Title

Trans-differentiation of outer hair cells into inner hair cells in the absence of INSM1

Journal

Nature, 563(7733)

ISSN

0028-0836

Authors

Wiwatpanit, Teerawat

Lorenzen, Sarah M

Cantú, Jorge A

et al.

Publication Date

2018-11-01

DOI

10.1038/s41586-018-0570-8

HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2019 April 10.

Published in final edited form as:

Nature. 2018 November ; 563(7733): 691–695. doi:10.1038/s41586-018-0570-8.

Trans-differentiation of outer hair cells into inner hair cells in the absence of INSM1

Teerawat Wiwatpanit[#1], **Sarah M. Lorenzen**[#1], **Jorge A. Cantú**[#1], **Chuan Zhi Foo**[1], **Ann K. Hogan**[1], **Freddie Márquez**[1], **John C. Clancy**[1], **Matthew J. Schipma**[2], **Mary Ann Cheatham**[3,2,4], **Anne Duggan**[1,#], and **Jaime García-Añoveros**[1,3,4,5,#]

[1]Department of Anesthesiology, Northwestern University, Feinberg School of Medicine, Chicago, Illinois 60611, USA

[2]Next Generation Sequencing Core, Northwestern University Feinberg School of Medicine, Chicago, IL, 60611, USA

[3]Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208, USA

[4]Hugh Knowles Center for Clinical and Basic Science in Hearing and its Disorders, Northwestern University, Chicago, IL 60611, USA

[5]Departments of Neurology and Physiology, Northwestern University, Feinberg School of Medicine, Chicago, Illinois 60611, USA

[#] These authors contributed equally to this work.

Abstract

The mammalian cochlea contains two types of mechanosensory hair cells (HCs) that play different and critical roles in hearing. Inner hair cells (IHCs), with an elaborate presynaptic apparatus, signal to cochlear neurons and communicate sound information to the brain. Outer hair cells (OHCs) mechanically amplify sound-induced vibrations, enabling enhanced sensitivity to sound and sharp tuning. Cochlear HCs are solely generated during development and their death, most often of OHCs, is the main cause of deafness. OHCs and IHCs, together with supporting cells, originate embryonically from the prosensory region of the otocyst, but how HCs differentiate into two different types is unknown[1–3]. Here we show that *Insm1*, which encodes a zinc finger protein transiently expressed in nascent OHCs, consolidates their fate by preventing trans-differentiation into IHCs. In the absence of INSM1 many HCs born embryonically as OHCs switch fates to

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors have no competing interests. Correspondence and requests should be addressed to J.G.A. (anoveros@northwestern.edu).
[#]Co-corresponding authors

Author Contributions J.G-A. and A.D. conceived the project. T.W., S.M.L., J.A.C., C.Z.F., A.K.H., F.M., J.C.C., M.A.C., A.D. and J.G-A. performed experiments. T.W., S.M.L., J.A.C., C.Z.F., M.J.S., M.A.C., A.D. and J.G-A. analyzed data. J.G-A. and T.W. wrote the manuscript.

Data Availability All data are available from the corresponding authors upon reasonable request. RNAseq data available for public view at the gEAR portal (https://umgear.org/).

Supplementary Information is available in the online version of the paper at www.nature.com/nature.

become mature IHCs. In order to identify the genetic mechanisms by which *Insm1* operates, we compared transcriptomes of immature IHCs vs OHCs, as well as OHCs with and without INSM1. OHCs lacking INSM1 upregulate a set of genes, most of which are normally preferentially expressed by IHCs. The homeotic cell transformation of OHCs without INSM1 into IHCs reveals for the first time a mechanism by which these neighboring mechanosensory cells begin to differ: INSM1 represses a core set of early IHC-enriched genes in embryonic OHCs and makes them unresponsive to an IHC-inducing gradient, so that they proceed to mature as OHCs. Without INSM1, some of the OHCs upregulating these few IHC-enriched transcripts trans-differentiate into IHCs, revealing the first candidate genes for IHC-specific differentiation.

OHCs express *Insm1* transiently from the onset of differentiation (E15.5) to ~P2[4]. Neuronal progenitors and nascent spiral ganglion (SG) neurons also express *Insm1*[4]. Because *Insm1* complete knockouts die embryonically by E19.5[5,6] we generated an allele (*Insm1*^{F}) in which the entire coding sequence may be deleted (Fig 1a, Extended Fig 1). We conditionally ablated *Insm1*^{F} with *Atoh1*^{Cre}, expressed starting at E13.5, three days before *Insm1*, and recombining in most cochlear hair cells (HCs) and some supporting cells (SCs), but not in SG neurons[7]. We also ablated *Insm1*^{F} with Tg*Pax2*^{Cre}, expressed earlier in the otocyst and recombining in most inner ear cells[8]. In these mice, *Insm1* was ablated prior to its expression in OHCs (Extended Fig 2). Both *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ and *TgPax2*$^{Cre/+}$*; Insm1*$^{F/F}$ (cKO) mice displayed ABR threshold shifts that can be accounted for by DPOAE shifts, a characteristic of OHC dysfunction (Fig 1b-d; extended Fig 3a,b). In their organs of Corti, many cells in the position of OHCs (the outer compartment) had IHC rather than OHC features. They had large stereocilia like IHCs and not the shorter, W-arranged stereocillia of OHCs (Fig 1e,f); expressed the IHC-enriched calcium buffer calmodulin and lacked OHC-specific oncomodulin (Fig 1e,f,h;extended Fig 3f); expressed the vesicular glutamate transporter 3 (VGLUT3), required for IHC presynaptic function, and lacked prestin, required for OHC electromotility (Fig 1g,m); had the flask shape of IHCs rather than the cylindrical of OHCs; had large nuclei like IHCs instead of the smaller of OHCs (Fig 1i-j;extended Fig 3j). These nuclei harbored transcription factor CtBP2 normally expressed in IHCs (Fig 1k), and the cells contained a number of presynaptic ribbon synapses (10.6 ± 2.1 SD, n=3 mice, 39 cells) closer to that of control IHCs (16.3 ± 0.7 SD on average, n=3 littermate controls, 30 cells), instead of the few of OHCs (1.8 ± 0.2 SD, n=3 mice, 90 cells) (Fig 1k). With rare exceptions (Fig 1g), these abnormal cells displayed all IHC features examined and lacked those of OHCs, so we termed them oc-IHCs (outer compartment IHCs).

The number of oc-IHCs in *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ (42.6% ± 10.9 SD, n=12 mice) and *TgPax2*$^{Cre/+}$*; Insm1*$^{F/F}$ (46.0% ± 5.64 SD, n=3 mice) was about half, the rest appearing as OHCs. This is not due to incomplete or delayed ablation of *Insm1* because we did not detect *Insm1* mRNA in any OHCs of *TgPax2*$^{Cre/+}$*; Insm1*$^{F/F}$ mice during or after the onset of expression (E16.5; Extended Fig 2a,b, bottom). Notably, the oc-IHCs were more frequent in the first than second and third HC rows of the outer compartment (Fig 1l,n). In principle these oc-IHCs in mature organs of Corti lacking INSM1 could be displaced IHCs, newly generated IHCs replacing lost OHCs, IHCs born in the outer compartment, or OHCs trans-differentiated into IHCs. They were not displaced IHCs, as the IHC row in both cKO mice

had a normal arrangement and density of IHCs (Extended Figs 4a,3c). Although during normal development cochlear HCs are all born during embryogenesis (E12-E16[9,10]), early HC death can trigger the generation of HCs from proliferating and trans-differentiating SCs in the early postnatal days[11–15]. This does not occur in the absence of INSM1, as HC density in the outer compartment (OHCs + oc-IHCs) is unaltered up to P34 (Extended Figs 4b,c, 3d,e), whereas oc-IHCs are present well before. Second, postnatally produced HCs derived from SCs initially express SOX2[12–14], whereas none of the oc-HCs of cKO pups expressed SOX2 (Extended Fig 4d). Third, some postnatally produced HCs result from proliferation of SCs[11–14], yet none of the oc-HCs in cKO mice derived from postnatal proliferations (Extended Fig 4e). These results reveal that oc-IHCs are not the result of OHC death followed by replacement from displaced IHCs or postnatally generated HCs. Instead, the oc-IHCs represent a homeotic transformation (of mechanosensory OHCs into IHCs) due to a developmental defect in the generation or differentiation of OHCs. Either IHCs are generated in place of OHCs, or OHCs trans-differentiate into IHCs.

We examined organs of Corti from conditional (*TgPax2*$^{Cre/+}$*; Insm1*$^{F/F}$ and *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$) and complete (*Insm1*$^{GFP.Cre/-}$) KOs of *Insm1* during late embryogenesis, when OHCs and IHCs begin to differentiate. At E16.5, all cells in the outer compartment begin to express the earliest markers of OHCs: the *Insm1* promoter in *Insm1*$^{GFP.Cre/-}$ embryos (which lack INSM1 protein but express GFP from the *Insm1* promoter[4]; Fig 2a), and BCL11B in nuclei (Fig 2b,c). While in control mice BCL11B expression is maintained past birth, in embryos lacking INSM1 it subsides in nearly half of the oc-HCs (from E18.5 to P2; Fig 2b-c). During the same period many oc-HCs express the early IHC marker *fgf8* (Fig 2d). Around birth two additional markers begin to be expressed in control mice, neuroplastin preferentially in OHC stereocillia[16] and VGLUT3 in IHCs[17]. By comparison, in both cKOs many oc-HCs expressed VGLUT3 and not neuroplastin (Fig 2f,g). Finally, although the orientation of IHCs vs OHCs is maintained by birth in cKOs, the disorganization of the OHC rows at the level of the nuclei already revealed alterations in cell shape (Fig 2e). We conclude that in the absence of INSM1, oc-HCs are generated with early OHC features, but soon after some of these cells lose them, express early IHC markers, and proceed to differentiate into mature IHCs. This trans-differentiation of early OHCs into IHCs reveals that INSM1 is not required to initiate commitment to the OHC fate, but acts subsequently by preventing it from switching to that of IHCs. *Insm1* acts by consolidating the OHC fate, making it permanent.

A brief expression of *Insm1* is sufficient to evade phenotypic conversion (Extended Fig 5). It appears that *Insm1* locks the OHC fate during a narrow developmental period. Curiously, despite complete absence of *Insm1* expression in KO OHCs since their birth, less than half of them will trans-differentiate into IHCs. This trans-differentiation in *TgPax2*$^{Cre/+}$*; Insm1*$^{F/F}$ and *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice is more frequent in HC rows closer to the IHCs and less further away (Fig 1l,n). This distribution reveals the existence of a gradient in the neural to abneural axis of the organ of Corti regulating cochlear HC types. This gradient could be inducing IHC-differentiation, and INSM1 acting by preventing embryonic OHCs from responding to it.

In other developing cell types, INSM1 functions as a transcriptional activator or repressor[18–24]. We hypothesized that *Insm1* directs OHCs to develop differentially from IHCs by activating OHC-specific genes or inhibiting IHC-specific genes. We first determined which genes are expressed in either differentiating HC type, when *Insm1* is expressed (Fig 3), and then searched for genes regulated by INSM1 in developing OHCs (Fig 4). For both approaches we used $Insm1^{GFP.Cre}$, in which the coding sequence of *Insm1* is replaced by that of a fusion protein between GFP and the Cre recombinase, thereby serving as a reporter as well as a null allele[4,23]. We generated $Insm1^{GFP.Cre/+}$; $Atoh1^{A1GFP/+}$; $R26R^{tdTomato/+}$ mice, in which all HCs express GFP (starting at E13.5 from $Atoh1^{A1GFP}$ and, in OHCs, from $Insm1^{GFP.Cre}$) but only OHCs also express tdTomato following $Insm1^{GFP.Cre}$ expression (throughout the cochlea by E18.5[4]; Fig 3a,b). We used these mice to sort OHCs and IHCs from neonatal (P0, ~E19.5) organs of Corti ((Fig 3c). With fluorescent activated cell sorting (FACS) we collected pools of IHC and OHCs RNAs (Fig 3d) and by RNA-seq obtained their transcriptomes (Supplementary table 1). We thus identified 922 IHC-enriched genes and 676 OHC-enriched genes (Fig 3e, supplementary tables 1–3). Among these are the 12 genes previously shown to be expressed preferentially in early IHCs or OHCs[4,17,25–29] (Fig 3e), indicating that our approach detects most differentially-expressed genes. A concern was whether genes with small expression differences (≤2×, e.g. *Zmat3*), or detected at very low levels in one cell type only (e.g., *Sox18* and *Msx1*), were truly differentially expressed. We selected 21 transcripts (Fig 3e), and tested by RT-qPCR for differential expression using additional IHC and OHC RNA pools. All 21 genes were confirmed to be differentially expressed, and the differences in expression were similar whether estimated by RNA-seq or RT-qPCR (Fig 3f). We also confirmed differential expression of additional genes by methods not susceptible to potential artifacts of cell sorting and mRNA detection: *Bcl11b* in OHCs by immunohistochemistry (Fig 2b); *Insm2* in OHCs with an $Insm2^{LacZ}$ mouse line (Fig 3e inset); and other genes by in situ hybridization (ISH) as preferentially expressed in OHCs (*Neurod6*, *Sez6l*) or IHCs (*Tbx2*, *Id4*, *Rprm*, *Smad3*, *Car13*, *Brip1*, *Lrrn1*, *Pink1*) (Fig 4g-o). All the above attests to the low prevalence of false positives among the genes we estimated as differentially-expressed between immature IHCs and OHCs.

The transcriptomes of perinatal cochlear hair vs SCs have been obtained, but these included a mixture of both OHCs and IHCs[25,26]. Although cell-type specific transcriptomes of mature IHCs and OHCs, obtained with microarray, were also reported[30], ours are the first transcriptomes of these cell types prior to maturity, during early differentiation. A comparison among differentiating vs mature IHC- or OHC-enriched genes reveals very little overlap (Fig 3g and supplementary table 4): only 5.9% IHC-enriched and 2% OHC-enriched genes are differentially expressed at both differentiating and mature stages. These include some known genes characteristic of the mature stage (*Vglut3* and *Otof* in IHCs and *Prestin* in OHCs) but whose expression is incipient at birth. However, the vast majority of the genes preferentially expressed in either cell type during differentiation (like *Insm1*, *Insm2* and *Bcl11b* in OHCs; and *Brip1*, *Car13* and *Fgf8* in IHCs), are not expressed upon maturation, and vice versa. These results reveal that a complex transcriptome, involving hundreds of genes, is transiently active during OHC- and IHC-specific differentiation. It is in this genetic

context that INSM1 locks the fate of OHCs so that they proceed to differentiate into mature OHCs and not IHCs.

To explore how INSM1 drives OHC differentiation, we looked by RNAseq for genes differentially expressed in differentiating OHCs with and without INSM1 (*Insm1*$^{GFP.Cre/+}$ vs *Insm1*$^{GFP.Cre/-}$) (extended Fig 6, Fig 4a, supplementary table 5). We identified 31 to 331 genes that could be differentially expressed, either up- or down-regulated by *Insm1* (supplementary tables 6, 7). Comparing these genes with those normally enriched in OHCs or IHCs (Fig 4b-e), combined with RT-qPCR retesting (supplementary table 8) and ISH (Figs 4f-o), reveals that in OHCs *Insm1* is not activating OHC genes but inhibiting IHC genes. No upregulated genes were confirmed by RT-qPCR and, of the 22 down-regulated genes confirmed, 21 are normally preferentially expressed by IHCs. The enrichment of these genes in wild type IHCs is similar to their upregulation in OHCs lacking INSM1 (Fig 4f,g, extended table 1). By contrast, most genes differentially expressed in OHCs vs IHCs were not affected by *Insm1*. We conclude that INSM1 down-regulates a specific subset of IHC-enriched genes in embryonic OHCs; without INSM1, those genes are expressed in embryonic OHCs, nearly half of which transdifferentiate into IHCs.

At E18.5, OHCs lacking INSM1 have not upregulated most of the early IHC-specific genes and still express early OHC-specific genes (Fig 4f), even though many of these cells will, once differentiated, express all examined features and markers of IHC and none of OHCs (Fig 1e-k,m, extended Fig 3f). The small set of early IHC-specific genes (21/922, ~2%) upregulated in embryonic OHCs lacking INSM1 likely represent an early step in the genetic cascade leading to their complete transformation into IHCs. Since oc-HCs expressing these few genes differentiate as IHCs, these genes are likely required for IHC differentiation. Hence, in addition to identifying *Insm1* as a critical gene for OHC differentiation, our results also identify the first candidate genes regulating the specific differentiation of IHCs. Because all OHCs express *Insm1*, yet in its absence less than half trans-differentiate into IHCs, we expected two patterns of miss-expression by ISH (Fig 4g-o,g). Some genes were upregulated in all OHCs lacking INSM1, as expected if repressed by INSM1. These (*Rprm*, *Id4*, *Lrrn1, Car13*, *Pink1* and *Brip1*; Fig 4h-n) must include the genes whose disinhibition in the absence of INSM1 renders embryonic OHCs susceptible to the gradient that induces IHC trans-differentiation. Other genes (*Fgf8* and *Tbx2*; Figs 2d,4o) were only upregulated in less than half of oc-HCs, presumably those trans-differentiating into IHCs. These genes are some of the earliest expressed in IHCs, and likely include regulators of IHC differentiation.

Our results reveal a homeotic transformation of OHCs into IHCs in the absence of INSM1, identify the genes initially miss-regulated by ablation of *Insm1*, and provide a genetic mechanism for how these two cell types differentiate: nascent OHCs transiently express *Insm1*, which represses (directly or indirectly) a core set of early IHC-specific genes and renders the cells insensitive to an IHC-inducing gradient; this consolidates the fate of OHCs by preventing their trans-differentiation into IHCs.

Materials and Methods

Ethics

All animal care and procedures were in strict accordance with the *Guide for the Care and Use of Laboratory Animals* published by the National Institutes of Health and were approved by Northwestern University's Institutional Animal Care and Use Committee (Animal Study Protocols IS00001281 and IS00000593).

Generation of the *Insm1* floxed allele for conditional ablation

The *Insm1* targeting construct was generated using a genomic BAC clone, 439G2, from the mouse 129/SvEv genomic BAC library, RPCI-22. The *Insm1* gene, including the coding sequence, 5' and 3' untranslated regions (UTRs), a 2,790bp 5' homologous sequence and a 4,098 bp 3' homologous sequence, was subcloned into the pL253 vector (IA1-pL253) using recombineering as described in[31]. The recombined clone, IA1-pL253 was further modified using recombineering to add a *LoxP* recombination site immediately downstream of the 5'UTR but before the Kozak sequence. A second *Frt-NEO-Frt-LoxP* site was recombined immediately downstream of the 3'UTR. The completed targeting vector was sequence verified and sent to the Northwestern Transgenic and Targeted Mutagenesis Laboratory (Chicago, IL) for electroporation into SvEv 129 mouse embryonic stem cells.

Using Q5 High-Fidelity Polymerase with GC Enhancer (NEB Catalog:M0491) and the primers: WT = TCTTAGATTCTGCCCTTTCTGACAG, CKO = CCAAGGAGATGACCACGCATAG, R2 = CTCTTGTAGGGCCTCCTGTG, we performed a PCR to identify recombinant clones. Conditions for Thermal-cycler are: Step 1: 98°C, 3:00 minutes, Step 2: 98°C, 0:10 minutes, Step 3: 65°C, 0:30 minutes, *−1deg/cycle*, Step 4: 72°C, 6:45 minutes, Repeat Step 2, 10×, Step 5: 98°C, 0:10 minutes, Step 6: 60°C, 0:30 minutes, Step 7: 72°C, 6:45 minutes, Repeat step 5, 25X, Step 8: 72°C, 10:00 minutes, Step 9: 4°C. Expected sizes for wild type allele using primers WT to R2 was 6163 bps. Expected size for recombinant clones using CKO-Reverse was 6145 bps. We screened a total of 439 clones and identified 5 recombinants.

We further screened these 5 ES cell clones for recombination upstream of the 5' LoxP site. DNA from selected recombinant clones was digested with the restriction enzyme, *Spe*I (NEB Cat:R0133) and homologous recombination was confirmed by Southern hybridization. DNA was visualized using a 168 bp radiolabeled probe (as described in[6]). The expected band sizes for wild type and conditional knockout alleles are 18,162 bps and 14,333 bps, respectively. All 5 ES clones contained a targeted allele of *Insm1*.

These clones were used for the generation of mosaic embryos and implantation into surrogate mothers by the Northwestern Transgenic and Targeted Mutagenesis Laboratory (Chicago, IL). From one of these clones (B3) we generated first chimeric mice, which were mated to mice expressing the FlpE recombinase (B6-Tg(CAG-FLPe)36[32] to deleted the NEO cassette flanked by FRT sites and thus generate mice with a floxed allele of *Insm1.*

Hearing Tests

During testing, mice of both sexes aged P25 to P31 were anesthetized with ketamine/xylazine (120 mg/kg; 10 mg/kg, IP) and their body temperature maintained using a heating blanket. In order to assay OHC function, distortion product otoacoustic emissions (DPOAE) were recorded using a custom probe equipped with a sensitive microphone (Knowles Electronics, FG-3652-CX). Responses were analyzed using Emission Averager (EMAV)[33]. Because the probe can be placed close to the eardrum, sound calibrations in the ear canal of each mouse were performed out to 48 kHz using a chirp stimulus generated in System Response (SysRes)[34]. All signals were generated using a CardDeluxe 24-bit sound card with a sampling rate of 96 kHz. Iso-input functions (f2/f1=1.2) at L1=50; L2=35 dB were recorded for f2 frequencies between 2 and 47 kHz, thereby covering most of the mouse audiogram. Input-output functions were also acquired for various f2 frequencies (6, 12, 27 kHz), where L1=L2+10dB. Thresholds for 2f1-f2 were then calculated and represent the level of f1 producing a DPOAE of 0 dB. After emission testing, neural responses were measured by collecting auditory brainstem responses (ABR) using tone-burst stimuli. Threshold was determined by noting the level at which the ABR waveform disappeared into the noise. For these experiments, sound calibration was obtained using a real pinna coupler[35].Further details are provided in a previous publication[36].

Tissue Collection and Preparation

For neonates, animals were euthanized by decapitation, and cochleae dissected in cold HBSS with calcium and magnesium (Gibco). For embryos, time pregnant dams were euthanized by isoflurane overdose followed by cervical dislocation. Their abdomens were opened to expose the uterus which was dissected in cold HBSS with calcium and magnesium (Gibco) and embryos harvested, and cochleae retrieved. After dissection, neonatal and embryonic cochleae were processed depending on future use. For immunohistochemistry, embryonic and neonatal cochleae were fixed in 4% paraformaldehyde for 2 hours at room temperature. For older tissues (>P20), the animals were cardiac-perfused with 4% paraformaldehyde, cochleae dissected, and post-fixed in 4% paraformaldehyde for 2 hours at room temperature. Cochleae from animals older than P5 were decalcified in 10% EDTA, pH 7.4, at 4˚C until needed. Organs of Corti were dissected out from the cochleae into one apical, two middle and one basal sections using a whole mount surface preparation method[37]. Frozen sections were processed as described in[4].

Immunohistochemistry

Whole mount organ of Corti sections were processed for immunohistochemistry as described previously[38]. Primary antibodies were mouse-anti-calmodulin (1:100, C-7055, Sigma Aldridge), goat-anti-oncomodulin (1:200, sc-7446, Santa Cruz), rabbit-anti-prestin (1:1000, from J Zheng, Northwestern University), guinea pig-anti-VGLUT3 (1:2500, from Robert Edwards, University of California, San Francisco), mouse-anti-CtBP2 (1:400, 612044, BD Biosciences), rabbit-anti-myosin7a (1:800, 25–6790, Proteus Biosciences), sheep-anti-neuroplastin (1:150, AF7818, R&D SYSTEMS), mouse-anti-BCL11B (1:400, ab18465, abcam), goat-anti-SOX2 (1:500, sc-17320, Santa Cruz). For BCL11B immune-labelling on whole mount cochlea, we performed antigen retrieval by incubating samples in

10 mM sodium citrate, pH 6 with 0.25% Triton X-100 for 20 minutes at 92°C and cooling for 30 minutes at room temperature prior to blocking. For CtBP2 and SOX2 immuno-labelling, samples were prepared using a freeze-thaw method. In brief, organ of Corti sections were incubated in 30% sucrose at room temperature for 20 minutes, put in −80°C for 5 minutes, thawed at room temperature for 20 minutes and rinsed off sucrose with PBS before blocking and incubating with primary antibodies at 37°C overnight. Nuclei were counterstained with 1:1000 DAPI or 1:2000 Hoechst 33342.

X-gal staining to detect β-galactosidase expression on sections of *Insm2*^{LacZ} embryos was performed as described[39].

Cell Proliferation Assay

In order to label hair and supporting cells generated from progenitors proliferating postnatally, neonates were injected twice daily from P0-P5 with the thymidine analog 5-ethynyl-2′-deoxyuridine (EdU; 50mg/kg in sterile saline). EdU incorporation into DNA was detected using Click-iT® Plus EdU Alexa Fluor® 555 Imaging Kit (Thermo Fisher Scientific) according to the manufacturer's manual. Following EdU detection, the samples were immune-labelled with antibodies as described above.

Image acquisition and analysis

We acquired images on either a Nikon A1 or A1R+ Confocal imaging system using a 100× objective. Three-dimensional renderings were generated using NIS Elements AR4.60.00 (Nikon) and Imaris X64 8.4.1 (BitPlane) software. Nuclear and ribbon synapse measurements were performed using built-in analysis functions on Imaris. Post-acquisition, we identically processed image pairs of control and knockout samples. This included adjustment for brightness, contrast and parameters for 3D volume and surface renderings of all images.

Fluorescent Activated Cell Sorting (FACS)

For collecting OHCs or IHCs, organs of Corti were dissected from E18.5 embryos (*Insm1*$^{GFP.Cre/-}$ or *Insm1*$^{GFP.Cre/+}$) or P0 neonates (*Insm1*$^{GFP.Cre/+}$; *Atoh1*$^{A1GFP/+}$; *R26R*$^{tdTomato/+}$) in ice-cold HBSS with calcium and magnesium (Gibco). A portion of tail from each embryo and neonate was collected for genotyping. Organs of Corti were washed three times in cold 1×PBS and then they were digested in 0.33U/ml papain, 0.5mM EDTA and 1mM L-cystine in EBSS for 10 minutes at 37°C, rinsed 3 times in 2% FBS and mechanically dissociated by gentle trituration (~100–150× with a P1000 pipet). Cell suspensions were kept on ice until FACS sorting on a BD FACS Aria 4 flow cytometer through a 100 μm nozzle at speed 2 (<100 events/sec). Hair cell populations were collected into RLT buffer (Qiagen, Valencia, CA). RNA was then isolated from cells using Qiagen RNeasy Plus Micro Kit or cells were stored at −80˚C until RNA isolation. Isolated RNA was evaluated for quality and concentration on BioAnylizer and stored at −80˚C.

qRT-PCR

RT-PCR was performed using either SYBR® Green or TaqMan systems. Total RNA was extracted from pools of HCs collected from E18.5 *Insm1*$^{GFP.Cre/-}$, E18.5 *Insm1*$^{GFP.Cre/+}$, and

P0 *Insm1*$^{GFP.Cre/+}$; *Atoh1*$^{A1GFP/+}$; *R26R*$^{tdTomato/+}$ animals through FACS. RNA extraction was done using RNeasy® Plus Micro Kit (Qiagen) according to the manufacturer's instructions. RNA quality was determined with a BioAnalyzer through NUSeq Core Facility, Northwestern University, Chicago, IL.

For SYBR® Green qRT-PCR, we used ~3000pg of total RNA from each HC pool for first strand cDNA synthesis using iScript™ reverse transcription supermix (Bio-Rad) according to the manufacturer's manual. We then performed RT-qPCR with ~200pg of first strand cDNA using SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad) in triplicates on a CFX Connect™ Real-Time PCR Detection System (Bio-Rad) using a 40-cycle protocol.

For TaqMan qRT-PCR, we used 1ug total RNA from each HC pool for first strand cDNA synthesis using SuperScript® VILO™ cDNA Synthesis Kit with ezDNase™ Enzyme (Applied Biosystems) according to the manufacturer's instructions. First strand cDNA was subjected to ezDNase™ inactivation using 1ul of 100mM DTT per reaction. Prior to qRT-PCR, we performed pre-amplification of first strand cDNA using TaqMan® PreAmp Master Mix (Applied Biosystems) according to the manufacturer's instructions. We then performed qRT-PCR using 1ug of diluted pre-amplified cDNA (1:20 in TE buffer) per reaction in triplicates on a QuantStudio™ 7 flex Real-Time system (Applied Biosystems) using a 14-cycle protocol at NUSeq Core Facility, Northwestern University, Chicago, IL.

All primers used for qRT-PCR in this study were designed and pre-mixed to their optimal concentrations by BioRad and Applied Biosystems. qRT-PCR reactions were performed according to the corresponding manufacturer's instructions.

RNA-Seq and transcriptome analysis of embryonic OHCs and P0 HCs

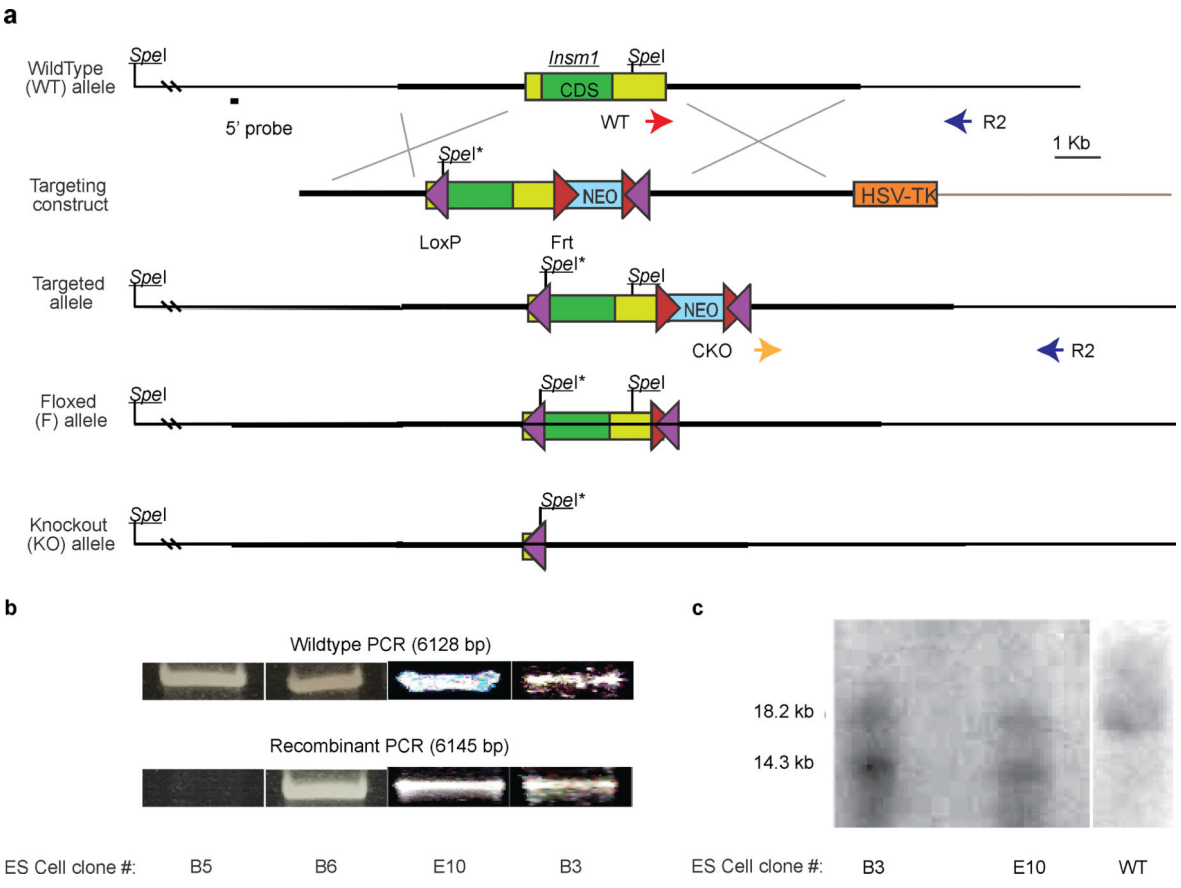
To purify enough RNA for deep sequencing and also analyze results statistically, or the determination of the IHC and OHC transcriptomes we collected by FACS six separate pools of IHCs (700–1100 cells per pool) and six of OHCs (2800–3700 cells per pool) from *Insm1*$^{GFP.Cre/+}$; *Atoh1*$^{A1GFP/+}$; *R26R*$^{tdTomato/+}$ at P0 (generated by timed pregnancies and found to correspond in most cases to E19.5 and in the rest to E20.5). For the determination of the transcriptomes of OHCs with and without INSM1 we collected OHCs by FACS into three separate pools per genotype (*Insm1*$^{GFP.Cre/-}$ and *Insm1*$^{GFP.Cre/+}$), each with 2200 to 5000 OHCs from 8 to 12 E18.5 embryos. We extracted 3 to 7.5 ng of RNA per E18.5 OHC pool, and ~3 ng for P0 IHC and 10–18 ng P0 OHC pools. We only used samples with an RNA integrity number (RIN) >8.

Beijing Genomics Institute (BGI) performed sample preparation and sequencing at their facility in the Children's Hospital of Philadelphia (CHOP). The total RNA samples were first treated with DNase I to degrade any possible DNA contamination followed by ribosomal RNA removal using RiboZero (Epicentre), converted it to cDNA and amplified with NuGEN Ovation® RNA-Amplification System V2. Mixed with the fragmentation buffer, the mRNA was fragmented into short fragments of about 200 bp. Then the first strand of cDNA was synthesized using random hexamer-primer. Buffer, dNTPs, RNase H, and DNA polymerase I were added to synthesize the second strand. Double stranded cDNA was purified with magnetic beads, performed end reparation and 3'-end single nucleotide A

(adenine) addition, and ligated sequencing adaptors to the fragments, which were enriched by PCR amplification. Libraries were qualified and quantified with an Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR. Individually barcoded 100 bp paired-end library products were sequenced on the Illumina HiSeq2000 (3 libraries from E18.5 *Insm1*$^{GFP.Cre/-}$ and 3 from *Insm1*$^{GFP.Cre/+}$ OHCs) or the HiSeq4000 (6 IHC plus 6 OHC libraries from P0 *Insm1*$^{GFP.Cre/+}$; *Atoh1*$^{A1GFP/+}$; *R26R*$^{tdTomato/+}$), multiplexed per lane, yielding 48–50 million (for each of the 6 E18.5 OHC libraries) and 92–116 million (for each of the 12 P0 IHC or OHC libraries) paired reads. DNA read quality was evaluated in fastq format using FastQC, adapters were trimmed, and reads of poor quality or aligning to rRNA sequences were filtered.

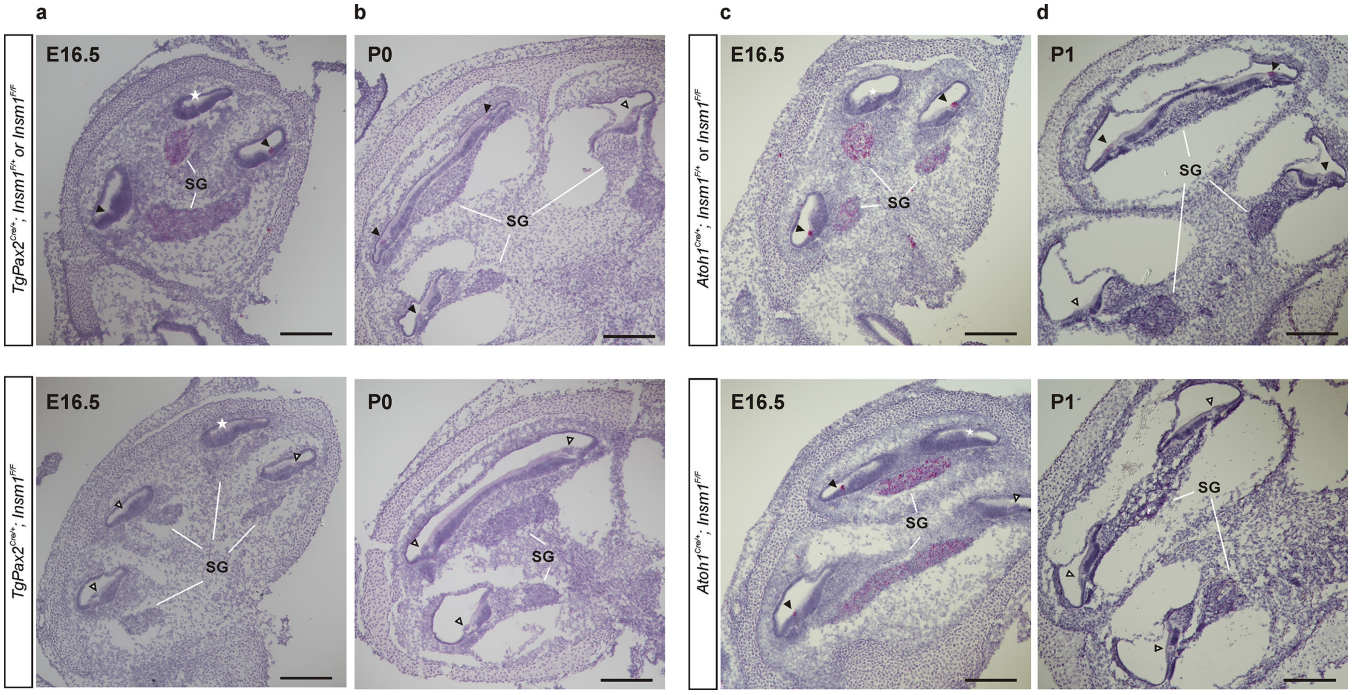
The cleaned reads were aligned to the Mus musculus genome (mm10) using STAR[40]. Read counts for each gene were calculated using htseq-count[41] in conjunction with a gene annotation file for mm10 obtained from UCSC (University of California Santa Cruz; http://genome.ucsc.edu). Differential expression was determined using DESeq2[42]. The cutoff for determining significantly differentially expressed genes was an FDR-adjusted p-value less than 0.05.

Extended Data



Extended Figure 1: Generation of a conditional KO allele of *Insm1*.

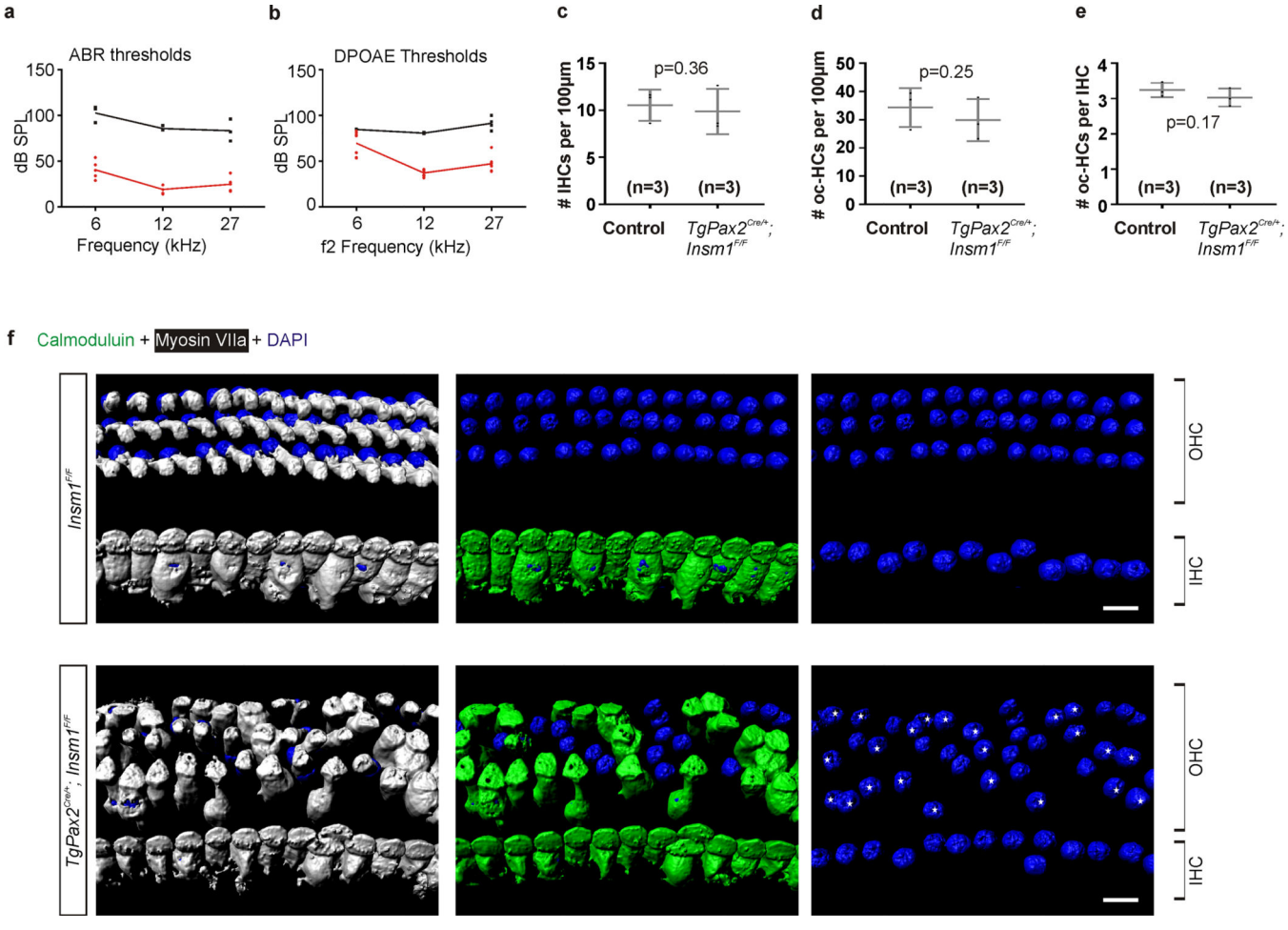
a We generated a targeting construct in which the sole exon of *Insm1* (green rectangle, with the coding sequence in dark green and the UTRs in light green) has a *loxP* site (purple triangle) inserted in a poorly-conserved area of its 5' UTR and another *loxP* site downstream of the *Insm1* gene. The construct also incorporates a neomycin resistance cassette (NEO, blue) surrounded by *FRT* sites (red triangles) and a Thymidine kinase cassette (HSV-TK; orange), which are used to select for recombination events after gene targeting. **b** We screened a total of 439 clones and identified 5 recombinants (1 non-recombinant wild type; B5, and 3 recombinants; B6, E10 and B3, are shown) with PCR using primers indicated in **a** (arrows). Expected sizes for wild type allele using primers WT to R2 is 6163 bp. Expected size for recombinant clones using CKO-Reverse is 6145 bp. **c** Selected ES cell clones were additionally screened for homologous recombination upstream of the first *loxP* site by Southern blot after digestion with *Spe*I and using the 5' probe indicated in (**a**)[6]. Southern blot was performed twice. The expected band sizes for wild type and conditional KO alleles are 18,162 bp and 14,333 bp, respectively. From one of these clones (B3) we generated first chimeric mice and then mice with floxed alleles of *Insm1 (*obtained by crossing the chimeras with mice expressing the FlpE recombinase (B6-Tg(CAG-FLPe)36, which deleted the NEO cassette flanked by FRT sites). Homozygous *Insm1*$^{F/F}$ mice are viable, demonstrating that the *loxP* insertions do not interfere with the vital functions of *Insm1* and hence may be used for its conditional ablation. Co-expression with Cre recombinase generates an *Insm1* KO allele lacking its entire coding sequence.



Extended Figure 2: Conditional ablation of *Insm1* in cochleae.

In situ hybridization for *Insm1* transcripts on cryosections of embryonic E16.5 and neonatal (P0 and P1) cochleae. **a**, **c** In control cochleae (top panels), *Insm1* is expressed in all OHCs (72/72 OHCs from 3 animals) and spiral ganglion (SG; white arrows) at E16.5. By postnatal age P0-P1, *Insm1* mRNA is present in 90% of OHCs (94/105 OHCs from 2 animals), and its

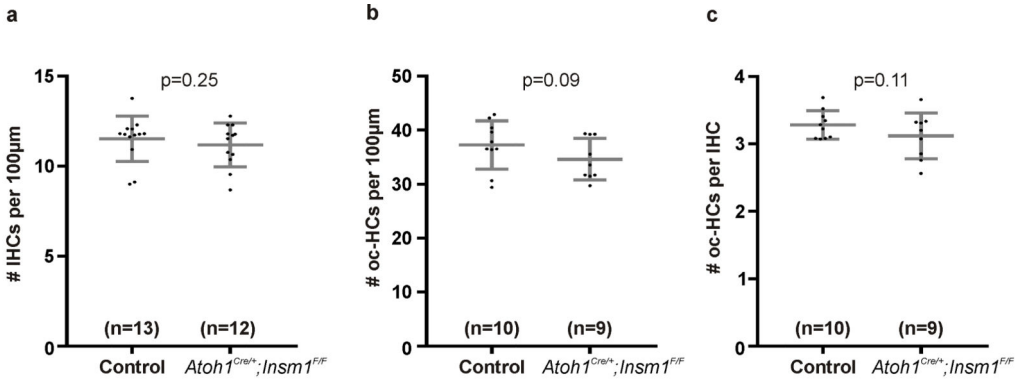
level is undetectable in SG. However, there is no *Insm1* detected in the organs of Corti from apical turns, in which recognizable hair cells have not yet appeared (**a**, **c**; asterisks). **a**, **b** (bottom panels) In *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice, *Insm1* mRNA is undetectable in SG and in all OHCs from E16.5 (0/69 OHCs from 2 mice; **a**) and P0 (0/42 OHCs from 1 mouse; **b**) cochleae. **c, d** (Bottom panels) In *Atoh1*$^{Cre/+}$; *Insm1*$^{Flox/Flox}$ cochleae, *Insm1* mRNA is present in SG and 43% of OHCs (18/42 OHCs from 1 animal; **c**) at E16.5, reduced to 7% (4/54) at E17.5 (not shown), and entirely absent all OHCs (0/60 OHCs from 1 animal) and SG by postnatal day P1. For quantification at E16.5, we did not include organs of Corti from apical turns, which do not yet have recognizable hair cells. Filled arrowheads indicate the organs of Corti with *Insm1* expression, while empty arrowheads indicate the organs of Corti without *Insm1* expression. Scale bars are 200 μm.

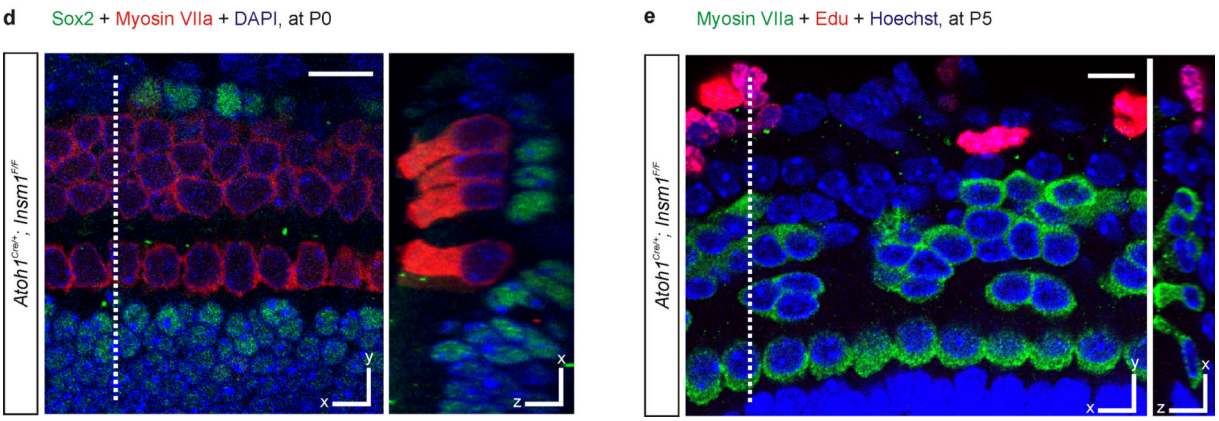


Extended Figure 3: Conditional ablation of *Insm1* in hair cells and spiral ganglia neurons using *TgPax2*^{Cre} causes hearing impairment and the appearance of IHC-like cells in place of OHCs.

a-b Hearing thresholds determined by auditory brainstem responses (ABRs; **a**) and distortion product otoacoustic emissions (DPOAEs; **b**) of *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice at age P35 to P46 (black traces; n=4, 4 females) and control littermates (red traces; n=5, 2 males + 3 females). The larger ABR than DPOAE threshold shifts may indicate an additional contribution to hearing impairment of the SG neurons lacking INSM1 in *TgPax2*$^{Cre/+}$;

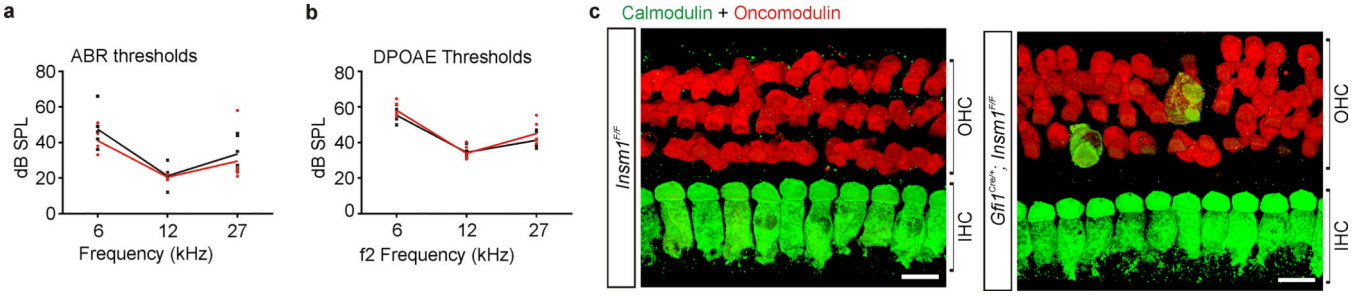
Insm1$^{F/F}$ cochlea. **c** Despite the prevalence of OHCs with IHC characteristics (oc-IHCs) in *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ cochlea (46.0% ± 5.64 SD, n=3 mice; Figure 1m,n), these mice have the same density of IHCs (9.87 ± 2.41 SD cells per 100 μm along the organ of Corti; n=3) as that of littermate controls (*TgPax2*$^{Cre/+}$; *Insm1*$^{F/+}$ and *Insm1*$^{F/F}$; 10.54 ± 1.67 SD cells per 100 μm; n=3) suggesting that oc-IHCs are not IHCs displaced from the inner to the outer compartment. **d** There is no OHC loss in *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice at ~P14–16. Densities of outer compartment hair cells (oc-HCs) do not differ between *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice (OHCs + oc-IHCs) and their littermate controls (OHCs only) (29.88 ± 7.45 SD cells per 100 μm along the organ of Corti of *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice, n=3; vs 34.3 ± 6.92 SD cells per 100 μm of *TgPax2*$^{Cre/+}$; *Insm1*$^{F/+}$ and *Insm1*$^{F/F}$ littermate controls, n=3). **e** The ratio of oc-HCs per IHC do not differ between *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice and their littermate controls (3.03 ± 0.25 SD OHCs plus oc-IHCs per IHC in *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ mice, n=3; vs 3.24 ± 0.21 SD OHCs per IHC in *TgPax2*$^{Cre/+}$; *Insm1*$^{F/+}$ and *Insm1*$^{F/F}$ littermate controls, n=3). One-tailed Student's t-tests were used in (**c**, **d**, **e**) based on n = number of mice. Average ± SD is shown on the plots. Statistically significance is defined as $p < 0.05$. **f** Immunofluorescence for the IHC-enriched calmodulin (green) and HC marker myosin VIIa (white) on whole mount organs of Corti from mid-cochlear positions at ages ~P14–16 further confirmed that many of *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$ oc-HCs have IHC characteristics, in addition to having the flask shape and large nuclei like those of IHCs (blue, DAPI, marked with asterisks), as well as lacking prestin and expressing VGLUT3 (Figure 1m). Scale bars are 10μm. Similar immunohistochemistry results obtained from ≥3 animals per genotype.





Extended Figure 4: IHC-like cells in the outer compartment (oc-IHCs) result from OHC misdifferentiation in the absence of INSM1, not from IHC displacement or from supporting cell trans-differentiation.

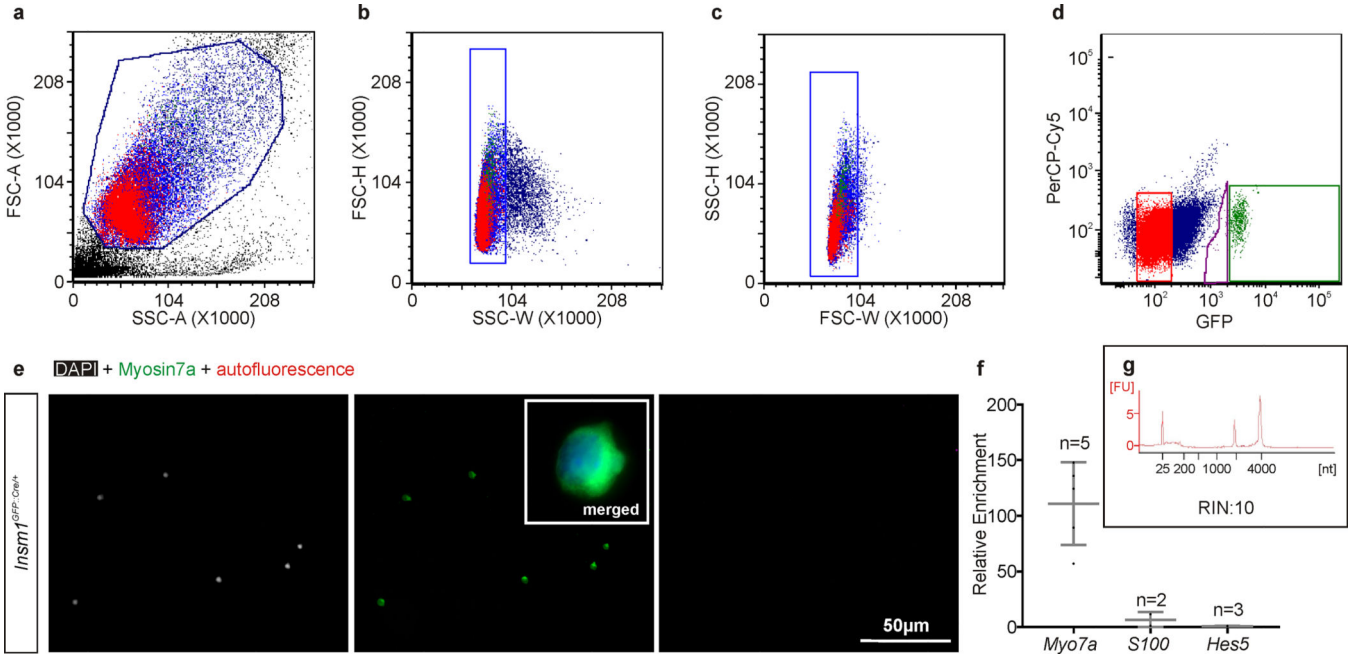
a *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice have the same density of IHCs (11.2 ± 1.2 SD cells per 100 μm; n=12) as that of littermate controls (*Atoh1*$^{Cre/+}$*; Insm1*$^{+/F}$ and *Insm1*$^{F/F}$; 11.5 ± 1.3 SD cells per 100 μm; n=13). **b** There is no loss of OHCs in *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice up to P34. Densities of oc-HCs do not differ between *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice (OHCs + oc-IHCs) and their littermate controls (OHCs only) (34.6 ± 3.8 SD cells per 100 μm in *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice, n=9; vs 37.3 ± 4.5 SD cells per 100 μm of *Atoh1*$^{Cre/+}$*; Insm1*$^{+/F}$ and *Insm1*$^{F/F}$ littermate controls, n=10). **c** The ratio of outer compartment HCs per IHC do not differ significantly between *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice and their littermate controls (3.1 ± 0.3 SD OHCs and oc-IHCs per IHC in *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice, n=9; vs 3.3 ± 0.2 SD OHCs per IHC in *Atoh1*$^{Cre/+}$*; Insm1*$^{+/F}$ and *Insm1*$^{F/F}$ littermate controls, n=10). The criteria to identify oc-IHCs in *Atoh1*$^{Cre/+}$*; Insm1*$^{F/F}$ mice were the presence in the outer compartment of hair cells expressing IHC markers (VGLUT3, high-levels of calmodulin and/or nuclear CtBP2), lacking OHC markers (oncomodulin and/or prestin) and with a shape (determined by myosin VIIa immunoreactivity) like that of IHCs. Mice used for hair cell counts were P0-P34. One-tailed Student's t-tests were used in (**a**, **b**, **c**) based on n = number of mice. Average ± SD is shown on the plots. Statistically significance is defined as $p < 0.05$. **d** SOX2-immunoreactivity, which labels the nuclei of cochlear supporting cells and, under certain conditions, of hair cells trans-differentiated from them postnatally, was not present in cells of the OHC region in *Atoh1*$^{Cre/+}$; *Insm1*$^{F/F}$ pups (0/95 OHCs at P0, 0/42 OHCs at P2, and 0/39 OHCs at P5). **e** To track postnatal cell proliferation in the organ of Corti, neonatal mice were injected twice daily with the thymidine analog 5-ethynyl-2′-deoxyuridine (EdU) from P0 to P5 or P8. The lack of EdU in any hair cell from in *Atoh1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice (0/77 OHCs at P5 and 0/40 OHCs at P8) confirmed that these, including oc-IHCs, were not produced from postnatally-dividing supporting cells. Unless otherwise noted, images are from mid cochlear positions. Hair cells were identified by myosin VIIa immunoreactivity, phalloidin, DAPI and Hoechst. Scale bars are 10μm. Similar immunohistochemistry results obtained from ≥3 animals per genotype.



Extended Figure 5: Conditional ablation of *Insm1* in hair cells using Gfi1-Cre causes no hearing impairment, and results in only very few oc-IHC.

We generated a conditional knockout of *Insm1* in hair cells using Gfi1-Cre in which the expression of cre recombinase coincided with that of *Insm1*. **a-b** Hearing thresholds determined by auditory brainstem responses (ABRs; **a**) and distortion product otoacoustic emissions (DPOAEs; **b**) of *Gfi1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice at age P30 to P35 (black traces; n=6, 4 males + 2 females) and control littermates (red traces; n=6; 1 *Insm1*$^{F/F}$, 1 *Insm1*$^{F/+}$ and 4

Gfi1$^{Cre/+}$; *Insm1*$^{F/+}$; 4 males + 2 females). There is no significant difference in ABR and DPOAE thresholds at all frequencies tested between *Gfi1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice and their control littermates. **c** Immunohistochemistry in whole mount organs of Corti from mid-cochlear positions of P34 mice tested for hearing in (**a-b**) revealed that *Gfi1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice had normal IHCs expressing high level of calmodulin. However, very few HCs in the outer compartment also expressed calmodulin at high levels, oncomodulin at low levels, and had a round, flask shape similar to that of IHCs (0.78%; 5/526 OHCs from 2 mice). Because in these *Gfi1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice the onset of Cre recombinase expression coincides with that of *Insm1* (E15.5 to E17.5)[4,44], their nascent OHCs will express *Insm1* for at least several hours. This result indicates that a brief expression of *Insm1* is sufficient to promote proper OHC differentiation. Scale bars are 10μm. Similar immunohistochemistry results obtained from ≥3 animals per genotype.



Extended Figure 6: OHC FACS purification and RNA extraction from an E18.5 *Insm1*$^{GFP.Cre/+}$ embryo.

a-c Forward and side light-scattering are used to exclude dead cells and debris (**a**) and aggregates (≥2 cells) (**b,c**). **d** Live cells are gated in green and red (PerCP-Cy5, to assess autofluorescence) channels to define the GFP+ (green dots) and GFP- (red dots) sorting windows. **e** Myosin VIIa immunoreactivity and DAPI stain of cells collected through cytospinning after FACS confirm that most of the sorted 547 GFP+ cells are hair cells. This verification was done on all HC pools sorted, 3 pools per genotype. Inset is a representative merged image of one sorted OHC at high magnification. In this pool, there were no autofluorescent cells collected. **f** RT-qPCR after cell sorting (average ± SEM) reveals that, compared with GFP- cells, GFP+ cells express HC marker gene *Myo7a* and not the supporting cell marker genes *S100* and *Hes5*. **g** To ensure the quality of the extracted RNA, the RNA Integrity (RIN) score was determined with a BioAnalyzer. **e**, **f**, **g** Two pools of

OHCs (6e) and similar RIN scores were obtained from all pools of OHCs examined (including the 3 per genotype used for RNAseq on Figure 4a)

Extended Table 1:

Confirmed genes missregulated in OHC lacking INSM1.

Differential expression of IHCs respect to OHCs and of OHCs without INSM1 (from *Insm1*$^{GFP.Cre/-}$ mice, referred to as KO) with respect to OHCs with INSM1 (from *Insm1*$^{GFP.Cre/+}$, referred to as Het). Differential expression between KO and Het OHCs estimated by RNAseq was confirmed by TaqMan RT-qPCR (n=5 pools of OHCs per genotype for Tbx2, Nhs, Lrrn1, Brip1 and Rin3; n=4 for all other genes). P values are for one-tailed t-tests on RT-qPCR values. All 22 genes increase their expression in OHCs lacking INSM1. Of these, all except Sez6l are normally preferentially expressed in IHCs. Note that for Id4, differential expression between IHCs and OHCs was not detected by RNAseq at P0, and it did not reach significance between KO and HET OHCs by RT-qPCR at E18.5. However, significance was achieved (p=0.013) at E16.5, at which time differential expression was confirmed and visualized by in RNAscope situ hibridization (Fig 4j). Hence, for Id4 the differential expression ocurs transiently and very early.

GeneSymbol	Description	RNAseq Change (IHC/OHC)	RNAseq Change (KO/Het	RT-qPCR Change (KO/Het)	RT-qPCR p-value (KO/Het)
Tbx2	T-box 2	19.28	13.75	102.31	0.000428
Fgf8	fibroblast growth factor 8	18.02	8.12	65.08	0.004112
Smad3	SMAD family member 3	10.69	14.04	46.43	4.14E-05
Nhs	Nance-Horan syndrome (human)	8.01	11.56	3.15	0.034775
Lrrnl	leucine rich repeat protein 1, neuronal	7.86	4.93	7.79	0.003956
Bripl	BRCA1 interacting protein C-terminal helicase 1	7.47	12.65	118.50	5.53E-05
Rin3	Ras and Rab interactor 3	5.71	10.98	1.88	0.011625
Pcdhl	protocadherin 1	4.47	3.53	2.84	0.009309
Spryd3	SPRY domain containing 3	3.64	6.70	1.53	0.033057
Pacsl	phosphofurin acidic cluster sorting protein 1	3.56	2.63	2.33	0.00229
Carl3	carbonic anhydrase 13	3.55	4.27	3.69	0.003681
Tmprss7	transmembrane serine protease 7	3.03	3.07	1.58	0.033105
Rprm	reprimo, TP53 dependent G2 arrest mediator candidate	2.89	1.83	1.67	0.039198
Zfp668	zinc finger protein 668	1.85	3.23	1.30	0.044668
Mtssl	metastasis suppressor 1	1.83	2.34	1.56	0.02911
Cuxl	cut-like homeobox 1	1.61	2.05	1.57	0.026945
Lrrc8b	leucine rich repeat containing 8family, member B	1.56	1.71	1.31	0.045587
Rail	retinoic acid induced 1	1.57	2.56	1.14	0.033262
Pinkl	PTEN induced putative kinase 1	1.36	6.52	4.42	0.003253

GeneSymbol	Description	RNAseq Change (IHC/OHC)	RNAseq Change (KO/Het	RT-qPCR Change (KO/Het)	RT-qPCR p-value (KO/Het)
Cmtm8	CKLF-like MARVEL transmembrane domain containing 8	1.21	8.29	1.84	0.002469
Id4	inhibitor of DNA binding 4	0.99	1.94	1.84	0.063034
Sez6l	seizure related 6 homolog like	0.23	7.47	5.04	0.020066

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

NU core facilities used were TTML (CA60553), CAM (CA060553), FC (CA060553) and NUSeq. We thank A. Groves for protocols and advice, D. He for original databases of microarray results. Supported by NIH Grants DC015903, DC000089 and DC012483.

Bibliography for main text

1. Puligilla C & Kelley MW Building the world's best hearing aid; regulation of cell fate in the cochlea. Curr Opin Genet Dev 19, 368–373 (2009). [PubMed: 19604683]
2. Groves AK, Zhang KD & Fekete DM The genetics of hair cell development and regeneration. Annu Rev Neurosci 36, 361–381 (2013). [PubMed: 23724999]
3. Basch ML, Brown RM, Jen H-I & Groves AK Where hearing starts: the development of the mammalian cochlea. J Anat 228, 233–254 (2016). [PubMed: 26052920]
4. Lorenzen SM, Duggan A, Osipovich AB, Magnuson MA & García-Añoveros J Insm1 promotes neurogenic proliferation in delaminated otic progenitors. Mech Dev 138 Pt 3, 233–245 (2015). [PubMed: 26545349]
5. Gierl MS, Karoulias N, Wende H, Strehle M & Birchmeier C The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. Genes Dev 20, 2465–2478 (2006). [PubMed: 16951258]
6. Rosenbaum JN, Duggan A & García-Añoveros J Insm1 promotes the transition of olfactory progenitors from apical and proliferative to basal, terminally dividing and neuronogenic. Neural Dev 6, 6 (2011). [PubMed: 21284846]
7. Yang H, Xie X, Deng M, Chen X & Gan L Generation and characterization of Atoh1-Cre knock-in mouse line. Genesis 48, 407–413 (2010). [PubMed: 20533400]
8. Ohyama T & Groves AK Generation of Pax2-Cre mice by modification of a Pax2 bacterial artificial chromosome. Genesis 38, 195–199 (2004). [PubMed: 15083520]
9. Ruben RJ Development of the inner ear of the mouse: a radioautographic study of terminal mitoses. Acta Otolaryngol Suppl 220:1–44 (1967).
10. Chen P & Segil N p27(Kip1) links cell proliferation to morphogenesis in the developing organ of Corti. Development 126, 1581–1590 (1999). [PubMed: 10079221]
11. Kelley MW, Talreja DR & Corwin JT Replacement of hair cells after laser microbeam irradiation in cultured organs of corti from embryonic and neonatal mice. J Neurosci 15, 3013–3026 (1995). [PubMed: 7722642]
12. Cox BC et al. Spontaneous hair cell regeneration in the neonatal mouse cochlea in vivo. Development 141, 816–829 (2014). [PubMed: 24496619]
13. Bramhall NF, Shi F, Arnold K, Hochedlinger K & Edge ASB Lgr5-positive supporting cells generate new hair cells in the postnatal cochlea. Stem Cell Reports 2, 311–322 (2014). [PubMed: 24672754]
14. Hu L et al. Diphtheria Toxin-Induced Cell Death Triggers Wnt-Dependent Hair Cell Regeneration in Neonatal Mice. J Neurosci 36, 9479–9489 (2016). [PubMed: 27605621]
15. White PM, Doetzlhofer A, Lee YS, Groves AK & Segil N Mammalian cochlear supporting cells can divide and trans-differentiate into hair cells. Nature 441, 984–987 (2006). [PubMed: 16791196]
16. Mburu P et al. Whirlin complexes with p55 at the stereocilia tip during hair cell development. Proc Natl Acad Sci U S A 103, 10973–10978 (2006). [PubMed: 16829577]
17. Flores EN et al. A non-canonical pathway from cochlea to brain signals tissue-damaging noise. Curr Biol 25, 606–612 (2015). [PubMed: 25639244]

18. Breslin MB, Zhu M, Notkins AL & Lan MS Neuroendocrine differentiation factor, IA-1, is a transcriptional repressor and contains a specific DNA-binding domain: identification of consensus IA-1 binding sequence. Nucleic Acids Res 30, 1038–1045 (2002). [PubMed: 11842116]
19. Jia S, Wildner H & Birchmeier C Insm1 controls the differentiation of pulmonary neuroendocrine cells by repressing Hes1. Dev Biol 408, 90–98 (2015). [PubMed: 26453796]
20. Lan MS & Breslin MB Structure, expression, and biological function of INSM1 transcription factor in neuroendocrine differentiation. FASEB J 23, 2024–2033 (2009). [PubMed: 19246490]
21. Liu W-D, Wang H-W, Muguira M, Breslin MB & Lan MS INSM1 functions as a transcriptional repressor of the neuroD/beta2 gene through the recruitment of cyclin D1 and histone deacetylases. Biochem J 397, 169–177 (2006). [PubMed: 16569215]
22. Wang H-W et al. Identification of an INSM1-binding site in the insulin promoter: negative regulation of the insulin gene transcription. J Endocrinol 198, 29–39 (2008). [PubMed: 18417529]
23. Osipovich AB et al. Insm1 promotes endocrine cell differentiation by modulating the expression of a network of genes that includes Neurog3 and Ripply3. Development 141, 2939–2949 (2014). [PubMed: 25053427]
24. Jia S et al. Insm1 cooperates with Neurod1 and Foxa2 to maintain mature pancreatic β-cell function. EMBO J 34, 1417–1433 (2015). [PubMed: 25828096]
25. Cai T et al. Characterization of the transcriptome of nascent hair cells and identification of direct targets of the Atoh1 transcription factor. J Neurosci 35, 5870–5883 (2015). [PubMed: 25855195]
26. Scheffer DI, Shen J, Corey DP & Chen Z-Y Gene Expression by Mouse Inner Ear Hair Cells during Development. J Neurosci 35, 6366–6380 (2015). [PubMed: 25904789]
27. Yang T et al. Expression and localization of cabp ca2+ binding proteins in the mouse cochlea. PLoS ONE 11, e0147495 (2016). [PubMed: 26809054]
28. Jacques BE, Montcouquiol ME, Layman EM, Lewandoski M & Kelley MW Fgf8 induces pillar cell fate and regulates cellular patterning in the mammalian cochlea. Development 134, 3021–3029 (2007). [PubMed: 17634195]
29. Pirvola U et al. FGFR1 is required for the development of the auditory sensory epithelium. Neuron 35, 671–680 (2002). [PubMed: 12194867]
30. Liu H et al. Characterization of transcriptomes of cochlear inner and outer hair cells. J Neurosci 34, 11085–11095 (2014). [PubMed: 25122905]

Bibliography for Extended Data and Methods

31. Liu P, Jenkins NA & Copeland NG A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res 13, 476–484 (2003). [PubMed: 12618378]
32. Kanki H, Suzuki H & Itohara S High-efficiency CAG-FLPe deleter mice in C57BL/6J background. Exp Anim 55, 137–141 (2006). [PubMed: 16651697]
33. Neely ST & Liu Z EMAV: Otoacoustic emission averager. Omaha, NE: Technical Memo No 17 Boy's Town National Research Hospital (1994).
34. Neely ST & Stevenson R SysRes. Omaha, NE: Tech Memo No. 1 BoysTown National Research Hospital NE (1992).
35. Pearce M, Richter CP & Cheatham MA A reconsideration of sound calibration in the mouse. J Neurosci Methods 106, 57–67 (2001). [PubMed: 11248341]
36. Cheatham MA et al. Loss of the tectorial membrane protein CEACAM16 enhances spontaneous, stimulus-frequency, and transiently evoked otoacoustic emissions. J Neurosci 34, 10325–10338 (2014). [PubMed: 25080593]
37. Montgomery SC & Cox BC Whole mount dissection and immunofluorescence of the adult mouse cochlea. J Vis Exp (2016). doi:10.3791/53561
38. Delmaghani S et al. Hypervulnerability to Sound Exposure through Impaired Adaptive Proliferation of Peroxisomes. Cell 163, 894–906 (2015). [PubMed: 26544938]
39. Nagy A, Gertsenstein M, Vintersten K & Behringer R Staining Frozen Mouse Embryo Sections for {beta}-Galactosidase (lacZ) Activity. CSH Protoc 2007, pdb.prot4726 (2007).

40. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21 (2013). [PubMed: 23104886]

41. Anders S, Pyl PT & Huber W HTSeq — a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015). [PubMed: 25260700]

42. Love MI, Huber W & Anders S Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550–550 (2014). [PubMed: 25516281]

43. Salic A & Mitchison TJ A chemical method for fast and sensitive detection of DNA synthesis in vivo. Proc Natl Acad Sci U S A 105, 2415–2420 (2008). [PubMed: 18272492]

44. Yang H et al. Gfi1-Cre knock-in mouse line: A tool for inner ear hair cell-specific gene deletion. Genesis 48, 400–406 (2010). [PubMed: 20533399]

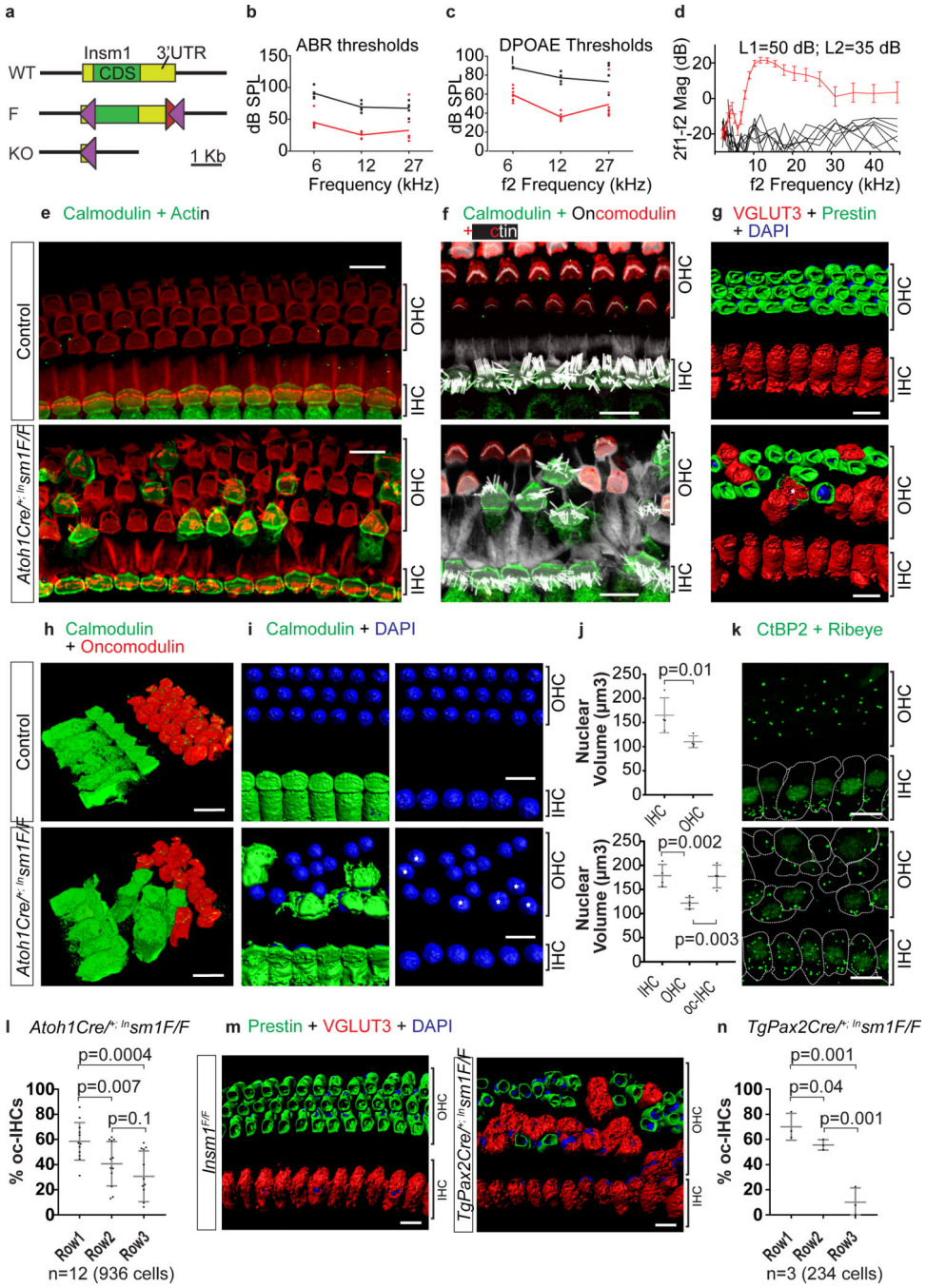


Figure 1: Conditional ablation of *Insm1* in hair cells results in IHC-like cells in place of OHCs.

a Wild type (WT), floxed (F) and conditionally deleted (cKO) alleles of *Insm1*. Co-expression with Cre recombinase generates an *Insm1* KO allele lacking its CDS and 3'UTR, leaving only part of the 5'UTR. Purple triangles represent loxP, and red FRT, sites. **b-d** Hearing tests (average with all values in **b,c** and SEM in **d**). Auditory brainstem responses (ABRs; **b**) and distortion product otoacoustic emissions (DPOAEs; **c**) thresholds, and iso-input functions for the DPOAE (**d**) of *Atoh1*$^{Cre/+}$; *Insm1*$^{F/F}$ mice at P25–31 (black; n=3 males + 2 females) and control littermates (red; n=4 *Insm1*$^{F/F}$ + 4 *Atoh1*$^{Cre/+}$; *Insm1*$^{F/+}$; 6

males + 2 females). Up arrow in **c** indicates that maximum sound was insufficient to reach threshold. **e-l** Immunohistochemistry in organs of Corti revealed that the *Atoh1* $^{Cre/+}$; *Insm1*$^{F/F}$ mice had normal IHCs but many OHCs expressed calmodulin and not oncomodulin (**e**, **f**, **h**), had stereocilliary bundles resembling those of IHCs (f-actin labelled with phalloidin; **e**, **f**), expressed VGLUT3 instead of Prestin (**g**; one rare cell, expressed both; asterisk), had cell shapes (**h**) and large nuclei like IHCs (asterisks; **i**, **j**). **j** Number of nuclei measured for each cell type are indicated under each bar. **k** oc-IHCs had nuclear CtBP2 and a number of presynaptic ribbons approaching that of IHCs. **l** In *Atoh1*$^{Cre/+}$; *Insm1*$^{F/F}$, oc-IHCs are more frequent in the first row of OHCs (closer to the IHCs) than the second and third. Data from 2 males + 3 females + 7 undetermined (936 cells). **m, n** Similar distribution of oc-IHCs in *TgPax2*$^{Cre/+}$; *Insm1*$^{F/F}$. n = 3 animals (234 cells). **j**,**l**,**n** Average ± SEM, one-tailed Student's t-tests, n = number of mice. Images and quantifications from mid-cochlear positions at P34 (**e**, **g**, **i**), P46 (**f**, **h**), P21, P23, P46 (**k**), P14 and P15 (**m**). Controls were *Insm1*$^{F/F}$ (**e**, **g**, **i, m**) or *Atoh1*^{Cre}; *Insm1*$^{F/+}$ (**f**, **j, k**) littermates. Scale bars = 10μm. Similar results obtained from ≥3 animals per genotype.

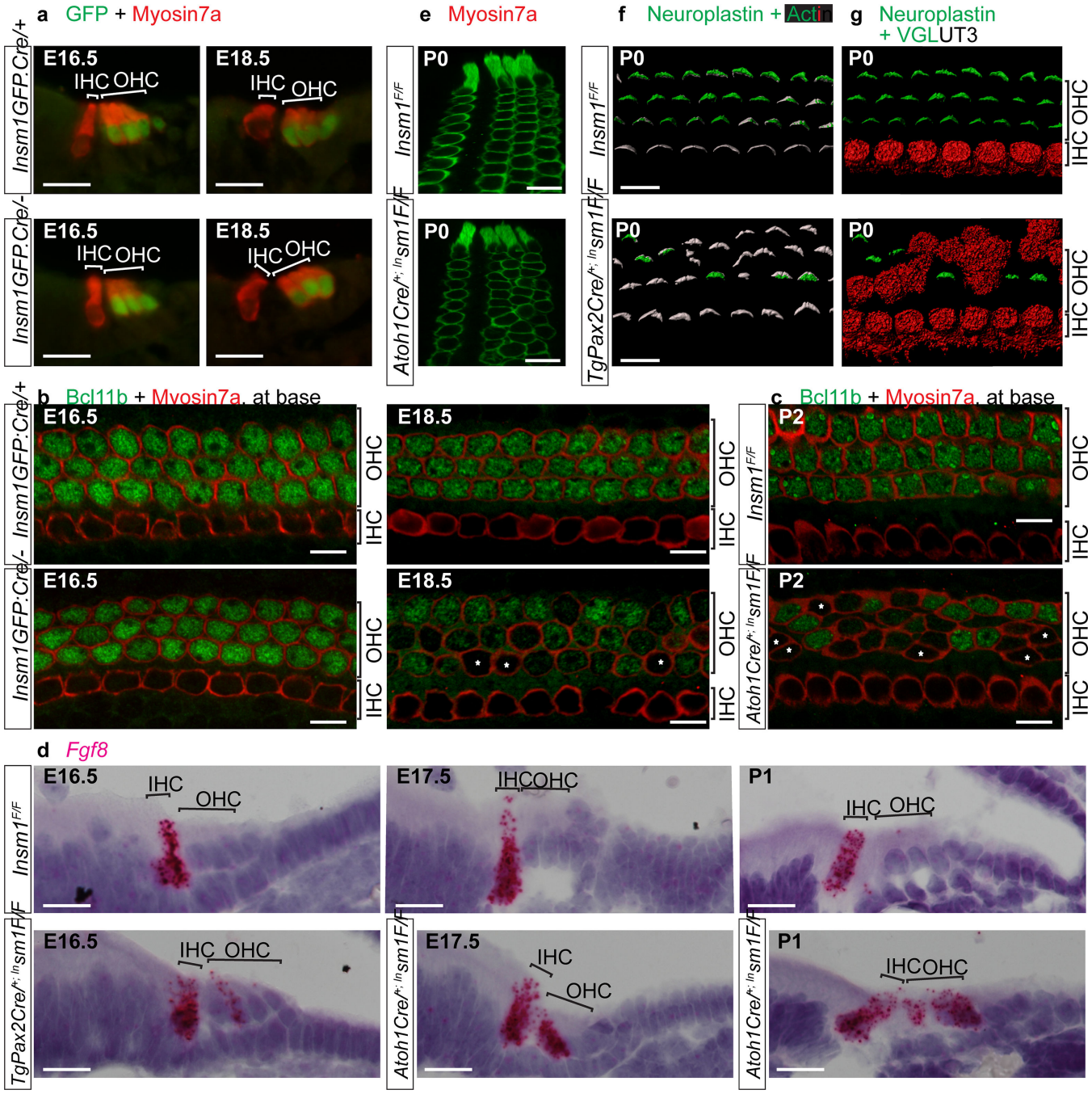


Figure 2: Trans-differentiation of embryonic OHCs into embryonic IHCs in the absence of INSM1.

GFP on embryonic organs of Corti indicated that the *Insm1* promoter (expressing GFP.Cre) is active in oc-HCs (OHCs) at E16.5 and E18.5 regardless of the presence (*Insm1*$^{GFP.Cre/+}$) or absence (*Insm1*$^{GFP.Cre/-}$) of INSM1. **b-c** Immunohistochemistry at E16.5 to P2. While BCL11B, an immature OHC-specific transcription factor, was in all oc-HCs from control *Insm1*$^{GFP.Cre/+}$ and *Insm1*$^{GFP.Cre/-}$ mice at E16.5 (**b**), its expression was diminished or undetectable (asterisks) in about half oc-HCs of mice lacking INSM1 at E18.5 (**b**) and P2

(**c**). **d** In situ hybridization revealed that in the absence of INSM1, a subset of oc-HCs began expressing *fgf8* weakly as early as E16.5 (32.5%, 110/338 OHCs from n=3 $TgPax2^{Cre/+}$;$Insm1^{F/F}$ or $Atoh1^{Cre/+}$; $Insm1^{F/F}$ animals) and strongly by E17.5-E18.5 (40%, 52/130 OHCs from n=2 $Insm1^{GFP.Cre/-}$ or $Atoh1^{Cre/+}$; $Insm1^{F/F}$ animals) and P0-P4 (46.6%, 61/131 OHCs from n=3 $TgPax2^{Cre/+}$;$Insm1^{F/F}$ or $Atoh1^{Cre/+}$; $Insm1^{F/F}$ animals). **e** Immunohistochemistry at P0 revealed that although in $Atoh1^{Cre/+}$; $Insm1^{F/F}$ there was no HC loss and OHCs retained their characteristic inclination, their cell bodies were disorganized at the nuclear level. **f** All OHCs at the base of cochleae from $Insm1^{F/F}$ mice displayed neuroplastin in stereocilia (visualized with phalloidin). However, some OHCs at the base of cochleae from $TgPax2^{Cre/+}$; $Insm1^{Flox/Flox}$ mice lacked neuroplastin (**f**). Conversely, the oc-HCs lacking neuroplastin expressed VGLUT3 (**g**). Images are from mid (**a**-**e**) or basal (**f**,**g**) cochlea. Scale bars = 10μm. Similar results obtained from ≥3 animals per genotype.

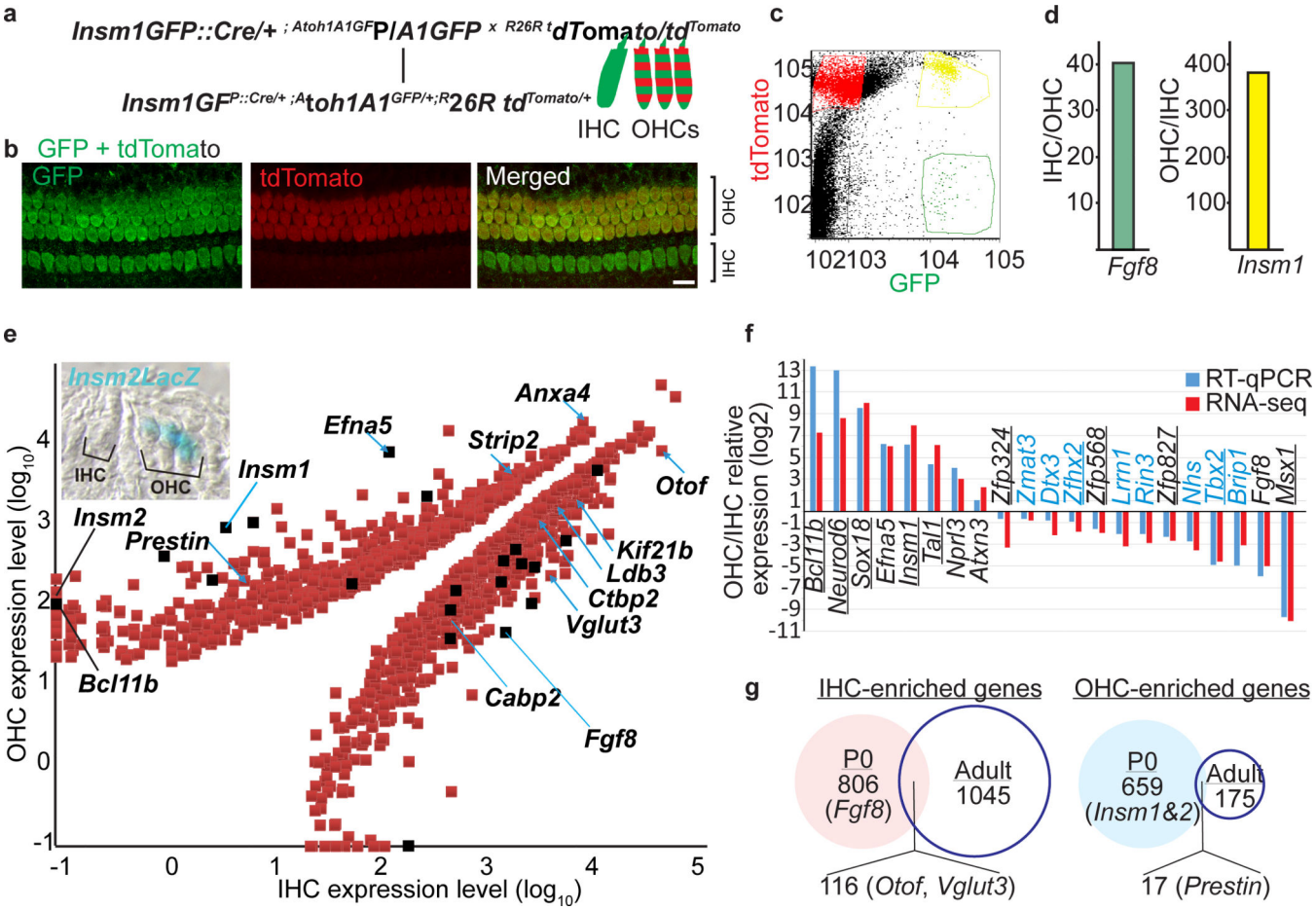


Figure 3: Genes preferentially expressed in immature IHCs or OHCs.

a-d Strategy for collecting separate pools of perinatal (P0) OHCs and IHCs by FACS. **a** Crosses for generating pups in which live IHC and OHCs can be fluorescently distinguished (Green, GFP; Red, tdTomato). **b** *Insm1*$^{GFP.Cre/+}$*; Atoh1*$^{A1GFP/+}$*; R26R*$^{tdTomat0/+}$ mice express GFP in IHCs and GFP plus tdTomato in OHCs. Neurons express tdTomato but GFP from *Insm1*$^{GFP.Cre}$ has subsided. **c** FACS separating IHCs (green) from OHCs (yellow) and neurons (red), done on 6 separate pools of IHCs and OHCs. **d** RT-qPCR for IHC-specific *Fgf8* and OHC-specific *Insm1* confirm that these pools of cells collected are enriched for IHCs or OHCs, respectively. **e** Logarithmic plot of genes preferentially expressed in either IHCs or OHCs based on RNA-seq values. Blue arrows indicate genes previously known to be HC subtype specific in neonates[4,17,25–29]. We additionally confirmed perinatal OHC or IHC-specific expression for *Insm2* with a knock in reporter line (E17.5; top inset). **f** Comparison of fold difference of mRNA expression levels determined (for genes indicated as black boxes in **e**) by RNA-seq and RT-qPCR. Transcription factor-coding genes are underlined. IHC-specific genes inhibited by INSM1 in embryonic OHCs are in blue. **g** Venn diagrams indicating the number of genes enriched in either IHCs or OHCs of neonates vs adults (estimated from[30]). Representative genes are named in parentheses. Although neonatal HCs of either type begin to show expression of some functional markers characteristic of mature cells (*Otoferlin* and *Vglut3* in IHCs and *Prestin* in OHCs), the

majority of HC-type specific genes at this early stage differ from those of the mature cells. Similar results obtained from ≥3 animals per genotype.

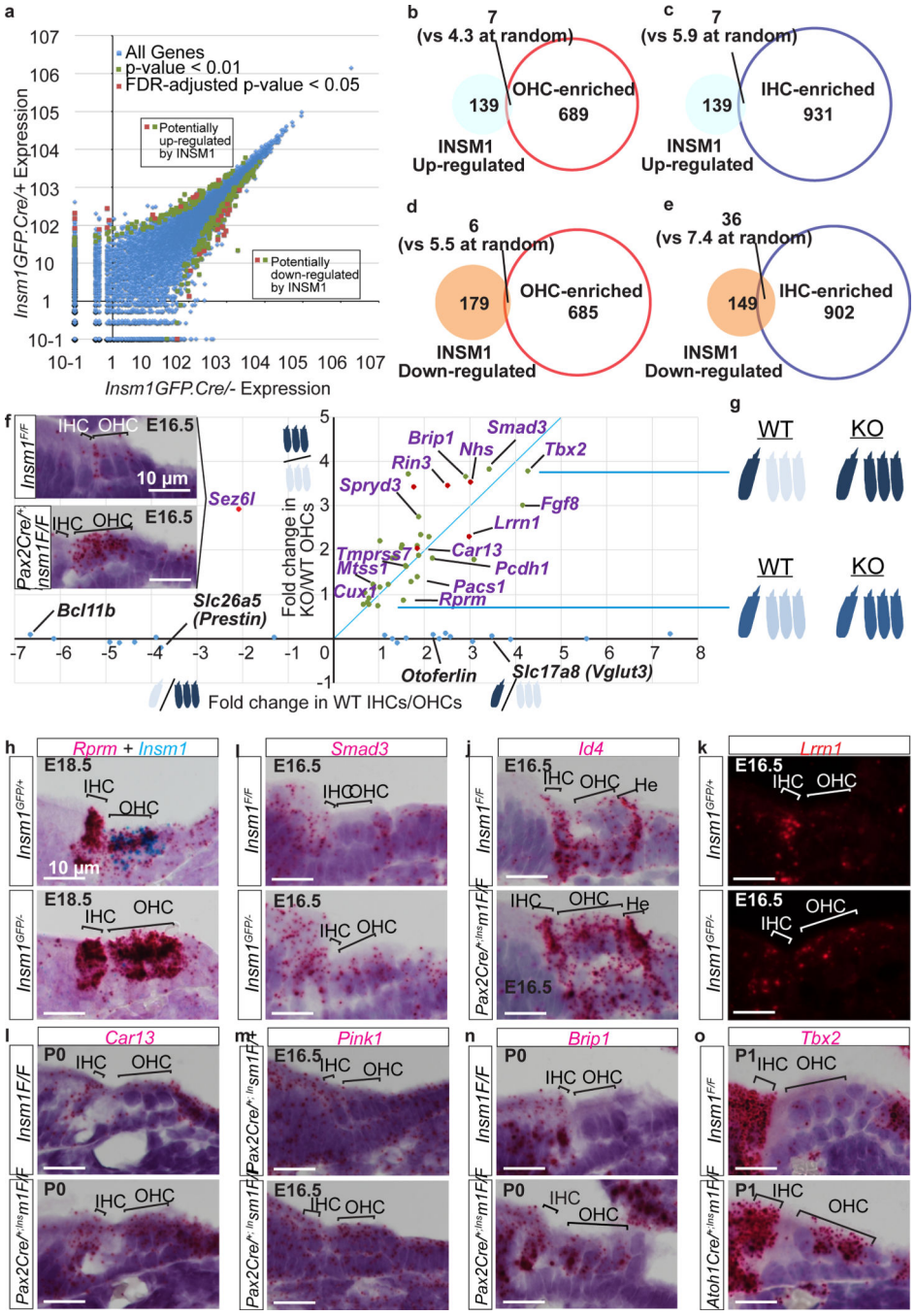


Figure 4: *Insm1* prevents expression of a subset of immature IHC-specific genes in embryonic OHCs.

a Identification of genes potentially regulated by INSM1 in embryonic to perinatal OHCs. Plot of average expression levels determined by RNA-seq from E18.5 OHCs expressing (*Insm1*$^{GFP.Cre/+}$) or lacking INSM1 (*Insm1*$^{GFP.Cre/-}$) (n = 3 pools of OHCs per genotype, each from 8–12 mice). Undetected transcripts are assigned an expression of 10^{-1}. We established as cutoff either an FDR-adjusted p <0.05 (red squares) or, for a less stringent selection, a raw p <0.01 (green squares). Blue diamonds represent all other transcripts. **b-e**

Venn diagrams indicating overlap between IHC or OHC-enriched genes with the genes presumably regulated by INSM1 in OHCs. Up and down-regulated genes are those overexpressed in OHCs with or without INSM1, respectively. Expected number of genes that would appear by random coincidence is in parentheses. Only (**e**) reveals a larger overlap than randomly expected, pointing to 36 IHC-specific genes that appear down-regulated by INSM1 in OHCs. **f** Plot of the differential expression of genes in OHCs vs IHCs and OHCs with vs without INSM1. Genes are plotted as dots with color-code corresponding to the p-value criteria used in (**a**). Expression levels in OHCs with and without INSM1 (from which KO/WT changes are estimated) are average RNA-seq values of 3 pools of OHCs per genotype. Expression levels in IHCs and OHCs (from which IHCs/OHCs changes are estimated) are average RNAseq values of 6 pools of each cell type. Blue dots along the X-axis are examples of the many genes enriched in either IHCs or OHCs that are not affected by INSM1. Differentially-expressed genes confirmed by RT-qPCR are named in purple. Each gene is upregulated to a similar extent in IHCs (vs OHCs) as in OHCs lacking INSM1 (vs IHCs with INSM1). **g** Graphic interpretation. Darker shades of blue indicate higher expression levels. **h-o** ISH confirms preferential expression in IHCs (and often other cells of the organ of Corti) compared with OHCs, and increased expression in OHCs lacking INSM1. Scale bars = 10μm. Similar results obtained from ≥3 animals per genotype.