PNAS

Supplementary Information for
SOX9 and SOX10 control fluid homeostasis in the inner ear for hearing through independent and cooperative mechanisms

Irene, Y.Y. Szeto[1]†[*], Daniel, K.H. Chu[1]¶, Peikai Chen[1]||¶, Ka Chi Chu[1], Tiffany, Y. K. Au[1], Keith, K.H. Leung[1], Yong-Heng Huang[2]‡, Sarah, L. Wynn[1], Angel, C.Y. Mak[1]§, Ying-Shing Chan[1], Wood Yee Chan[3], Ralf Jauch[1,2], Bernd Fritzsch[4], Mai Har Sham[1,¥], Robin Lovell-Badge[5], Kathryn, S.E. Cheah[1*]

[1]School of Biomedical Sciences, The University of Hong Kong, Li Ka Shing Faculty of Medicine, 21 Sassoon Rd, Hong Kong, China.
[2]Genome Regulation Laboratory, CAS Key Laboratory of Regenerative Biology, Joint School of Life Sciences, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China; Guangzhou Medical University, Guangzhou 511436, China.
[3]School of Biomedical Sciences, The Chinese University of Hong Kong, Lo Kwee-Seong Integrated Biomedical Sciences Building, Shatin, Hong Kong, China.
[4]Department of Biology, College of Liberal Arts & Sciences, University of Iowa, Iowa City, IA 52242, USA.
[5]Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK.

Present addresses: †Department of Molecular Medicine, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany; ‡ Center for Cell Lineage and Atlas (CCLA), Bioland Laboratory (Guangzhou Regenerative Medicine and Health GuangDong Laboratory), Guangzhou 510005, China; § Department of Medicine, University of California, San Francisco, 1550 4th St, Box 2911, San Francisco, CA 94143; ||Hong Kong University – Shenzhen Hospital, 1 Haiyuan Rd, Shenzhen, China; ¥School of Biomedical Sciences, The Chinese University of Hong Kong, Lo Kwee-Seong Integrated Biomedical Sciences Building, Shatin, Hong Kong, China

¶Equal contribution.
[*]Corresponding author

Email: kathycheah@hku.hk (K.S.E.C)

This PDF file includes:

Supplementary text methods
Figures S1 to S6
Tables S1 to S2
Legends for Movies S1 to S5
Legends for Datasets S1 to S2
SI References

Other supplementary materials for this manuscript include the following:

Movies S1 to S5
Datasets S1 to S2

Supplementary Methods

Endocochlear potential measurement

Two small bone screws (1 mm diameter, FHC) used as recording electrodes were fixed above the cerebellum on either side of the midline. Two stainless steel pins (0.5 mm diameter) were inserted subcutaneously into the left and right pinnae to serve as reference electrodes. One stainless steel pin was inserted subcutaneously at the mid-position of the lower jaw to serve as the ground electrode.

Sound stimulation with 0.1 ms rectangle pulse and sound pressure level of 35-120 dB was used. Signals collected from recording electrodes were amplified (x 10,000) by amplifier (Grass P55) with a band pass filter 100-3,000 Hz and were then sampled by PowerLab system (PowerLab 4SP, AD Instruments) at 40 kHz. Auditory brainstem responses (ABR) obtained from 512 or 1,024 stimuli were averaged by software (Scope). Auditory threshold was defined as the lowest stimulus intensity at which an ABR was observed in 2 averaged runs.

To measure the endocochlear potential, a small hole just big enough for the recording electrode was opened on the basal turn of the cochlea. A glass micropipette filled with 150 mM KCl was advanced up to the outside of spiral ligament. The baseline of the recording was set to zero. Stable endocochlear potential was recorded by inserting the probe into the endolymphatic space. A silver wire was placed under the skin on the vertex of the skull as reference electrode.

Scanning electron microscopy (SEM), analyses of sound conducting chambers and cochlear lumen

Inner ears from *B2-Sox9*$^{cY440X/+}$ mice and control littermates at P4 were processed for SEM as follows. Inner ears from fetuses were dissected in 0.1 M sodium phosphate buffer quickly and fixed for 5-6 hours in 2.5% glutaraldehyde (EM grade Sigma) in 0.1 M sodium phosphate buffer. Fixed cochleae were further dissected to expose the organ of Corti, followed by removal of the tectorial membrane. The samples were then post-fixed in 1% osmium tetroxide for 1 hour, dehydrated in a graded series of ethanol solutions and dried by Critical Point Drying. The samples were mounted, coated by sputtering (Edwards sputter coater S150B) and ready to be observed by scanning electron microscope (JEOL JSM 6301F).

Quantitation of cochlea lumen

Adult inner ear of both *B2-Sox9*$^{cY440X/+}$ and control adult mice were analysed by three-dimensional reconstruction of 5 μm sections and measurement of sound-conducting chambers. The sound-conducting chambers, namely, scala vestibuli, scala media and scala tympani were outlined manually and color-coded. One in every three consecutive sections was outlined and 100-150 sections were outlined for each sample. The area of the outlined chambers was calculated by the software Neurolucida. The volume of the chambers was calculated as follows: sum of chamber area of all sections x no. of interval slides (i.e. 3) x thickness (i.e. 5 μm). Heads of both E15.5 mutant and control embryos were sectioned in sagittal plane at 7 μm thickness. five sections of basal cochlear lumens in each embryo (n=4-5) were manually outlined and the outlined areas were calculated using ImageJ software.

Single cell RNA-seq (scRNAseq)

Endolymphatic sacs of WT and mutant were harvested together with surrounding tissues as described (1) and incubated in 1 mg/ml collagenase (LS004194, Worthington, Lakewood, NJ) and 1 mg/ml dispase (LS02100, Worthington) in DMEM/F12 at 37°C for 10 min. The epithelia were dissected from the rest of the tissues in 10% fetal bovine serum (FBS) in DMEM/F12 (Gibco). Finally, the isolated epithelia were digested in 50 µl of 0.125% trypsin/EDTA/PBS for 15 min at 37°C. 50 µl of 10% FBS in DMEM/F12 was subsequently added to stop the reaction. The epithelia were gently triturated with 200 µl pipette tips and transferred to a 6 cm^2 tissue culture dish. Single cells were dissociated from E14.5 ESs of WT (n=3) and *Sox9*$^{Y440X/+}$ embryos (n=2) for RNA-seq as described in (2). ScRNAseq was performed on manually picked single cells (246 cells: 91 WT and 155 mutant cells) using the Smartseq protocol.

scRNAseq libraries were constructed according to Picelli *et al* in two batches (3), before being sequenced with Illumina HiSeq 1500 (first batch) and NovaSeq 6000 (second batch), at the Centre for PanorOmic Sciences (CPOS) of the University of Hong Kong.
Each cell has on average 4.5 million read-pairs, of which 89% have Phred quality scores over 30. The reads were mapped to mouse genome mm10 by HISAT2 aligner (4) and sorted by Samtools (v1.8) (5) with an average overall mapping rate of 81.8%. Raw read counts were performed by the HTSeq-count package (6) and expression levels were normalized by Cufflinks (7), using gene annotation file from GENCODE (8). A median of 3,638 protein coding genes were detected per cell, which were comparable among the subsequently identified populations.

Bioinformatics analyses

Informative markers were unbiasedly detected based on over-dispersion using the PAGODA package (9) before principal component analysis (PCA) was performed using the built-in PCA functions of R package. Genome-wide PCA was also formed. Dimension reduction methods such as tSNE was applied showing comparable results (10). Given the small cell number, PCA patterns were adopted for discussion. Differentially expressed genes (DEGs) were detected by MAST (11). To detect DEGs experiencing ectopic differences, and not level differences, a further cut-off of 'ratio-gap', the difference of expressed percentages in two groups of cells, greater than 20% was applied, on top of the FDR<0.05 measure. Gene-ontology (GO) and pathway analysis were performed by GSEA (12). Pseudotime trajectory was conducted using Monocle (13); regulon analyses were carried out using the SCENIC package (14). Cell maturity was predicted by the "pseudotime quantity" and empirical entropy. Higher entropy indicates more randomness (thus less regulated) in expression profiles, and thus higher chance of being progenitor-like. Empirical entropy for cell i was calculated based per-cell expression profile: $S_i = -\sum_j p_j \log_2 p_j$ where p_j is the frequency (number of genes) at level j (rounded to integer for logarithm of raw counts) for all levels greater than zero.

RT-qPCR *In situ* hybridization, hybridization chain reaction

In situ hybridization using digoxigenin-labeled riboprobes was performed as previously described (15). HCR were performed with reference to the HCR v3.0 protocol (Molecular Technologies).

For RT-qPCR, total RNA was extracted from cochlea tissue dissected from P3 neonates using TRIZOL according to the manufacturer's instructions. cDNAs were synthesized by Superscript II (Invitrogen) and targets were amplified using qPCR master mix (TAKARA) on the StepOne Plus

real time PCR system (Life technologies). Primers are described in *SI Appendix*, Dataset S2. Expression levels were normalized to the expression of the housekeeping gene *HPRT or Gapdh*.
HCR experiments were performed with reference to the HCR v3.0 protocol (Molecular Technologies). Embryos were fixed in 4% paraformaldehyde overnight at 4ºC, dehydrated with ethanol and embedded in paraffin for sectioning. Sections were rehydrated and permeabilized with 10 µg/ml proteinase K in PBS at 37°C for 10 min, followed by washing, pre-hybridization and hybridization with DNA probes overnight at 37°C. Excessive probes were removed by sequential washings. Sections were incubated with hairpin solutions overnight at room temperature, washed with 5X SSCT and mounted with anti-fading medium with DAPI (H-1200, Vector Laboratories). Probe sets are described in *SI Appendix*, Dataset S2.

Protein production

pDEST-hisMBP-SOX9-NHMG was kindly provided by Dr. Saravanan Vivendakan and verified by DNA sequencing (Life Technologies). SOX9-HMG, hisSOX9-HMG and hisSOX10-HMG were purified as described (16). The SOX10-DIM constructs were PCR amplified from mouse *Sox10* cDNA [AccNo. BC023356] using SOX10-DIM-for & SOX10-rev. pENTR/SOX10-DIM and pDEST-hisMBP/SOX10-DIM constructs were generated using Gateway BP and LR Cloning systems (Life Technologies). Proteins were essentially purified with established protocols (16) with the following modification. The destination plasmids were transformed into Escherichia coli BL21 (DE3) cells (TIANGEN BIOTECH). Positive clones were picked and cultured in LB (Luria-Bertani) medium containing 100 µg/ml Ampicillin at 37°C. The overnight inoculum was diluted 100-fold with fresh TB (Terrific Broth) containing 100 µg/ml Ampicillin and 0.2% glucose and further grown at 37°C and 220 rpm. When the OD600 reached 0.6-0.9, protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 18 hrs at 18°C. Bacterial cells were harvested by centrifugation at 8,000 g for 30 min at 4°C, and the pellet resuspended in 50 ml Buffer A (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM Imidazole) per 10 g pellet. Cells were disrupted using a high-pressure device (Guangzhou JuNeng Biology & Technology, JN3000 PLUS Low Temperature Ultra-High Pressure Continuous Flow Cell Disruption) at 4°C and 1200-1500 bar for 3-6 times. Lysate and debris were separated by centrifugation at 20,000 g for 30 min at 4°C. The supernatant was further cleared using MILLEX GP 0.22 µm filter unit (Millipore). Fusions proteins were bound to Ni-NTA Agarose beads (Qiagen) for 1-2 h at 4°C. Beads were washed with Buffer B (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 30 mM Imidazole) until 10 µl of the wash flow through does not cause a change of color to 1 ml of Bradford reagent (GE Healthcare). HisMBP-Sox9-NHMG and HisMBP-Sox10-DIM fusing proteins were eluted with buffer C (20 mM Tris-HCl pH8.0, 100 mM NaCl, 300 mM Imidazole) and the tobacco etch virus (TEV) protease was added at a 1:100 weight per weight ratio. The Sox9-NHMG and Sox10-DIM were further purified using a Resource S column (GE Healthcare) that was connected to AKTA Express system (GE Healthcare) and eluted by gradually increasing the NaCl concentration from 0.1 M to 1 M at a flow rate of 1 ml/min. SOX9-NHMG OR SOX10-DIM containing fractions were pooled, buffer exchanged into Desalt Buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl) using PD-10 desalting columns (GE Healthcare), concentrated using VS 2011 Ultrafiltration Concentration Devices (Sartorius Stedim Biotech GmbH), aliquoted and stored at -80°C.

Electrophoretic mobility shift assays (EMSA)

For EMSA Recombinant SOX9 and SOX10 protein constructs were expressed using Escherichia coli BL21 (DE3) and purified by immobilized metal affinity chromatography, ion exchange chromatography and desalting as described previously (17). EMSAs were carried out as described

using DNA oligos (Beijing Genome Institute) labelled with 5′ Cy5 on the forward strand. The canonical motif sequence was promoter sequence of the CD-Rap gene (18).
Binding reactions containing 25 nM dsDNA with varying concentrations of Sox proteins in 1× Reaction Buffer (10 mM Tris-HCl pH 8.0, 0.1 mg/mL BSA, 50 μM $ZnCl_2$, 100 mM KCl, 10% Glycerol, 0.10% NP-10, 2 mM 2-mercaptoethanol) were incubated for 4 h at 4°C in the dark. 10 μl (for 10 wells gel) or 6.7 μl (for 15 wells gel) of the reactions were loaded onto a pre-run 12% native 1× Tris-glycine (25 mM Tris pH 8.3; 192 mM glycine) gel, and electrophoresed at 200 V for 30-45 min at 4°C. Gels were imaged using a Fuji FL-7000 scanner (GE healthcare) and band intensities were quantified using ImageQuant TL. Homodimer cooperativity factors were estimated as described (19)

Chromatin immunoprecipitation assay

E14.5 inner ear tissue lysates were crosslinked, lysed and sonicated to yield DNA fragments of 200-500 bp. Immunoprecipitation on the chromatin was performed using anti-SOX9 (AB5535, Millipore), anti-SOX10 (gift of Dr. M. Wegner) and a serum control. The target fragments in *Aqp3* and *Crygb* were amplified by standard PCR or real time PCR.

Luciferase reporter assays

mIMCD3 cells (ATCC CRL-2123) were transfected using GeneJuice (Millipore) according to the manufacturer's instructions. Co-transfection experiments were performed with 80 ng each of the reporter, 20 ng each of the expression constructs and 10 ng of pRL-TK Renilla luciferase (Promega) as internal standard. The amount of total DNA per transfection was kept constant by adding empty vector pCALL2 together with Sox9, Sox9-Y440X and Sox10 expression construct. Luciferase activities (relative light unit RLU) were measured using the Dual Luciferase Reporter Assay System (Promega) after 48 hours of transfection. The relative fold of suppression or activation is expressed as the normalized value of relative light unit (RLU) of reporter and expression construct to reporter alone. All data and statistical tests are the results of at least 3 independent transfection experiments.

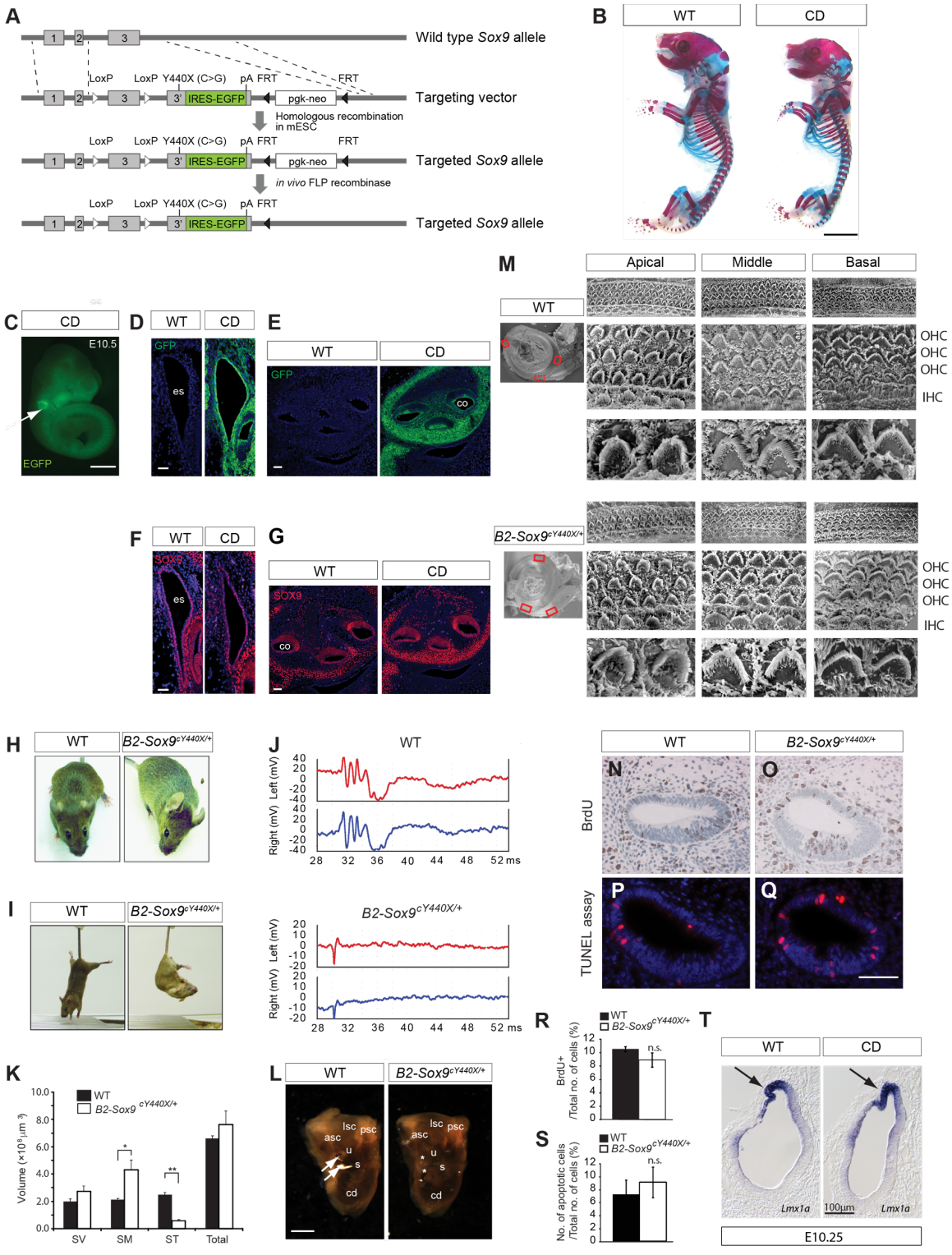

A
Wild type Sox9 allele
LoxP
LoxP Y440X (C>G)
pA FRT
FRT
IRES-EGFP
pgk-neo
Targeting vector
Homologous recombination in mESC
Targeted Sox9 allele
in vivo FLP recombinase
B
WT
CD
C
CD
E10.5
EGFP
D
WT
CD
GFP
es
E
WT
CD
GFP
co
M
Apical
Middle
Basal
WT
OHC
OHC
OHC
IHC
B2-Sox9cY440X/+
F
WT
CD
SOX9
es
G
WT
CD
SOX9
co
H
WT
B2-Sox9cY440X/+
I
WT
B2-Sox9cY440X/+
J
WT
Right (mV) Left (mV)
28 32 36 40 44 48 52 ms
B2-Sox9cY440X/+
K
WT
B2-Sox9 cY440X/+
Volume (×10⁸ μm ³)
SV
SM
ST
Total
L
WT
B2-Sox9cY440X/+
lsc
psc
asc
u
s
cd
N
O
P
Q
BrdU
TUNEL assay
R
WT
B2-Sox9cY440X/+
BrdU+ /Total no. of cells (%)
n.s.
S
WT
B2-Sox9cY440X/+
No. of apoptotic cells /Total no. of cells (%)
n.s.
T
WT
CD
Lmx1a
100μm
Lmx1a
E10.25

Fig. S1. Conditional *Sox9-floxed-Y440X* allele and mutant phenotype in the inner ear. (*A)* Schematic diagram of the targeting strategy. *Pgk-neo* was removed by crossing to FLP-recombinase mice (20). *(B*) Alizarin red and alcian blue stained skeletons of WT and CD *Sox9*$^{Y440X/+}$(CD) newborn mice. Note skeletal abnormalities characteristic of CD e.g. campomelia (bent bones) of the appendicular skeleton, shortened craniofacial and axial skeleton. (*C-E)* EGFP expression in the otic vesicle (arrow) of *Sox9*$^{Y440X/+}$(CD) embryo at E10.5 (n=3) (*C*), endolymphatic sac (n=3) (*D*) and cochlea (n=3) (*E*) at E14.5 (n=2). (*F* and G) SOX9 expression in WT and CD endolymphatic sac (F) and cochlea (*G*) at E14.5 (n=2). (*H*) *B2-Sox9*$^{cY440X/+}$ mutants showed a tilted head and body posture (*I*) failed tests for reaching response (n=5). (*J*) A characteristic waveform of auditory brainstem response recorded from both sides (left in red and right in blue) after a 0.1 ms rectangular pulse at 65 dB and 110 dB was applied to a WT and *B2-Sox9*$^{cY440X/+}$ adult mouse at 6-month respectively (n=5). (*K*) The volume of sound-conducting chamber equals to section area (μm²) x thickness (μm) x no. of interval slides in μm (n=3). The total volume of the cochlea was not significantly different but the scala media and scala tympani were larger (2-fold) (*p<0.05) and smaller (one-third) (**p<0.01) in the mutant, respectively. (*L*) Otoconia deficiency (*) in *B2-Sox9*$^{cY440X/+}$ mutant at P9 (n=3, both ears). U, utricle, S, saccule, asc, lsc, psc, anterior, lateral and posterior semi-circular canals. (*M*) Normal hair cell and stereocilium arrangement in apical, middle and basal regions of the *B2-Sox9*$^{cY440X/+}$ mutant cochlea at P4 (n=4) by scanning electron microscopy. *(N-Q*) Proliferation and apoptosis *B2-Sox9*$^{cY440X/+}$ cochlea at E14.5. (*N, O* and *R*) The number of BrdU positive cells (brown) counted against the total number of epithelial cells (blue) (n=5). (*P* and *Q*) TUNEL assay. (S) The number of apoptotic cells (red) counted against the total no. of epithelial cells (blue). (n=5). The total number of cells within the otic epithelium is not significantly different between the WT and mutant cochlea. n.s. not significant. (*T*) Prospective endolymphatic sac (arrow) marked by *Lmx1a* RNA probes at E10.25 (n=2, both ears). ch, chondrocyte; co, cochlea; es, endolymphatic sac. *p<0.05 and **p<0.01. Scale bar = 0.5 cm (B), 0.5 mm (*C)*, 5 mm (*L-M*), 100 μm (*T*).

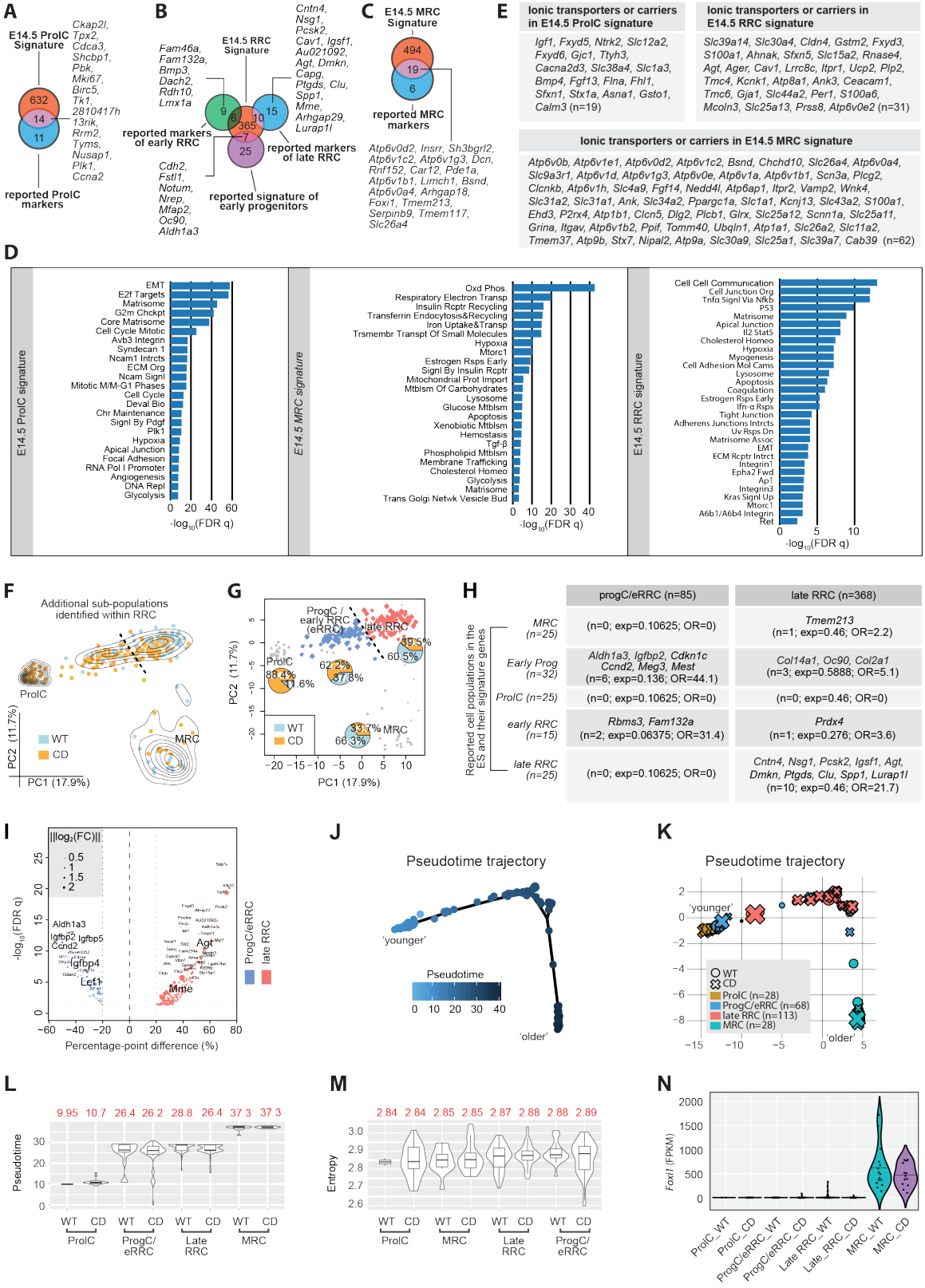

A
E14.5 ProlC Signature
632
14
11
Ckap2l, Tox2, Cdca3, Shcbp1, Pbk, Mki67, Birc5, Tk1, 2810417h 13rik, Rrm2, Tyms, Nusap1, Plk1, Ccna2
reported ProlC markers
B
E14.5 RRC Signature
Fam46a, Fam132a, Bmp3, Dach2, Rdh10, Lmx1a
Cntn4, Nsg1, Pcsk2, Cav1, Igsf1, Au021092, Agt, Dmkn, Capg, Ptgds, Clu, Spp1, Mme, Arhgap29, Lurap1l
9
6
10
15
365
7
25
reported markers of early RRC
reported markers of late RRC
Cdh2, Fstl1, Notum, Nrep, Mfap2, Oc90, Aldh1a3
reported signature of early progenitors
C
E14.5 MRC Signature
494
19
6
reported MRC markers
Atp6v0d2, Insrr, Sh3bgrl2, Atp6v1c2, Atp6v1g3, Dcn, Rnf152, Car12, Pde1a, Atp6v1b1, Limch1, Bsnd, Atp6v0a4, Arhgap18, Foxi1, Tmem213, Serpinb9, Tmem117, Slc26a4
E
Ionic transporters or carriers in E14.5 ProlC signature
Igf1, Fxyd5, Ntrk2, Slc12a2, Fxyd6, Gjc1, Ttyh3, Cacna2d3, Slc38a4, Slc1a3, Bmp4, Fgf13, Flna, Fhl1, Sfxn1, Stx1a, Asna1, Gsto1, Calm3 (n=19)
Ionic transporters or carriers in E14.5 RRC signature
Slc39a14, Slc30a4, Cldn4, Gstm2, Fxyd3, S100a1, Ahnak, Sfxn5, Slc15a2, Rnase4, Agt, Ager, Cav1, Lrrc8c, Itpr1, Ucp2, Plp2, Tmc4, Kcnk1, Atp8a1, Ank3, Ceacam1, Tmc6, Gja1, Slc44a2, Per1, S100a6, Mcoln3, Slc25a13, Prss8, Atp6v0e2 (n=31)
Ionic transporters or carriers in E14.5 MRC signature
Atp6v0b, Atp6v1e1, Atp6v0d2, Atp6v1c2, Bsnd, Chchd10, Slc26a4, Atp6v0a4, Slc9a3r1, Atp6v1d, Atp6v1g3, Atp6v0e, Atp6v1a, Atp6v1b1, Scn3a, Plcg2, Clcnkb, Atp6v1h, Slc4a9, Fgf14, Nedd4l, Atp6ap1, Itpr2, Vamp2, Wnk4, Slc31a2, Slc31a1, Ank, Slc34a2, Ppargc1a, Slc1a1, Kcnj13, Slc43a2, S100a1, Ehd3, P2rx4, Atp1b1, Clcn5, Dlg2, Plcb1, Glrx, Slc25a12, Scnn1a, Slc25a11, Grina, Itgav, Atp6v1b2, Ppif, Tomm40, Ubqln1, Atp1a1, Slc26a2, Slc11a2, Tmem37, Atp9b, Stx7, Nipal2, Atp9a, Slc30a9, Slc25a1, Slc39a7, Cab39 (n=62)
D
E14.5 ProlC signature
EMT
E2f Targets
Matrisome
G2m Chckpt
Core Matrisome
Cell Cycle Mitotic
Avb3 Integrin
Syndecan 1
Ncam1 Intrcts
ECM Org
Ncam Signl
Mitotic M/M-G1 Phases
Cell Cycle
Deval Bio
Chr Maintenance
Signl By Pdgf
Plk1
Hypoxia
Apical Junction
Focal Adhesion
RNA Pol I Promoter
Angiogenesis
DNA Repl
Glycolysis
0 20 40 60
-log10(FDR q)
E14.5 MRC signature
Oxd Phos.
Respiratory Electron Transp
Insulin Rcptr Recycling
Transferrin Endocytosis&Recycling
Iron Uptake&Transp
Trsmembr Transpt Of Small Molecules
Hypoxia
Mtorc1
Estrogen Rsps Early
Signl By Insulin Rcptr
Mitochondrial Prot Import
Mtblsm Of Carbohydrates
Lysosome
Glucose Mtblsm
Apoptosis
Xenobiotic Mtblsm
Hemostasis
Tgf-β
Phospholipid Mtblsm
Membrane Trafficking
Cholesterol Homeo
Glycolysis
Matrisome
Trans Golgi Netwk Vesicle Bud
0 10 20 30 40
-log10(FDR q)
E14.5 RRC signature
Cell Cell Communication
Cell Junction Org
Tnfa Signl Via Nfkb
P53
Matrisome
Apical Junction
Il2 Stat5
Cholesterol Homeo
Hypoxia
Myogenesis
Cell Adhesion Mol Cams
Lysosome
Apoptosis
Coagulation
Estrogen Rsps Early
Ifn-α Rsps
Tight Junction
Adherens Junctions Intrcts
Uv Rsps Dn
Matrisome Assoc
EMT
ECM Rcptr Intrct
Integrin1
Epha2 Fwd
Ap1
Integrin3
Kras Signl Up
Mtorc1
A6b1/A6b4 Integrin
Ret
0 5 10
-log10(FDR q)
F
Additional sub-populations identified within RRC
ProlC
MRC
WT
CD
PC2 (11.7%)
PC1 (17.9%)
G
ProgC / early RRC (eRRC)
late RRC
ProlC
MRC
39.5%
60.5%
62.2%
37.8%
88.4%
11.6%
33.7%
66.3%
WT
CD
PC2 (11.7%)
PC1 (17.9%)
H
Reported cell populations in the ES and their signature genes
progC/eRRC (n=85)
late RRC (n=368)
MRC (n=25)
(n=0; exp=0.10625; OR=0)
Tmem213 (n=1; exp=0.46; OR=2.2)
Early Prog (n=32)
Aldh1a3, Igfbp2, Cdkn1c Ccnd2, Meg3, Mest (n=6; exp=0.136; OR=44.1)
Col14a1, Oc90, Col2a1 (n=3; exp=0.5888; OR=5.1)
ProlC (n=25)
(n=0; exp=0.10625; OR=0)
(n=0; exp=0.46; OR=0)
early RRC (n=15)
Rbms3, Fam132a (n=2; exp=0.06375; OR=31.4)
Prdx4 (n=1; exp=0.276; OR=3.6)
late RRC (n=25)
(n=0; exp=0.10625; OR=0)
Cntn4, Nsg1, Pcsk2, Igsf1, Agt, Dmkn, Ptgds, Clu, Spp1, Lurap1l (n=10; exp=0.46; OR=21.7)
I
||log2(FC)||
0.5
1
1.5
2
-log10(FDR q)
Aldh1a3
Igfbp2
Igfbp5
Ccnd2
Igfbp4
Lef1
Agt
Mme
ProgC/eRRC
late RRC
Percentage-point difference (%)
J
Pseudotime trajectory
'younger'
Pseudotime
0 10 20 30 40
'older'
K
Pseudotime trajectory
'younger'
'older'
WT
CD
ProlC (n=28)
ProgC/eRRC (n=68)
late RRC (n=113)
MRC (n=28)
L
9.95 10.7 26.4 26.2 28.8 26.4 37.3 37.3
Pseudotime
WT CD
ProlC
ProgC/ eRRC
Late RRC
MRC
M
2.84 2.84 2.85 2.85 2.87 2.88 2.88 2.89
Entropy
WT CD
ProlC
MRC
Late RRC
ProgC/ eRRC
N
Foxi1 (FPKM)
ProlC_WT
ProlC_CD
ProgC/eRRC_WT
ProgC/eRRC_CD
Late RRC_WT
Late_RRC_CD
MRC_WT
MRC_CD

Fig. S2. scRNA-seq (E14.5 WT and *Sox9*$^{Y440X/+}$ (CD) endolymphatic sac) bioinformatics analysis. (*A*-*C*) Venn diagram comparing signatures of cell populations and reported markers. (*D*) Gene-ontology (GO) term analysis of the genes up-regulated in each of the three major populations against the remaining cells (signatures of each cell population). (*E*) Lists of ion channels and transporters found in each the signatures of each cell population. (*F*) Density contour plot of the 246 single cells identifying 2 additional sub-population in the RRC. (*G*) The proportion of CD cells in each of the four populations. The percentages were normalized by taking account into consideration the total WT and total CD cells. (*H*) Table comparing DEGs of ProgC/eRRCs and late RRCs, with reported population signatures. n, number observed; exp, expected; OR, odds ratio. (*I*) Volcano plot showing DEGs between ProgC/eRRC and late RRC. Percentage-point difference is the difference between the percentage of cells expressed in one group minus that in the other. (*J* and *K*) Pseudotime trajectory showing ProlC and ProgC/eRRCs are on one end, followed by late RRCs and MRCs. (*L* and *M*) Violin plots showing the predicted pseudotime quantities and entropy of the four cell populations in the two genotypes. (*N*) Violin plots showing expression of *Foxi1* in the scRNA-seq data.

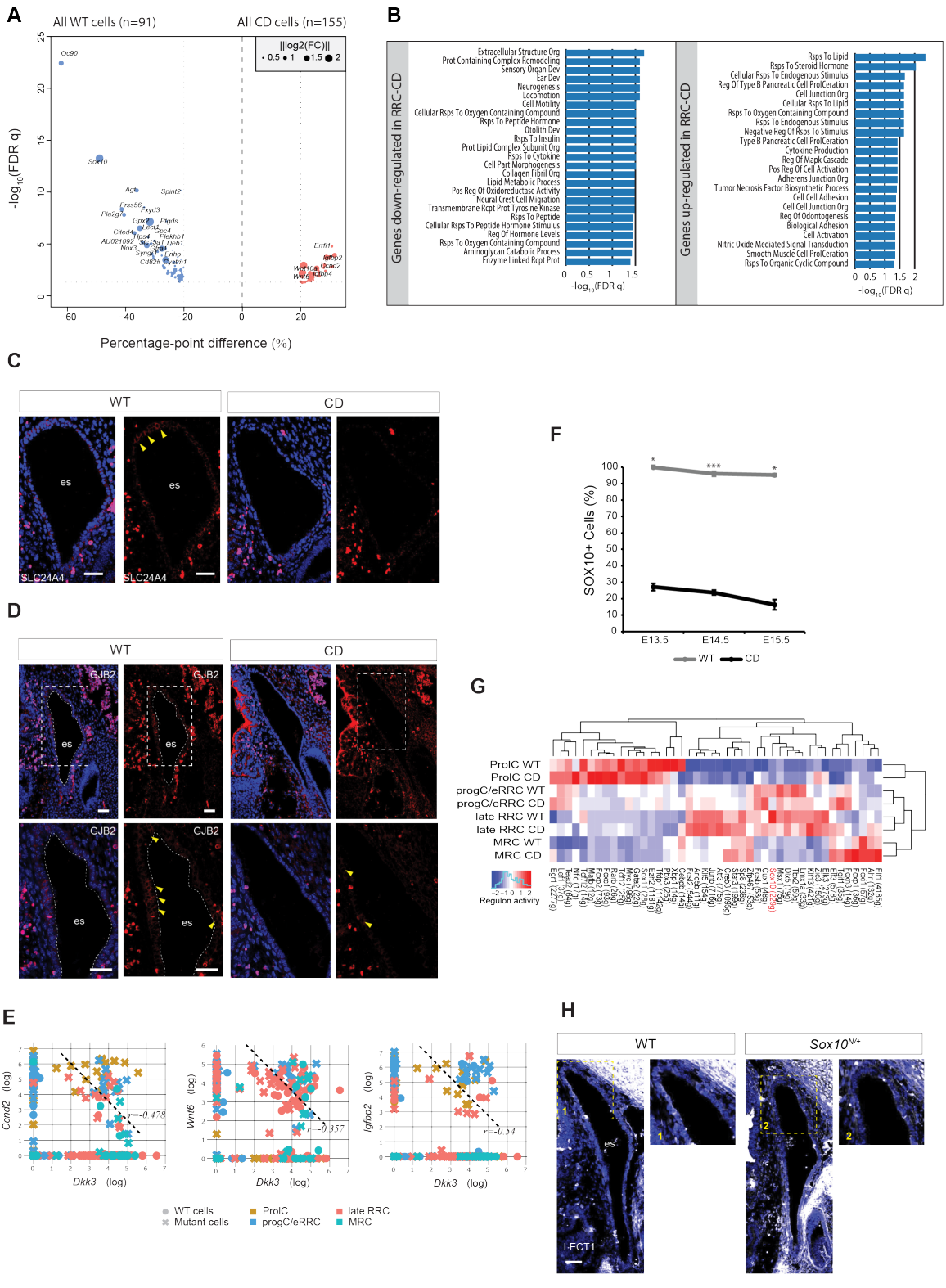
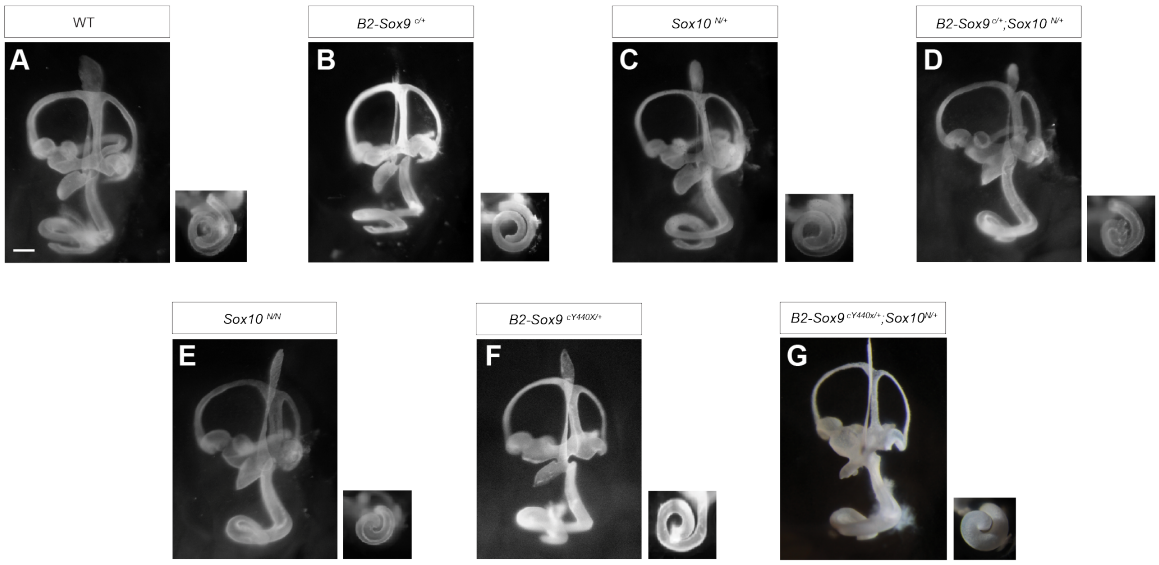
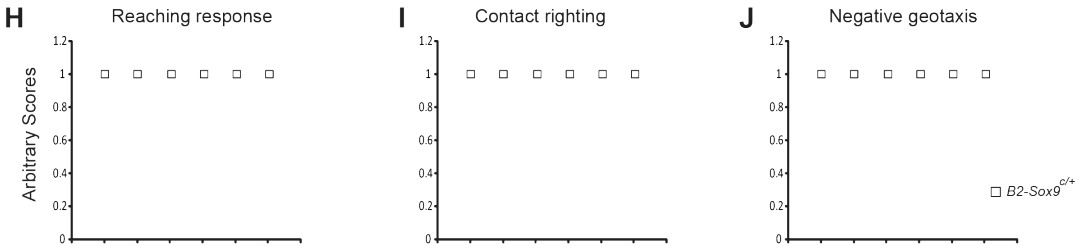
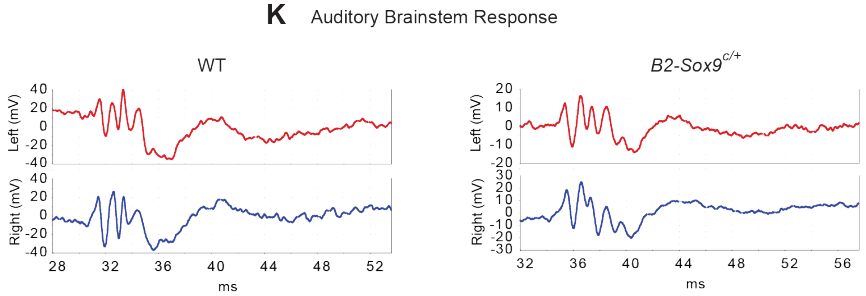

A
All WT cells (n=91)
All CD cells (n=155)
||log2(FC)||
0.5 1 1.5 2
-log10(FDR q)
Percentage-point difference (%)
B
Genes down-regulated in RRC-CD
Extracellular Structure Org
Prot Containing Complex Remodeling
Sensory Organ Dev
Ear Dev
Neurogenesis
Locomotion
Cell Motility
Cellular Rsps To Oxygen Containing Compound
Rsps To Peptide Hormone
Otolith Dev
Rsps To Insulin
Prot Lipid Complex Subunit Org
Rsps To Cytokine
Cell Part Morphogenesis
Collagen Fibril Org
Lipid Metabolic Process
Pos Reg Of Oxidoreductase Activity
Neural Crest Cell Migration
Transmembrane Rcpt Prot Tyrosine Kinase
Rsps To Peptide
Cellular Rsps To Peptide Hormone Stimulus
Reg Of Hormone Levels
Rsps To Oxygen Containing Compound
Aminoglycan Catabolic Process
Enzyme Linked Rcpt Prot
-log10(FDR q)
Genes up-regulated in RRC-CD
Rsps To Lipid
Rsps To Steroid Hormone
Cellular Rsps To Endogenous Stimulus
Reg Of Type B Pancreatic Cell ProlCeration
Cell Junction Org
Cellular Rsps To Lipid
Rsps To Oxygen Containing Compound
Rsps To Endogenous Stimulus
Negative Reg Of Rsps To Stimulus
Type B Pancreatic Cell ProlCeration
Cytokine Production
Reg Of Mapk Cascade
Pos Reg Of Cell Activation
Adherens Junction Org
Tumor Necrosis Factor Biosynthetic Process
Cell Cell Adhesion
Cell Cell Junction Org
Reg Of Odontogenesis
Biological Adhesion
Cell Activation
Nitric Oxide Mediated Signal Transduction
Smooth Muscle Cell ProlCeration
Rsps To Organic Cyclic Compound
-log10(FDR q)
C
WT
CD
es
SLC24A4
D
WT
CD
GJB2
es
E
Cond2 (log)
Wnt6 (log)
Igfbp2 (log)
Dkk3 (log)
r=-0.478
r=-0.357
r=-0.54
WT cells
Mutant cells
ProlC
progC/eRRC
late RRC
MRC
F
SOX10+ Cells (%)
E13.5
E14.5
E15.5
WT
CD
G
ProlC WT
ProlC CD
progC/eRRC WT
progC/eRRC CD
late RRC WT
late RRC CD
MRC WT
MRC CD
Regulon activity
H
WT
Sox10N/+
es
LECT1

Fig. S3. Analysis on differential expressions revealed from scRNA-seq (E14.5 WT and *Sox9*$^{Y440X/+}$ (CD) endolymphatic sac (ES)) and downstream investigation. (*A*) Volcano plot showing DEGs between all WT and CD cells. Expressed ratio margin is the difference between the percentage of cells expressed in one group minus that in the other. (*B*) GO enrichment analysis of the DEGs between WT and CD RRC cells. (*C*) Immunostaining of SLC24A4 in the ES at E14.5 of WT and CD (n=2). (*D*) Immunostaining of GJB2 in the ES at E14.5 of WT and CD (n=2). (*E*) Scatter plots showing co-expression of *Ccnd2* (left), *Wnt6* (middle) and *Igfbp2* (right), with respect to *Dkk3*. r, Pearson correlation coefficient. (*F*) Quantification of percentage of SOX10-expressing cells in ES from E13.5 to E15.5 in WT and CD (n=2-3). (*G*) Regulon expression analysis reveals population-specific regulatory activities. (*H*) Immunostaining of LECT1 in the ES at E14.5 of WT and *Sox10*$^{N/+}$ (n=2). es, endolymphatic sac. *p<0.05; ***p<0.001. Scale bar = 40 μm.


WT
A
B2-Sox9 c/+
B
Sox10 N/+
C
B2-Sox9 c/+;Sox10 N/+
D
Sox10 N/N
E
B2-Sox9 cY440X/+
F
B2-Sox9 cY440X/+;Sox10N/+
G


H
Reaching response
I
Contact righting
J
Negative geotaxis
Arbitrary Scores
1.2
1
0.8
0.6
0.4
0.2
0
B2-Sox9c/+


K
Auditory Brainstem Response
WT
B2-Sox9c/+
Left (mV)
Right (mV)
ms

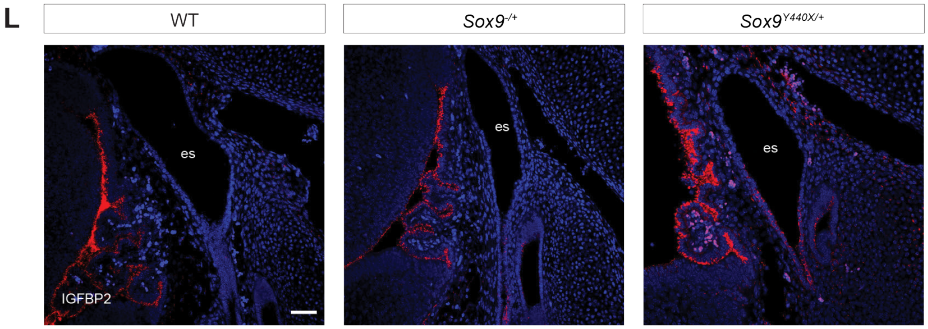

L
WT
Sox9-/+
Sox9Y440X/+
es
IGFBP2

Fig. S4. Phenotype analysis of SoxE mutants. (*A*-*G*) Paint-fill analyses of inner ears of WT and mutants of *Sox9* and *Sox10* at E15.5. Paint-filling of the inner ears of these mutants revealed the cochlea of $Sox10^{N/N}$ is shortened but gross truncations were not observed for the other genotypes (*H*-*K*) *B2-*$Sox9^{c/+}$ mutants passed all the tests for vestibular and hearing function (n=6). (*H*) Reaching response, (*I*) Contact righting, (*J*) Negative geotaxis, (*K*) Auditory brainstem response measured on both sides (left in red and right in blue) after a 0.1 ms rectangular pulse at 60 dB was given to both WT and *B2-*$Sox9^{c/+}$ adult mice (n=3). (*L*) IGFBP2 staining on E14.5 ES of WT, $Sox9^{-/+}$ and $Sox9^{Y440X/+}$. Scale bar = 200 μm (*A*-*G*), 40 μm (*L*).

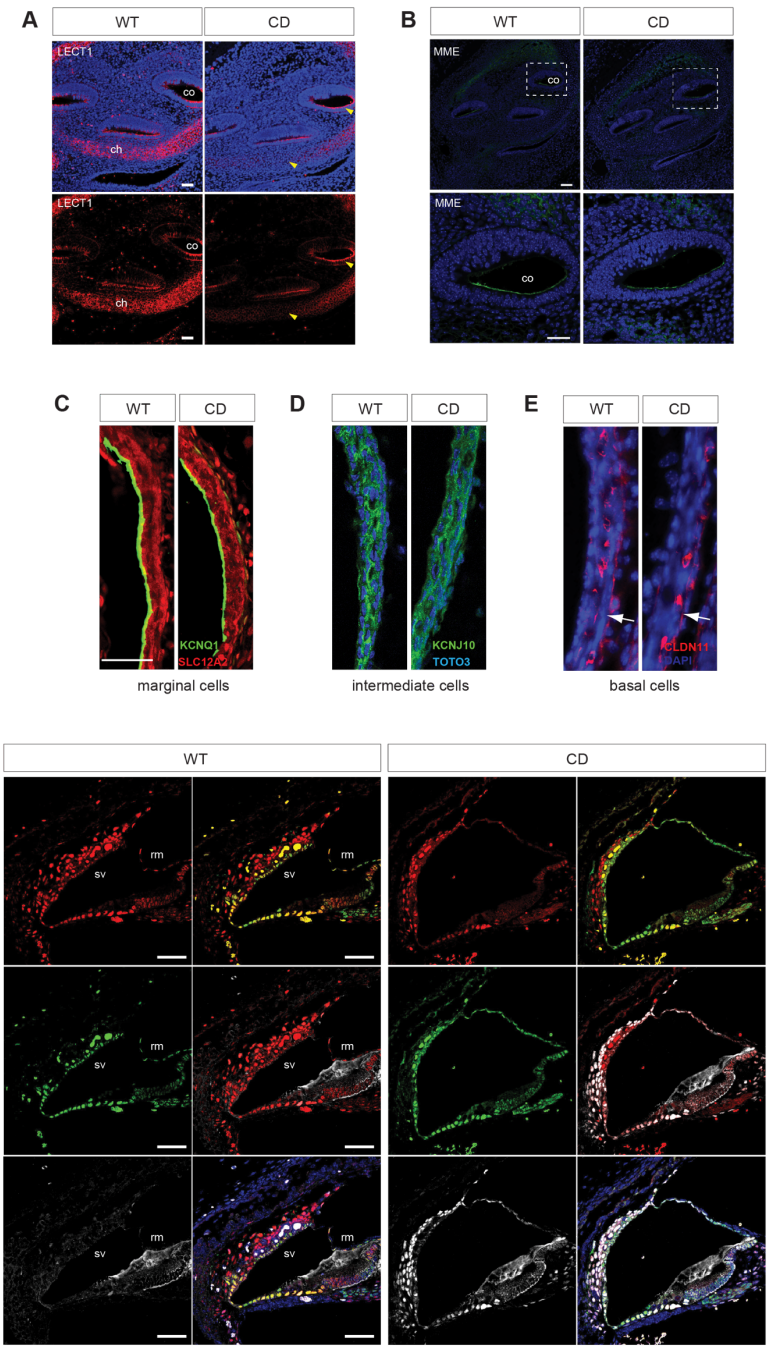

A
WT
CD
LECT1
co
ch
B
MME
C
KCNQ1
SLC12A2
marginal cells
D
KCNJ10
TOTO3
intermediate cells
E
CLDN11
DAPI
basal cells
F
sv
rm
SOX9
SOX10
SOX9Y440X
DAPI

Fig. S5. Targeted analysis on embryonic cochlea (CO) and characterization of post-natal stria vascularis (SV) in WT and *Sox9*$^{Y440X/+}$ (CD). (*A*) Immunostaining of LECT1 in the CO at E14.5 of WT and CD (n=2). (*B*) Immunostaining of MME in the CO at E14.5 of WT and CD (n=2). (*C*-*E*) No morphological changes of stria vascularis in CD mutant at P3 (n=5-6). (*C*) Apical membrane of the marginal cells stained by KCNQ1 antibody and basolateral membrane of marginal cells stained by SLC12A2 antibody (*D*) Intermediate cells stained by KCNJ10 antibody (*E*) Basal cells (arrow) stained by CLDN11 antibody. (*F*) Immunostaining of SOX9, SOX10 and SOX9^{Y440X} in the P5 CO (n=2). ch, chondrocyte; co, cochlea; rm, Reissner's membrane; sv, stria vascularis. Scale bar = 40 μm (*A*, *B* and *F*), 50 μm (*C*-*E*).

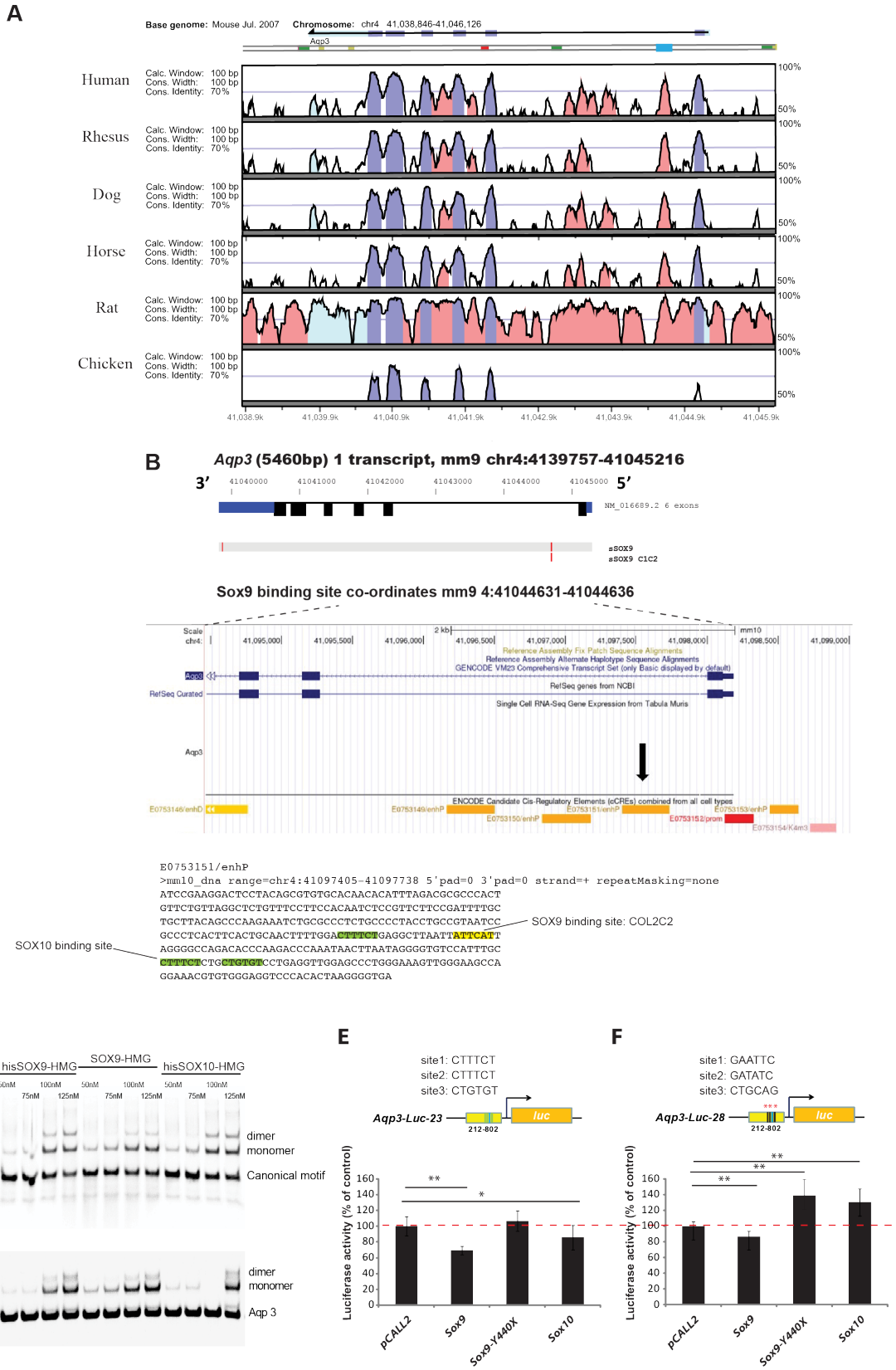

A
Base genome: Mouse Jul. 2007 Chromosome: chr4 41,038,846-41,046,126
Human
Rhesus
Dog
Horse
Rat
Chicken
Calc. Window: 100 bp
Cons. Width: 100 bp
Cons. Identity: 70%
100%
50%
41,038.9k
41,039.9k
41,040.9k
41,041.9k
41,042.9k
41,043.9k
41,044.9k
41,045.9k
B
Aqp3 (5460bp) 1 transcript, mm9 chr4:4139757-41045216
3'
5'
NM_016689.2 6 exons
sSOX9
sSOX9 C1C2
Sox9 binding site co-ordinates mm9 4:41044631-41044636
RefSeq Curated
Aqp3
ENCODE Candidate Cis-Regulatory Elements (cCREs) combined from all cell types
E0753151/enhP
>mm10_dna range=chr4:41097405-41097738 5'pad=0 3'pad=0 strand=+ repeatMasking=none
ATCCGAAGGACTCCTACAGCGTGTGCACAACACATTTAGACGCGCCCACT
GTTCTGTTAGGCTCTGTTTCCTTCCACAATCTCCGTTCTTCCGATTTTGC
TGCTTACAGCCCAAGAAATCTGCGCCCTCTGCCCCTACCTGCCGTAATCC
GCCCTCACTTCACTGCAACTTTTGGACTTTCTGAGGCTTAATTATTCATT
AGGGGCCAGACACCCAAGACCCAAATAACTTAATAGGGGTGTCCATTTGC
CTTTCTCTGCTGTGTCCTGAGGTTGGAGCCCTGGGAAAGTTGGGAAGCCA
GGAAACGTGTGGGAGGTCCCACACTAAGGGGTGA
SOX10 binding site
SOX9 binding site: COL2C2
C
hisSOX9-HMG
SOX9-HMG
hisSOX10-HMG
50nM
100nM
75nM
125nM
K
dimer
monomer
Canonical motif
D
dimer
monomer
Aqp 3
E
site1: CTTTCT
site2: CTTTCT
site3: CTGTGT
Aqp3-Luc-23
luc
212-802
Luciferase activity (% of control)
pCALL2
Sox9
Sox9-Y440X
Sox10
F
site1: GAATTC
site2: GATATC
site3: CTGCAG
Aqp3-Luc-28
luc
212-802

Fig. S6. Genomic analyses of the *Aqp3* locus and supplementary experimental information for demonstration of the interaction between SOX9 and SOX10 and their suppression on *Aqp3*. (*A*) Conserved SOX9 and SOX10 binding sites predicted by bsConserve and Jasper over a 205 bp conserved non-coding region (Cyan box) by Vista in intron 1 of *Aqp3*. (*B*) Human-mouse perfectly conserved matches to *Sox9* are indicated as vertical red lines. Scanning this region in HOCOMOCO v11 identified the region (chr4:41,044,565-41,045,962, accessible in the cochlea ATAC-seq) which contains an ENCODE candidate cis regulatory element (cCRE, (ENCODE candidate cis regulatory element) E0753151/enhP) with proximal enhancer features such as DNase hypersensitive sites, histone modification and CTCF binding sites. A SOX consensus (C[AT]TTG[AT][AT])86-89 was found in the 3' untranslated region while a COL2C2 (ATTCAT) *in vivo* SOX9 binding site (21) was found in intron 1. The region also contained the SOX10 binding consensus (mm9 chr4:41,045,517-41,045,527). (*C, D*) The SOX9-HMG, His-SOX9-HMG, His-SOX10-HMG proteins without dimerization (DIM) domains dimerized ineffectively on the (*C*) canonical and (*D*) *Aqp3* DNA (n=4). (*E* and *F*) Inhibition of *Aqp3* expression through SOX10 binding sites (n=3). (*E*) The activity of the WT regulatory element of *Aqp3* linked to a luciferase reporter (Aqp3-luc-23) was suppressed by SOX9 and SOX10. (*F*) This inhibition was alleviated when SOX10 binding sites were mutated (Aqp3-luc-28). Site 1 from CTTTCT to GAATTC; Site 2 from CTTTCT to GATATC; Site 3 from CTGTGT to CTGCAG. *p<0.05; ***p<0.001.

Table S1. Frequency of ES/ED cells expressing marker genes at E13.5, E14.5 and E15.5 in WT and *Sox9*$^{Y440X/+}$ (CD).

Gene	Genotype	Stage	% Positive Cell	Fetuses analysed (n)	Number of Cell Quantified	t-test p-value (WT vs CD)
IGFBP2	WT	E13.5	27.81	2	331	0.055451086
IGFBP2	CD	E13.5	40.29	2	342	0.055451086
IGFBP2	WT	E14.5	5.98	3	841	0.000221164
IGFBP2	CD	E14.5	26.98	3	773	0.000221164
IGFBP2	WT	E15.5	0	2	242	0.001382771
IGFBP2	CD	E15.5	24.42	2	487	0.001382771
SOX11	WT	E14.5	1.02	2	372	0.018721789
SOX11	CD	E14.5	28.52	2	328	0.018721789
AGT	WT	E14.5	98.52	2	353	0.000520816
AGT	CD	E14.5	13.26	2	330	0.000520816
MME	WT	E14.5	91.04	2	212	0.005800691
MME	CD	E14.5	24.88	2	258	0.005800691
FOXII1	WT	E14.5	15.72	3	462	0.036439376
FOXII1	CD	E14.5	7.75	3	415	0.036439376
Atpdv0d2	WT	E14.5	12.44	3	433	0.003702273
Atpdv0d2	CD	E14.5	4.33	3	184	0.003702273
SOX10	WT	E13.5	100	2	422	0.000441837
SOX10	CD	E13.5	27.17	2	225	0.000441837
SOX10	WT	E14.5	95.97	3	984	4.68218E-07
SOX10	CD	E14.5	23.73	3	618	4.68218E-07
SOX10	WT	E15.5	95.24	2	424	0.000861387
SOX10	CD	E15.5	16.33	2	455	0.000861387
SLC24A4	WT	E14.5	36.27	2	207	0.011515065
SLC24A4	CD	E14.5	8.43	2	222	0.011515065
GJB2	WT	E14.5	14.35	2	202	0.000225722
GJB2	CD	E14.5	3.78	2	209	0.000225722

The proportions (%) of cells labelled with marker genes in ES/ED from CD mutant and WT control litter mates were determined. n, no. of fetuses analyzed.

Table S2. Frequency of ES/ED cells expressing marker genes at E14.5 in WT and *Sox10*$^{N/+}$ (N+).

Gene	Genotype	Stage	% Positive Cell	Fetuses analysed (n)	Number of Cell Quantified	t-test p-value (WT vs N+)
IGFBP2	WT	E14.5	4.2	3	419	0.028656019
IGFBP2	N+	E14.5	13.54	3	406	0.028656019
AGT	WT	E14.5	95.55	2	305	0.024962483
AGT	N+	E14.5	33.77	2	252	0.024962483
FOXII1	WT	E14.5	23.36	2	217	0.150457916
FOXII1	N+	E14.5	17.39	2	229	0.150457916

The proportions (%) of cells labelled with marker genes in ES/ED of E14.5 N+ mutants and WT control litter mates were determined. n, no. of fetuses analyzed.

Movie S1. Circling behavior of (B) surviving mouse line *Sox9*$^{Y440X/+}$ (n=6) versus (A) WT.

Movie S2. Swimming tests of (A) WT and (B) surviving mouse line *Sox9*$^{Y440X/+}$ (n=8).

Movie S3. Heading tossing and circling behavior of *B2-Sox9*$^{cY440X/+}$ compared to WT.

Movie S4. *B2-Sox9*$^{cY440X/+}$ failed the test of contact righting versus WT mouse.

Movie S5. *B2-Sox9*$^{cY440X/+}$ failed the test of negative geotaxis versus WT mouse.

Dataset S1. Lists of differentially expressed genes (DEGs) from comparisons of different cell populations.

Dataset S2. List of probes for Hybridization Chain Reactions (HCR), primers for qPCR and antibodies

SI References

1. K. Honda, H. J. Lee, A. J. Griffith, I. Roux, Dissection of the Endolymphatic Sac from Mice. *J Vis Exp* 10.3791/62375 (2021).
2. K. Honda *et al.*, Molecular architecture underlying fluid absorption by the developing inner ear. *Elife* **6** (2017).
3. S. Picelli *et al.*, Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-181 (2014).
4. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360 (2015).
5. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
6. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).
7. C. Trapnell *et al.*, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578 (2012).
8. J. Harrow *et al.*, GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**, 1760-1774 (2012).
9. J. Fan *et al.*, Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat Methods* **13**, 241-244 (2016).
10. D. Kobak, P. Berens, The art of using t-SNE for single-cell transcriptomics. *Nat Commun* **10**, 5416 (2019).
11. G. Finak *et al.*, MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* **16**, 278 (2015).
12. A. Subramanian *et al.*, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
13. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386 (2014).
14. S. Aibar *et al.*, SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083-1086 (2017).

15. A. Chotteau-Lelievre, P. Dolle, F. Gofflot, Expression analysis of murine genes using in situ hybridization with radioactive and nonradioactively labeled RNA probes. *Methods Mol Biol* **326**, 61-87 (2006).
16. C. K. Ng *et al.*, Deciphering the Sox-Oct partner code by quantitative cooperativity measurements. *Nucleic acids research* **40**, 4933-4941 (2012).
17. Y. H. Huang, A. Jankowski, K. S. Cheah, S. Prabhakar, R. Jauch, SOXE transcription factors form selective dimers on non-compact DNA motifs through multifaceted interactions between dimerization and high-mobility group domains. *Sci Rep* **5**, 10398 (2015).
18. E. Sock *et al.*, Loss of DNA-dependent dimerization of the transcription factor SOX9 as a cause for campomelic dysplasia. *Hum Mol Genet* **12**, 1439-1447 (2003).
19. D. S. Tan *et al.*, Directed Evolution of an Enhanced POU Reprogramming Factor for Cell Fate Engineering. *Mol Biol Evol* **38**, 2854-2868 (2021).
20. S. M. Dymecki, Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc Natl Acad Sci U S A* **93**, 6191-6196 (1996).
21. D. M. Bell *et al.*, SOX9 directly regulates the type-II collagen gene. *Nature genetics* **16**, 174-178 (1997).