

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

Teerawat Wiwatpanit^{1,2,10}, Sarah M. Lorenzen^{1,3,10}, Jorge A. Cantú^{1,9,10}, Chuan Zhi Foo^{1,2}, Ann K. Hogan^{1,2}, Freddie Márquez¹, John C. Clancy¹, Matthew J. Schipma⁴, Mary Ann Cheatham^{5,6}, Anne Duggan^{1,11*} & Jaime García-Añoveros^{1,6,7,8,11*}

The mammalian cochlea contains two types of mechanosensory hair cell that have different and critical functions in hearing. Inner hair cells (IHCs), which have an elaborate presynaptic apparatus, signal to cochlear neurons and communicate sound information to the brain. Outer hair cells (OHCs) mechanically amplify sound-induced vibrations, providing enhanced sensitivity to sound and sharp tuning. Cochlear hair cells are solely generated during development, and hair cell death—most often of OHCs—is the most common cause of deafness. OHCs and IHCs, together with supporting cells, originate in embryos from the prosensory region of the otocyst, but

how hair cells differentiate into two different types is unknown^{1–3}. Here we show that *Insm1*, which encodes a zinc finger protein that is transiently expressed in nascent OHCs, consolidates their fate by preventing trans-differentiation into IHCs. In the absence of INSM1, many hair cells that are born as OHCs switch fates to become mature IHCs. To identify the genetic mechanisms by which *Insm1* operates, we compared the transcriptomes of immature IHCs and OHCs, and of OHCs with and without INSM1. In OHCs that lack INSM1, a set of genes is upregulated, most of which are normally preferentially expressed by IHCs. The homeotic cell transformation of OHCs without INSM1 into IHCs reveals a mechanism by which these neighbouring 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 in which these few IHC-enriched transcripts are upregulated trans-differentiate into IHCs, identifying candidate genes for IHC-specific differentiation.

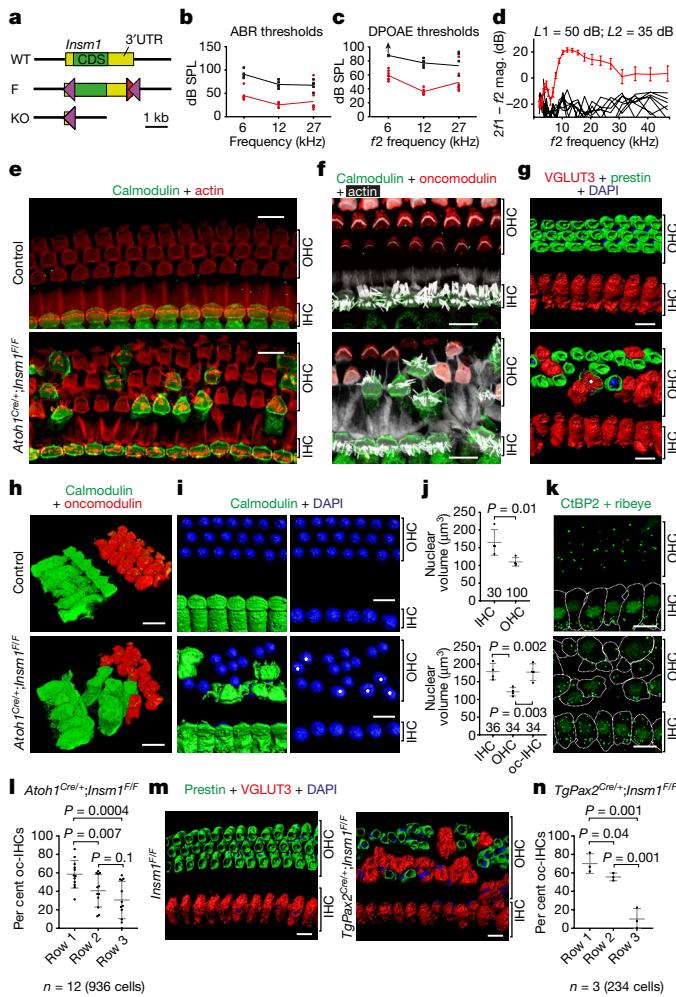


Fig. 1 | Conditional ablation of *Insm1* in hair cells results in IHC-like cells in place of OHCs. **a**, Wild-type (WT), floxed (F) and conditionally deleted (cKO) alleles of *Insm1*. Co-expression with Cre recombinase generates an *Insm1* knockout allele lacking its coding sequence (CDS) and 3' untranslated region (UTR), leaving only part of the 5'UTR. Purple triangles, *LoxP* sites; red triangles, *Fr*t sites. **b–d**, Hearing tests (b, c, mean with all values; d, mean \pm s.e.m.). ABR thresholds (b), DPOAE thresholds (c), and iso-input functions (d) for the DPOAE of *Atoh1*^{Cre/+}; *Insm1*^{F/F} mice at P25–31 (black; $n = 3$ males and 2 females) and control littermates (red; $n = 4$ *Insm1*^{F/F} and 4 *Atoh1*^{Cre/+}; *Insm1*^{F/F}; 6 males and 2 females). Up arrow (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 stereociliary bundles resembling those of IHCs (f-actin labelled with phalloidin; e, f), expressed VGLUT3 instead of prestin (g; asterisk, one rare cell expressed both), had cell shapes (h) and large nuclei like IHCs (i, j, asterisks). **j**, Number of nuclei measured for each cell type are indicated under each data set. **k**, oc-IHCs had nuclear CtBP2 and a number of presynaptic ribbons (ribeye) approaching that of IHCs. **l**, In *Atoh1*^{Cre/+}; *Insm1*^{F/F} mice, oc-IHCs are more frequent in the first row of OHCs (closer to the IHCs) than the second or third rows. Data from two males, three females and seven undetermined (936 cells). **m, n**, Similar distribution of oc-IHCs in *TgPax2*^{Cre/+}; *Insm1*^{F/F} mice. $n = 3$ mice (234 cells). **j, l, n**, Mean \pm s.e.m., 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 and P46 (k), and P14 and P15 (m). Controls were *Insm1*^{F/F} (e, g, i, m) or *Atoh1*^{Cre/+}; *Insm1*^{F/F} (f, j, k) littermates. Scale bars, 10 μm . Biological replicates were used for all experiments and similar results were obtained from three or more mice per genotype.

¹Department of Anesthesiology, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. ²Driskill Graduate Program in Life Sciences, Northwestern University, Chicago, IL, USA.

³NUNI Graduate Program, Northwestern University, Chicago, IL, USA. ⁴Next Generation Sequencing Core, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. ⁵Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA. ⁶Hugh Knowles Center for Clinical and Basic Sciences in Hearing and Its Disorders, Northwestern University, Chicago, IL, USA. ⁷Department of Neurology, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. ⁸Department of Physiology, Northwestern University, Feinberg School of Medicine, Chicago, IL, USA. ⁹Present address: Department of Biology, Northeastern Illinois University, Chicago, IL, USA. ¹⁰These authors contributed equally: Teerawat Wiwatpanit, Sarah M. Lorenzen, Jorge A. Cantú. ¹¹These authors jointly supervised this work: Anne Duggan, Jaime García-Añoveros. *e-mail: a-duggan@northwestern.edu; anoveros@northwestern.edu

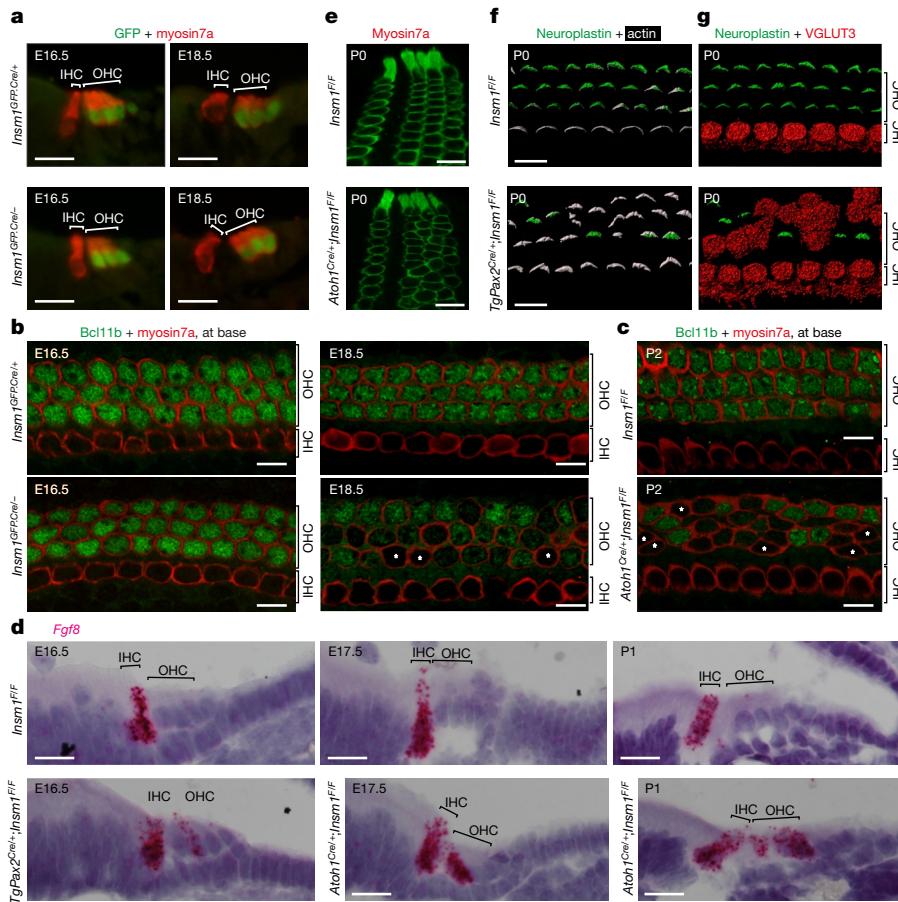


Fig. 2 | Trans-differentiation of embryonic OHCs into embryonic IHCs in the absence of INSM1. **a**, GFP on embryonic organs of Corti indicates that the *Insm1* promoter (expressing GFP.Cre) was active in oc-HCs (OHCs) at E16.5 and E18.5, regardless of the presence (*Insm1*^{GFP;Cre+}) or absence (*Insm1*^{GFP;Cre-}) of INSM1. **b, c**, Immunohistochemistry at E16.5–P2. While Bcl11b, an immature OHC-specific transcription factor, was found 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 of oc-HCs in 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 to express *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} mice) and strongly by E17.5–E18.5 (40%, 52/130 OHCs from $n = 2$

Insm1^{GFP;Cre-} or *Atoh1*^{Cre/+;Insm1}^{F/F} mice) and P0–P4 (46.6%, 61/131 OHCs from $n = 3$ *TgPax2*^{Cre/+;Insm1}^{F/F} or *Atoh1*^{Cre/+;Insm1}^{F/F} mice). **e**, Immunohistochemistry at P0 revealed that although in *Atoh1*^{Cre/+;Insm1}^{F/F} mice 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 bases of cochleae from *Insm1*^{F/F} mice had neuroplastin in stereocilia (visualized with phalloidin). However, some OHCs at the base of cochleae from *TgPax2*^{Cre/+;Insm1}^{F/F} mice lacked neuroplastin. **g**, Conversely, the oc-HCs that lacked neuroplastin expressed VGLUT3. Images are from mid (**a–e**) or basal (**f, g**) cochleae. Scale bars, 10 μ m. Biological replicates were used for all experiments and similar results obtained from three or more mice per genotype.

OHCs express *Insm1* transiently from the onset of differentiation (embryonic day 15.5 (E15.5)) to approximately postnatal day 2 (P2). Neuronal progenitors and nascent spiral ganglion neurons also express *Insm1*⁴. Because mice in which *Insm1* is completely knocked out die embryonically by E19.5^{5,6}, we generated an allele (*Insm1*^F) in which the entire coding sequence can be deleted (Fig. 1a, Extended Data Fig. 1). We conditionally ablated *Insm1*^F with *Atoh1*^{Cre}, expressed from E13.5 (three days before *Insm1*) and recombining in most cochlear hair cells and some supporting cells, but not in spiral ganglion neurons⁷. We also ablated *Insm1*^F with *TgPax2*^{Cre}, expressed earlier in the otocyst and recombining in most inner ear cells⁸. In these mice, *Insm1* was ablated before its expression in OHCs (Extended Data Fig. 2). Both *Atoh1*^{Cre/+;Insm1}^{F/F} and *TgPax2*^{Cre/+;Insm1}^{F/F} (cKO) mice displayed alterations in auditory brainstem response (ABR) thresholds that can be accounted for by shifts in distortion product otoacoustic emissions (DPOAEs), a characteristic of OHC dysfunction (Fig. 1b–d, Extended Data Fig. 3a, b). In the organs of Corti of these mice, many cells in the positions of OHCs (the outer compartment) had features of IHCs. They had large stereocilia, like IHCs, and not the shorter, W-arranged stereocilia of OHCs (Fig. 1e, f); expressed the IHC-enriched calcium buffer calmodulin and lacked OHC-specific oncomodulin (Fig. 1e, f, h,

Extended Data Fig. 3f); expressed the vesicular glutamate transporter 3 (VGLUT3), which is required for IHC presynaptic function, and lacked prestin, which is required for OHC electromotility (Fig. 1g, m); had the flask shape of IHCs rather than the cylindrical shape of OHCs; and had large nuclei, like IHCs, instead of the smaller nuclei of OHCs (Fig. 1j, Extended Data Fig. 3j). These nuclei harboured the transcription factor CtBP2, which is normally expressed in IHCs (Fig. 1k), and the cells contained a number of presynaptic ribbon synapses (10.6 ± 2.1 (mean \pm s.d.), $n = 3$ mice, 39 cells) closer to that found in control IHCs (16.3 ± 0.7 , $n = 3$ littermate controls, 30 cells), instead of the smaller number found in OHCs (1.8 ± 0.2 , $n = 3$ 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 proportion of oc-IHCs in *Atoh1*^{Cre/+;Insm1}^{F/F} mice ($42.6 \pm 10.9\%$, $n = 12$ mice) and *TgPax2*^{Cre/+;Insm1}^{F/F} mice ($46.0 \pm 5.64\%$, $n = 3$ 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 Data Fig. 2a, b, bottom). Notably, the oc-IHCs were more prevalent in the first hair

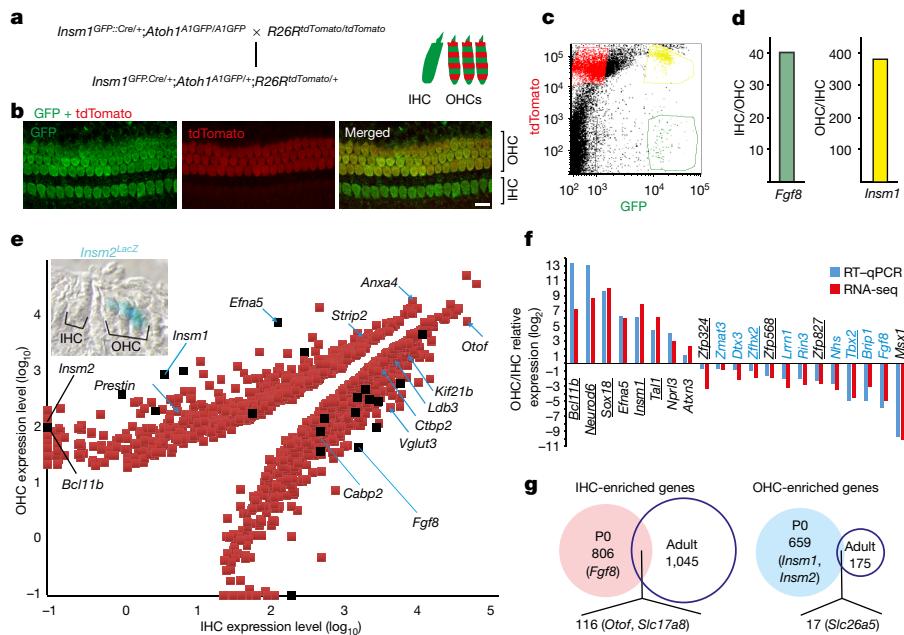


Fig. 3 | Genes preferentially expressed in immature IHCs or OHCs. **a–d**, Strategy for collecting separate pools of perinatal (P0) OHCs and IHCs by FACS. **a**, Crosses for generating pups in which live IHC and OHCs can be fluorescently distinguished (green, GFP; red, tdTomato). **b**, *Insm1*^{GFP,Cre/+}; *Atoh1*^{A1GFP/+}; *R26R*^{tdTomato/+} 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 six separate pools of IHCs and OHCs. **d**, RT-qPCR for IHC-specific *Fgf8* and OHC-specific *Insm1* confirm that these pools of cells 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 of *Insm2* with a knock-in reporter line (E17.5; top inset).

cell row of the outer compartment than in the second or third rows (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 that had 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 Data Figs. 3c, 4a). Although during normal development cochlear hair cells are all born during embryogenesis (E12–E16)^{9,10}, early hair cell death can trigger the generation of hair cells from proliferating and trans-differentiating supporting cells in the first few days after birth^{11–15}. This does not occur in the absence of INSM1, as hair cell density in the outer compartment (OHCs + oc-IHCs) was unaltered up to P34 (Extended Data Figs. 3d, e, 4b, c), whereas oc-IHCs were present well before that. Second, hair cells that are derived postnatally from supporting cells initially express SOX2^{12–14}, whereas none of the oc-IHCs of cKO pups expressed SOX2 (Extended Data Fig. 4d). Third, some postnatally produced HCs result from proliferation of supporting cells^{11–14}, but none of the oc-IHCs in cKO mice derived from postnatal proliferation (Extended Data Fig. 4e). These results show that oc-IHCs do not result from OHC death followed by replacement from displaced IHCs or postnatally generated hair cells. Instead, the oc-IHCs represent 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 mice with conditional (*TgPax2*^{Cre/+}; *Insm1*^{F/F} and *Atoh1*^{Cre/+}; *Insm1*^{F/F}) or complete (*Insm1*^{GFP,Cre–}) *Insm1* knockout 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 but express GFP from the *Insm1* promoter⁴; Fig. 2a), and BCL11B in nuclei (Fig. 2b, c). Whereas in control mice, BCL11B expression was maintained past birth, in embryos lacking INSM1 it subsided in nearly half of outer compartment hair cells (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 stereocilia¹⁶, and VGLUT3 in IHCs¹⁷. By comparison, in both cKO mice, many oc-HCs expressed VGLUT3 and not neuroplastin (Fig. 2f, g). Finally, although the orientation of IHCs and OHCs is maintained at birth in cKO mice, 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 thereafter some of these cells lose these features, 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.

Brief expression of *Insm1* is sufficient to evade phenotypic conversion (Extended Data Fig. 5). It appears that *Insm1* locks the OHC fate during a narrow developmental period. Curiously, although KO OHCs completely lack *Insm1* from their birth, fewer than half of these cells trans-differentiate into IHCs. This trans-differentiation in *TgPax2*^{Cre/+}; *Insm1*^{F/F} and *Atoh1*^{Cre/+}; *Insm1*^{F/F} mice is more frequent in hair cell rows closer to the IHCs than in those further away (Fig. 1l, n). This distribution reveals the existence of a gradient in the neural to abneural axis of the organ of Corti that regulates cochlear hair cell types. This gradient might induce IHC differentiation, and

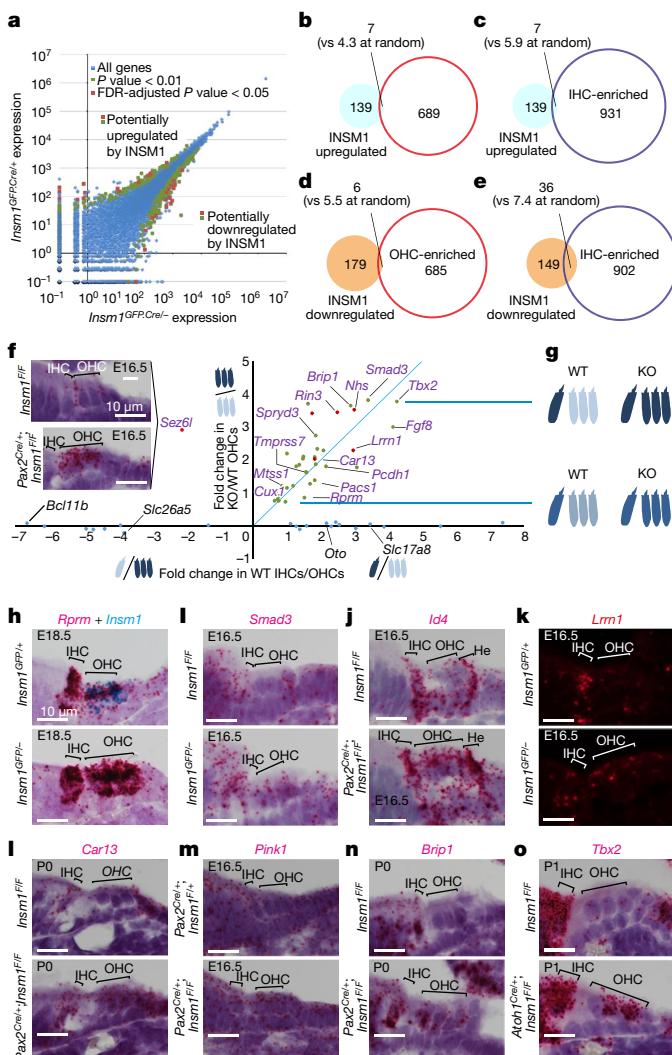


Fig. 4 | *Insm1* prevents expression of a subset of immature IHC-specific genes in embryonic OHCs. a, Identification of genes potentially regulated by INSM1 in embryonic to perinatal OHCs. Plot of average expression levels determined by RNA-seq from E18.5 OHCs expressing (*Insm1*^{GFP,Cre/+}) or lacking (*Insm1*^{GFP,Cre/-}) INSM1 ($n = 3$ 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 genes that are presumably regulated by INSM1 in OHCs. Upregulated and downregulated genes are those overexpressed in OHCs with or without INSM1, respectively. The expected number of genes that would appear by random coincidence is in parentheses. Only e shows a larger overlap than randomly expected, pointing to 36 IHC-specific genes that appear to be downregulated by INSM1 in OHCs. **f**, Plot of the differential expression of genes in OHCs versus IHCs and OHCs with INSM1 versus OHCs without INSM1. Genes are plotted as dots with colours 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 three pools of OHCs per genotype. Expression levels in IHCs and OHCs (from which IHCs/OHCs changes are estimated) are average RNA-seq values of six 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 labelled in purple. Each gene is upregulated to a similar extent in IHCs (versus OHCs) as in OHCs lacking INSM1 (versus IHCs with INSM1). **g**, Graphic interpretation. Darker shades of blue indicate higher expression. **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, 10 μ m. Biological replicates were used for all experiments and similar results obtained from three or more mice per genotype.

INSM1 could act 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 were expressed in either differentiating hair cell type when *Insm1* was 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*^{AIGFP/+}; *R26R*^{tdTomato/+} mice, in which all hair cells express GFP (starting at E13.5 from *Atoh1*^{AIGFP} and, in OHCs, from *Insm1*^{GFP,Cre}) but only OHCs also express tdTomato following *Insm1*^{GFP,Cre} expression (throughout the cochlea by E18.5⁴; Fig. 3a, b). We used these mice to sort OHCs and IHCs from neonatal (P0, approximately E19.5) organs of Corti (Fig. 3c). Using fluorescence-activated cell sorting (FACS), we collected pools of RNA from IHCs and OHCs (Fig. 3d), and then used RNA sequencing (RNA-seq) to obtain 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 were 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. One concern was whether genes that showed small differences in expression (twofold or less; for example, *Zmat3*), or that were detected at very low levels in one cell type only (for example, *Sox18* and *Msx1*), were truly differentially expressed. We selected 21 transcripts (Fig. 3e), and used quantitative PCR with reverse transcription (RT-qPCR) to test for differential expression using additional pools of RNA from IHCs and OHCs. 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 artefacts of cell sorting and mRNA extraction: *Bcl11b* in OHCs by immunohistochemistry (Fig. 2b); *Insm2* in OHCs using 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). These results attest to the low prevalence of false positives among the genes we estimated as being differentially expressed between immature IHCs and OHCs.

The transcriptomes of perinatal cochlear hair cells and supporting cells 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 using microarrays, have also been published³⁰, we have obtained transcriptomes of these cell types before maturity, during early differentiation. A comparison of genes expressed in differentiating and mature IHCs and OHCs reveals very little overlap (Fig. 3g and Supplementary Table 4): only 5.9% of IHC-enriched and 2% of OHC-enriched genes are differentially expressed at differentiating and mature stages. These include some genes that are 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 genes that were preferentially expressed in either cell type during differentiation (such as *Insm1*, *Insm2* and *Bcl11b* in OHCs, and *Brip1*, *Car13* and *Fgf8* in IHCs) are not expressed upon maturation and vice versa. Thus 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 investigate how INSM1 drives OHC differentiation, we used RNA-seq to look for genes that were differentially expressed in differentiating OHCs with and without INSM1 (*Insm1*^{GFP,Cre/+} versus *Insm1*^{GFP,Cre/-}) (Fig. 4a, Extended Data Fig. 6, Supplementary Table 5). We identified between 31 and 331 genes that could be differentially expressed (either upregulated or downregulated) by INSM1 (Supplementary Tables 6, 7).

Comparison of these genes with those normally enriched in OHCs or IHCs (Fig. 4b–e), combined with RT–qPCR retesting (Supplementary Table 8) and ISH (Fig. 4f–o), showed that, in OHCs, INSM1 does not activate OHC genes but rather inhibits IHC genes. No upregulated genes were confirmed by RT–qPCR and, of the 22 downregulated 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 Data Table 1). By contrast, most genes that are differentially expressed in OHCs versus IHCs were not affected by INSM1. We conclude that INSM1 downregulates 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 IHCs and none of OHCs (Fig. 1e–k, m, Extended Data Fig. 3f). The small number of early IHC-specific genes (21 of 922, about 2%) that were upregulated in embryonic OHCs lacking INSM1 are likely to represent an early step in the genetic cascade that leads to their complete transformation into IHCs. As oc-HCs expressing these few genes differentiate into IHCs, these genes are likely to be required for IHC differentiation. Hence, in addition to identifying *Insm1* as a critical gene for OHC differentiation, our results also identify candidate genes for regulating the specific differentiation of IHCs. Because all OHCs express *Insm1*, but in its absence fewer than half trans-differentiate into IHCs, we expected two patterns of misexpression by ISH (Fig. 4g–o). Some genes (*Rprm*, *Id4*, *Lrrn1*, *Car13*, *Pink1* and *Brip1*; Fig. 4h–n) were upregulated in all OHCs lacking INSM1, as expected if they were repressed by INSM1. These 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 upregulated only in fewer than half of oc-HCs—presumably those that would trans-differentiate into IHCs. These genes are some of the earliest expressed in IHCs, and are likely to include regulators of IHC differentiation.

Our results reveal homeotic transformation of OHCs into IHCs in the absence of INSM1, identify the genes initially misregulated by ablation of *Insm1*, and provide a genetic mechanism for differentiation of these two cell types: 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.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41586-018-0570-8>.

Received: 8 May 2017; Accepted: 14 August 2018;

Published online 10 October 2018.

1. Puligilla, C. & Kelley, M. W. Building the world's best hearing aid: regulation of cell fate in the cochlea. *Curr. Opin. Genet. Dev.* **19**, 368–373 (2009).
2. Groves, A. K., Zhang, K. D. & Fekete, D. M. The genetics of hair cell development and regeneration. *Annu. Rev. Neurosci.* **36**, 361–381 (2013).
3. Basch, M. L., Brown, R. M. II, Jen, H.-I. & Groves, A. K. Where hearing starts: the development of the mammalian cochlea. *J. Anat.* **228**, 233–254 (2016).
4. Lorenzen, S. M., Duggan, A., Osipovich, A. B., Magnuson, M. A. & García-Añoveros, J. *Insm1* promotes neurogenic proliferation in delaminated otic progenitors. *Mech. Dev.* **138**, 233–245 (2015).
5. Gierl, M. S., 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).
6. Rosenbaum, J. N., 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).

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).
8. Ohyama, T. & Groves, A. K. Generation of Pax2-Cre mice by modification of a Pax2 bacterial artificial chromosome. *Genesis* **38**, 195–199 (2004).
9. Ruben, R. J. Development of the inner ear of the mouse: a radioautographic study of terminal mitoses. *Acta Otolaryngol.* **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).
11. Kelley, M. W., Talreja, D. R. & Corwin, J. T. Replacement of hair cells after laser microbeam irradiation in cultured organs of corti from embryonic and neonatal mice. *J. Neurosci.* **15**, 3013–3026 (1995).
12. Cox, B. C. et al. Spontaneous hair cell regeneration in the neonatal mouse cochlea *in vivo*. *Development* **141**, 816–829 (2014).
13. Bramhall, N. F., Shi, F., Arnold, K., Hochedlinger, K. & Edge, A. S. B. Lgr5-positive supporting cells generate new hair cells in the postnatal cochlea. *Stem Cell Reports* **2**, 311–322 (2014).
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).
15. White, P. M., Doetzlhofer, A., Lee, Y. S., Groves, A. K. & Segil, N. Mammalian cochlear supporting cells can divide and trans-differentiate into hair cells. *Nature* **441**, 984–987 (2006).
16. Mburu, P. et al. Whirlin complexes with p55 at the stereocilia tip during hair cell development. *Proc. Natl. Acad. Sci. USA* **103**, 10973–10978 (2006).
17. Flores, E. N. et al. A non-canonical pathway from cochlea to brain signals tissue-damaging noise. *Curr. Biol.* **25**, 606–612 (2015).
18. Breslin, M. B., Zhu, M., Notkins, A. L. & Lan, M. S. 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).
19. Jia, S., Wildner, H. & Birchmeier, C. *Insm1* controls the differentiation of pulmonary neuroendocrine cells by repressing *Hes1*. *Dev. Biol.* **408**, 90–98 (2015).
20. Lan, M. S. & Breslin, M. B. Structure, expression, and biological function of INSM1 transcription factor in neuroendocrine differentiation. *FASEB J.* **23**, 2024–2033 (2009).
21. Liu, W.-D., Wang, H.-W., Muguira, M., Breslin, M. B. & Lan, M. S. INSM1 functions as a transcriptional repressor of the neuroD/β2 gene through the recruitment of cyclin D1 and histone deacetylases. *Biochem. J.* **397**, 169–177 (2006).
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).
23. Osipovich, A. B. 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).
24. Jia, S. et al. *Insm1* cooperates with *Neurod1* and *Foxa2* to maintain mature pancreatic β-cell function. *EMBO J.* **34**, 1417–1433 (2015).
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).
26. Scheffer, D. I., Shen, J., Corey, D. P. & Chen, Z.-Y. Gene expression by mouse inner ear hair cells during development. *J. Neurosci.* **35**, 6366–6380 (2015).
27. Yang, T. et al. Expression and localization of CaBP Ca²⁺ binding proteins in the mouse cochlea. *PLoS One* **11**, e0147495 (2016).
28. Jacques, B. E., Montcouquiol, M. E., Layman, E. M., Lewandoski, M. & Kelley, M. W. *Fgf8* induces pillar cell fate and regulates cellular patterning in the mammalian cochlea. *Development* **134**, 3021–3029 (2007).
29. Pirvola, U. et al. FGFR1 is required for the development of the auditory sensory epithelium. *Neuron* **35**, 671–680 (2002).
30. Liu, H. et al. Characterization of transcriptomes of cochlear inner and outer hair cells. *J. Neurosci.* **34**, 11085–11095 (2014).

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

Reviewer information *Nature* thanks S. Heller and the other anonymous reviewer(s) for their contribution to the peer review of this work.

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. analysed data. J.G.-A. and T.W. wrote the manuscript.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41586-018-0570-8>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-018-0570-8>.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Correspondence and requests for materials should be addressed to A.D. and J.G.-A. **Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

Ethics and animals. 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). *Mus musculus* were used in this study. Animals were of CD1 and C57BL/6 genetic background. Similar numbers of male and female animals were used for all analyses, and these are stated in the manuscript. When sex was undetermined, as in the case of embryos, all the embryos were blindly selected. For embryonic samples, we used animals at E16.5, E17.5 and E18.5. For neonatal animals, the average age was 9.27 ± 5.6 (s.d.) days old. For adult animals, the average age was 34.32 ± 6.3 days old. All tests were performed in mutant and littermate control mice, which were otherwise randomly chosen from the litters. Although the phenotype is obvious, quantifications of nuclear size, ribbon synapse densities, cell densities (Fig. 1, Extended Data Figs. 3, 4) and hearing tests (Fig. 1, Extended Data Figs. 3, 5) were performed without knowledge of genotype. Transcriptomic analyses were carried out by a biostatistician from a core facility with no knowledge of genotype.

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' UTRs, a 2,790-bp 5' homologous sequence and a 4,098-bp 3' homologous sequence, was subcloned into the pL253 vector (IA1-pL253) using recombineering, as described previously³¹. 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: TCTTAGATTCTGCCCTTCTGACAG; CKO: CCAAGGAGATGACCACCGCATAG; and R2: CTCTTGAGGGCCTCTGTG, we performed a PCR to identify recombinant clones. Conditions for Thermal-cycler were: step 1: 98°C, 3:00 min; step 2: 98°C, 0:10 min; step 3: 65°C, 0:30 min, -1°C per cycle; step 4: 72°C, 6:45 min; repeat step 2, 10×; step 5: 98°C, 0:10 min; step 6: 60°C, 0:30 min; step 7: 72°C, 6:45 min, repeat step 5, 25×; step 8: 72°C, 10:00 min; step 9: 4°C. Expected sizes for wild-type allele using primers WT to R2 was 6,163 bp. Expected size for recombinant clones using CKO-Reverse was 6,145 bp. We screened a total of 439 clones and identified 5 recombinants.

We further screened these five ES cell clones for recombination upstream of the 5' *LoxP* site. DNA from selected recombinant clones was digested with the restriction enzyme SpeI (NEB Cat:R0133) and homologous recombination was confirmed by Southern hybridization. DNA was visualized using a 168-bp radio-labelled probe (as described⁶). The expected band sizes for wild-type and conditional knockout alleles are 18,162 bp and 14,333 bp, respectively. All five ES clones contained a targeted allele of *Insm1*.

These clones were used for the generation of mosaic embryos, which were implanted into surrogate mothers by the Northwestern Transgenic and Targeted Mutagenesis Laboratory (Chicago, IL). From one of these clones (B3) we first generated chimeric mice, which were mated to mice expressing the FlpE recombinase (B6-Tg(CAG-FLPe)36³² to delete 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–P31 were anaesthetized with ketamine and xylazine (120 mg/kg and 10 mg/kg, respectively, intraperitoneal (IP)) and their body temperature maintained using a heating blanket. In order to assay OHC function, DPOAEs were recorded using a custom probe equipped with a sensitive microphone (Knowles Electronics, FG-3652-CX). Responses were analysed 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 ($f_2/f_1 = 1.2$) at $L_1 = 50$ and $L_2 = 35$ dB were recorded for f_2 frequencies between 2 and 47 kHz, thereby covering most of the mouse audiogram. Input–output functions were also acquired for various f_2 frequencies (6, 12, 27 kHz), where $L_1 = L_2 + 10$ dB. Thresholds for $2f_1-f_2$ were then calculated and represent the level of f_1 that produces a DPOAE of 0 dB. After emission testing, neural responses were measured by collecting ABRs using tone-burst stimuli. The 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 coupler³⁵. Further details were provided in a previous publication³⁶.

Tissue collection and preparation. Neonatal mice were killed by decapitation, and cochleae dissected in cold HBSS with calcium and magnesium (Gibco). For embryos, timed pregnant dams were killed 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); embryos were then collected and their cochleae removed. 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 h at room temperature. For older tissues (>P20), mice were cardiac-perfused with 4% paraformaldehyde, and cochleae were dissected and post-fixed in 4% paraformaldehyde for 2 h 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 method³⁷. Frozen sections were processed as described⁴.

Immunohistochemistry. Whole-mount organ of Corti sections were processed for immunohistochemistry as described previously³⁸. Primary antibodies were mouse anti-calmodulin (1:100, C-7055, Sigma Aldridge), goat anti-oncomodulin (1:200, sc-7446, Santa Cruz), rabbit anti-prestin (1:1,000, from J. Zheng, Northwestern University), guinea pig anti-VGLUT3 (1:2,500, from R. 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), and 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 min at 92°C and cooling for 30 min at room temperature before 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 min, put in -80°C for 5 min and thawed at room temperature for 20 min; sucrose was rinsed off with PBS before blocking and incubation with primary antibodies at 37°C overnight. Nuclei were counterstained with 1:1,000 DAPI or 1:2,000 Hoechst 33342.

X-gal staining. X-gal staining to detect β-galactosidase expression on sections of *Insm2*^{LacZ} embryos was performed as described³⁹.

Cell proliferation assay. In order to label hair and supporting cells generated from progenitors that proliferated postnatally, neonates were injected twice daily from P0 to P5 with the thymidine analogue 5-ethynyl-2'-deoxyuridine (EdU; 50 mg/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. 3D renderings were generated using NIS Elements AR4.60.00 (Nikon) and Imaris X64 8.4.1 (BitPlane) software. Nuclei and ribbon synapses were measured using built-in analysis functions on Imaris. After 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.

FACS. For collection of 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 0.33 U/ml papain, 0.5 mM EDTA and 1 mM L-cysteine in EBSS for 10 min 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/s). 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 a BioAnalyzer 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 hair cells collected from E18.5 *Insm1*^{GFP,Cre−}, E18.5 *Insm1*^{GFP,Cre+}, and P0 *Insm1*^{GFP,Cre+}; *Atoh1*^{A1GFP+}; *R26R*^{tdTomato+} mice through FACS. RNA was extracted using an 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 ~3,000 pg total RNA from each hair cell 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 ~200 pg of first strand cDNA using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) in triplicate on a CFX Connect Real-Time PCR Detection System (Bio-Rad) using a 40-cycle protocol.

For TaqMan qRT-PCR, we used 1 μg 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 1 µl of 100 mM 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 1 µg diluted pre-amplified cDNA (1:20 in TE buffer) per reaction in triplicate on a QuantStudio 7 flex Real-Time system (Applied Biosystems) using a 14-cycle protocol at NUSed 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 hair cells.** To purify enough RNA for deep sequencing and to analyse results statistically in order to determine the IHC and OHC transcriptomes, we collected by FACS six separate pools of IHCs (700–1,100 cells per pool) and six of OHCs (2,800–3,700 cells per pool) from *Insm1*^{GFP,Cre/+}; *Atoh1*^{A1GFP/+;R26R^{tdTomato/+} mice at P0 (generated by timed pregnancies and found to correspond in most cases to E19.5 and in the rest to E20.5). To determine 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 2,200–5,000 OHCs from 8–12 E18.5 embryos. We extracted 3–7.5 ng of RNA per E18.5 OHC pool, ~3 ng per P0 IHC pool and 10–18 ng per P0 OHC pool. We used only 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 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, end reparation and 3'-end single nucleotide A (adenine) addition were performed, and sequencing adaptors were ligated 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 (three libraries from E18.5 *Insm1*^{GFP,Cre/-} and three from *Insm1*^{GFP,Cre/+} OHCs) or the HiSeq4000 (six libraries each for IHCs and OHCs from P0 *Insm1*^{GFP,Cre/+}; *Atoh1*^{A1GFP/+;R26R^{tdTomato/+} mice) and multiplexed per lane, yielding 48–50 million (for each of the six E18.5 OHC libraries) and 92–116 million (for each of the twelve 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.

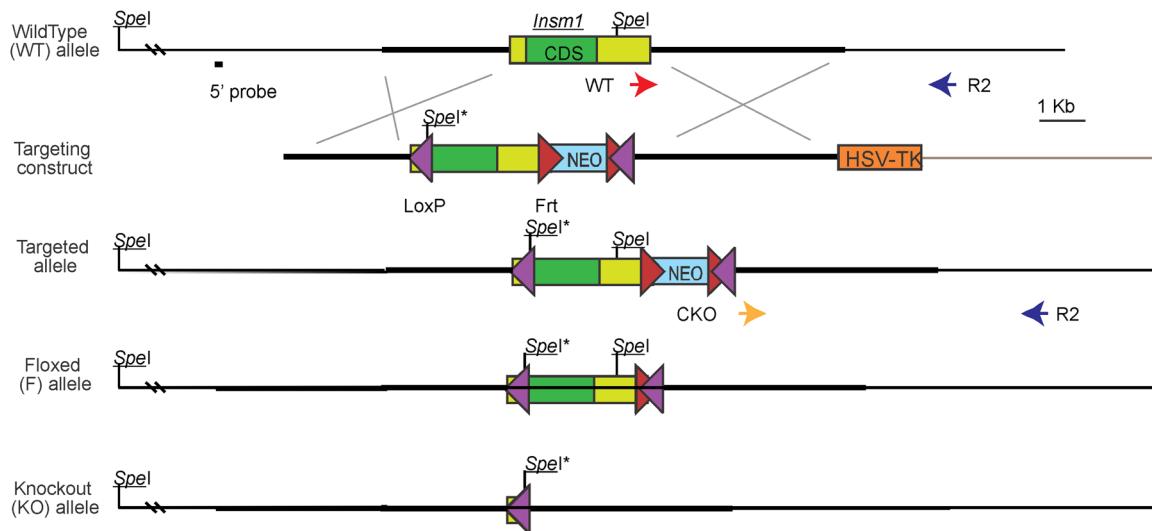
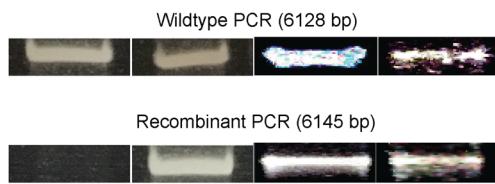
The cleaned reads were aligned to the *Mus musculus* genome (mm10) using STAR⁴⁰. Read counts for each gene were calculated using htseq-count⁴¹ in conjunction with a gene annotation file for mm10 obtained from UCSC (University of California Santa Cruz; <http://genome.ucsc.edu>). Differential expression was determined using DESeq2⁴². The cutoff for determining significantly differentially expressed genes was an FDR-adjusted *P* value less than 0.05.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

All data are available from the corresponding authors upon reasonable request. RNA-seq data are available for public view at the gEAR portal (<https://umgear.org/>).

31. Liu, P., Jenkins, N. A. & Copeland, N. G. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**, 476–484 (2003).
32. Kanki, H., Suzuki, H. & Itohara, S. High-efficiency CAG-FLPe deleter mice in C57BL/6J background. *Exp. Anim.* **55**, 137–141 (2006).
33. Neely, S. T. & Liu, Z. EMAV: Otoacoustic emission averager. Omaha, NE: Technical Memo No 17. *Boy's Town National Research Hospital* (1994).
34. Neely, S. T. & Stevenson, R. SysRes. Omaha, NE: Tech Memo No. 1. *BoysTown National Research Hospital NE* (1992).
35. Pearce, M., Richter, C. P. & Cheatham, M. A. Reconsideration of sound calibration in the mouse. *J. Neurosci. Methods* **106**, 57–67 (2001).
36. Cheatham, M. A. et al. Loss of the tectorial membrane protein CEACAM16 enhances spontaneous, stimulus-frequency, and transiently evoked otoacoustic emissions. *J. Neurosci.* **34**, 10325–10338 (2014).
37. Montgomery, S. C. & Cox, B. C. Whole mount dissection and immunofluorescence of the adult mouse cochlea. *J. Vis. Exp.* **107**, e53561 (2016).
38. Delmaghani, S. et al. Hypervulnerability to sound exposure through impaired adaptive proliferation of peroxisomes. *Cell* **163**, 894–906 (2015).
39. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. Staining frozen mouse embryo sections for β-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).
41. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
42. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
43. 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).

a**b**

ES Cell clone #: B5 B6 E10 B3

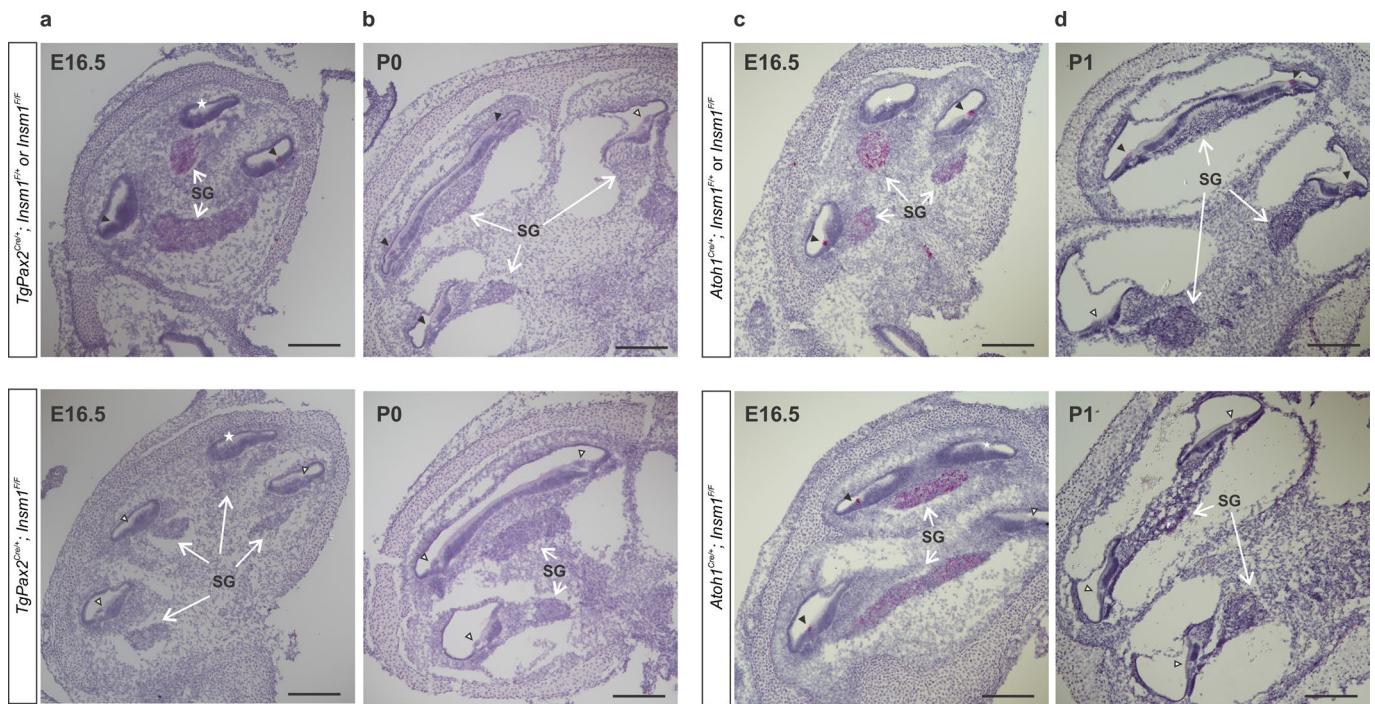
c

ES Cell clone #: B3 E10 WT

Extended Data Fig. 1 | Generation of a conditional KO allele of *Insm1*.

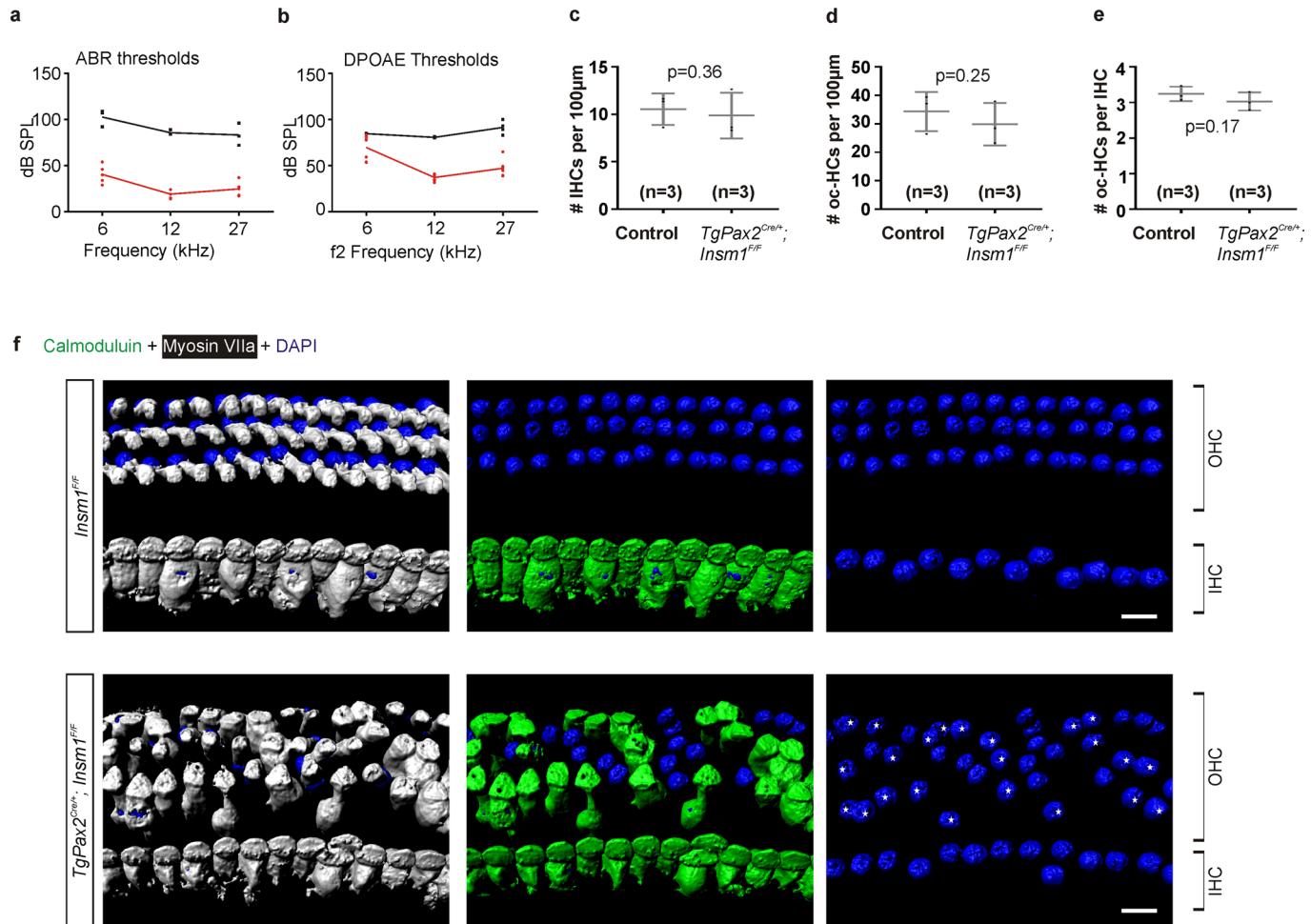
a. We generated a targeting construct in which the sole exon of *Insm1* (green rectangle, with the coding sequence in dark green and the UTRs in light green) has a *loxP* site (purple triangle) inserted in a poorly conserved area of its 5'UTR and another *loxP* site downstream of the *Insm1* gene. The 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 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). The expected size for the wild-type allele using primers WT to R2 is 6,163 bp. The expected size for recombinant clones using CKO-forward and R2 is 6,145 bp.

c. Selected embryonic stem (ES) cell clones were additionally screened for homologous recombination upstream of the first *loxP* site by Southern blotting after digestion with *SpeI* and using the 5' probe indicated in **a**. Southern blotting 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 chimaeras 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 Data Fig. 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), *Insm1* is expressed in all OHCs (72/72 OHCs from 3 animals) and spiral ganglion (SG; white arrows) at E16.5. However, no *Insm1* was detected in the organs of Corti from apical turns, in which recognizable hair cells have not yet appeared (**a, c**; asterisks). By postnatal age P0–P1, *Insm1* mRNA is present in 90% of OHCs (94/105 OHCs from 2 animals), and it is undetectable in spiral ganglion (**b, d**). **a, b**, Bottom, in *TgPax2Cre/+; Insm1F/F* mice, *Insm1* mRNA is undetectable in spiral ganglion 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, in *Atoh1Cre/+; Insm1F/F* cochleae, *Insm1* mRNA is present in spiral ganglion and 43% of OHCs (18/42 OHCs from 1 animal) at E16.5 (**c**), reduced to 7% (4/54) at E17.5 (not shown), and entirely absent all OHCs (0/60 OHCs from 1 animal) and spiral ganglion by postnatal day P1 (**d**). 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 organs of Corti with *Insm1* expression, and empty arrowheads indicate organs of Corti without *Insm1* expression. Scale bars, 200 μ m.

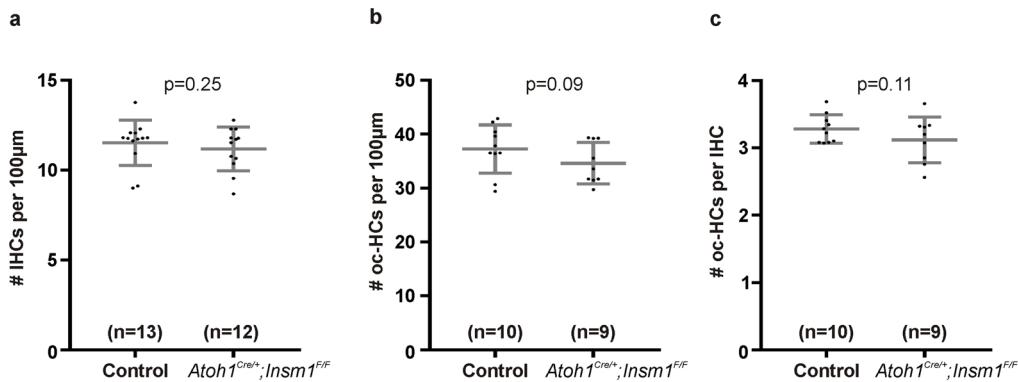
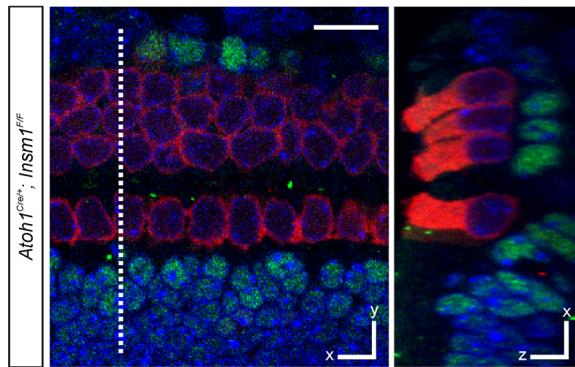
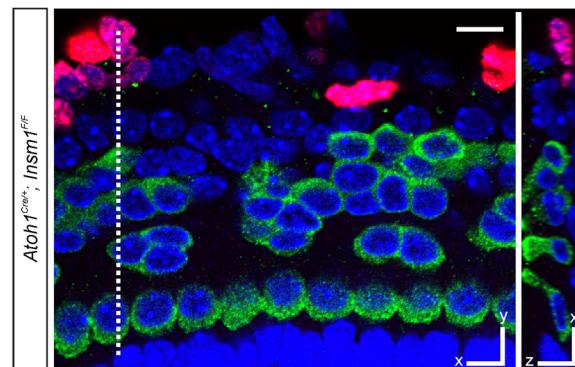


Extended Data Fig. 3 | Conditional ablation of *Insm1* in hair cells and spiral ganglia neurons using *TgPax2*^{Cre} causes hearing impairment and the appearance of IHC-like cells in place of OHCs.

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

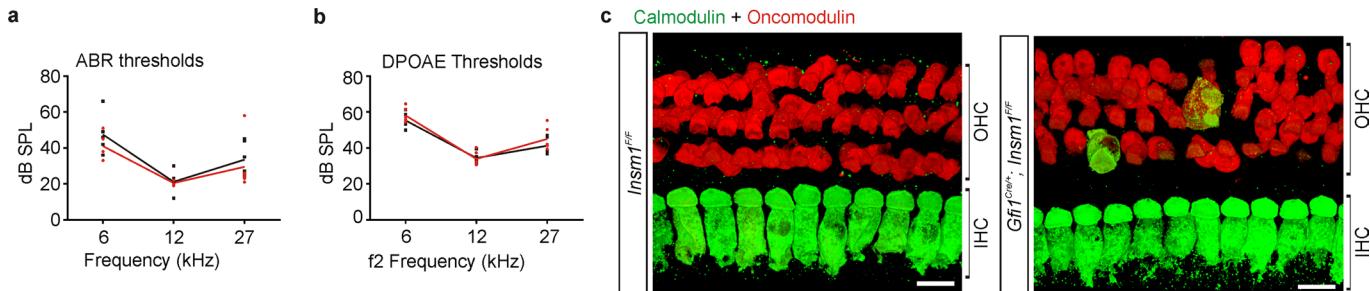
c, Despite the prevalence of OHCs with IHC characteristics (oc-IHCs) in *TgPax2*^{Cre/+}; *Insm1*^{F/F} cochleae ($46.0 \pm 5.64\%$ (mean \pm s.d.), $n = 3$ mice; Fig. 1m, n), these mice have the same density of IHCs (9.87 ± 2.41 cells per $100\mu\text{m}$ along the organ of Corti; $n = 3$) as littermate controls (*TgPax2*^{Cre/+}; *Insm1*^{F/+} and *Insm1*^{F/F}; 10.54 ± 1.67 cells per $100\mu\text{m}$; $n = 3$) 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 \sim P14–P16. Densities of oc-HCs do not differ between *TgPax2*^{Cre/+}; *Insm1*^{F/F} mice (OHCs and oc-IHCs) and their littermate

controls (OHCs only) (29.88 ± 7.45 cells per $100\mu\text{m}$ along the organ of Corti of *TgPax2*^{Cre/+}; *Insm1*^{F/F} mice, $n = 3$; 34.3 ± 6.92 cells per $100\mu\text{m}$ in *TgPax2*^{Cre/+}; *Insm1*^{F/+} and *Insm1*^{F/F} littermate controls, $n = 3$). **e**, The number of oc-HCs per IHC does not differ between *TgPax2*^{Cre/+}; *Insm1*^{F/F} mice and their littermate controls (3.03 ± 0.25 OHCs plus oc-IHCs per IHC in *TgPax2*^{Cre/+}; *Insm1*^{F/F} mice, $n = 3$; 3.24 ± 0.21 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–e**, n is number of mice. Statistical significance is defined as $P < 0.05$. **f**, Immunofluorescence for the IHC-enriched calmodulin (green) and hair cell marker myosin VIIa (white) on whole-mount organs of Corti from mid-cochlear positions at ages \sim P14–P16 confirmed that many *TgPax2*^{Cre/+}; *Insm1*^{F/F} oc-HCs had IHC characteristics, in addition to having the flask shape and large nuclei of IHCs (blue, DAPI, marked with asterisks), as well as lacking prestin and expressing VGLUT3 (Fig. 1m). Scale bars, $10\mu\text{m}$. Biological replicates were used for all experiments and similar immunohistochemistry results obtained from three or more mice per genotype.

**d** Sox2 + Myosin VIIa + DAPI, at P0**e** Myosin VIIa + Edu + Hoechst, at P5

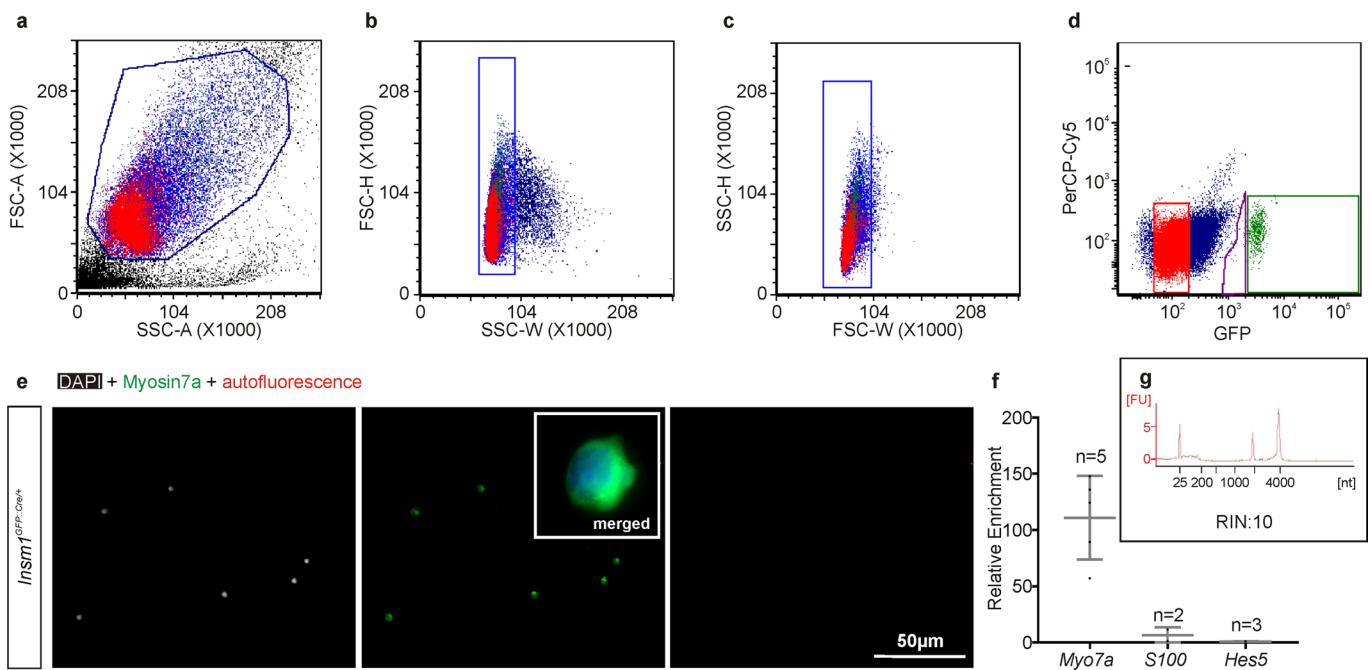
Extended Data Fig. 4 | IHC-like cells in the outer compartment (oc-IHCs) result from OHC misdifferentiation in the absence of INSM1, not from IHC displacement or from trans-differentiation of supporting cells. **a**, *Atoh1^{Cre/+};Insm1^{F/F}* mice have the same density of IHCs (11.2 ± 1.2 (mean \pm s.d.) cells per $100\mu\text{m}$; $n = 12$) as littermate controls (*Atoh1^{Cre/+};Insm1^{+/F}* and *Insm1^{F/F}*; 11.5 ± 1.3 cells per $100\mu\text{m}$; $n = 13$). **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 cells per $100\mu\text{m}$ in *Atoh1^{Cre/+};Insm1^{F/F}* mice, $n = 9$; 37.3 ± 4.5 cells per $100\mu\text{m}$ in *Atoh1^{Cre/+};Insm1^{+/F}* and *Insm1^{F/F}* littermate controls, $n = 10$). **c**, The number of oc-HCs per IHC do not differ significantly between *Atoh1^{Cre/+};Insm1^{F/F}* mice and their littermate controls (3.1 ± 0.3 OHCs and oc-IHCs per IHC in *Atoh1^{Cre/+};Insm1^{F/F}* mice, $n = 9$; 3.3 ± 0.2 OHCs per IHC in *Atoh1^{Cre/+};Insm1^{+/F}* and *Insm1^{F/F}* littermate controls, $n = 10$). The criteria for identification of 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–c**; *n* is number of mice. Statistical 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 analogue 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 oc-HCs at P5 and 0/40 oc-HCs at P8) confirmed that these cells, 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, $10\mu\text{m}$. Biological replicates were used for all experiments and similar immunohistochemistry results obtained from three or more mice per genotype.



Extended Data Fig. 5 | Conditional ablation of *Insm1* in hair cells using *Gfi1*-Cre causes no hearing impairment, and results in only very few oc-IHCs. 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 ABRs (**a**) and DPOAEs (**b**) of *Gfi1*^{Cre/+}; *Insm1*^{F/F} mice at age P30–P35 (black traces; $n = 6$, 4 males and 2 females) and control littermates (red traces; $n = 6$; 1 *Insm1*^{F/F}, 1 *Insm1*^{F/+} and 4 *Gfi1*^{Cre/+}; *Insm1*^{F/+}; 4 males and 2 females). There is no significant difference in ABR and DPOAE thresholds at any frequency 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 levels of calmodulin. However, very few oc-HCs 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–E17.5)⁴³, their nascent OHCs will express *Insm1* for at least several hours. This result indicates that brief expression of *Insm1* is sufficient to promote proper OHC differentiation. Scale bars, 10 μm. Biological replicates were used for all experiments and similar immunohistochemistry results obtained from three or more mice per genotype.



Extended Data Fig. 6 | FACS purification of and RNA extraction from OHCs from an E18.5 *Insm1*^{GFP.Cre/+} embryo. **a–c**, Forward and side light scattering were used to exclude dead cells and debris (a) and aggregates (≥ 2 cells) (b, c). **d**, Live cells were 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 cytopinning after FACS confirm that most of the 547 sorted GFP^+ cells are hair cells. This verification was done on all hair cell pools sorted (three pools per genotype). Inset is a

representative merged image of one sorted OHC at high magnification. In this pool, no autofluorescent cells were collected. **f**, RT-qPCR after cell sorting (mean \pm s.e.m.) reveals that, compared with GFP^- cells, GFP^+ cells express the hair cell marker gene *Myo7a* and not the supporting cell marker genes *S100* and *Hes5*. **g**, To ensure the quality of the extracted RNA, the RIN score was determined using a BioAnalyzer. **g**, Similar RIN scores were obtained from all pools of OHCs examined (including the three per genotype used for RNA-seq in Fig. 4a).

Extended Data Table 1 | Confirmed genes misregulated in OHCs lacking INSM1

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
Lrrn1	leucine rich repeat protein 1, neuronal	7.86	4.93	7.79	0.003956
Brip1	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
Pcdh1	protocadherin 1	4.47	3.53	2.84	0.009309
Spry3	SPRY domain containing 3	3.64	6.70	1.53	0.033057
Pacs1	phosphofuran acid cluster sorting protein 1	3.56	2.63	2.33	0.00229
Car13	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	reproto, 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
Mtss1	metastasis suppressor 1	1.83	2.34	1.56	0.02911
Cux1	cut-like homeobox 1	1.61	2.05	1.57	0.026945
Lrrc8b	leucine rich repeat containing 8 family, member B	1.56	1.71	1.31	0.045587
Rai1	retinoic acid induced 1	1.57	2.56	1.14	0.033262
Pink1	PTEN induced putative kinase 1	1.36	6.52	4.42	0.003253
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

Differential expression of IHCs with 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 RNA-seq 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 RNA-seq 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 RNAscope in situ hybridization (Fig. 4j). Hence, for *Id4* the differential expression occurs transiently and very early.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

Hearing data were collected using System Response (SysRes) (Neely et al., 2005, Tech. Memo. 19, Boys Town, Nat. Res. Hosp.). Images were acquired using either a Nikon A1 or A1R+ Confocal imaging system with NIS Elements AR4.60.00 software. Flow Cytometry data was collected using BD FACSDiva Version 8.0. qRT-PCR data were collected using either a CFX ConnectTM Real-Time PCR Detection System (Bio-rad) or QuantStudioTM 7 flex Real-Time system (Applied Biosystems). RNA-seq data were obtained using either a Illumina HiSeq2000 sequencing system (Illumina) or Illumina HiSeq4000 sequencing system (Illumina).

Data analysis

NIS Elements AR4.60.00 (Nikon)
 Imaris X64 8.4.1 (BitPlane)
 Microsoft Excel (Microsoft)
 BD FACSDiva 8.0 (BD Biosciences)
 Agilent 2100 Bioanalyzer system (Agilent)
 Illumina HiSeq2000 sequencing system (Illumina)
 Illumina HiSeq4000 sequencing system (Illumina)
 FastQC (Babraham Bioinformatics)
 STAR (Dobin et al., 2013, Bioinformatics)
 htseq-count (Anders et al., 2015, Bioinformatics)
 DESeq2 (Love et al., 2014, Genome Biol)
 Emission Averager (EMAV) (Audres)
 System Response (SysRes) (Neely et al., 2005, Tech. Memo. 19, Boys Town, Nat. Res. Hosp.)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data supporting the findings of this current study are available upon reasonable request. In addition, all RNAseq data will be deposited for public view at the gEAR portal (<https://umgear.org/>).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

For statistical analyses, we calculated all analyses based on n = the number of animals used. We used a minimum of 3 animals for each analyses based on the availability of out KO animals.
 For the RNAseq transcriptome comparison of outer hair cells with vs without INSM1 (Fig 4 and supplementary table 5), the sample size was 3 per genotype, the minimum for statistical analyses.
 For the RNAseq transcriptome comparison of outer vs inner hair cells (Fig 3 and supplementary tables 1-3) we were able to collect more cells and hence pooled them into six samples, so as to be able to exclude any outlier. However, all six samples per cell type were analyzed, since we found no outliers.
 For both RNAseq experiments the limitation was to use 3 ng of RNA per pool, which were extracted from ~1000 FACS sorted cells. This limitation was based on recommendation from the company performing the RNAseq, BGI.

Data exclusions

No data were excluded from the statistical analyses in this study.

Replication

All attempts at replication were successful. The sample sizes are indicated in the figure legends and methods.

Randomization

All tests were performed in mutant and litter-mate control mice, which were otherwise randomly chosen from the litters.

Blinding

Quantifications of cell densities (Extended Fig 3 and 4) were made without knowledge of genotype, although the phenotype is obvious. Quantifications of nuclear size (Fig 1) were made by a computer, not by an individual. Hearing tests (Fig 1, Extended figure 3 and 5) were made without knowledge of genotype. Transcriptomic analysis were made by a biostatistician from our core facility with no knowledge of genotype.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

Primary Antibody:

mouse-anti-calmodulin (used at 1:100 dilution, catalog # C3545, Lot # 067M4880V, Sigma)
 goat-anti-oncomodulin (used at 1:200 dilution, catalog # sc-7446, Lot # L0814, Santa Cruz)
 rabbit-anti-prestin (used at 1:1000 dilution, a gift from J Zheng, Northwestern University)
 guinea pig-anti-VGLUT3 (used at 1:2500 dilution, a gift from Robert Edwards, University of California, San Francisco)
 mouse-anti-CtBP2 (used at 1:400 dilution, catalog # 612044, Lot # 7069947, BD Biosciences)
 rabbit-anti-myosin7a (used at 1:800 dilution, catalog # 25-6790, Lot # 091501, Proteus Biosciences)
 sheep-anti-neuroplastin (used at 1:150 dilution, catalog # AF7818, Lot # CHIV0116011, R&D SYSTEMS)
 rat-anti-BCL11B (also known as rat-anti-CTIP2, used at 1:400 dilution, catalog # ab18465, Lot # GR322373-6, abcam)
 goat-anti-SOX2 (used at 1:500 dilution, catalog # sc-17320, Lot # H0516, Santa Cruz)

Secondary Antibody:

AlexaFluor 488 D anti Mouse (used at 1:200 dilution, Catalog # 715-545-150, Lot # 120901, Jackson ImmunoResearch)
 AlexaFluor 488 D anti Guinea Pig (used at 1:200 dilution, Catalog # 706-546-148, Lot # 128097, Jackson ImmunoResearch)
 FITC D anti Sheep (used at 1:200 dilution, Catalog # 713-095-147, Lot # 44672, Jackson ImmunoResearch)
 AlexaFluor 488 D anti Goat (used at 1:200 dilution, Catalog # 705-545-147, Lot # 136089, Jackson ImmunoResearch)
 AlexaFluor 488 D anti Rabbit (used at 1:200 dilution, Catalog # 711-545-152, Lot # 127725, Jackson ImmunoResearch)
 AlexaFluor 594 G anti Mouse (used at 1:200 dilution, Catalog # 115-585-072, Lot # 111296, Jackson ImmunoResearch)
 AlexaFluor 594 D anti Guinea Pig (used at 1:200 dilution, Catalog # 706-586-148, Lot # 128708, Jackson ImmunoResearch)
 Cy3 D anti Goat (used at 1:200 dilution, Catalog # 705-165-147, Lot # 81185, Jackson ImmunoResearch)
 AF594 D anti Rabbit (used at 1:200 dilution, Catalog # 711-585-152, Lot # 113078, Jackson ImmunoResearch)

Validation

We included "no primary antibody control" (no primary antibody added) for all immunohistochemistry stainings described here, and these controls did not exhibit signals above the background fluorescence. Detailed validation of each commercially available antibody is available on the manufacturer's product description. All these antibodies are used as markers of either hair cells, IHCs, OHCs, or supporting cells, which we confirm in each case with control litter-mates (Fig 1e,f,g,h,i,k,m, Fig 2a-c,e-g, Extended Fig 3f, Extended Fig 4d,e, and Extended Fig 5c). In addition, we confirmed lack of immunoreactivity in the inner ears of KO mice for anti-BCL11b (unpublished), anti-VGLUT3 (Flores et al., 2015), and anti-Prestin (Cheatham et al., Audiol Neurorootol. 2007 Jul 27; 12(6):378-390), as others did for anti-neuroplastin (Zeng et al., J Neurosci. 2016 Aug 31; 36(35):9201-9216), anti-SOX2 (kempfle et al 2016, Scientific Reports). The anti-CtBP2 labels in hair cells puncta that correspond to the number of ribbon synapses determined ultrastructurally, and hence are assumed to recognized the ribbon-component CtBP2 (Lberman et al., J Neurosci. 2011 Jan 19; 31(3):801-808).

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

Mus musculus were used in this study. Animals were of CD1 and C57BL/6 genetic background. Similar number of male and female animals were used for all analysis, and there are stated in the manuscript. When sex was undetermined, as in the case of embryos, we used all the embryos blindly with regard to their sex. For embryonic samples, we used animals at E16.5, E17.5 and E18.5 as clearly stated in the manuscript. For neonatal animals, the average age was 9.27 ± 5.6 SD days old. For adult animals, the average age was 34.32 ± 6.3 SD days old.

Wild animals

The study did not involve wild animals.

Field-collected samples

N/A

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

The organ of Corti was dissected out of the inner ear of mouse embryos at E18.5 or neonates at P0, and stored in PBS on ice. The organ was dissociated using 0.33U/ml papain, 0.5mM EDTA and 1mM L-cystine in EBSS at 37°C for 10 min, rinsed in 2% FBS and mechanically dissociated by trituration. Cell suspensions were kept on ice in PBS prior to sorting.

Instrument

BD FACSAria SORP

Software

BD FACSDiva Version 8.0

Cell population abundance

Not done by flow cytometry due to the limited number of cells. Cells were sorted directly into lysis buffer for RNA extraction.

Gating strategy

FSC and SSC were used to remove doublets. PerCP was used to eliminate autofluorescent cells. Cells were then sorted for a bright tdTomato and/or GFP signals.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.