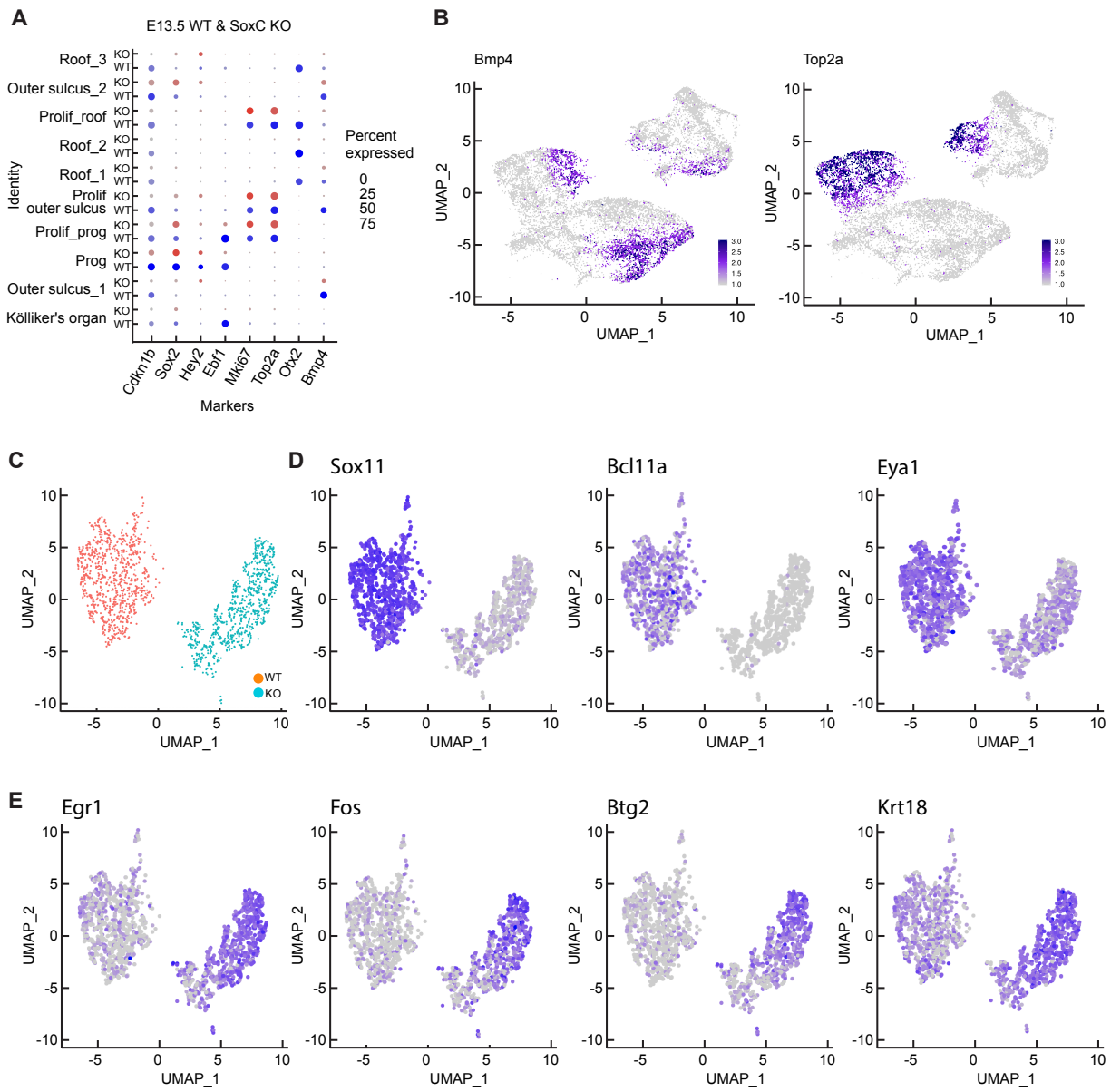


**Figure S4.** Single cell RNA-seq indicates that SoxC transcription factors promote hair cell differentiation by regulating key sensory lineage genes. (A) Dot plot demonstrates the gene expression profiles of selected cluster-specific markers in previously known cell types identified in the WT or *SoxC* KO cochlea. The organ of Corti progenitor cell population (Prog) expresses high levels of Cdkn1b, Sox2, Hey2 and Ebf1; The Kölliker's organ cells are also positive for Sox2 and Ebf1, but are negative for Cdkn1b and Hey2; Otx2 is a specific marker for the roof domain of the cochlea duct (Roof); Bmp4 expression is detected specifically in the outer sulcus domain (outer sulcus). Mki67 and Top2a expression indicates the proliferating populations of the Kölliker's organ and the organ of Corti (Prolif_prog), outer sulcus (Prolif_ outer sulcus) and roof domains (Prolif_roof). A full list of top markers for each cell cluster can be found in Dataset S1. (B) UMAP feature plots demonstrate transcript levels for cell markers specific to the known cochlear cell types. The outer sulcus is identified based on the Bmp4 expression. The active proliferating cells in the cochlea duct are identified by high level of Top2a expression. (C) UMAP visualization of WT (orange) and *SoxC* KO (green) organ of Corti progenitor cell population at E13.5. (D) UMAP feature plots

demonstrate that essential sensory lineage genes such as Sox11, Bcl11a, and Eya1 are significantly downregulated in *SoxC* KO progenitor cells compared to control. A full list of differentially expressed genes can be found in Dataset S2. (E) UMAP feature plots demonstrate that several genes such as Egr1, Fos, Btg2 and Krt18 are significantly upregulated in *SoxC* KO progenitor cells compared to control.

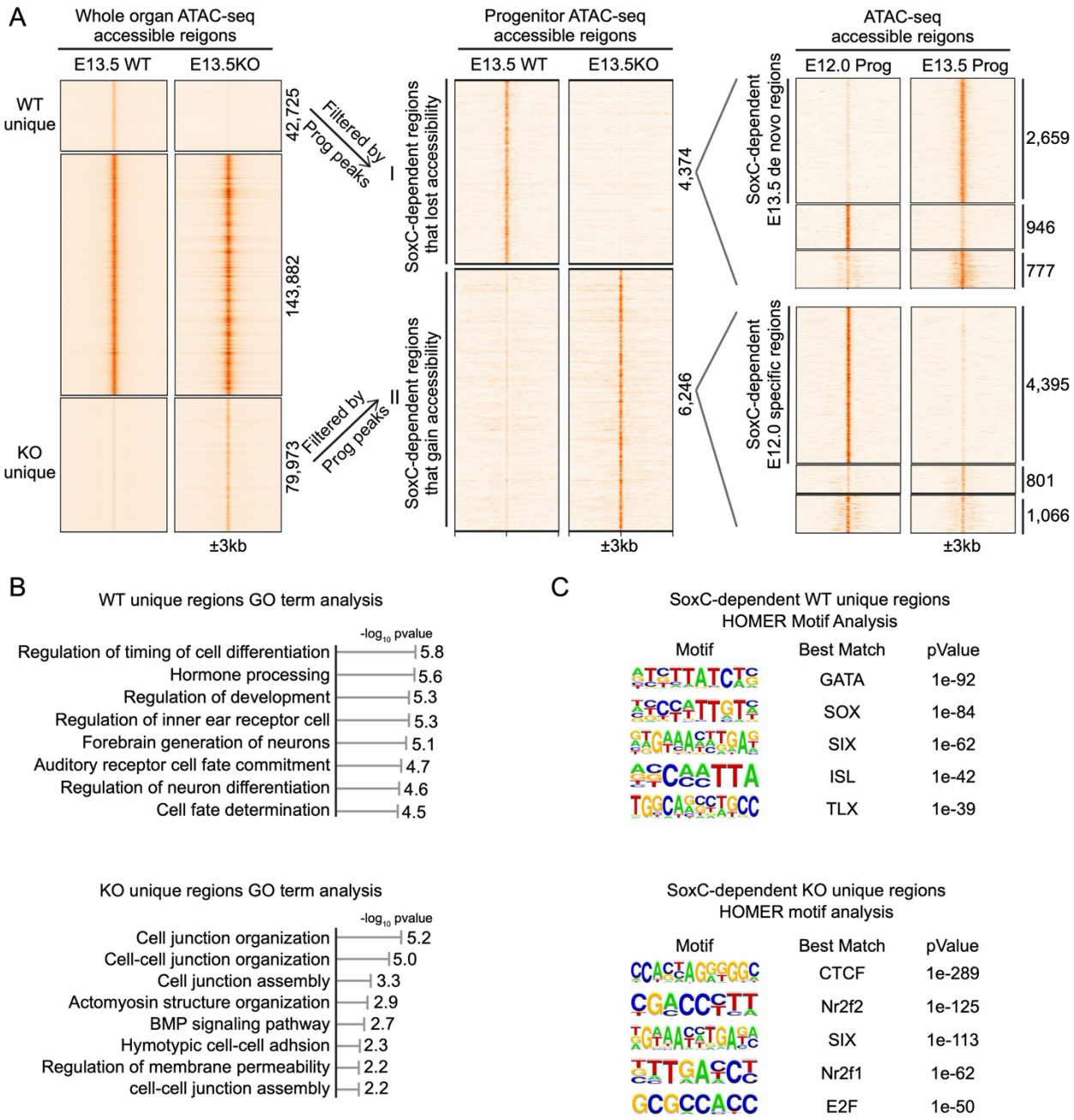


**Figure S5.** SoxC transcription factors control chromatin accessibility to promote hair cell differentiation in the E13.5 organ of Corti. (A) A Heatmap on the left shows all accessible genomic regions identified by ATAC-seq in the WT and *SoxC* KO cochlea at E13.5. In the middle heatmap, regulatory elements that lose (42,725; SoxC-dependent regions, WT unique) or gain (79,973; SoxC-dependent regions, KO unique) accessibility in the *SoxC* KO cochlea (SoxC-dependent regions) are filtered by their accessibility in both E12.0 and E13.5 purified progenitors (Prog). Out of 4,374 peaks in Cluster I, 2,659 peaks are E13.5-specific. Out of 6,246 peaks in Cluster II, 4,395 were E12.0-specific. (B) Gene ontology (GO) terms identified using GREAT show the top 10 biological processes most-enriched in the genes associated with Cluster I and Cluster II – SoxC-dependent regions that lose and gain accessibility, respectively. (C) HOMER analysis shows top five most enriched DNA-binding motifs in SoxC-dependent peaks lost their accessibility (Cluster I) and gained accessibility (Cluster II), respectively.

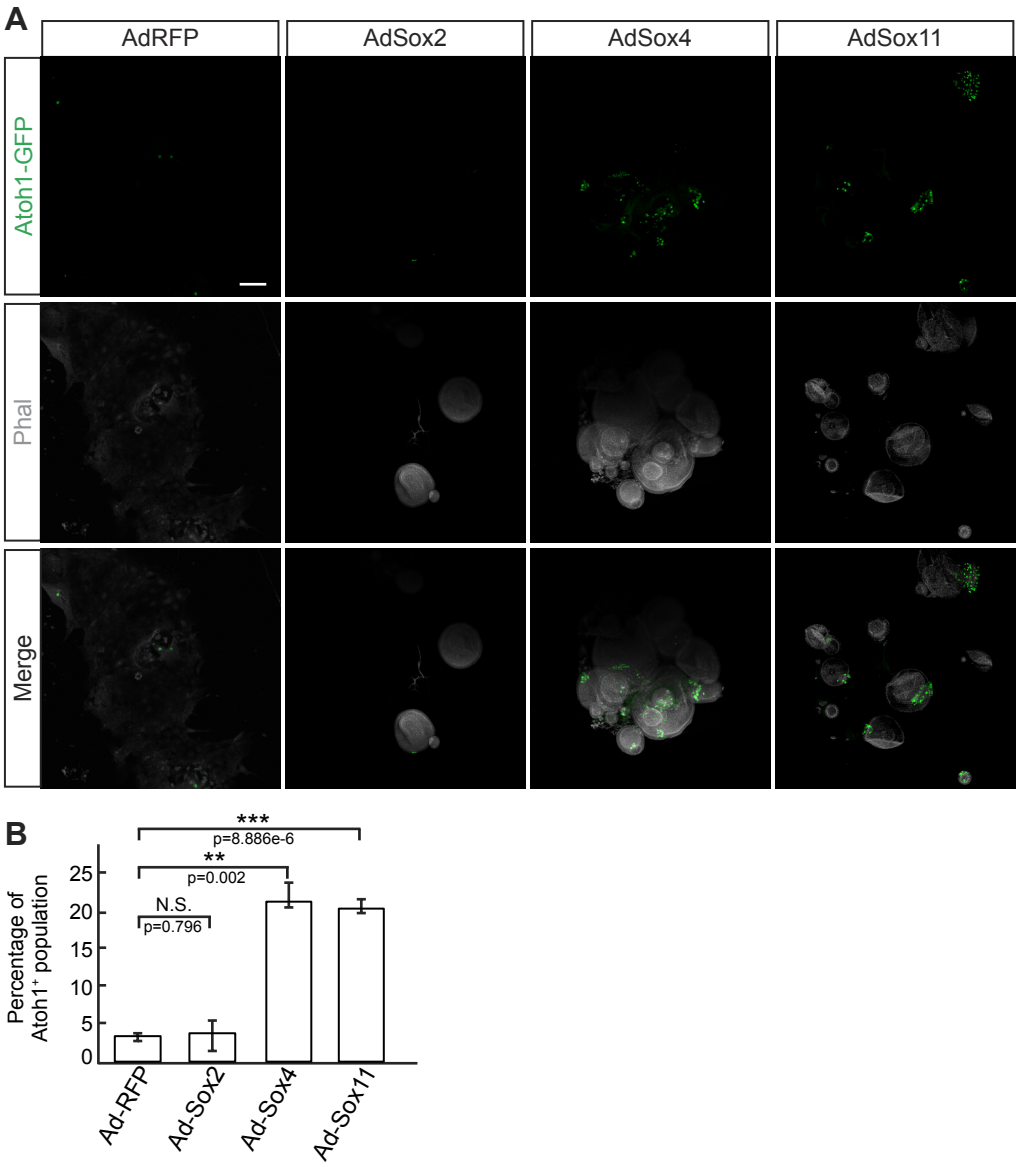


**Figure S6.** SoxC transcription factors trigger new hair cell generation in the non-competent organ of Corti progenitor cells *in vitro*. (A) Representative low magnification immunofluorescent images demonstrate that overexpression of Sox2 in the dissociated cochlea progenitor cells isolated from *Atoh1-GFP* reporter mice fails to induce sensory differentiation. In contrast, overexpression of either Sox4 or Sox11 promotes hair cell differentiation under the same conditions. Atoh1-GFP-positive hair cells (green) in the *Sox4* and *Sox11* overexpressing cochlea cultures and are organized into the sensory rosettes. Actin is labeled in white (phalloidin). Scale bar = 100 μm. (B) Bar graph shows quantitative analysis of the cultures in A. The percentages of Atoh1-GFP-positive hair cell are low in RFP control and *Sox2* overexpression conditions compared to the SoxC overexpression groups (n=3 for each condition).

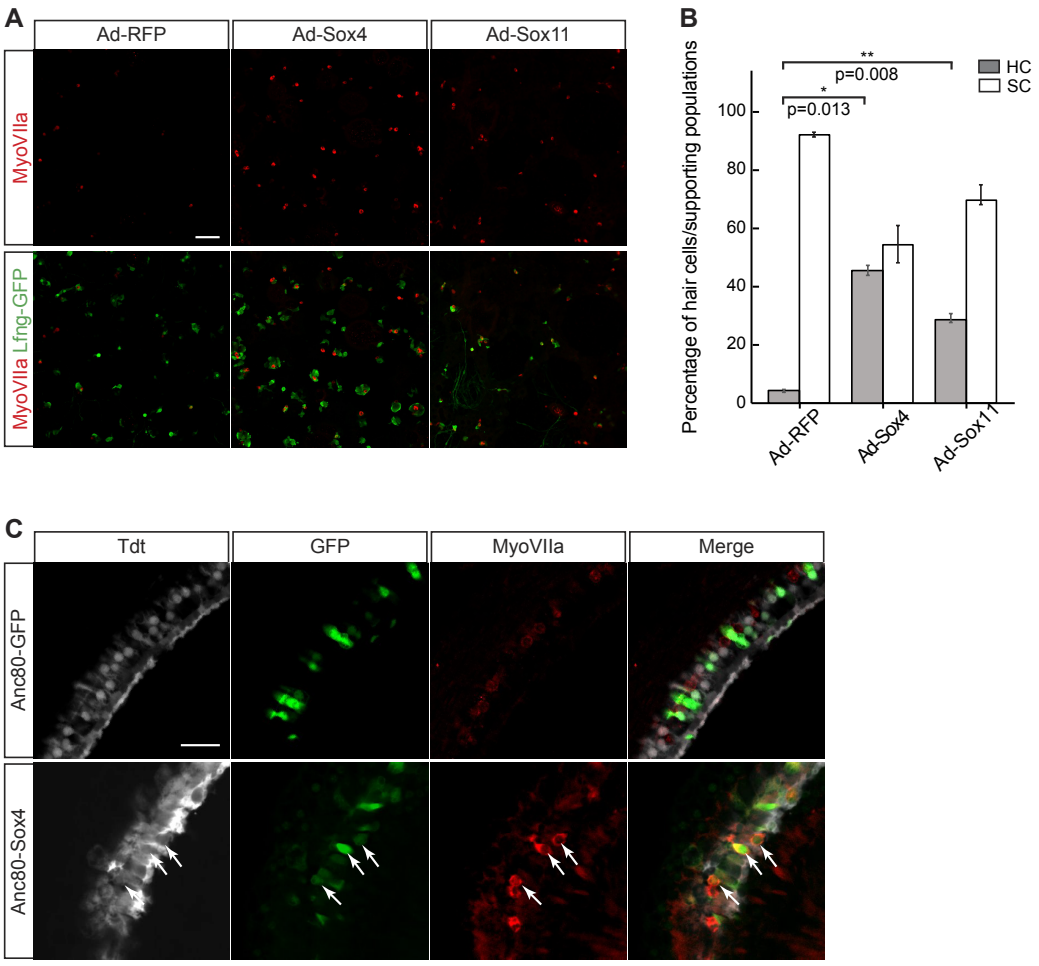


**Figure S7.** SoxC transcription factors trigger new hair cell generation in the postnatal organ of Corti supporting cells in vitro and in vivo. (A) Representative immunofluorescent images show that FACS purified P6 Lfng-GFP-positive cochlea supporting cells (green) were transdifferentiated into MyoVIIa-positive hair cells (red) after 7 days in culture upon *Sox4* or *Sox11* overexpression (Scale bars = 50 μm). (B) Bar graphs show quantitative analysis of cultures in A. A significant increase in percentage of MyoVIIa-positive hair cells is observed in *Sox4* or *Sox11* overexpression conditions compared to the *RFP*-control (n=3 for each condition). (C) Representative immunofluorescent images show the P15 whole-mount cochlea from *Sox4* overexpression conditions. Lfng-positive supporting cells are labeled in grey, infected cells are labeled in green, hair cells are stained with MyoVIIa in red. White arrows indicate transdifferentiated triple labeled cells. Scale bar = 20 μm.

**Datasets S1.** Top20 expressed genes in each cluster identified in WT and SoxC cKO cochlea.

**Dataset S2.** Genes differentially expressed between WT and SoxC KO postmitotic progenitor cells at E13.5 (cutoff: logFC > 0.6).

**Dataset S3.** Sox4-occupied genomic loci that gain accessibility between E12.0 and E13.5 and lose accessibility in SoxC cKO.

## Reference


1. K. Gnedeva, A. J. Hudspeth, N. Segil, Three-dimensional Organotypic Cultures of Vestibular and Auditory Sensory Organs. *J. Vis. Exp.* (2018) https:/doi.org/10.3791/57527.
2. A. Doetzlhofer, P. M. White, J. E. Johnson, N. Segil, A. K. Groves, In vitro growth and differentiation of mammalian sensory hair cell progenitors: a requirement for EGF and periotic mesenchyme. *Dev. Biol.* **272**, 432–47 (2004).
3. D. C. Thein, *et al.*, The closely related transcription factors Sox4 and Sox11 function as survival factors during spinal cord development. *J. Neurochem.* **115**, 131–41 (2010).
4. K. Gnedeva, A. J. Hudspeth, SoxC transcription factors are essential for the development of the inner ear. *Proc. Natl. Acad. Sci.* **112**, 14066–14071 (2015).
5. L. D. Landegger, *et al.*, A synthetic AAV vector enables safe and efficient gene transfer to the mammalian inner ear. *Nat. Biotechnol.* **35**, 280–284 (2017).
6. K. Gnedeva, *et al.*, Organ of Corti size is governed by Yap/Tead-mediated progenitor self-renewal. *Proc. Natl. Acad. Sci.* **117**, 13552–13561 (2020).
7. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
8. T. Stuart, *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
9. S. Aibar, *et al.*, SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
10. P. J. Skene, S. Henikoff, An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6** (2017).
11. A. Dobin, *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
12. F. Ramírez, *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
13. J. T. Robinson, *et al.*, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
14. S. Heinz, *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–89 (2010).