

Helios is a key transcriptional regulator of outer hair cell maturation

Lauren Chessim^{1,11}, Maggie S. Matern^{2,11}, Michael C. Kelly³, Stuart L. Johnson⁴, Yoko Ogawa², Beatrice Milon², Mark McMurray², Elizabeth C. Driver³, Andrew Parker¹, Yang Song⁵, Gemma Codner⁶, Christopher T. Esapa¹, Jack Prescott¹, Graham Trent², Sara Wells⁶, Abigail K. Dragich⁷, Gregory I. Frolenkov⁷, Matthew W. Kelley³, Walter Marcotti⁴, Steve D. M. Brown¹, Ran Elkon^{8,9}, Michael R. Bowl^{1,12*} & Ronna Hertzano^{2,5,10,12*}

The sensory cells that are responsible for hearing include the cochlear inner hair cells (IHCs) and outer hair cells (OHCs), with the OHCs being necessary for sound sensitivity and tuning¹. Both cell types are thought to arise from common progenitors; however, our understanding of the factors that control the fate of IHCs and OHCs remains limited. Here we identify *Ikzf2* (which encodes Helios) as an essential transcription factor in mice that is required for OHC functional maturation and hearing. Helios is expressed in postnatal mouse OHCs, and in the *cello* mouse model a point mutation in *Ikzf2* causes early-onset sensorineural hearing loss. *Ikzf2*^{cello/cello} OHCs have greatly reduced prestin-dependent electromotile activity, a hallmark of OHC functional maturation, and show reduced levels of crucial OHC-expressed genes such as *Slc26a5* (which encodes prestin) and *Ocm*. Moreover, we show that ectopic expression of *Ikzf2* in IHCs induces the expression of OHC-specific genes; reduces the expression of canonical IHC genes; and confers electromotility to IHCs, demonstrating that *Ikzf2* can partially shift the IHC transcriptome towards an OHC-like identity.

The mature mammalian cochlea contains two distinct types of sensory cells, IHCs and OHCs, each of which are highly specialized and, in humans, do not regenerate once they are damaged or lost². Progressive loss of these cells, particularly the OHCs, underlies much of the aetiology of age-related hearing loss—a worldwide epidemic^{3,4}. Although these two cell types were first described by Retzius in the 1800s, the mechanisms that underlie the specification of their common progenitor cells to functional inner versus outer hair cells remain poorly understood. In addition, attempts to direct stem cells towards hair cell fates have, so far, resulted only in the formation of immature cells that lack many of the markers of mature IHCs or OHCs⁵. Given the vulnerability of the OHCs, identifying factors that specify OHC fate is crucial, not only for understanding the biology of this unique cell type, but also for ultimately working towards regenerative therapies for hearing loss.

To define a set of high-confidence OHC-expressed genes for downstream gene regulation analyses, we crossed the knock-in prestin-CreER^{T2} mouse, which can be induced to express Cre recombinase specifically in OHCs, with a transgenic RiboTag mouse, to enable OHC-specific ribosome immunoprecipitation^{6,7}. Cochlear ducts from the resulting *RiboTag*^{HAI+};prestin^{creERT2/+} mice were collected at five postnatal time points (postnatal day (P) 8, 14 and 28, and 6 and 10 weeks), and actively translated OHC transcripts were enriched for by ribosome immunoprecipitation, followed by RNA sequencing (RNA-seq) of all immunoprecipitated and paired input RNA (Extended Data Fig. 1a, b, Supplementary Table 1). We calculated an OHC enrichment factor based on the immunoprecipitated/input RNA log₂ fold change for each gene at each time point (Supplementary Table 2). Reassuringly, known

postnatal hair cell-enriched and OHC-expressed genes such as *Pou4f3*, *Gfi1*, *Strc*, *Ocm* and *Slc26a5* generally had high enrichment factor values across all time points (enrichment factor (EF) > 1), whereas prominent IHC marker genes such as *Otof*, *Atp2a3* and *Slc17a8* were generally depleted from the immunoprecipitated samples (EF < -1). In addition, marker genes for supporting cells, neurons and otic mesenchyme were also depleted (Extended Data Fig. 1c). Further informatics analyses of our RiboTag OHC dataset demonstrated a systematic enrichment of OHC markers and a depletion of IHC markers previously identified in an adult mouse OHC and IHC transcriptomic dataset⁸, and classified the OHC-enriched transcripts into three clusters (Extended Data Fig. 1d-f, Supplementary Table 3). Intersecting genes with transcripts that were enriched in OHCs in our most mature RiboTag OHC data point (10 weeks, EF > 0.5) compared with the published dataset⁸ resulted in a list of 100 highly confident postnatal OHC markers that are significantly and consistently enriched in postnatal OHCs (Fig. 1a, Supplementary Table 4). We and others have previously shown that relevant transcriptional regulators can be discovered by analysing the promoters of cell-type-specific genes to identify statistically over-represented transcription factor-binding motifs^{9,10}. A transcription factor-binding motif prediction analysis of the 100 OHC marker genes identified several enriched motifs in the 20-kb regions that centred around the transcription start site, the top five of which correspond to the transcription factors HNF4A, MZF1, POU3F2, Helios and RFX3¹¹. Of these, only *Ikzf2* (which encodes Helios) was included in the list of 100 OHC marker genes, and was found to be markedly enriched in OHCs at all time points (Fig. 1b, c), with an approximately fourfold enrichment in OHCs compared to IHCs in the previously published dataset⁸ (Supplementary Table 4). Further characterization of Helios protein expression in the inner ear confirmed that it is restricted to the OHC nuclei starting from P4, and persists in functionally mature OHCs (Fig. 1d-f, Extended Data Fig. 2a). Together, these data suggest an important role for Helios in regulating the OHC transcriptome from early postnatal to adult stages.

A recent phenotype-driven N-ethyl-N-nitrosourea (ENU) mutagenesis screen, undertaken at the MRC Harwell Institute, identified a C-to-A transversion at nucleotide 1551 of *Ikzf2* in the *cello* mouse mutant, causing a non-synonymous histidine-to-glutamine substitution (p.H517Q) in the encoded Helios transcription factor¹² (Fig. 1g, Extended Data Fig. 2b-d). A combination of in silico mutation analyses, structural 3D modelling, immunolabelling of Helios in the *cello* mutant mice, and in vitro assays predicted and validated a deleterious effect of the *cello* mutation on the ability of Helios to dimerize, without impairing its cellular localization (Fig. 1g, Extended Data Figs. 2e and 3). We further investigated the functional role of Helios in hearing by

¹Mammalian Genetics Unit, MRC Harwell Institute, Oxfordshire, UK. ²Department of Otorhinolaryngology Head and Neck Surgery, University of Maryland School of Medicine, Baltimore, MD, USA.

³National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD, USA. ⁴Department of Biomedical Science, University of Sheffield, Sheffield, UK.

⁵Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA. ⁶Mary Lyon Centre, MRC Harwell Institute, Oxfordshire, UK. ⁷Department of Physiology, College of Medicine, University of Kentucky, Lexington, KY, USA. ⁸Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. ⁹Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. ¹⁰Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, MD, USA. ¹¹These authors contributed equally: Lauren Chessim, Maggie S. Matern. ¹²These authors jointly supervised this work: Michael R. Bowl, Ronna Hertzano. *e-mail: m.bowl@har.mrc.ac.uk; rhertzano@som.umaryland.edu

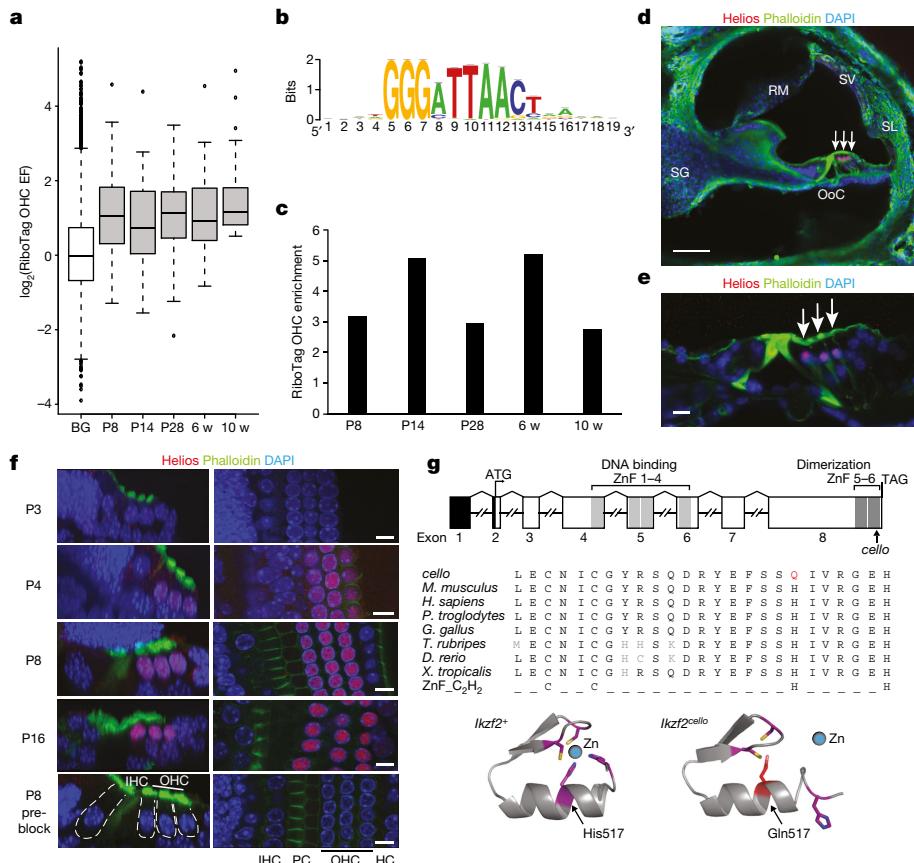


Fig. 1 | Helios is a candidate regulator of OHC genes. **a**, The 100 OHC marker genes ($n = 100$) are enriched in OHCs at all RiboTag OHC dataset time points compared to the expression of all other genes detected (background, BG) ($n = 13,044$). P values: P8 = 1.73×10^{-17} , P14 = 6.55×10^{-12} , P28 = 1.60×10^{-18} , 6 weeks (w) = 7.79×10^{-18} , 10 weeks = 1.43×10^{-33} (two-sided Wilcoxon's test). Centre line represents median enrichment factor (EF; \log_2 fold change), box demarcates first and third quartiles, whiskers demarcate first and third quartile $\pm 1.5 \times$ interquartile range (IQR) values, dots denote single outliers. **b**, Transcription factor-binding motif analysis using the 100 highly confident OHC marker genes identifies the binding signature for Helios as significantly overrepresented. Normalized enrichment score (NES) = 3.85; NES ≥ 3.0 corresponds to a false discovery rate (FDR) of 3–9%; see ref. ¹¹. **c**, *Iκzf2* transcript enrichment in OHCs as measured by RiboTag OHC RNA-seq. **d**, Specific expression of Helios in the nuclei of wild-type P8 OHCs (white arrows). $n = 3$ biologically independent samples. Scale bar, 50 μm . OoC, organ of Corti; RM, Reissner's membrane;

assessing auditory brainstem response (ABR) thresholds in wild-type and *cello* mice across several time points. Results show that *Iκzf2*^{cello/cello} mice have progressive deterioration of hearing function that starts as early as P16 (>60 dB sound pressure level (SPL)), with a threshold of ≥ 85 dB SPL by 9 months (Fig. 2a, b, Extended Data Fig. 4a–c). Using scanning electron microscopy, we show that the ultrastructure of the cochlear sensory epithelia and hair cell stereocilia bundles in the *cello* mice appear normal up to 1 month of age, after which the OHC bundles, and later the IHC bundles, begin to degenerate (Extended Data Figs. 4d, 5a–d, Supplementary Tables 5, 6). These data indicate that the hearing impairment in *cello* mice precedes the loss of hair cell bundles, and suggest that the Helios mutation instead leads to a functional deficit in OHCs. Furthermore, by using a second *Iκzf2* mutant allele (*Iκzf2*^{del890}), which leads to an in-frame deletion of the third coding exon, we confirm *Iκzf2* as the causative gene that underlies the auditory dysfunction in the *cello* mutants. At 1 month of age, *Iκzf2*^{cello/del890} compound heterozygotes display increased ABR thresholds (up to 40 dB SPL) compared to heterozygotes and wild-type mice (Extended Data Fig. 5e, f), confirming *Iκzf2*^{cello} as the causative allele in the *cello* mutant.

SG, spiral ganglion; SL, spiral ligament; SV, stria vascularis. **e**, Helios expression is maintained in wild-type OHCs at 1 month (white arrows). $n = 3$ biologically independent samples. Scale bar, 10 μm . **f**, Helios is detected in wild-type OHCs from P4 and is maintained in mature P16 OHCs. $n = 2$ (P3) and $n = 4$ (P4, P8 and P16) biologically independent samples. Specificity is confirmed by the loss of labelling when the anti-Helios antibody is 'pre-blocked' with its immunizing peptide. $n = 5$ biologically independent samples. Scale bars, 10 μm . HC, Hensen's cells; PC, pillar cells. **g**, Top, the genomic and domain structure of *Iκzf2*. Black, 5' untranslated region; light grey, N-terminal DNA-binding domain; dark grey, C-terminal dimerization domain. The *Iκzf2*^{cello} mutation lies in ZnF6. Bottom, further alignment of the Helios ZnF sequence with its paralogues and the classical Cys₂His₂ ZnF motif shows that the H517Q *cello* mutation causes substitution of a highly conserved zinc-coordinating histidine residue. 3D modelling of wild-type *Iκzf2*⁺ ZnF6 and mutant *Iκzf2*^{cello} ZnF6 illustrates the requirement of residue His517 for zinc coordination, which is not possible when residue Gln517 is substituted.

To explore the effect of the *cello* mutation on OHC physiology, we investigated the basolateral properties of OHCs. We found that the mechanoelectrical transducer (MET) current (Extended Data Fig. 6a–c) and the adult-like potassium (K^+) current $I_{K,n}$ (Extended Data Fig. 6d–h) have normal biophysical characteristics in *Iκzf2*^{cello/cello} OHCs. The resting membrane potential (V_m) of OHCs is also similar between genotypes ($Iκzf2$ ^{cello/+}: -68 ± 2 mV (mean \pm s.e.m.); $Iκzf2$ ^{cello/cello}: -70 ± 1 mV). We then investigated whether Helios regulates OHC electromotile activity. We found that stepping the membrane potential from -64 mV to $+56$ mV causes the OHCs from both genotypes to shorten (Fig. 2c, d), as previously described^{13–15}. However, *Iκzf2*^{cello/cello} OHCs show significantly reduced movement compared to *Iκzf2*^{cello/+} control OHCs (Fig. 2e), even when the values are normalized to their reduced surface area (Fig. 2f). We also found that young adult *Iκzf2*^{cello/cello} mice have significantly reduced distortion product oto-acoustic emission (DPOAE) responses (≤ -15 dB SPL) compared to littermate controls (Fig. 2g), further demonstrating impaired OHC function.

To identify genes regulated by Helios in OHCs, we compared gene expression from the cochleae of P8 *Iκzf2*^{cello/cello} and their wild-type

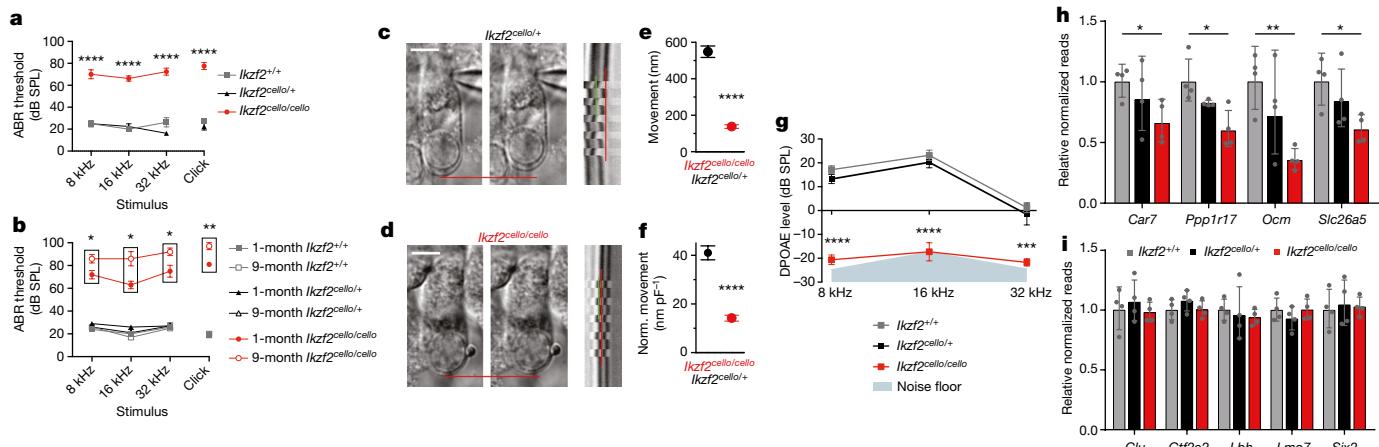


Fig. 2 | Helios is required for hearing and OHC electromotility.

a, b, Averaged ABR thresholds for *cello* mice at P16 (**a**) and 1 and 9 months of age (**b**). Age-matched *Iκzf2^{+/+}* and *Iκzf2^{cello/+}* controls display thresholds within the expected range (15–30 dB SPL) at all time points tested. $n = 4$ (**a**) and $n = 5 (**b**) biologically independent animals per genotype at each time period. Data are mean \pm s.e.m. $****P < 0.0001$ (*P16 Iκzf2^{cello/cello}* vs *Iκzf2^{+/+}* and vs *Iκzf2^{cello/+}* at 8 kHz, 16 kHz, 32 kHz, and click stimulus); $*P = 0.0284$ (1- vs 9-month *Iκzf2^{cello/cello}* at 8 kHz); $*P = 0.0166$ (1- vs 9-month *Iκzf2^{cello/cello}* at 16 kHz); $*P = 0.0303$ (1- vs 9-month *Iκzf2^{cello/cello}* at 32 kHz); $**P = 0.0042$ (1- vs 9-month *Iκzf2^{cello/cello}* click stimulus) (one-way ANOVA with Tukey post hoc test (**a**) or two-sided Welch's *t*-test (**b**)). See also Extended Data Fig. 4. **c, d,** Left, images show a patch pipette attached to an OHC from control *Iκzf2^{cello/+}* (**c**) and mutant *Iκzf2^{cello/cello}* (**d**) cochleae at P16–P18. Red lines indicate the position of the OHC basal membrane before (left) and during (right) a depolarizing voltage step from -64 mV to $+56$ mV, highlighting the shorting of the cells. Scale bars, 5 μ m. Right, time-based *z*-stack projections, in which red lines indicate the resting position of the basal membrane and the green lines indicate the movement. $n = 10$ (*Iκzf2^{cello/cello}*)$

and $n = 21$ (*Iκzf2^{cello/cello}*) *z*-stack projections (one set per OHC) from 5 biologically independent animals per genotype. **e, f,** Average movement was significantly reduced in *Iκzf2^{cello/cello}* OHCs compared to *Iκzf2^{cello/+}* at P16–P18 (**e**), even after normalization to respective membrane capacitance (**f**) (for this set of recordings, *Iκzf2^{cello/+}*: 13.6 ± 0.4 pF; *Iκzf2^{cello/cello}*: 10.0 ± 0.3 pF). Data are mean \pm s.e.m. $n = 10$ (*Iκzf2^{cello/+}*) and $n = 21$ (*Iκzf2^{cello/cello}*) OHCs from 5 biologically independent animals per genotype. $****P < 0.0001$ (two-sided Welch's *t*-test). **g,** Average DPOAE responses for *cello* mice at 1 month of age ($n = 5$ biologically independent animals per genotype). Data are mean \pm s.e.m. $****P < 0.0001$ (*Iκzf2^{cello/cello}* vs *Iκzf2^{+/+}* and vs *Iκzf2^{cello/+}* at 8 kHz, 16 kHz); $***P = 0.0004$ (*Iκzf2^{cello/cello}* vs *Iκzf2^{+/+}* at 32 kHz); $***P = 0.0012$ (*Iκzf2^{cello/cello}* vs *Iκzf2^{cello/+}* at 32 kHz) (one-way ANOVA with Tukey post hoc test). **h, i,** NanoString validations of genes downregulated in *Iκzf2^{cello/cello}* cochleae at P8 (**h**) and results showing no change in expression of other OHC transcription factors (**i**). Data are normalized to wild-type (*Iκzf2^{+/+}*) and shown as mean \pm s.d. ($n = 4$ biologically independent samples per genotype). $*P = 0.028$ (*Car7*; *Iκzf2^{cello/cello}* vs *Iκzf2^{+/+}*); $*P = 0.017$ (*Ppp1r17*); $**P = 0.006$ (*Ocm*); $*P = 0.017$ (*Slc26a5*) (two-sided Welch's *t*-test).

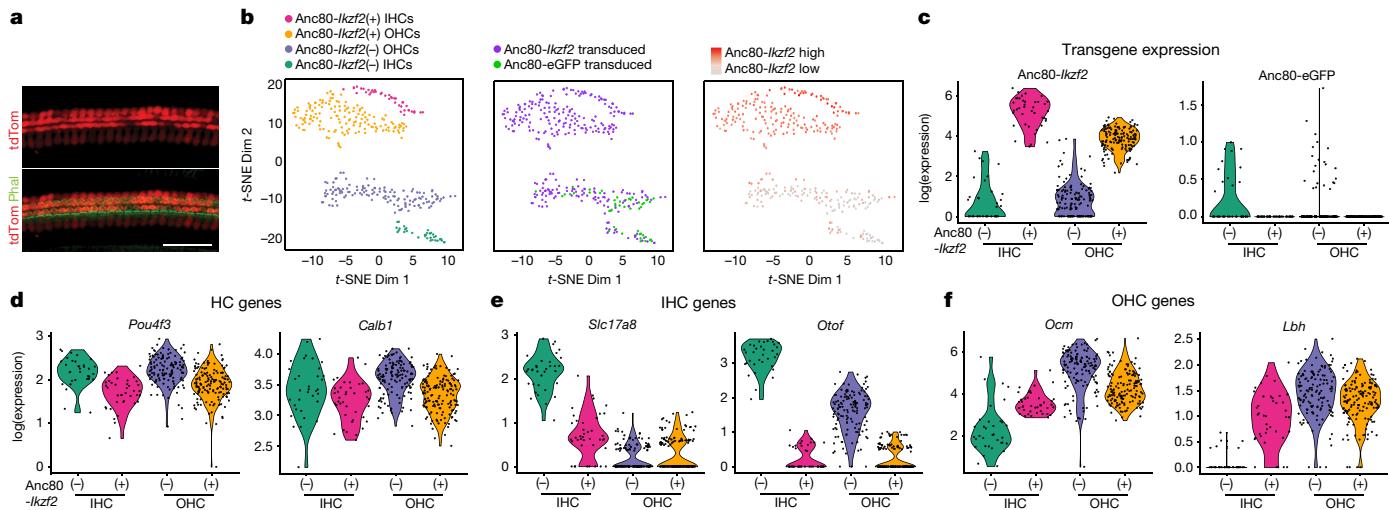


Fig. 3 | Partial transcriptional conversion of Anc80-*Iκzf2* transduced IHCs identified by scRNA-seq. **a,** Representative *Myo15^{cre/+};ROSA26^{CAG-tdTomato}* cochlear whole-mount staining. *Myo15^{cre}*-driven *tdTomato* expression is hair cell specific at P6 ($n = 3$ biologically independent samples with similar results). Scale bar, 50 μ m. **b,** *t*-distributed stochastic neighbour embedding (*t*-SNE) plots of all cochlear hair cells profiled by scRNA-seq, including the cluster to which each cell was assigned, the experimental origin of each cell (cochlea injected with *Anc80-*Iκzf2** or *Anc80-eGFP*), and the relative transcript abundance of *Anc80-*Iκzf2** measured in each cell. **c,** *Anc80-*Iκzf2** is highly expressed in the *Anc80-*Iκzf2(+)** IHCs and OHCs, whereas *Anc80-eGFP* expression is only seen in the cells assigned to the *Anc80-*Iκzf2(-)** IHC and OHC clusters. Dots represent the expression values of individual cells, with width of violins summarizing

overall relative distribution of expression. **d,** Canonical hair cell (HC) markers are highly expressed in all clusters, and not notably changed as a result of *Anc80-*Iκzf2** expression. **e,** IHC-enriched genes that are highly expressed in control IHCs vs control OHCs, but are significantly reduced in *Anc80-*Iκzf2(+)** IHCs. *Anc80-*Iκzf2(-)** IHC ($n = 34$) vs *Anc80-*Iκzf2(+)** IHC ($n = 40$) FDR: *Slc17a8* = 2.25×10^{-12} , *Otof* = 6.76×10^{-14} (Kruskal–Wallis test followed by post hoc pairwise Wilcoxon ranked sum test adjusted for multiple comparisons). **f,** OHC-enriched genes that are induced in *Anc80-*Iκzf2(+)** IHCs. *Anc80-*Iκzf2(-)** IHC ($n = 34$) vs *Anc80-*Iκzf2(+)** IHC ($n = 40$) FDR: *Ocm* = 3.65×10^{-8} , *Lbh* = 1.81×10^{-10} (Kruskal–Wallis test followed by post hoc pairwise Wilcoxon ranked sum test adjusted for multiple comparisons). See also Extended Data Figs. 8 and 9.

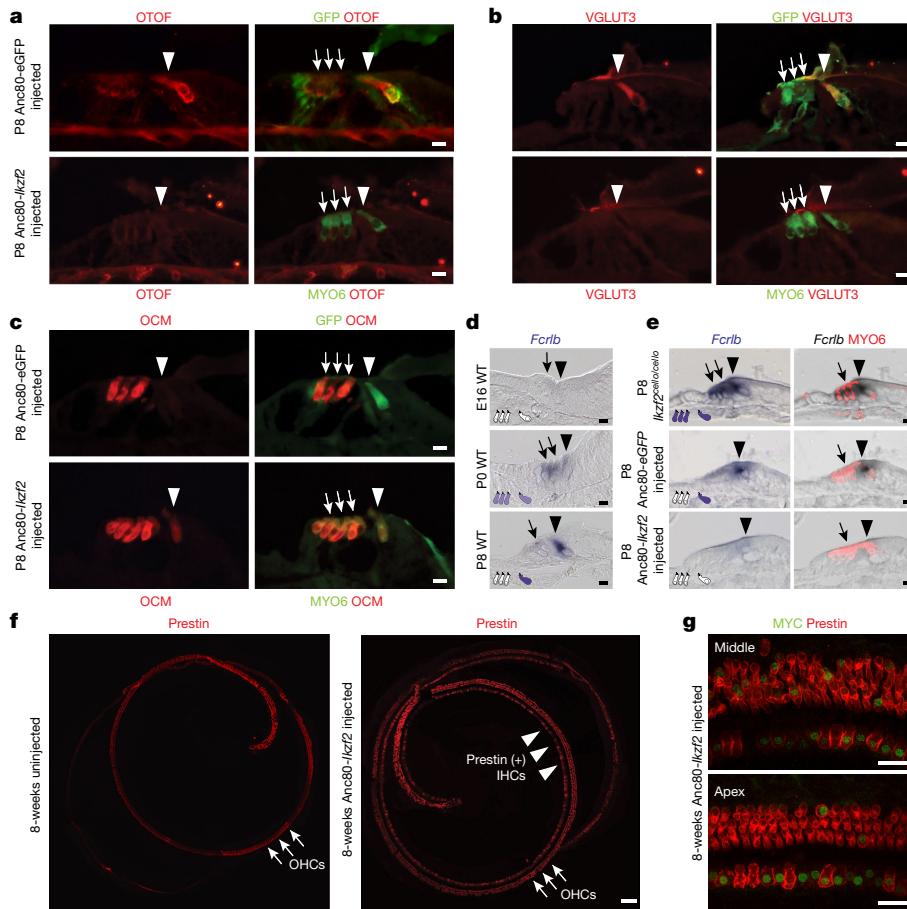


Fig. 4 | Helios overexpression modulates expression of hair cell markers. **a, b**, The IHC markers OTOF and VGLUT3 are downregulated in *Anc80-Ikzf2*-transduced IHCs ($n = 3$ biologically independent samples). **c**, The OHC marker OCM is expressed in *Anc80-Ikzf2*-transduced IHCs ($n = 3$ biologically independent samples per condition). **d**, *Fcrlb* expression during wild-type (WT) mouse inner ear development as detected by *in situ* hybridization. At embryonic day (E) 16, *Fcrlb* expression is not detected in the inner ear, but by P0 it is detected in both IHCs and OHCs, and is largely restricted to the IHCs by P8 ($n = 3$ biologically independent samples per time point). **e**, In the absence of

functional Helios (*Ikzf2^{cello/cello}* mouse), *Fcrlb* is robustly expressed in IHCs and OHCs at P8. IHC expression of *Fcrlb* is not affected by *Anc80-eGFP* transduction, whereas *Fcrlb* expression is lost in *Anc80-Ikzf2*-transduced hair cells ($n = 3$ biologically independent samples per condition). **f, g**, Prestin expression can be seen in *Anc80-Ikzf2*-transduced IHCs up to 8 weeks of age ($n = 3$ biologically independent samples at 6–8 weeks) (**f**), and overlaps with MYC staining (**g**). Scale bars, 10 μm (**a–e**), 100 μm (**f**) and 20 μm (**g**). Arrows denote OHCs, arrowheads denote IHCs. See also Extended Data Fig. 10.

littermate controls by RNA-seq. We identified 105 upregulated and 36 downregulated genes in *Ikzf2^{cello/cello}* cochleae (Supplementary Table 7), including downregulation of the canonical OHC markers *Slc26a5* and *Ocm*, which was confirmed by NanoString validation (Fig. 2h). Furthermore, we did not observe modulation of other OHC-expressed transcription factors selected from a previously published dataset¹⁶ (Fig. 2i), suggesting that the observed dysregulation in OHC genes results from disruption of a specific transcriptional cascade. Notably, by P16, the transcript levels of *Car7*, *Ocm* and *Slc26a5*, but not *Ppp1r17*, in *Ikzf2^{cello/cello}* cochleae are similar to the levels of wild-type littermate controls, suggesting that other factors may be compensating for the functional loss of Helios by this time point (Extended Data Fig. 6i).

To characterize the transcriptional cascade downstream of Helios, we performed *in vivo* *Anc80L65* adeno-associated virus (AAV) gene delivery of a Myc-tagged *Ikzf2* or enhanced green fluorescent protein (eGFP) (hereafter termed *Anc80-Ikzf2* or *Anc80-eGFP*, respectively) to neonatal inner ears of *Myo15^{cre/+}; ROSA26^{CAG-tdTomato}* mice, sorted the cochlear hair cells at P8, and measured resultant changes in gene expression using single-cell RNA sequencing (scRNA-seq)^{17,18} (Fig. 3a, Extended Data Fig. 7). The hair cells from inner ears injected with *Anc80-Ikzf2* separated into two distinct sets of clusters, containing both IHCs and OHCs. One set of IHCs and OHCs completely overlapped with the hair cells from the control ears injected with *Anc80-eGFP* (Fig. 3b, bottom clusters), whereas the other set clustered separately

(Fig. 3b, top clusters). Separation of the two sets of clusters showed a clear correlation with expression of the *Anc80-Ikzf2* transgene (Fig. 3b), in which hair cells in the bottom clusters had lower expression of *Anc80-Ikzf2*, and the hair cells in the top clusters had higher expression of *Anc80-Ikzf2* (hereafter defined as *Anc80-Ikzf2* low (−) and high (+), respectively). Because the hair cells defined as *Anc80-Ikzf2* (−) clustered together with the hair cells transduced with *Anc80-eGFP*, these two groups of hair cells were merged and named *Anc80-Ikzf2* (−) IHCs and OHCs for all downstream analyses (Fig. 3b, c).

Although the overexpression of *Ikzf2* in IHCs and OHCs did not change the expression of hair cell markers such as *Pou4f3* and *Calb1* (Fig. 3d), it led to a significant downregulation of many genes whose transcripts were identified as IHC-enriched in the control hair cell populations, including *Slc17a8*, *Otof*, *Rprm*, *Atp2a3* and *Fgf8* (Fig. 3e, Extended Data Fig. 8, Supplementary Tables 8–10). Notably, some of the genes that are downregulated in both *Anc80-Ikzf2* (+) IHCs and OHCs are genes that are normally expressed in both cell types in early postnatal development, and that later become IHC-specific^{19,20} (for example, *Pvalb* and *Otof*, Supplementary Table 10). This suggests that the overexpression of *Ikzf2* in OHCs results in an accelerated downregulation of these genes. Furthermore, *Ikzf2* overexpression in IHCs results in the upregulation of genes that are normally enriched in OHCs, such as *Ocm*, *Pde6d*, *Ldhb* and *Lbh* (Fig. 3f, Extended Data Fig. 8). Overall, these data suggest that during normal OHC development, Helios

functions to decrease the expression of early pan-hair-cell markers, such as *Ot6*, in the maturing OHCs, as well as to upregulate OHC marker genes. A correlation analysis further validates the role of Helios in regulating OHC-related gene expression (Extended Data Figs. 8, 9, Supplementary Table 11). The effect of *Ikzf2* transduction on IHC gene expression was also validated by immunolabelling for OTOF, VGLUT3, OCM or prestin and by *in situ* hybridization for *Fcrlb* (Fig. 4, Extended Data Fig. 10a, b). Analysis of the surface characteristics of the transduced IHCs does not show a change from an IHC-like to an OHC-like stereociliary bundle, consistent with a partial role for Helios in regulating OHC-fate (Extended Data Fig. 10c). However, *Ikzf2* transduction resulted in the appearance of prominent voltage-dependent (nonlinear) capacitance in IHCs (Extended Data Fig. 10d, e), which is an electrical ‘signature’ of prestin-dependent OHC electromotility^{21,22}. These data indicate that *Anc80-Ikzf2*-transduced IHCs start to acquire the major function of normal OHCs.

In conclusion, our study demonstrates that Helios is necessary for hearing and is a crucial regulator of gene expression in the maturing postnatal OHC. In particular, our results suggest that Helios functions to suppress IHC and early pan-hair-cell gene expression in OHCs, as well as to upregulate canonical OHC marker genes. It further shows that Helios is sufficient to induce the essential functional characteristic of electromotility and many of the molecular characteristics of OHCs when expressed in early postnatal IHCs, albeit not all of them, supporting the notion that additional OHC-expressed transcription factors are involved in postnatal OHC development. To our knowledge, this is the first study to demonstrate functional shifts in postnatal hair-cell molecular identities via viral gene delivery, and it suggests that the delivery of combinations of transcription factors may lead to successful regeneration of functional OHCs in the deafened cochlea.

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/s41586-018-0728-4>.

Received: 15 August 2018; Accepted: 25 October 2018;

Published online: 21 November 2018

- Dallos, P. et al. Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron* **58**, 333–339 (2016).
- Ehret, G. Development of absolute auditory thresholds in the house mouse (*Mus musculus*). *J. Am. Audiol. Soc.* **1**, 179–184 (1976).
- Bielefeld, E. C., Tanaka, C., Chen, G. D. & Henderson, D. Age-related hearing loss: is it a preventable condition? *Hear. Res.* **264**, 98–107 (2010).
- World Health Organization. *Deafness and Hearing Loss*; available at: <http://www.who.int/mediacentre/factsheets/fs300/en/> (2018).
- Mittal, R. et al. Recent advancements in the regeneration of auditory hair cells and hearing restoration. *Front. Mol. Neurosci.* **10**, 236 (2017).
- Fang, J. et al. Outer hair cell-specific prestin-CreER^{T2} knockin mouse lines. *Genesis* **50**, 124–131 (2012).
- Sanz, E. et al. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc. Natl. Acad. Sci. USA* **106**, 13939–13944 (2009).
- Liu, H. et al. Characterization of transcriptomes of cochlear inner and outer hair cells. *J. Neurosci.* **34**, 11085–11095 (2014).
- Elkon, R. et al. RFX transcription factors are essential for hearing in mice. *Nat. Commun.* **6**, 8549 (2015).

- Hertzano, R. et al. Cell type-specific transcriptome analysis reveals a major role for Zeb1 and miR-200b in mouse inner ear morphogenesis. *PLoS Genet.* **7**, e1002309 (2011).
- Janky, R. et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLOS Comput. Biol.* **10**, e1003731 (2014).
- Potter, P. K. et al. Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nat. Commun.* **7**, 12444 (2016).
- Ashmore, J. F. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol. (Lond.)* **388**, 323–347 (1987).
- Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227**, 194–196 (1985).
- Marcotti, W. & Kros, C. J. Developmental expression of the potassium current $I_{K,n}$ contributes to maturation of mouse outer hair cells. *J. Physiol. (Lond.)* **520**, 653–660 (1999).
- Li, Y. et al. Transcription factors expressed in mouse cochlear inner and outer hair cells. *PLoS One* **11**, e0151291 (2016).
- Zinn, E. et al. In silico reconstruction of the viral evolutionary lineage yields a potent gene therapy vector. *Cell Reports* **12**, 1056–1068 (2015).
- Landegger, L. D. et al. A synthetic AAV vector enables safe and efficient gene transfer to the mammalian inner ear. *Nat. Biotechnol.* **35**, 280–284 (2017).
- Roux, I. et al. Otoferlin, defective in a human deafness form, is essential for exocytosis at the auditory ribbon synapse. *Cell* **127**, 277–289 (2006).
- Simmons, D. D., Tong, B., Schrader, A. D. & Hornak, A. J. Oncomodulin identifies different hair cell types in the mammalian inner ear. *J. Comp. Neurol.* **518**, 3785–3802 (2010).
- Santos-Sacchi, J. Reversible inhibition of voltage-dependent outer hair cell motility and capacitance. *J. Neurosci.* **11**, 3096–3110 (1991).
- Zheng, J. et al. Prestin is the motor protein of cochlear outer hair cells. *Nature* **405**, 149–155 (2000).

Acknowledgements We thank L. Vizor, J. Sanderson and W. Chien for technical help and Z. Ahmed for comments on the manuscript. This work was supported by Action on Hearing Loss (G65 to M.R.B., R.H., W.M. and S.D.M.B.), Medical Research Council (MC_U142684175 to S.D.M.B.), Wellcome Trust (102892 to W.M.), NIDCD/NIH R01DC013817 and R01DC03544 (R.H.), DOD CDMRP MR130240 (R.H.), NIDCD/NIH T32DC00046 and F31DC016218 (M.S.M.), the Intramural Program at NIDCD DC000059 (M.W.K.), and NIDCD/NIH R01DC014658 (G.I.F.). S.L.J. is a Royal Society University Research Fellow. R.E. is a Faculty Fellow of the Edmond J. Safra Center for Bioinformatics at Tel Aviv University.

Reviewer information *Nature* thanks M. Montcouquiol, B. Walters and the anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions L.C., M.S.M., M.C.K., Y.O., B.M., M.M., G.C., C.T.E., G.I.F., M.W.K., W.M., S.D.M.B., M.R.B. and R.H. designed and interpreted the experiments. L.C., M.S.M., M.C.K., S.L.J., Y.O., B.M., M.M., A.P., J.P., R.H., E.C.D., G.T., A.K.D., G.I.F. and W.M. performed the experiments. Y.S. and R.E. analysed the gene expression data. S.W. aided in the management of the *cello* colony. M.R.B. and R.H. conceived and coordinated the study. L.C., M.S.M., M.C.K., W.M., S.D.M.B., R.E., M.R.B. and R.H. 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/s41586-018-0728-4>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-018-0728-4>.

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

Correspondence and requests for materials should be addressed to M.R.B. or R.H.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment unless stated otherwise.

Animal procedures. Animal procedures performed at the University of Maryland School of Medicine were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and have been approved by the Institutional Animal Care and Use Committee at the University of Maryland, Baltimore (protocol numbers 1112005 and 1015003). The RiboTag (maintained on a C57BL/6N background), *prestin*^{creERT2} and *Myo15*^{cre} mouse models (maintained on a C57BL/6J background) have been described previously^{6,7,23}, and were provided by M. K. Lobo (RiboTag), J. Zuo (*prestin*^{creERT2}), and C. Petit and T. Friedman (*Myo15*^{cre}). CBA/CaJ mice (stock 000654) and B6.Cg-Gt(ROSA)26Sor^{tm1(CAG-tdTOMO)Hze}/J mice (stock 007914, referred to as ROSA26^{CAG-tdTOMO}) were procured from the Jackson Laboratory. The specificity of *prestin*^{creERT2} was determined by crossing *prestin*^{creERT2/creERT2} mice to ROSA26^{CAG-tdTOMO} mice, and the resulting offspring were dissected at P21 for whole-mount immunohistochemistry. To generate animals for the RiboTag OHC RNA-seq dataset, RiboTag^{HA/HA} mice were crossed to *prestin*^{creERT2/creERT2} mice to produce RiboTag^{HA/+;prestin}^{creERT2/+} mice. These mice were further intercrossed to obtain double homozygous RiboTag^{HA/HA;prestin}^{creERT2/creERT2} animals, which were then crossed to CBA/CaJ mice to generate F₁ RiboTag^{HA/+;prestin}^{creERT2/+} offspring on a mixed CBA/C57BL/6 background, avoiding the recessively inherited age-related hearing loss phenotype inherent to C57BL/6 mice²⁴. Recombination was induced by tamoxifen injection (3 mg per 40 g body weight in mice younger than 21 days, 9 mg per 40 g body weight in mice 21 days or older), and cochlear tissues were collected at the following ages: P8, P14, P28, 6 weeks and 10 weeks. For the *cello* RNA-seq and NanoString experiments, cochlear ducts from *Ikzf2*^{+/+}, *Ikzf2*^{cello/+} and *Ikzf2*^{cello/cello} mice were dissected at P8 and P16. CD-1 or C57BL/6 pregnant females were procured from Charles River or the University of Maryland School of Medicine Veterinary Resources. Resulting neonates were injected with Anc80L65 virus between P1 and P3, and dissected for later analyses between P8 and 8 weeks. For the Anc80L65-transduced IHC scRNA-seq experiment, *Myo15*^{cre/cre} mice were crossed to ROSA26^{CAG-tdTOMO} mice, and resulting offspring were injected with Anc80L65 virus between P1 and P3, and the cochlear epithelium was collected at P8. Additionally, several litters with Anc80-*Ikzf2*-injected pups and their control littermates (aged P7–P8), together with a mother, were sent to the University of Kentucky for the measurements of nonlinear (voltage-dependent) capacitance, an electrical ‘signature’ of electromotility. All animal procedures for these experiments were approved by the Institutional Animal Care and Use Committee at the University of Kentucky (protocol 00903M2005). Both male and female animals were used for all experiments.

Animal procedures performed at the MRC Harwell Institute were licensed by the Home Office under the Animals (Scientific Procedures) Act 1986, UK and additionally approved by the relevant Institutional Ethical Review Committees. The *cello* mutant mouse was originally identified from the MRC Harwell Institute phenotype-driven N-ethyl-N-nitrosourea (ENU) Ageing Screen¹². In this screen, ENU-mutagenized C57BL/6J males were mated with wild-type sighted C3H⁺ (C3H.Pde6b+) females²⁵. The resulting G₁ males were crossed with C3H.Pde6b+ females to produce G₂ females, all of which were screened for the *Cdh23*^{ahl} allele²⁴. *Cdh23*^{ahl} G₂ females were then backcrossed to their G₁ fathers to generate recessive G₃ pedigrees, which entered a longitudinal phenotyping pipeline. Auditory phenotyping comprised clickbox testing at 3, 6, 9 and 12 months of age and ABR at 9 months of age. The *Ikzf2*^{del1890} mutant line was generated by the Molecular and Cellular Biology group at the MRC Harwell Institute using a CRISPR-Cas9-mediated deletion approach. Both male and female mice were used for experiments.

RiboTag immunoprecipitations. RiboTag immunoprecipitations were performed as described previously⁷. In brief, for one biological sample, 10 cochlear ducts from 5 mice were pooled and homogenized in 1 ml of supplemented homogenization buffer (50 mM Tris-HCl, pH 7, 100 mM KCl, 12 mM MgCl₂, 1% Nonidet P-40, 1 mM 1,4-dithiothreitol, 1× protease inhibitor cocktail, 200 U ml⁻¹ RNaseOUT, 100 µg ml⁻¹ cycloheximide, 1 mg ml⁻¹ heparin). Homogenates were spun down (9,400g for 10 min at 4 °C) to remove particulates. Then, 40 µl of homogenate was reserved for total RNA isolation (input control), and the remaining homogenate was incubated with 5 µg haemagglutinin (HA) antibody (BioLegend) at 4 °C under gentle rotation for 4–6 h. The supernatant was then added to 300 µl of rinsed Invitrogen Dynabeads Protein G magnetic beads (Thermo Fisher), and incubated overnight at 4 °C under gentle rotation. The next day, bound beads were rinsed three times with 800 µl high-salt buffer (50 mM Tris-HCl, pH 7, 300 mM KCl, 12 mM MgCl₂, 1% Nonidet P-40, 1 mM 1,4-dithiothreitol, 100 µg ml⁻¹ cycloheximide) at 4 °C for 10 min, rotating. Buffer RLT (350 µl) from the RNeasy Plus Micro kit (Qiagen) was then added to the beads or reserved input sample, and vortexed for 30 s to release bound ribosomes and RNA. RNA was extracted

according to the manufacturer’s instructions for the RNeasy Plus Micro kit (Qiagen), using 16 µl of nuclease free water for elution as described previously¹⁰. This method yielded an average of 10.9 ng of immunoprecipitated RNA (average concentration = 0.68 ng µl⁻¹) and 185.6 ng of input RNA (average concentration = 10.9 ng µl⁻¹) for downstream analyses. All RNA samples used for RNA-seq had a minimum RNA integrity number (RIN) of 8.

cello cochlear RNA extractions. For the *cello* RNA-seq, cochlear ducts from P8 *Ikzf2*^{+/+} and *Ikzf2*^{cello/cello} mice were dissected and pooled (6 cochlear ducts per sample) to generate two biological replicates per genotype. For the NanoString validations, cochlear ducts from P8 *Ikzf2*^{cello/cello}, *Ikzf2*^{cello/+} and *Ikzf2*^{+/+} mice were dissected and pooled (2–4 cochlear ducts per sample) to generate four biological replicates per genotype. RNA was extracted using the Direct-zol RNA MiniPrep kit (Zymo Research) following the manufacturer’s instructions. RNA quality and concentration were assessed using the Agilent RNA Pico kit (Agilent Technologies). All RNA samples used for RNA-seq had a minimum RNA integrity number (RIN) of 8.

RNA-seq and normalization. RiboTag OHC RNA-seq libraries were prepared using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs), and samples were sequenced in at least biological duplicates on a HiSeq 4000 system (Illumina) using a 75-bp paired end read configuration. P8 *Ikzf2*^{+/+} and *Ikzf2*^{cello/cello} RNA libraries were prepared using the TruSeq RNA Sample Prep kit (Illumina), and samples were sequenced in biological duplicates on a HiSeq 2000 system (Illumina) and a 125-bp paired-end read configuration. Reads were aligned to the *Mus musculus* reference genome (assembly GRCm38.87 (RiboTag) or GRCm38.84 (P8 *cello*)) using TopHat v.2.0.8²⁶, and HTSeq was used to quantify the number of reads aligning to predicted coding regions²⁷. See Supplementary Table 12 for alignment statistics. Expression levels were normalized using quantile normalization. In downstream analyses, only genes covered by at least 20 reads in a minimum of two samples from the same biological condition were considered as expressed. Significant differential gene expression between samples was assessed using DEseq²⁸. In addition to statistical significance between samples (FDR ≤ 0.05), we also required a complete separation of expression levels between compared conditions for a gene to be called as differentially expressed. That is, for a gene to be called downregulated in condition A compared to condition B, we required that all normalized expression levels measured in the samples of condition A to be lower than all normalized expression levels measured in the samples of condition B. To avoid inflation of fold change estimates for lowly expressed genes, a floor level equal to the tenth percentile of the distribution of the expression levels was applied (that is, all expression values below the tenth percentile were set to the tenth percentile value). The OHC enrichment factors were calculated for each gene and time point by comparing the RiboTag immunoprecipitated samples to matched input samples, and are defined as the log₂ ratio of expression levels between the immunoprecipitated and input samples. Inspection of these enrichment factors revealed a systematic association to transcripts length (Supplementary Fig. 2a). Therefore, we used a locally weighted regression, implemented by the R lowess function, to remove this systematic effect (Supplementary Fig. 2b).

Gene expression analyses. Genes with a changed level of expression in OHC immunoprecipitated samples at any time point relative to P8 were subjected to a clustering analysis using the CLICK algorithm, implemented in the EXPANDER package^{29,30}. Gene Ontology (GO) enrichment analysis was carried out using the EXPANDER implemented tool TANGO²⁹. The adult mouse IHC and OHC transcriptomic dataset used for comparisons was generated previously⁸ and can be accessed through the GEO database (accession number GSE111348)⁸. The expanded motif prediction analysis was performed using iRegulon¹¹ through the Cytoscape visualization tool³¹. The analysis was performed on the putative regulatory region of 20 kb centred around the transcription start site using default settings.

Immunohistochemistry. For cochlear sections, mice were euthanized by cervical dislocation and inner ears fixed in 4% paraformaldehyde (PFA) overnight at 4 °C then decalcified in 4% EDTA in PBS. Ears were positioned in 4% low melting temperature agarose (Sigma-Aldrich) in upturned BEEM capsules (Agar Scientific) at a 45° diagonal angle, with the apex of the cochlea facing down and the vestibular system uppermost. Once set, the agarose block was removed from the BEEM capsule and 200 µm sections were cut through the mid-modiolar plane of the cochlea using a Leica VT1000S Vibratome. Sections were simultaneously permeabilized and blocked with 10% donkey serum (Sigma) in 0.3% Triton-X for 30 min at room temperature then labelled with primary antibodies for 3 h at room temperature. To enable detection, samples were incubated with fluorophore-coupled secondary antibodies for 2 h at room temperature then stained with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; 1:2,500, Thermo Fisher) for 5 min. Sections were transferred to WillCo glass bottom dishes (Intracel) and visualized free-floating in PBS using a Zeiss 700 inverted confocal microscope (10–40× magnification). Primary antibodies: goat anti-Helios M-20 (1:400, Santa Cruz Biotechnology) and mouse anti-β-actin (1:500, Abcam). Secondary antibodies: Alexa Fluor 568 donkey

anti-goat (Invitrogen, 1:200) and Alexa Fluor 488 donkey anti-mouse (Invitrogen, 1:200).

For cochlear whole-mounts, mice were euthanized by cervical dislocation and inner ears fixed in 2% PFA for 30 min at 4 °C. After fixation, ears were fine dissected to expose the sensory epithelium then immediately permeabilized in 0.2% Triton-X for 10 min and blocked with 10% donkey serum (Sigma) for 1 h at room temperature. Cochleae were immunolabelled with goat anti-Helios M-20 (1:400, Santa Cruz Biotechnology) overnight at 4 °C then incubated with Alexa Fluor 568 donkey anti-goat secondary (1:200, Invitrogen) and the F-actin marker Alexa Fluor 488 Phalloidin (1:200, Invitrogen) for 1 h at room temperature. Samples were washed with DAPI (1:2,500, Thermo Fisher) for 60 s to stain nuclei then mounted onto slides with SlowFade Gold (Life Technologies) and visualized using a Zeiss LSM 710 fluorescence confocal microscope and 63× oil magnification.

Identification of the *cello* mutation. DNA was extracted from ear biopsies of affected G₃ mice using the DNeasy Blood and Tissue Kit (Qiagen) and used for an initial genome-wide linkage study, using SNP markers polymorphic between the parental strains C57BL/6J and C3H.Pde6b+ (Tepnel Life Sciences). Following linkage to a 21.57 Mb region on chromosome 1, additional SNP markers were identified and genotyped using standard PCR and restriction endonuclease protocols to delineate an 8.4 Mb critical interval between SNPs rs31869113 and rs13475914. Subsequently, high-quality DNA was extracted from the tail of an affected G₃ mouse using the Illustra Nucleon BAC2 Genomic DNA Extraction Kit (GE Healthcare) and sequenced by the Oxford Genomics Centre (Wellcome Trust Centre for Human Genetics) using the HiSeq system (Illumina). Sequencing reads were aligned to the mouse reference genome (assembly GRCm38) and known C57BL/6J and C3H.Pde6b+ SNPs were filtered out, leaving variants that were then given a quality score based on their sequencing read depth. Variants within the 8.4 Mb critical region which were deemed heterozygous, low-confidence (quality score < 200), non-coding or synonymous were discounted. The putative *Ikzf2* lesion was amplified by standard PCR (see Supplementary Table 13 for genotyping primers) and validated by Sanger sequencing, using DNA from an affected G₃ animal, as well as an unaffected G₃ (control). Sequence gaps that spanned coding regions were amplified by PCR using DNA from an affected G₃ mouse and analysed by Sanger sequencing. In all cases, sequence data were assessed for variation using DNASTAR Lasergene software (version 12.0.0).

In silico analyses. Three independent online tools were used to predict the functional effect of the *cello* mutation in silico: Sorting Intolerant From Tolerant (SIFT); Polymorphism Phenotyping version 2 (PolyPhen-2); and Protein Variation Effect Analyser (PROVEAN)^{32–34}. Structural 3D representations of wild-type and H517Q helios ZnF6 were predicted with RaptorX³⁵, using peptide sequences as input, and visualized using pyMOL software (version 1.7).

In vitro analyses. A full-length *Ikzf2*⁺ Helios construct was prepared using the pGEM-T Vector System II Kit (Promega) and used as a template for the generation of an *Ikzf2*^{cello} Helios construct with the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). Plasmid DNA was prepared using the Wizard Plus SV Miniprep Purification System (Promega) and validated by Sanger sequencing. Sequence-verified *Ikzf2*⁺ and *Ikzf2*^{cello} constructs were subcloned in-frame into pCMV-Myc and pEGFP-C3 mammalian expression vectors (provided by C. Esapa), to yield N-terminally tagged *Ikzf2*⁺ and *Ikzf2*^{cello} Helios. See Supplementary Table 13 for cloning and mutagenesis oligonucleotide sequences.

Constructs were subsequently used for subcellular localization studies using male *Cercopithecus aethiops* SV40 transformed kidney cells (Cos-7) cells that had been seeded onto 22 × 22 mm glass coverslips in six-well plates, at a volume of 1 × 10⁵ cells per well. After 24 h (or when 50–60% confluent), cells were transiently transfected with 1 µg DNA of the *Ikzf2*⁺-Myc or *Ikzf2*^{cello}-Myc Helios construct using JetPEI DNA Transfection Reagent (Polyplus Transfection). At 24 h after transfection, cells were fixed in 4% PFA for 10 min and permeabilized with 1% Triton-X for 15 min at room temperature. After blocking in 10% donkey serum (Sigma) for 1 h at room temperature, cells were immunolabelled with goat anti-Helios M-20 primary antibody (1:600, Santa Cruz Biotechnology) overnight at 4 °C, then incubated with Alexa Fluor 488 donkey anti-goat secondary antibody (1:200, Invitrogen) and F-actin marker Texas Red-X Phalloidin (1:200, Invitrogen) for 1 h at room temperature. Cells were washed with DAPI (1:2,500, Thermo Fisher) for 60 s. Coverslips were mounted onto slides with SlowFade Gold (Life Technologies) and cells were visualized using a Zeiss LSM 710 multiphoton fluorescence confocal microscope and 63× oil magnification.

Constructs were also used for co-immunoprecipitation studies using human embryonic kidney (HEK293T) cells that had been seeded directly onto six-well plates at a volume of 5 × 10⁵ cells per well. Cells were transiently co-transfected 24 h later with a total of 2 µg plasmid DNA to mimic the wild-type (1 µg *Ikzf2*⁺-Myc Helios + 1 µg *Ikzf2*⁺-GFP Helios), heterozygous (1 µg *Ikzf2*⁺-Myc Helios + 1 µg *Ikzf2*^{cello}-GFP Helios; 1 µg *Ikzf2*^{cello}-Myc Helios + 1 µg *Ikzf2*⁺-GFP Helios) or homozygous (1 µg *Ikzf2*^{cello}-Myc Helios + 1 µg *Ikzf2*^{cello}-GFP Helios) states using JetPEI DNA Transfection Reagent (Polyplus Transfection). Single transfections

with either 1 µg *Ikzf2*⁺-GFP Helios or 1 µg *Ikzf2*⁺-Myc Helios were also carried out for negative controls. Cells were lysed in 250 µl of 1× RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 50 mM Tris, pH 7.5, in milliQ water) 48 h after transfection, then incubated with Protein G Sepharose Beads (Sigma) for 2 h at 4 °C. The beads were pelleted by centrifugation and the supernatant incubated with either 1 µg of mouse anti-cMyc 9E10 antibody (Developmental Studies Hybridoma Bank) or 1–2 µg of custom-made rabbit anti-GFP antibody overnight at 4 °C. The immunoprecipitation complexes were captured using Protein G beads, washed with RIPA buffer and released by incubation with NuPAGE Reducing Agent (Novex). Immunoprecipitation reactions and their corresponding reduced cell lysate were analysed by western blotting. Samples were electrophoresed on NuPage 4–12% Bis-Tris gels (Invitrogen) and transferred onto nitrocellulose membranes using the iBlot system (Invitrogen). Membranes were incubated with mouse anti-cMyc 9E10 antibody (1:5,000, Developmental Studies Hybridoma Bank) and custom-made rabbit anti-GFP (1:1,000, CUK-1819 MGU-GFP-FL) primary antibodies. Mouse 12G10 anti-α-tubulin (1:10,000, Developmental Studies Hybridoma Bank) was also used as a loading control. For detection, membranes were incubated with goat anti-mouse IRDye 680RD (1:15,000, LI-COR) and goat anti-rabbit IRDye 800CW secondary antibodies (1:15,000, LI-COR) and imaged using the Odyssey CLx Infrared Imaging System (LI-COR). For quantification, band intensities were determined using Image Studio Lite Ver 5.2 software and used to calculate the relative ratio of the co-immunoprecipitation to immunoprecipitation signal. Cos-7 and HEK293T cell lines used in this study were provided by C. Esapa, were not authenticated, but were tested and confirmed to be free of mycoplasma contamination. Cells were grown at 37 °C under 5% CO₂ conditions in DMEM (Invitrogen) containing 10% heat-inactivated fetal bovine serum (FBS) (Invitrogen) and 1× penicillin/streptomycin (Invitrogen).

ABR. ABR tests were performed using a click stimulus in addition to frequency-specific tone-burst stimuli to screen mice for auditory phenotypes and investigate auditory function³⁶. Mice were anaesthetized by intraperitoneal injection of ketamine (100 mg ml⁻¹ at 10% v/v) and xylazine (20 mg ml⁻¹ at 5% v/v) administered at the rate of 0.1 ml per 10 g body mass. Animals were placed on a heated mat inside a sound-attenuated chamber (ETS Lindgren) and electrodes were placed subdermally over the vertex (active), right mastoid (reference) and left mastoid (ground). ABR responses were collected, amplified and averaged using TDT System 3 (Tucker Davies Technology) in conjunction with either BioSig RP (version 4.4.11) or BioSig RZ (v5.7.1) software. The TDT system click ABR stimuli comprised clicks of 0.1 ms broadband noise spanning approximately 2–48 kHz, presented at a rate of 21.1 s⁻¹ with alternating polarity. Tone-burst stimuli were of 7 ms duration, inclusive of 1 ms rise/fall gating using a Cos2 filter, presented at a rate of 42.5 s⁻¹ and were measured at 8, 16 and 32 kHz. All stimuli were presented free-field to the right ear of the mouse, starting at 90 dB SPL and decreasing in 5 dB increments. Auditory thresholds were defined as the lowest dB SPL that produced a reproducible ABR trace pattern and were determined manually. All ABR waveform traces were viewed and re-scored by a second operator blinded to genotype. Animals were recovered using 0.1 ml of anaesthetic reversal agent atipamezole (Antisedan, 5 mg ml⁻¹ at 1% v/v), unless aged P16, when the procedure was performed terminally.

Generation of *Ikzf2*^{del890} mice. The *Ikzf2*^{del890} mutant line was generated by the Molecular and Cellular Biology group at the Mary Lyon Centre, MRC Harwell Institute using CRISPR–Cas9 gene editing, as described previously³⁷ (see Supplementary Table 13 for single-guide RNA (sgRNA) sequences, donor oligonucleotide sequences and genotyping primers). For construction of each sgRNA plasmid, a pair of single-stranded donor oligonucleotides (IDT) was hybridized and cloned using Gibson Assembly Master Mix (NEB) into linearized p_1.1 plasmid digested with StuI and AflII to express sgRNAs under the T7 promoter.

The p_1.1_sgRNA plasmids were linearized with XbaI, purified with phenol-chloroform, and the products were used as templates from which sgRNAs were in vitro transcribed. sgRNAs were synthesized using MEGAshortscript T7 Transcription Kit (Ambion). RNAs were purified using MEGAClear Transcription Clean-Up Kit (Ambion). RNA quality was assessed using a NanoDrop (Thermo Scientific) and by electrophoresis on 2% agarose gel containing Ethidium Bromide (Fisher Scientific).

As this exon deletion mutant was generated as part of an experiment to generate a floxed mutant, a *Ikzf2* flox long single-stranded DNA (lssDNA) donor was also synthesized as described previously for inclusion in the microinjection mix³⁸.

For microinjections, the pronucleus of one-cell stage C57BL/6NTac embryos were injected with a mix containing Cas9 mRNA (5meC,Ψ, Tebu-Bio/TriLink Biotechnologies) at 100 ng µl⁻¹, the four *Ikzf2* sgRNAs, each at 50 ng µl⁻¹ and the *Ikzf2* flox lssDNA donor at 50 ng µl⁻¹ prepared in microinjection buffer. Injected embryos were re-implanted in pseudo-pregnant CD-1 females, which were allowed to litter and rear F₀ progeny.

For genotyping, genomic DNA was extracted from ear biopsies of F₀ and F₁ mice using DNA Extract All Reagents Kit (Applied Biosystems) and amplified by PCR using high fidelity Expand Long Range dNTPack (Roche) and specific genotyping

primers (see Supplementary Table 13). PCR products were further purified using QIAquick Gel Extraction Kit (Qiagen) and analysed by Sanger sequencing. Copy counting experiments by droplet digital PCR (ddPCR) against a known two copy reference (*Dot1l*) were also carried out to confirm the exon deletion and that there were no additional integrations of the lssDNA donor. Mice carrying the *del890* deletion allele were subsequently mated with mice carrying the *celfo* mutation to generate *Ikzf2^{cello/del890}* compound heterozygotes for complementation testing. **Scanning electron microscopy.** Mice were euthanized by cervical dislocation and inner ears were removed and fixed in 2.5% glutaraldehyde (TAAB Laboratories Equipment Ltd) in 0.1 M phosphate buffer for 4 h at 4°C. After decalcification in 4.3% EDTA, cochleae were dissected to expose the organ of Corti, and subjected to 'OTO' processing (1 h incubation in 1% osmium tetroxide (TAAB Laboratories Equipment), 30 min incubation in 1% thiocarbohydrazide (Sigma), 1 h incubation in 1% osmium tetroxide), before dehydration in increasing concentrations of ethanol (25%, 40%, 60%, 80%, 95%, 2 × 100%) at 4°C. Samples were critical point dried with liquid CO₂ using an Emitech K850 (EM Technologies), then mounted on stubs using silver paint (Agar Scientific) and sputter coated with platinum using a Quorum Q150R S sputter coater (Quorum Technologies). Samples were examined using a JEOL JSM-6010LV Scanning Electron Microscope. Hair cell bundle counts were performed by counting the number of OHC and IHC bundles adjacent to ten pillar cells in the apical (<180° from apex), mid (180–450° from apex) and basal (>450° from apex) regions of the cochlea. At least three ears (one ear per mouse) were analysed for each genotype at each time point.

Electrophysiological analyses. Electrophysiological recordings were made from OHCs of *celfo* mice aged P9–P18. Cochleae were dissected in normal extracellular solution (in mM): 135 NaCl, 5.8 KCl, 1.3 CaCl₂, 0.9 MgCl₂, 0.7 NaH₂PO₄, 5.6 D-glucose, 10 HEPES-NaOH. Sodium pyruvate (2 mM), MEM amino acids solution (50×, without L-glutamine) and MEM vitamins solution (100×) were added from concentrates (Fisher Scientific). The pH was adjusted to 7.5 (osmolality approximately 308 mmol kg⁻¹). The dissected cochleae were transferred to a microscope chamber, immobilized as previously described³⁹ and continuously perfused with a peristaltic pump using the above extracellular solution. The organs of Corti were viewed using an upright microscope (Nikon FN1) with Nomarski optics (60× objective).

MET currents were elicited by stimulating the hair bundles of P9 OHCs in the excitatory and inhibitory direction using a fluid jet from a pipette (tip diameter 8–10 μm) driven by a piezoelectric disc³⁹. The pipette tip of the fluid jet was positioned near to the bundles to elicit a maximal MET current. Mechanical stimuli were applied as 50 Hz sinusoids (filtered at 0.25 kHz, 8-pole Bessel) with driving voltages of ±40 V. MET currents were recorded with a patch pipette solution containing (in mM): 106 Cs-glutamate, 20 CsCl, 3 MgCl₂, 1 EGTA-CsOH, 5 Na₂ATP, 0.3 Na₂GTP, 5 HEPES-CsOH, 10 sodium phosphocreatine (pH 7.3). Membrane potentials were corrected for the liquid junction potential (-11 mV).

Patch clamp recordings were performed using an Optopatch (Cairn Research) amplifier. Patch pipettes were made from soda glass capillaries (Harvard Apparatus) and had a typical resistance in extracellular solution of 2–3 MΩ. To reduce the electrode capacitance, patch electrodes were coated with surf wax (Mr. Zog's SexWax). Potassium current recordings were performed at room temperature (22–24°C) and the intracellular solution contained (in mM): 131 KCl, 3 MgCl₂, 1 EGTA-KOH, 5 Na₂ATP, 5 HEPES-KOH, 10 Na₂-phosphocreatine (pH 7.3; osmolality approximately 296 mmol kg⁻¹). Data acquisition was controlled by pClamp software (version 10) using Digidata 1440A boards (Molecular Devices). Recordings were low-pass filtered at 2.5 kHz (8-pole Bessel), sampled at 5 kHz and stored on computer for off-line analysis (Origin, OriginLab). Membrane potentials in voltage clamp were corrected for the voltage drop across the uncompensated residual series resistance and for a liquid junction potential (-4 mV).

The presence of electromotile activity in P16–P18 OHCs was estimated by applying a depolarizing voltage step from the holding potential of -64 mV to +56 mV. Changes in cell length were viewed and recorded with a Nikon FN1 microscope (75× magnification) with a Flash 4.0 SCCD camera (Hamamatsu). Cell body movement was tracked using Fiji software. Lines were drawn across the basal membrane of patched OHCs, perpendicular to the direction of cell motion, and a projected time-based z-stack of the pixels under the line was made. Cell movement was measured with Photoshop as a pixel shift and then converted to nanometres (290 pixels = 10 μm).

Nonlinear (voltage-dependent) capacitance of IHCs in *Anc80-Ikzf2*-injected mice and their non-injected littermates was studied at P12–P16 using conventional whole-cell patch clamp recordings. Apical turn of the organ of Corti was carefully dissected in Leibovitz's L-15 cell culture medium (21083027, Gibco/ThermoFisher) containing the following inorganic salts (in mM): 137 NaCl, 5.4 KCl, 1.26 CaCl₂, 1.0 MgCl₂, 1.0 Na₂HPO₄, 0.44 KH₂PO₄ and 0.81 MgSO₄ and placed into a custom-made recording chamber, where it was held by two strands of dental floss. The organ of Corti explants were viewed with an upright microscope (BX51WIF, Olympus), equipped with a high numerical aperture (NA) objective (100×, 1.0

NA). To block voltage-gated ion channels in IHCs, the bath solution was made of L-15 medium supplemented with 10 mM tetraethylammonium-Cl, 2 mM CoCl₂, 10 mM CsCl and 0.1 mM nifedipine (all from Sigma), while the intrapipette solution contained (in mM): 140 CsCl, 2.5 MgCl₂, 2.5 Na₂ATP, 1.0 EGTA and 5 HEPES. During recordings, the organs of Corti were continuously perfused with the above extracellular bath solution. Whole-cell current responses were recorded with MultiClamp 700B patch clamp amplifier (Molecular Devices), controlled by jClamp software (SciSoft). Membrane capacitance was measured during the voltage ramp with a dual sinusoidal, FFT-based method⁴⁰. The recorded capacitance was fitted to the first derivative of a two-state Boltzmann function that is typically used to fit nonlinear capacitance of OHCs plus a small correction for the membrane area changes between expanded and contracted states of prestin⁴¹, as follows:

$C_m = C_v + C_{lin}$, in which C_m is the total membrane capacitance, C_v is a voltage-dependent (nonlinear) component, and C_{lin} is a voltage-independent (linear) component.

$$C_v = Q_{\max} \frac{ze}{kT} \frac{b}{(1+b)^2} + \frac{\Delta C_{sa}}{(1+b^{-1})}; b = \exp\left(\frac{-ze(V-V_{pk})}{kT}\right)$$

in which Q_{\max} is the maximum nonlinear charge moved, V_{pk} is a voltage at peak capacitance, V is membrane potential, z is valence, e is electron charge, k is Boltzmann's constant, T is absolute temperature, and ΔC_{sa} is the maximum increase in capacitance that occurs when all prestin molecules change from compact to expanded state. To account for some variability in sizes of IHCs, statistical data are shown as the maximum of voltage-dependent component of capacitance (C_v) normalized to the linear capacitance of the cell (C_v/C_{lin}).

DPOAEs. DPOAE tests were performed using frequency-specific tone-burst stimuli at 8, 16 and 32 kHz with the TDT RZ6 System 3 hardware and BioSig RZ (version 5.7.1) software (Tucker Davis Technology). An ER10B+ low noise probe microphone (Etymotic Research) was used to measure the DPOAE near the tympanic membrane. Tone stimuli were presented via separate MF1 (Tucker Davis Technology) speakers, with f_1 and f_2 at a ratio of $f_2/f_1 = 1.2$ (L1 = 65 dB SPL, L2 = 55 dB SPL), centred around the frequencies of 8, 16 and 32 kHz. Surgical anaesthesia was achieved by intraperitoneal injection of ketamine (100 mg ml⁻¹ at 10% v/v), xylazine (20 mg ml⁻¹ at 5% v/v) and acepromazine (2 mg ml⁻¹ at 8% v/v) administered at a rate of 0.1 ml per 10 g body mass. Once the required depth of anaesthesia was confirmed by the lack of the pedal reflex, a section of pinna was removed to allow unobstructed access to the external auditory meatus. Mice were then placed on a heated mat inside a sound-attenuated chamber (ETS-Lindgren) and the DPOAE probe assembly was inserted into the ear canal using a pipette tip to aid correct placement. In-ear calibration was performed before each test. The f_1 and f_2 tones were presented continuously and a fast-Fourier transform was performed on the averaged response of 356 epochs (each approximately 21 ms). The level of the $2f_1 - f_2$ DPOAE response was recorded and the noise floor calculated by averaging the four frequency bins either side of the $2f_1 - f_2$ frequency.

NanoString validation. Cochlear RNA extracted from biological triplicates of *Ikzf2^{cello/cell}*, *Ikzf2^{cello/+}* and *Ikzf2^{+/+}* animals at P8 were processed for NanoString validation at the UMSOM Institute for Genome Sciences using the nCounter Master Kit per manufacturer's instructions, and quantified using the NanoString nCounter platform. See Supplementary Table 13 for NanoString probe sequences. Data were analysed using nSolver 4.0 software (NanoString).

Anc80L65 AAV vector construction. The *Anc80L65-Myc-Ikzf2⁺* (*Anc80-Ikzf2*) expression vector was designed to drive expression of a Myc-tagged *Ikzf2* construct followed by a bovine Growth Hormone poly-adenylation (BGH pA) site under control of the cytomegalovirus (CMV) promoter. The *Anc80L65-eGFP* (*Anc80-eGFP*) expression construct also contained a Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) preceding the BGH pA site. *Anc80L65* AAV vectors^{17,18} were produced by the Gene Transfer Vector Core, Grousbeck Gene Therapy Center at the Massachusetts Eye and Ear Infirmary (<http://vector.meei.harvard.edu/>).

Inner ear gene delivery. For *in vivo* hair cell transductions, mice were injected with *Anc80L65* AAVs between P1 and P3 via the posterior semicircular canal using the injection method described previously⁴². In brief, animals were anesthetized on ice before a post-auricular incision was made on either the left or right side. Tissues were further dissected to reveal the posterior semicircular canal, and a Nanolitre 2010 microinjection system (World Precision Instruments) equipped with a loaded glass needle was used to inject 700 nl of 1.13×10^{13} genome copies (GC) per ml *Anc80-Ikzf2* or 500 nl of 4.85×10^{12} GC per ml *Anc80-eGFP*. Injections into the inner ear were performed in 50 nl increments over the course of 2 min. The needle was then removed, the incision sutured, and animals were placed on a 37°C heating pad to recover before being returned to their cage.

FACS. For the scRNA-seq analysis of *Anc80-Ikzf2* transduced hair cells, inner ears of neonatal *Myo15^{cre/+};ROSA26^{CAG-tdTomato}* mice were injected with *Anc80-Ikzf2* (4 mice) or control *Anc80-eGFP* (2 mice) via the posterior semicircular canal.

Cochlear tissues from both injected and uninjected ears were obtained at P8 and further dissected to reveal the sensory epithelium. Inclusion of the uninjected ear in the single cell analysis allowed for the study of changes in gene expression that occur in response to a gradient of transgene expression. This is because, in mice, inner ear gene delivery often results in transduction in the contralateral ear, albeit at a lower intensity¹⁸. Cochlear tissues were then dissociated for fluorescence activated cell sorting (FACS) using the method described previously⁹. In brief, the sensory epithelia from *Anc80-eGFP*- and *Anc80-Ikzf2*-injected mice were pooled separately into two wells of a 48-well plate containing 0.5 mg ml⁻¹ thermolysin (Sigma). Tissues were incubated at 37 °C for 20 min, after which the thermolysin was removed and replaced with accutase enzyme (MilliporeSigma). After a 3-min incubation at 37 °C, tissues were mechanically disrupted using a 23G blunt ended needle connected to a 1 ml syringe. This step was performed twice. After confirming tissue dissociation by direct visualization, the dissociation reaction was stopped by adding an equal volume of IMDM supplemented with 10% heat-inactivated FBS to the Accutase enzyme solution. Cells were passed through a 40 mm cell strainer (BD) to remove cell clumps. tdTomato-expressing hair cells were sorted into ice-cold tubes containing IMDM with 10% FBS on a BD FACSAria II (BD Biosciences) and processed for scRNA-seq. Flow cytometry analyses were performed with assistance from X. Fan at the University of Maryland Marlene and Stewart Greenebaum Comprehensive Cancer Center Flow Cytometry Shared Service.

scRNA-seq. tdTomato-positive sorted hair cells were pelleted once (300g at 4 °C) and resuspended in a minimal remaining volume (around 30 µl). Hair cell-enriched single-cell suspensions were then used as input on the 10× Genomics Chromium platform with 3' Single Cell v2 chemistry (10× Genomics). After capture and library preparation, scRNA-seq libraries were sequenced on a NextSeq 500 (Illumina) in collaboration with the NIDCD Genomics and Computational Biology Core. Samples were sequenced to an average depth of over 300,000 reads per cell, which resulted in detection of a median of >3,000 genes (*Anc80-eGFP*) and >4,000 genes (*Anc80-Ikzf2*) per cell, ensuring maximal transcriptional complexity and detection of low-abundance transcripts (see Extended Data Fig. 9b, c). Reads were aligned to a modified mm10 mouse reference containing the sequences for the *Ai14* locus, as well as *Anc80-eGFP* and *Anc80-Ikzf2* viral sequences (Extended Data Fig. 9a) using the 10× Genomics Cell Ranger (version 2.0.2) package to generate the read counts matrix files. Read counts from viral and *Ai14* loci were removed from the expression matrix before dimensionality reduction so as to not influence data clustering. Cells from these hair cell clusters were determined to be *Anc80-Ikzf2*(+) versus *Anc80-Ikzf2*(−), and IHCs versus OHCs, based on their expression of *Anc80-Ikzf2* and *Slc17a8*, respectively (Fig. 3, Extended Data Figs. 8 and 9, Supplementary Table 9). *Slc26a5* was not well detected in the scRNA-seq dataset and was therefore not used as an OHC marker. After clustering, four hair cells were excluded based on co-expression of a contaminating cell type. Secondary analyses, including shared nearest neighbour (SNN) clustering, t-SNE embedding, and differential expression testing (using either Wilcoxon ranked sum for marker gene identification or MAST for pairwise comparison between control IHCs and OHCs) were performed in R with Seurat (version 2.1.0)^{43,44}. Non-parametric analysis of variance between the four classified groups of HCs (IHCs and OHCs with either high or low *Anc80-Ikzf2* expression) using a Kruskal–Wallis test was performed to help qualify genes that had statistical difference across these cell populations. This was followed by post hoc pairwise Wilcoxon ranked sum comparisons to assess multiple-comparison-adjusted *P* values. Additional plots were generated by NMF (version 0.20.6) and ggplot2 (version 2.2.1)^{45,46}. These analyses used the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

Immunohistochemistry of AAV-injected cochleae. Mouse inner ears injected with either *Anc80-Ikzf2* or *Anc80-eGFP* were between P8 and 8 weeks, fixed in 4% PFA in PBS overnight at 4 °C, and decalcified in a solution of 5% EDTA in RNAlater (Invitrogen). Decalcified ears were processed by sucrose gradient and embedded in OCT compound (Tissue-Tek) for cryosectioning, or fine dissected for whole-mount immunohistochemistry. Cryosections (10 µm) on positively charged glass slides were used for *in situ* hybridization (ISH) and section immunohistochemistry. For whole-mount immunolabelling at 6–8 weeks, hair cell loss was observed in the injected ear and therefore the contralateral ear, expressing a lower level of the *Anc80-Ikzf2* virus, was used. Primary antibodies: goat anti-prestin N-20 (1:200, Santa Cruz Biotechnology); goat anti-oncomodulin N-19 (1:100, Santa Cruz Biotechnology); rabbit anti-myosinVI (1:1,000, Proteus BioSciences); rabbit anti-GFP (1:100, Life Technologies); mouse anti-cMyc 9E10 (1:100, Santa Cruz Biotechnology) and mouse anti-otoferlin (1:100, Abcam). The guinea pig anti-VGLUT3 antibody (1:5,000) used in this study was donated by R. Seal. Corresponding Alexa Fluor 488 and 546 (1:800, Invitrogen) were used for secondary detection, Alexa Fluor 488 Phalloidin (1:1,000, Invitrogen) was used to mark F-actin, and DAPI (1:20,000, Thermo Fisher) was used to mark cell nuclei. Images were acquired using a Nikon Eclipse E600 microscope (Nikon) equipped with a Lumenera Infinity 3 camera. Whole-mount images were acquired using a Zeiss

LSM DUO confocal microscope, located at the UMSOM Confocal Microscopy Core, at 63× oil magnification. Images were processed using Infinity Capture and Infinity Analyze software (Lumenera), and ImageJ software.

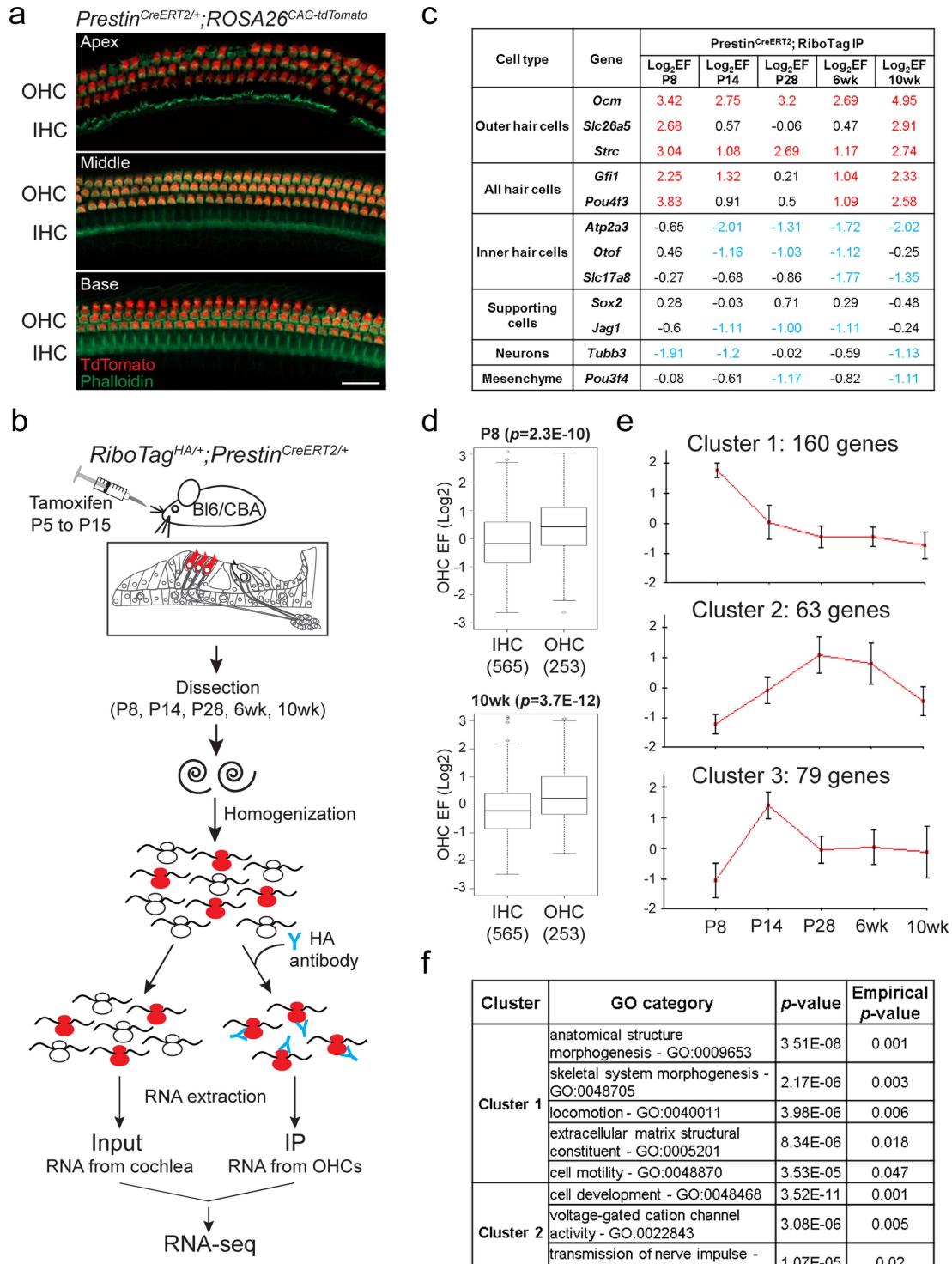
RNA *in situ* hybridization. *In situ* hybridization was performed as described previously⁴⁷. In brief, slides were re-fixed in 4% PFA, and then treated with 2 µg ml⁻¹ Proteinase-K for 10 min. Proteinase K reaction was stopped by soaking slides again in 4% PFA, followed by acetylation and permeabilization. Hybridization for the digoxigenin labelled *Fcrlb* probe was performed overnight at 65 °C (see Supplementary Table 13 for *Fcrlb* probe primers). After a series of washes in saline sodium citrate, slides were incubated with sheep-anti-digoxigenin antibody conjugated to alkaline phosphatase (Sigma-Aldrich, 1:100) overnight at 4 °C. Slides were then incubated in BM purple AP substrate precipitating solution (Roche) to localize bound anti-digoxigenin antibody.

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

Data availability

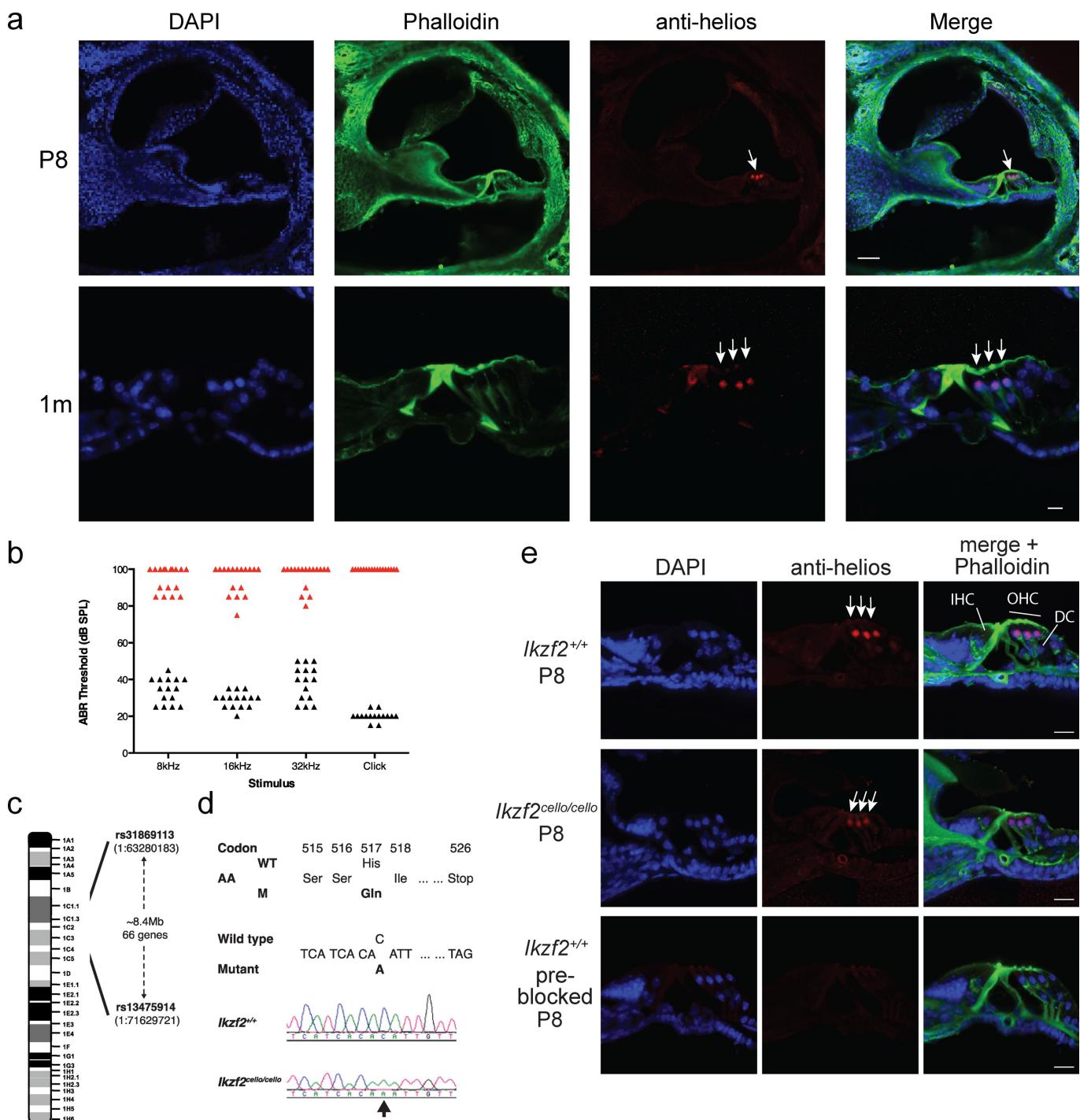
The RiboTag OHC RNA-seq, P8 cello cochlea RNA-seq, and P8 *Anc80-Ikzf2* and *Anc80-eGFP* injected cochlea scRNA-seq data have been submitted to the Gene Expression Omnibus (GEO) database under accessions GSE116703, GSE116702 and GSE120462, and are also available for viewing through the gEAR portal (<https://umgear.org/>).

23. Caberlotto, E. et al. Usher type 1G protein sans is a critical component of the tip-link complex, a structure controlling actin polymerization in stereocilia. *Proc. Natl. Acad. Sci. USA* **108**, 5825–5830 (2011).
24. Johnson, K. R., Zheng, Q. Y. & Noben-Trauth, K. Strain background effects and genetic modifiers of hearing in mice. *Brain Res.* **1091**, 79–88 (2006).
25. Hoelter, S. M. et al. "Sighted C3H" mice—a tool for analysing the influence of vision on mouse behaviour? *Front. Biosci.* **13**, 5810–5823 (2008).
26. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
27. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
28. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
29. Ulitsky, I. et al. Expander: from expression microarrays to networks and functions. *Nat. Protocols* **5**, 303–322 (2010).
30. Sharan, R., Maron-Katz, A. & Shamir, R. CLICK and EXPANDER: a system for clustering and visualizing gene expression data. *Bioinformatics* **19**, 1787–1799 (2003).
31. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
32. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
33. Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the functional effect of amino acid substitutions and indels. *PLoS One* **7**, e46688 (2012).
34. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protocols* **4**, 1073–1081 (2009).
35. Källberg, M. et al. Template-based protein structure modeling using the RaptorX web server. *Nat. Protocols* **7**, 1511–1522 (2012).
36. Hardisty-Hughes, R. E., Parker, A. & Brown, S. D. M. A hearing and vestibular phenotyping pipeline to identify mouse mutants with hearing impairment. *Nat. Protocols* **5**, 177–190 (2010).
37. Mianné, J. et al. Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. *Genome Med.* **8**, 16 (2016).
38. Codner, G. F. et al. Application of long single-stranded DNA donors in genome editing: generation and validation of mouse mutants. *BMC Biol.* **16**, 70 (2018).
39. Corns, L. F., Johnson, S. L., Kros, C. J. & Marcotti, W. Calcium entry into stereocilia drives adaptation of the mechanoelectrical transducer current of mammalian cochlear hair cells. *Proc. Natl. Acad. Sci. USA* **111**, 14918–14923 (2014).
40. Santos-Sacchi, J. Determination of cell capacitance using the exact empirical solution of partial $\delta Y/\delta C_m$ and its phase angle. *Biophys. J.* **87**, 714–727 (2004).
41. Santos-Sacchi, J. & Navarrete, E. Voltage-dependent changes in specific membrane capacitance caused by prestin, the outer hair cell lateral membrane motor. *Pflügers Arch.* **444**, 99–106 (2002).
42. Isgrig, K. et al. Gene therapy restores balance and auditory functions in a mouse model of usher syndrome. *Mol. Ther.* **25**, 780–791 (2017).
43. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
44. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
45. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2009).
46. Gaujoux, R. & Seoighe, C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* **11**, 367 (2010).
47. Geng, R. et al. Comprehensive expression of Wnt signaling pathway genes during development and maturation of the mouse cochlea. *PLoS One* **11**, e0148339 (2016).



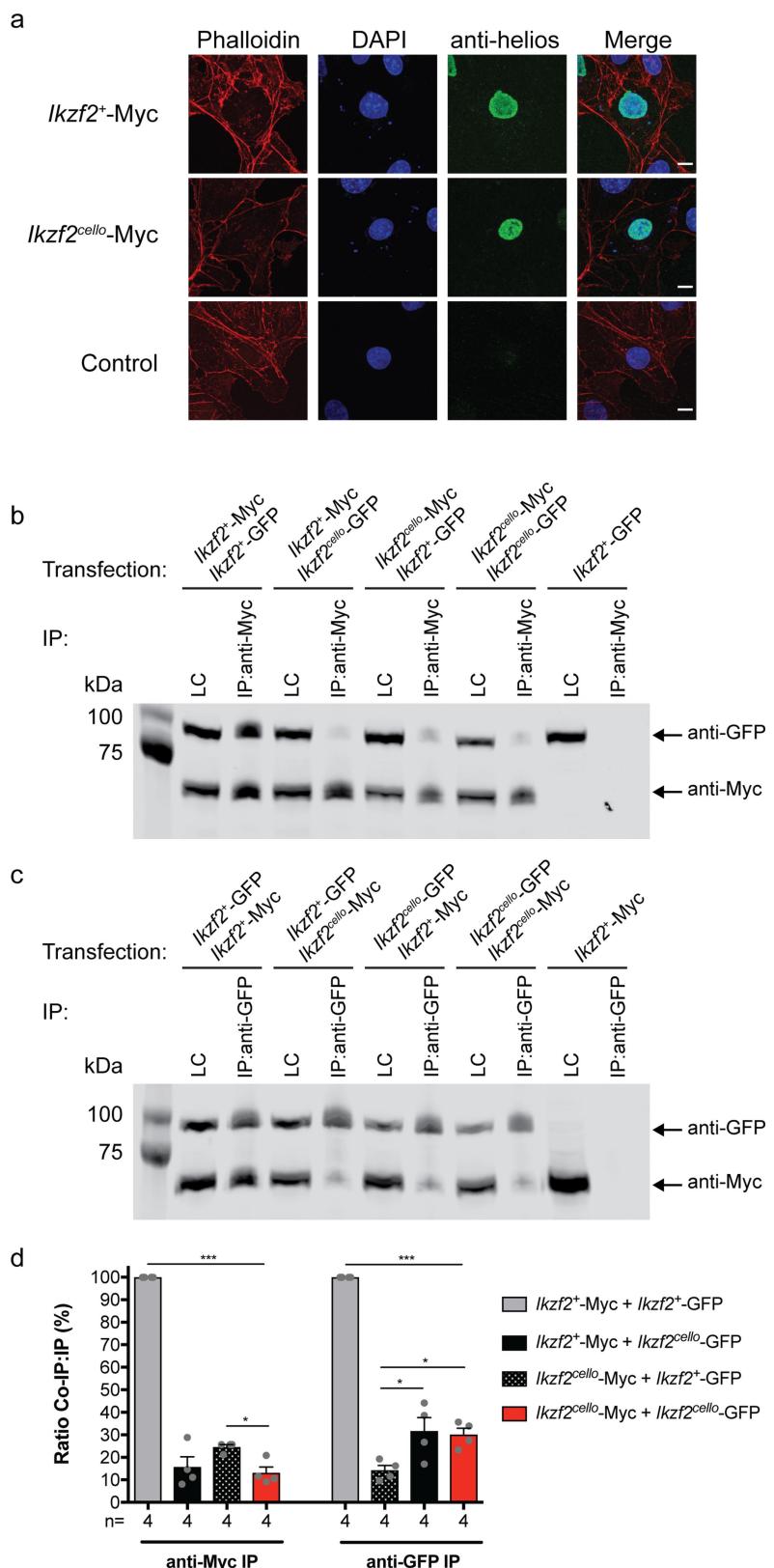
Extended Data Fig. 1 | RiboTag immunoprecipitation enriches for known OHC-expressed transcripts. **a**, Representative *prestin*^{CreERT2/+}; ROSA26^{CAG-tdTomato} cochlear whole-mount. The *prestin*^{CreERT2}-driven tdTomato expression is OHC-specific at P21 ($n = 1$). Scale bar, 20 μ m. **b**, Schematic of the RiboTag immunoprecipitation protocol. Red OHCs represent Cre/HA-tagged ribosome expression. **c**, RiboTag RNA-seq log₂ enrichment and depletion of transcripts for known inner ear cell type markers (EF = log₂(IP/input)). **d**, Genes at least two-fold enriched in IHCs ($n = 565$ genes) or OHCs ($n = 253$ genes) in the published dataset⁸ are significantly depleted or enriched, respectively, by the RiboTag OHC immunoprecipitation at all time points examined

(two-sided Wilcoxon's test). Black line represents median EF, box demarcates first and third quartiles, whiskers demarcate first and third quartiles $\pm 1.5 \times$ IQR values, dots represent single outliers. **e**, Clustering of genes differentially expressed across OHC postnatal development. Error bars denote s.d. Before clustering, expression levels were standardized to mean = 0 and s.d. = 1. **f**, Enriched Gene Ontology (GO) functional categories identified for the gene clusters in **e** (cluster 1 $n = 160$ genes, cluster 2 $n = 63$ genes). No significantly enriched GO categories were found for cluster 3 ($n = 79$ genes). Enrichment and statistical analyses were performed using the EXPANDER implemented tool TANGO.



Extended Data Fig. 2 | Auditory phenotyping, SNP mapping and whole-genome sequencing of mouse pedigree MPC173, subsequently named *cello*. **a**, Specific expression of Helios can be seen in the nuclei of wild-type P8 OHCs (white arrow), and is maintained in wild-type OHCs at 1 month (white arrows). Scale bars, 50 μ m (P8) and 10 μ m (1 month). $n = 3$ biologically independent samples for each time point. **b**, Auditory brainstem response phenotyping of pedigree MPC173 at 9 months of age identified 17 biologically independent animals with increased hearing thresholds (red triangles) compared to their normal hearing colony mates (black triangles). $n = 15$ biologically independent animals. **c**, The mutation mapped to an 8.4-Mb region on chromosome 1 between single nucleotide polymorphism (SNP) rs31869113 and

rs13475914 (chr1: 63280183–71629721), containing 66 genes. **d**, Detection of a non-synonymous mutation in *cello*. DNA sequencing identified a nucleotide transversion (c.1551C>A) in the *Ikzf2* gene at codon 517, thus altering the wild-type sequence CAC, encoding histidine, to the mutant (M) sequence CAA, encoding glutamine. Electropherograms derived from a *cello* mutant mouse (*Ikzf2*^{cello/cello}) and a wild-type colony mate (*Ikzf2*^{+/+}) control showing the sequence surrounding *Ikzf2* nucleotide 1551 (indicated by an arrow). **e**, Helios is expressed in the OHC nuclei of both *Ikzf2*^{+/+} and *Ikzf2*^{cello/cello} mice at P8. $n = 3$ biologically independent samples per genotype. Loss of labelling when the anti-Helios antibody is ‘pre-blocked’ confirms specificity. $n = 1$ biologically independent sample. Scale bars, 20 μ m. DC, Deiters’ cells.



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | The *Ikzf2^{cello}* mutation disrupts

homodimerization of Helios. **a,** Cos-7 cells transfected with *Ikzf2⁺*- or *Ikzf2^{cello}*-Myc. Nuclear localization is unaffected by the *Ikzf2^{cello}* mutation. $n = 2$ biologically independent experiments. Scale bars, 10 μm .

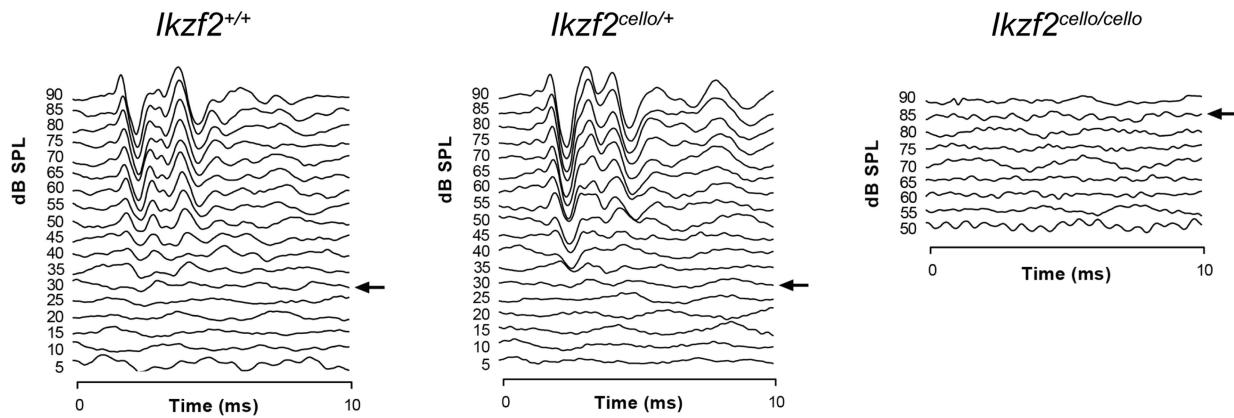
b, Co-immunoprecipitation (IP) of Myc-tagged (62 kDa) and GFP-tagged (88 kDa) *Ikzf2⁺* and *Ikzf2^{cello}* constructs. Transfected cell lysates were immunoprecipitated using an anti-Myc antibody and analysed by western blotting with both anti-Myc and anti-GFP antibodies. Results show that wild-type *Ikzf2⁺* Helios can dimerize, but that dimerization is impaired by the *cello* mutation. LC, cell lysate loading control.

c, Reciprocal immunoprecipitation reactions using an anti-GFP antibody confirm dimerization of wild-type *Ikzf2⁺* Helios and reduced dimerization of mutant *Ikzf2^{cello}* Helios. **d,** Quantification of co-immunoprecipitation western blots. Band intensities were determined and used to calculate

the relative ratio of the co-immunoprecipitation to immunoprecipitation signal. $n = 4$ biologically independent experiments. Data are mean \pm s.e.m. Anti-Myc IP: *** $P < 0.0001$ (*Ikzf2⁺*-Myc + *Ikzf2⁺*-GFP vs *Ikzf2⁺*-Myc + *Ikzf2^{cello}*-GFP), vs *Ikzf2^{cello}*-Myc + *Ikzf2⁺*-GFP and vs *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP). * $P = 0.0476$ (*Ikzf2^{cello}*-Myc + *Ikzf2⁺*-GFP vs *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP). $P = 0.1488$ (*Ikzf2⁺*-Myc + *Ikzf2^{cello}*-GFP vs *Ikzf2^{cello}*-Myc + *Ikzf2⁺*-GFP). $P = 0.9020$ (*Ikzf2⁺*-Myc + *Ikzf2^{cello}*-GFP vs *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP). Anti-GFP IP: *** $P < 0.0001$ (*Ikzf2⁺*-Myc + *Ikzf2⁺*-GFP vs *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP), * $P = 0.0202$ (*Ikzf2^{cello}*-Myc + *Ikzf2⁺*-GFP vs *Ikzf2⁺*-Myc + *Ikzf2^{cello}*-GFP) * $P = 0.0346$ (*Ikzf2^{cello}*-Myc + *Ikzf2⁺*-GFP vs *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP). $P = 0.9894$ (*Ikzf2⁺*-Myc + *Ikzf2^{cello}*-GFP vs + *Ikzf2^{cello}*-Myc + *Ikzf2^{cello}*-GFP) (one-way ANOVA with Tukey post hoc test). See Supplementary Fig. 1 for source images.

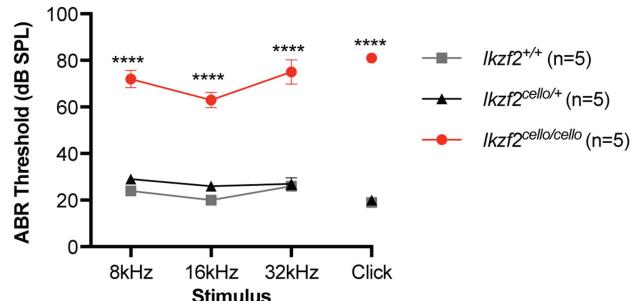
a

P16 Click ABR waveforms



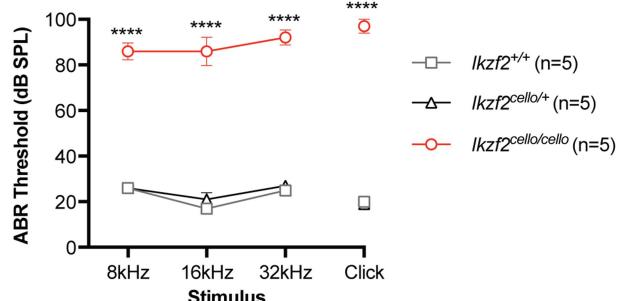
b

1-month ABR

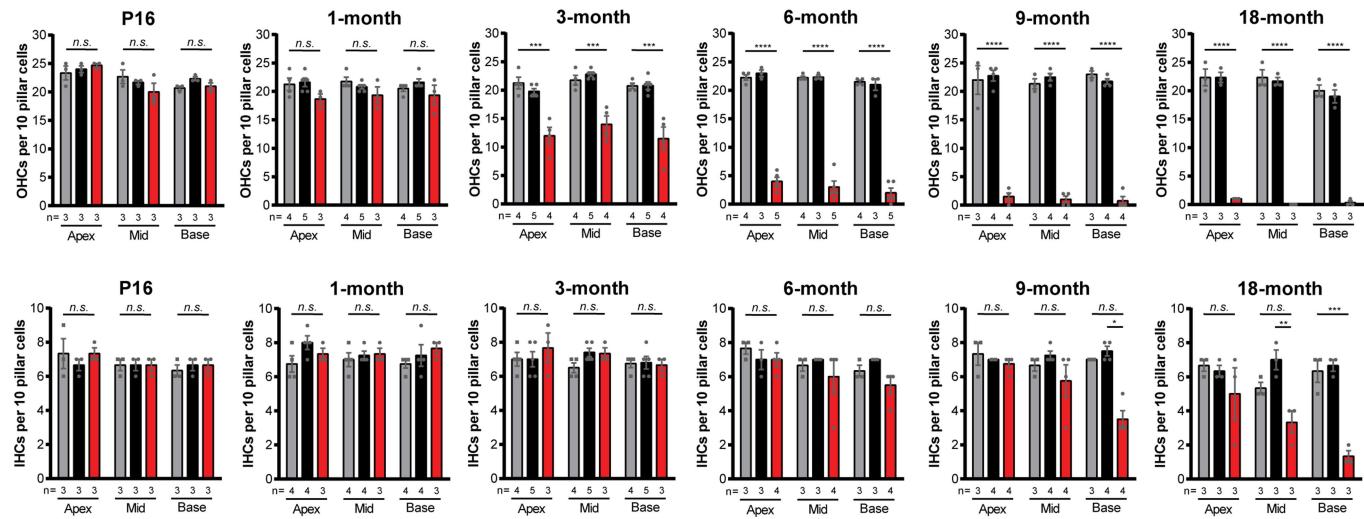


c

9-month ABR

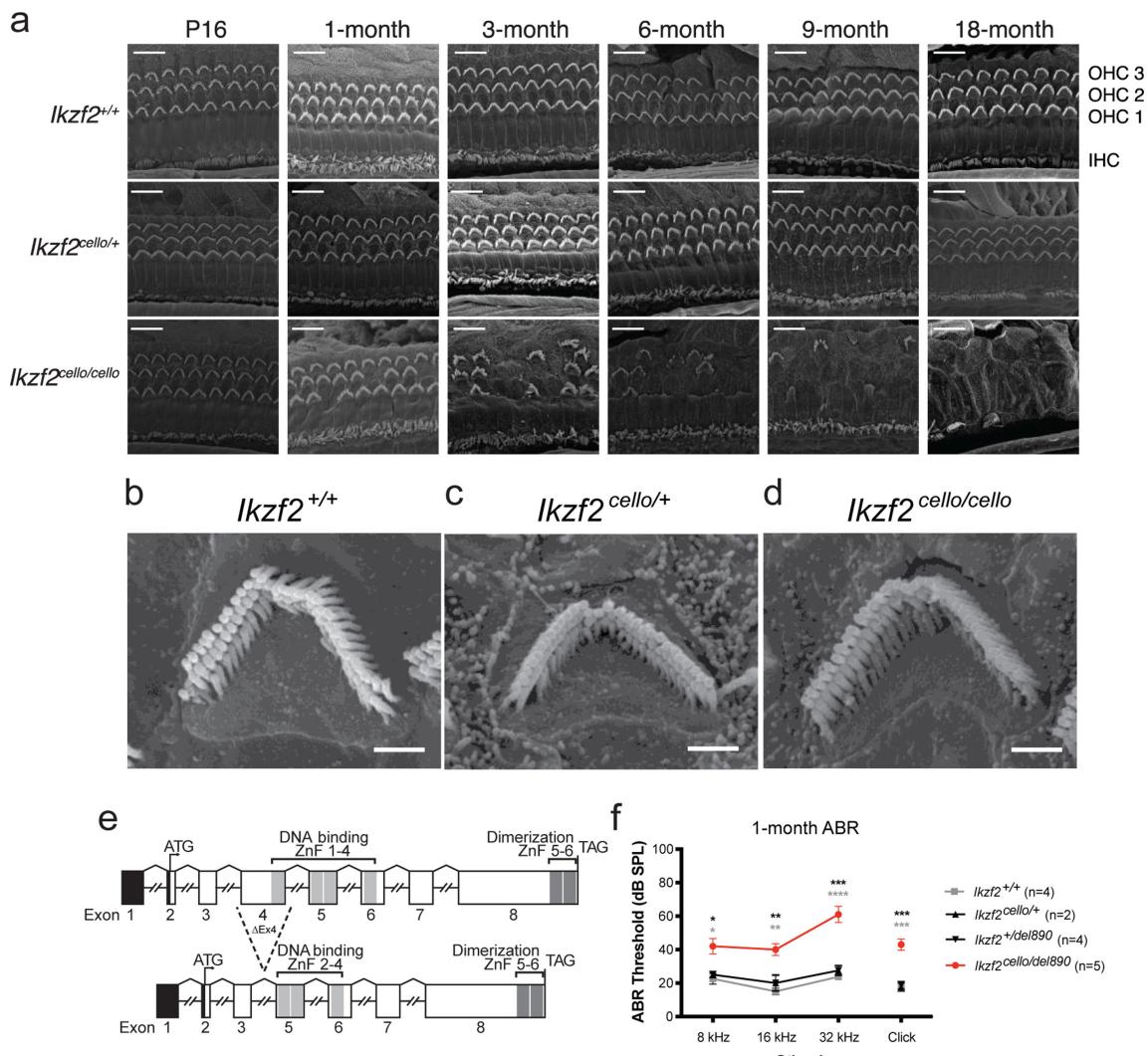


d



Extended Data Fig. 4 | Auditory function and HC bundle survival in *cello* mice. a, Representative click ABR waveforms for *Iκzf2^{+/+}*, *Iκzf2^{cello/+}* and *Iκzf2^{cello/cello}* littermates at P16. $n = 4$ biologically independent animals per genotype. b, c, Averaged ABR thresholds for *cello* mice at 1-month of age (b) and 9 months of age (c). Age-matched *Iκzf2^{+/+}* and *Iκzf2^{cello/+}* controls display thresholds within the expected range (15–30 dB SPL) at all time points tested. $n = 5$ biologically independent animals per genotype. Data are mean thresholds \pm s.e.m. 1-month *Iκzf2^{cello/cello}* vs *Iκzf2^{+/+}*: *** $P < 0.0001$ (8 kHz, 16 kHz, 32 kHz, click). 1-month *Iκzf2^{cello/cello}* vs 1 *Iκzf2^{cello/+}*: *** $P < 0.0001$ (8 kHz, 16 kHz, 32 kHz, click). 9-month

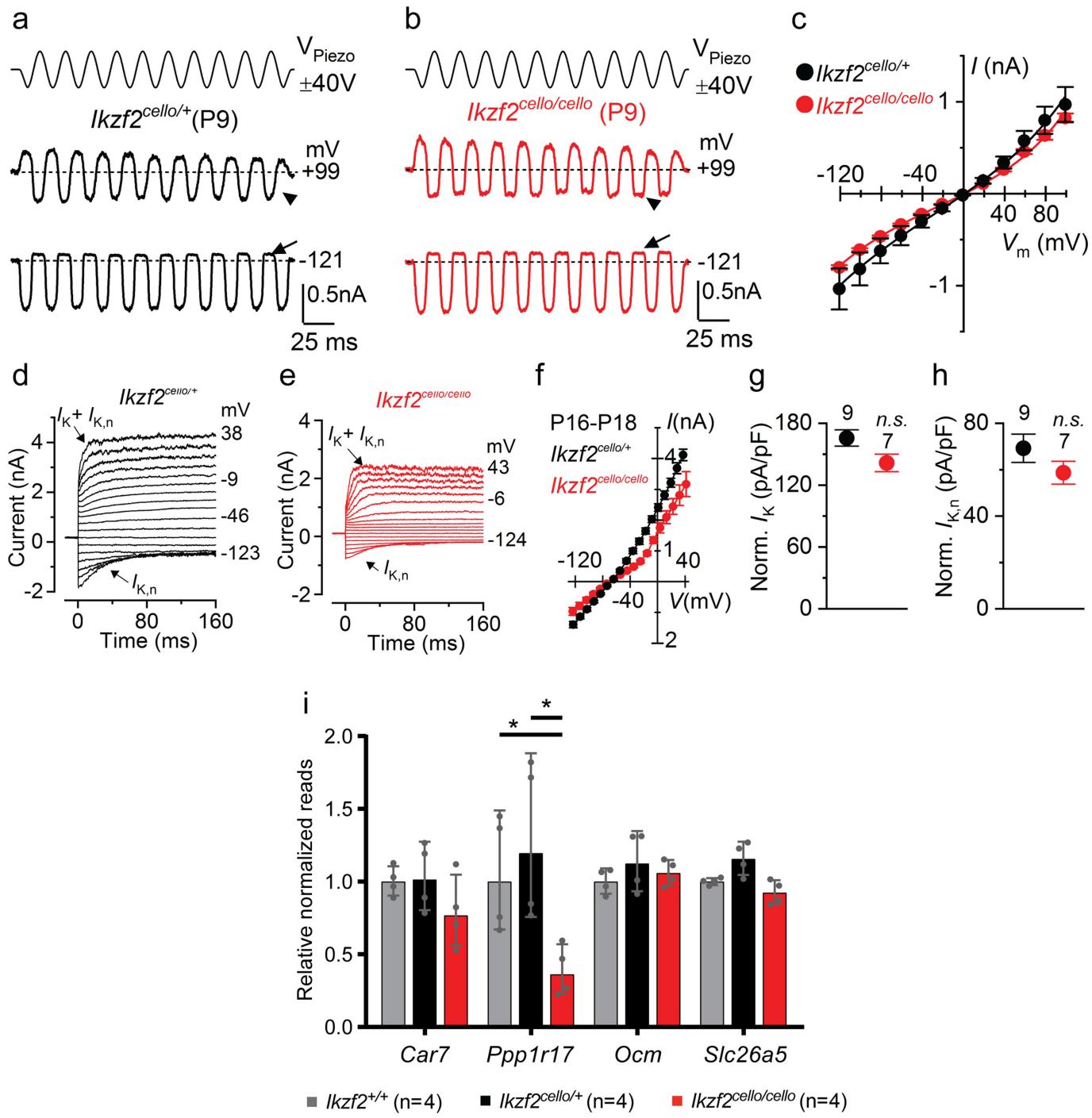
Iκzf2^{cello/cello} vs *Iκzf2^{+/+}*: *** $P < 0.0001$ (8 kHz, 16 kHz, 32 kHz, click). 9-month *Iκzf2^{cello/cello}* vs *Iκzf2^{cello/+}*: *** $P < 0.0001$ (8 kHz, 16 kHz, 32 kHz, click) (one-way ANOVA with Tukey post hoc test). d, OHC and IHC bundle counts for *cello* mice from P16 to 18 months of age. Grey, *Iκzf2^{+/+}*; black, *Iκzf2^{cello/+}*; red, *Iκzf2^{cello/cello}*. Data are mean \pm s.e.m. n.s., non-significant. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ (one-way ANOVA with Tukey post hoc test). Number of biologically independent samples for OHC and IHC bundle counts are shown. See also Supplementary Table 5 and 6.



Extended Data Fig. 5 | Scanning electron microscopy of *cello* mice and auditory function of *Ikzf2*^{cello/del890} compound heterozygotes.

a, Scanning electron micrographs of the organ of Corti of *cello* mice from P16 to 18 months of age. Representative images from the mid-region of the cochlear spiral are shown. Scale bars, 10 μ m. $n = 3$ (P16 *Ikzf2*^{cello/+}, P16 *Ikzf2*^{cello/cello}, 1-m *Ikzf2*^{cello/cello}, 9-m *Ikzf2*^{+/+}, 18-m *Ikzf2*^{+/+}, 18-m *Ikzf2*^{cello/+}, 18-m *Ikzf2*^{cello/cello}), $n = 4$ (P16 *Ikzf2*^{+/+}, 1-m *Ikzf2*^{+/+}, 3-m *Ikzf2*^{+/+}, 3-m *Ikzf2*^{cello/+}, 6-m *Ikzf2*^{+/+}, 6-m *Ikzf2*^{cello/cello}, 9-m *Ikzf2*^{cello/+}, 9-m *Ikzf2*^{cello/cello}) and $n = 5$ (1-m *Ikzf2*^{cello/+}, 3-m *Ikzf2*^{cello/+}, 6-m *Ikzf2*^{cello/+}) biologically independent samples. **b-d**, Scanning electron micrographs of OHC stereocilia bundles of *cello* mice at P16, showing that wild-type *Ikzf2*^{+/+} (**b**), *Ikzf2*^{cello/+} (**c**) and mutant *Ikzf2*^{cello/cello} (**d**) mice display overall expected bundle patterning. Images are from the mid-region of the cochlear spiral. Scale bars, 1 μ m. $n = 3$ biologically independent samples for each genotype. **e**, The genomic and domain

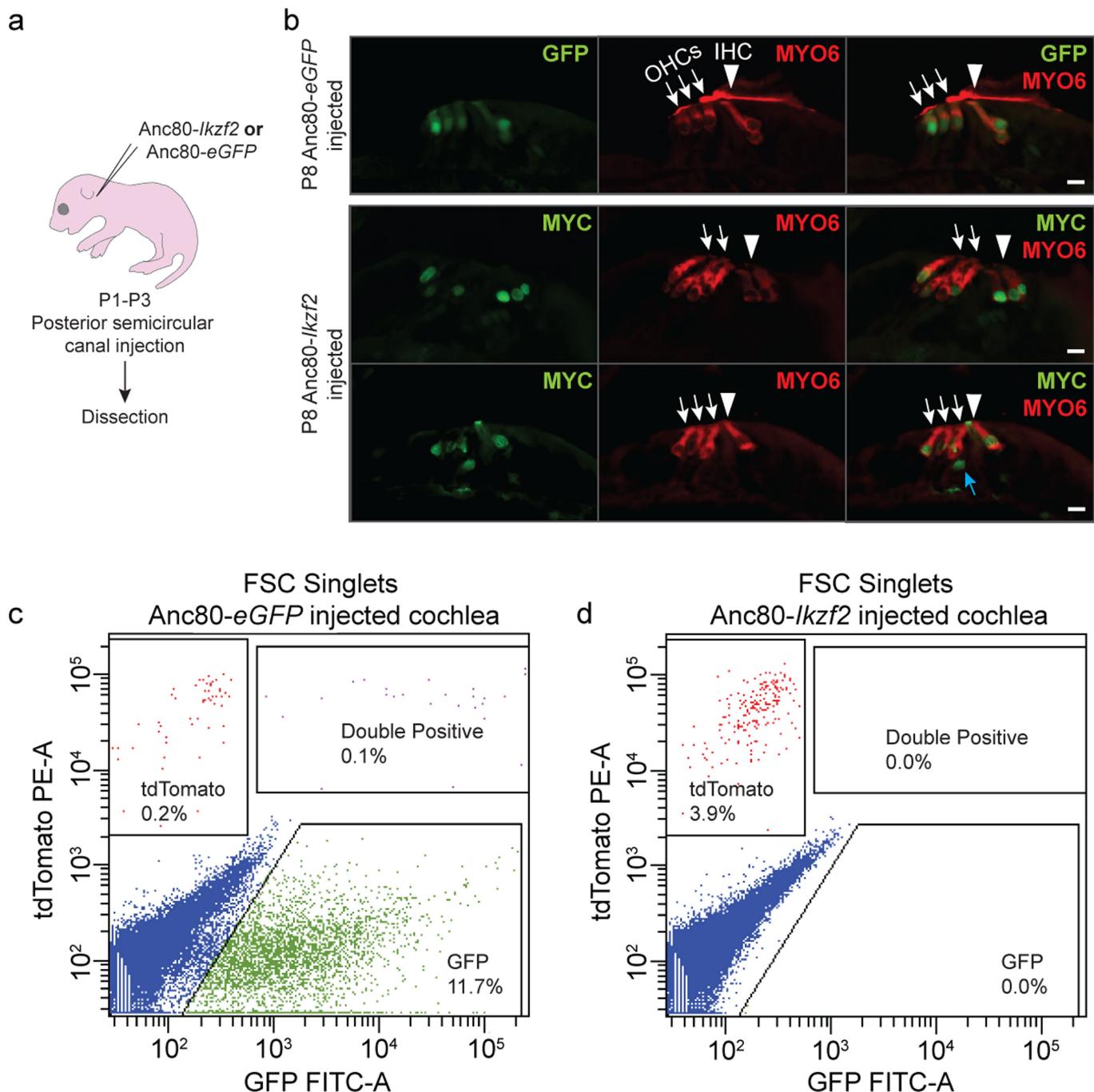
structure of *Ikzf2*^{del890}. Black, 5' untranslated region; light grey, N-terminal DNA-binding domain; dark grey, C-terminal dimerization domain. The *Ikzf2*^{cello} mutation lies in ZnF6. The *del890* mutation deletes exon 4 and the surrounding intronic sequence. **f**, Averaged ABR thresholds for *Ikzf2*^{cello/del890} compound heterozygotes at 1 month of age, showing increased thresholds (≥ 40 dB SPL) at all frequencies tested compared to *Ikzf2*^{+/+}, *Ikzf2*^{cello/+} and *Ikzf2*^{+/+del890} control colony mates. Data are mean \pm s.e.m. $n = 4$ (*Ikzf2*^{+/+}, *Ikzf2*^{+/+del890}), $n = 2$ (*Ikzf2*^{cello/+}) and $n = 5$ (*Ikzf2*^{cello/del890}) biologically independent samples. *Ikzf2*^{cello/del890} vs *Ikzf2*^{+/+}: * $P = 0.011$ (8 kHz), ** $P = 0.002$ (16 kHz), *** $P < 0.0001$ (32 kHz), *** $P = 0.0001$ (click); *Ikzf2*^{cello/del890} vs *Ikzf2*^{cello/+}: $P = 0.078$ (8 kHz), * $P = 0.034$ (16 kHz), ** $P = 0.001$ (32 kHz), ** $P = 0.001$ (click); *Ikzf2*^{cello/del890} vs *Ikzf2*^{+/+del890}: * $P = 0.025$ (8 kHz), ** $P = 0.009$ (16 kHz), *** $P = 0.0002$ (32 kHz), *** $P = 0.0002$ (click) (one-way ANOVA with Tukey post hoc test).



Extended Data Fig. 6 | See next page for caption.

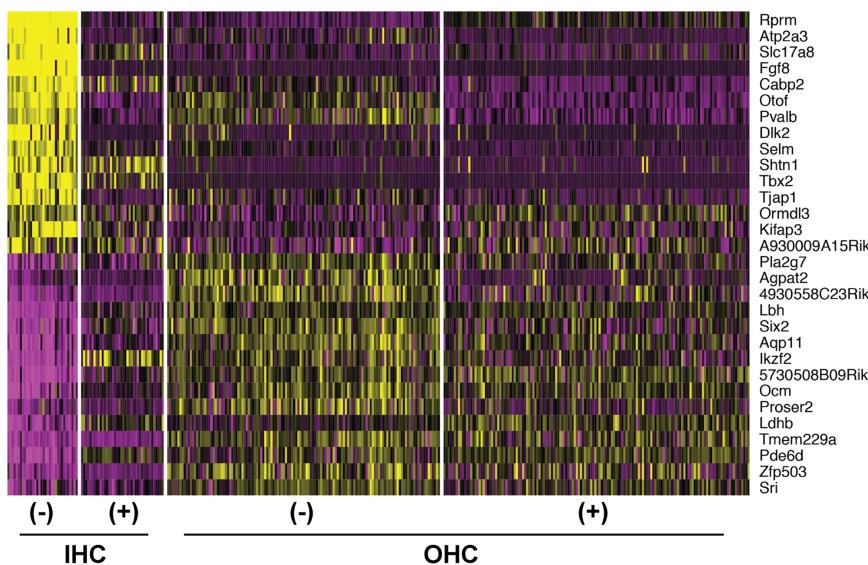
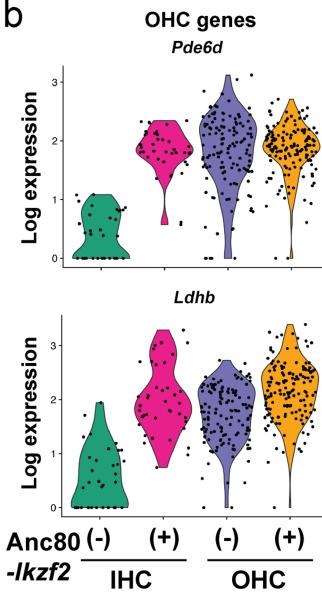
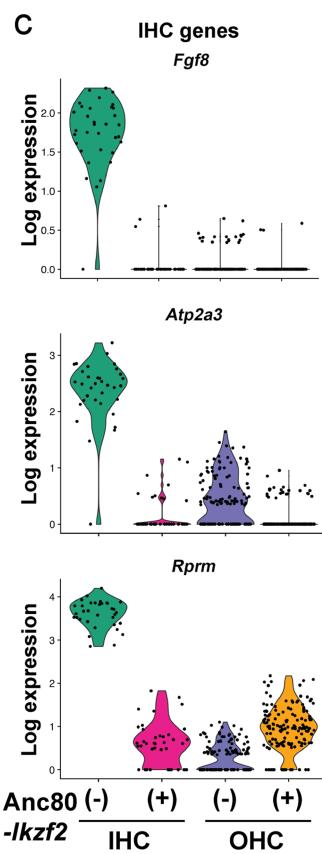
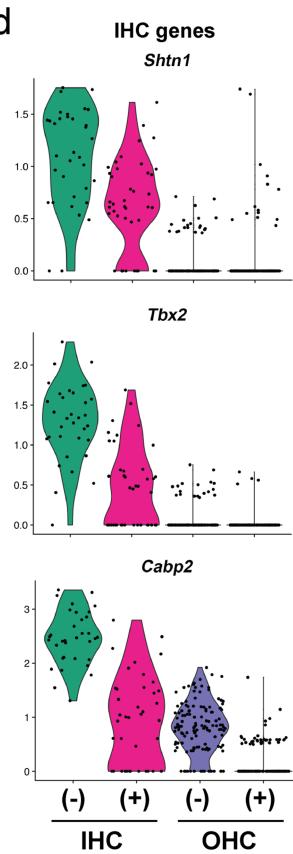
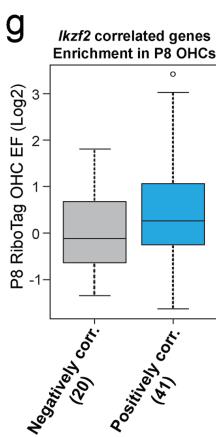
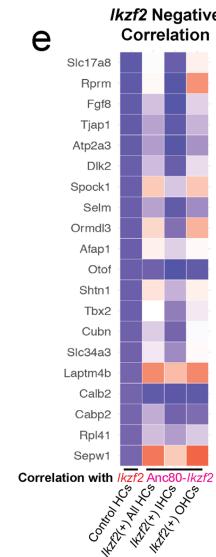
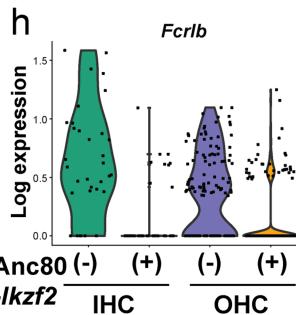
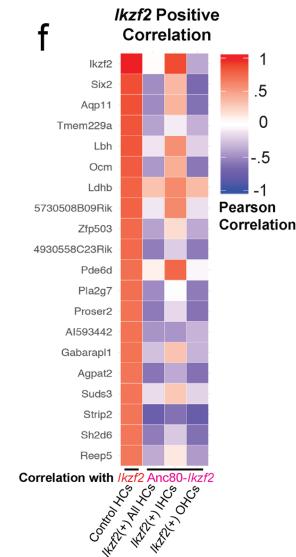
Extended Data Fig. 6 | The MET and adult-like potassium currents are normal in *Ikzf2^{cello}* mice. **a, b**, MET currents were recorded from OHCs of P9 *Ikzf2^{cello/cello}* and *Ikzf2^{cello/+}* (control) littermates. During voltage steps, hair bundles were displaced by applying a 50-Hz sinusoidal force stimuli (the driver voltage to the fluid jet is shown above the traces)³⁹. At hyperpolarized membrane potentials (-121 mV), saturating excitatory bundle stimulation (that is, towards the taller stereocilia) elicited a large inward MET current from both *Ikzf2^{cello/+}* and *Ikzf2^{cello/cello}* OHCs, whereas inhibitory bundle stimulation (that is, away from the taller stereocilia) closed the MET channels and reduced the resting current. Because the MET current reverses near 0 mV, it became outward when excitatory bundle stimulation was applied during voltage steps positive to its reversal potential. At positive membrane potentials ($+99$ mV), excitatory bundle stimulation now elicited similar outward MET currents with larger resting amplitudes. Arrows indicate closure of the MET channels (that is, disappearance of the resting current) during inhibitory bundle displacements, arrowheads indicate the larger resting MET current at $+99$ mV compared to -121 mV. **c**, Peak-to-peak current–voltage curves obtained from *Ikzf2^{cello/+}* ($n = 10$ biologically independent samples) and *Ikzf2^{cello/cello}* ($n = 8$ biologically independent samples) OHCs at P9.

The maximal MET current and the resting open probability of the MET channel were found to be similar between the two genotypes. Data are mean \pm s.e.m. **d, e**, Total K⁺ currents recorded from P18 *Ikzf2^{cello/+}* control (**d**) and *Ikzf2^{cello/cello}* mutant (**e**) OHCs. The size of the K⁺ current, which is mainly due to the negatively activated $I_{K,n}$ (in addition to a small delayed rectifier I_K^{15}), was smaller in *Ikzf2^{cello/cello}* OHCs. **f**, Average peak current–voltage relationship for the total K⁺ current recorded from the OHCs of *Ikzf2^{cello/+}* ($n = 9$ OHCs from 6 biologically independent animals) and *Ikzf2^{cello/cello}* ($n = 7$ OHCs from 5 biologically independent animals) mice at P16–P18. Data are mean \pm s.e.m. **g, h**, After normalization to the significantly reduced surface area of *Ikzf2^{cello/cello}* OHCs (for this set of experiments: *Ikzf2^{cello/+}*, 14.2 ± 0.4 pF; *Ikzf2^{cello/cello}*, 11.2 ± 0.5 pF; $P < 0.0005$), both the total I_K (**g**) and isolated $I_{K,n}$ (**h**) were not significantly different between the two genotypes at P16–P18 (two-sided Welch's *t*-test). Data are mean \pm s.e.m. **i**, NanoString validations of genes downregulated in P8 *Ikzf2^{cello/cello}* cochleae at P16, normalized to wild-type reads. Data are mean \pm s.d. ($n = 4$ biologically independent samples per genotype). * $P = 0.038$ (*Ppp17r1* in *Ikzf2^{cello/cello}* vs *Ikzf2^{+/+}*), * $P = 0.037$ (*Ppp17r1* in *Ikzf2^{cello/cello}* vs *Ikzf2^{cello/+}*) (two-sided Welch's *t*-test).



Extended Data Fig. 7 | Transduction of cochlear hair cells using Anc80L65 and hair cell enrichment by flow cytometry. **a**, Schematic representation of inner ear viral gene delivery via the posterior semicircular canal of CD-1 mice for hair cell marker immunolabelling. **b**, Immunolabelling for GFP in the Anc80-eGFP injected, and MYC in the Anc80-*Ikzf2* injected ears, showing mainly hair cell transduction, although some MYC staining could also be observed in supporting cells (blue arrow). $n = 3$ biologically independent samples per condition. Nuclear MYC staining suggests proper trafficking of the MYC-tagged Helios protein in transduced cells. White arrows indicate OHCs, white arrowheads indicate IHCs. Scale bars, 10 μm . **c, d**, Flow cytometry of

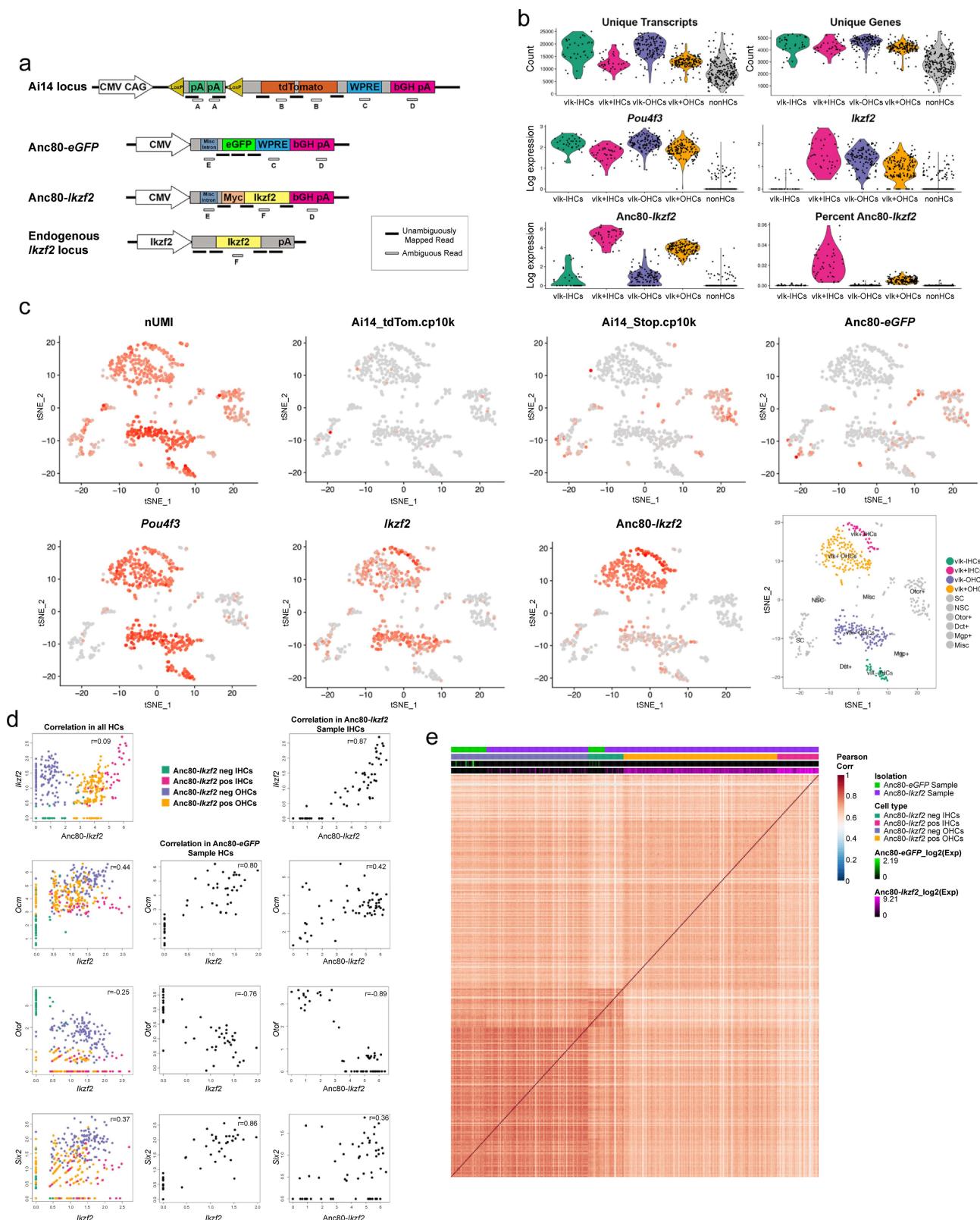
dissociated cochlear GFP-positive and tdTomato-positive cells from P8 *Myo15^{cre/+};ROSA26^{CAG-tdTomato}* mice injected with either Anc80-eGFP (**c**, 2 mice) or Anc80-*Ikzf2* (**d**, 4 mice). Cells were first gated by forward and side scatter to exclude doublets. For the Anc80-eGFP-transduced cochlear sample, transduced cells were identified based on GFP expression, and hair cells were further identified by tdTomato expression. tdTomato single-positive, GFP single-positive and tdTomato and GFP double-positive cells were collected. For the Anc80-*Ikzf2*-transduced cochlear sample, hair cells were gated based on tdTomato single-positive expression and collected.

a**b****c****d****e** *Ikzf2* Negative Correlation**f** *Ikzf2* Positive Correlation

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Transcriptional conversion of *Anc80-Ikzf2*-transduced IHCs. **a**, Heat map for the top 30 differently expressed genes between all hair cells profiled. Scaled expression values shown as *z*-scores, with yellow indicating higher and purple indicating lower expression than the mean. **b**, OHC enriched genes that are induced in *Anc80-Ikzf2(+)* IHCs. *Anc80-Ikzf2(−)* IHC ($n = 34$) vs *Anc80-Ikzf2(+) IHC* ($n = 40$) FDR: $Pde6d = 2.03 \times 10^{-12}$, $Ldhb = 3.74 \times 10^{-11}$. Dots represent the expression values of individual cells, with width of violins summarizing overall relative distribution of expression. **c**, IHC enriched genes that are highly expressed in control IHCs vs control OHCs, but are significantly reduced in *Anc80-Ikzf2(+) IHCs*. *Anc80-Ikzf2(−)* IHC ($n = 34$) vs *Anc80-Ikzf2(+) IHC* ($n = 40$) FDR: $Fgf8 = 3.30 \times 10^{-14}$, $Atp2a3 = 2.46 \times 10^{-13}$, $Rprm = 2.27 \times 10^{-13}$ (Kruskal–Wallis test followed by post hoc pairwise Wilcoxon ranked sum test adjusted for multiple comparisons). **d**, IHC-enriched genes that show only moderately reduced expression in *Anc80-Ikzf2(+) IHCs*. *Anc80-Ikzf2(−)* IHC ($n = 34$) vs *Anc80-Ikzf2(+) IHC* ($n = 40$) FDR: $Shtn1 = 8.59 \times 10^{-5}$, $Tbx2 = 3.88 \times 10^{-8}$, $Cabp2 = 1.40 \times 10^{-10}$ (Kruskal–Wallis test followed by post hoc pairwise Wilcoxon ranked sum adjusted for multiple comparisons). **e, f**, Top 20 genes negatively (e) or positively (f) correlated with *Ikzf2* expression in control hair cells, shown alongside corresponding correlations of gene

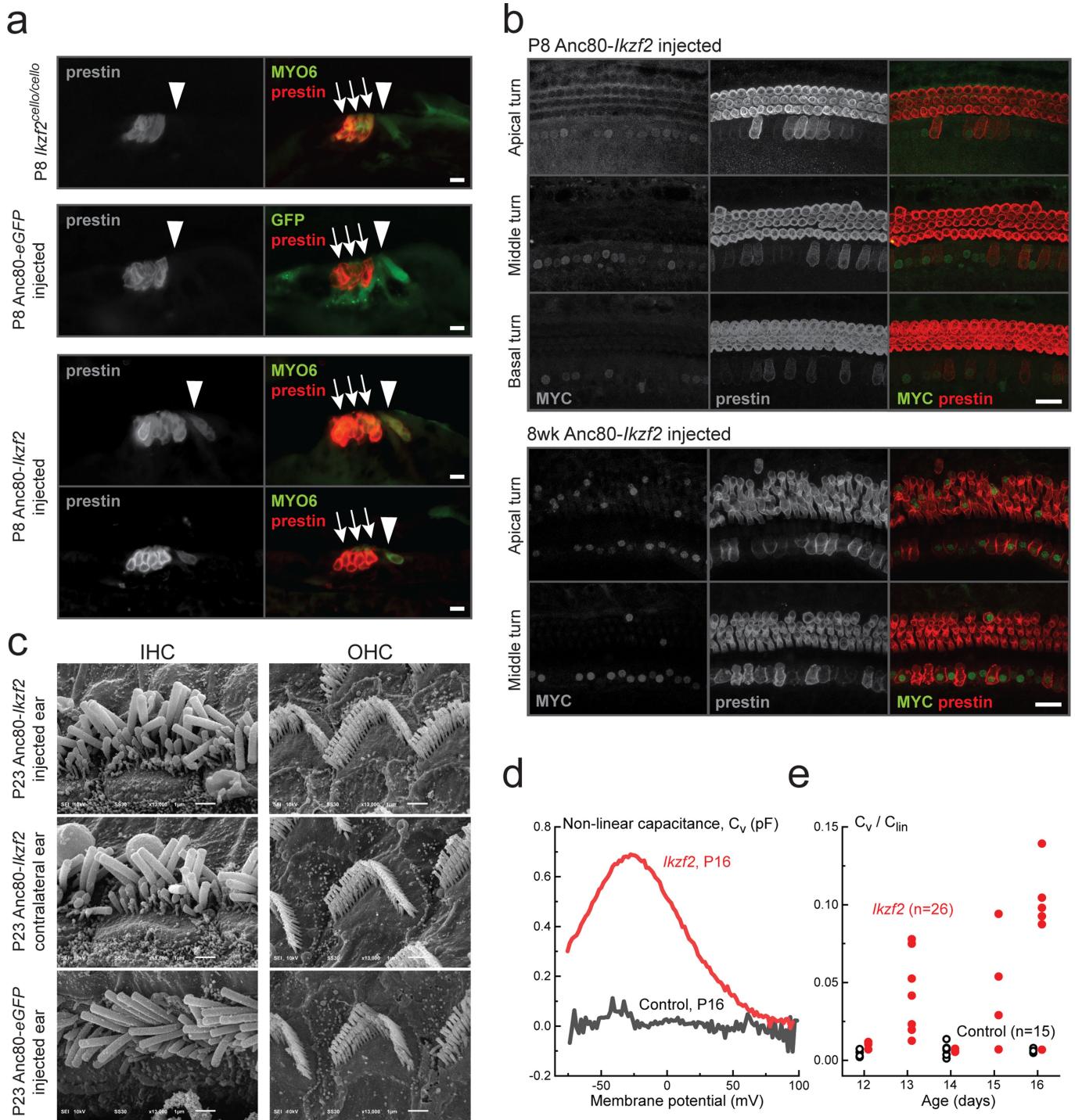
expression within all *Anc80-Ikzf2*-transduced hair cells, *Anc80-Ikzf2*-transduced IHCs, or *Anc80-Ikzf2* transduced-OHCs. See also Extended Data Fig. 9. **g**, Genes that are negatively correlated with *Ikzf2* ($n = 20$, Pearson correlation < -0.6) are not enriched in OHCs at P8 compared to all other genes detected in the RiboTag OHC dataset (background genes, $n = 13,124$). Genes that are positively correlated with *Ikzf2* ($n = 41$, Pearson correlation > 0.6) are significantly enriched in OHCs at P8 compared to background genes ($n = 13,103$) ($P = 0.025$, two-sided Wilcoxon's test). Black line represents median enrichment factor (\log_2 fold change), box demarcates first and third quartiles, whiskers demarcate first and third quartile $\pm 1.5 \times$ IQR values, dots represent single outliers. **h**, One of the most differentially expressed genes observed in our scRNA-seq experiment was *Fcrlb*, a gene which encodes an Fc receptor like protein, and the expression of which has not been previously described in the ear. *Fcrlb* is significantly downregulated in *Anc80-Ikzf2(+) hair cells*. *Anc80-Ikzf2(−)* IHC ($n = 34$) vs *Anc80-Ikzf2(+) IHC* ($n = 40$) FDR = 4.89×10^{-6} . *Anc80-Ikzf2(−)* OHC ($n = 132$) vs *Anc80-Ikzf2(+) OHC* ($n = 148$) FDR = 6.88×10^{-8} (Kruskal–Wallis test followed by post hoc pairwise Wilcoxon ranked sum test adjusted for multiple comparisons). See also Supplementary Tables 8–11.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | scRNA-seq allows for high-resolution discrimination of cell types and their transcriptional changes due to overexpression of *Ikzf2*. **a**, Custom annotation strategy with theoretical reads mapping to unambiguous regions of the various custom viral loci, as well as those regions that get discarded because of endogenous sequence similarity (that is, ambiguous reads). **b**, Violin plots of the overall scRNA-seq detection metrics, including number of unique molecules detected in each of the major cell type cluster identified (low *Anc80-Ikzf2* expressing IHCs: viral *Ikzf2* (*vIk*)⁻ IHCs $n = 34$; low *Anc80-Ikzf2* expressing OHCs: *vIk*⁻ OHCs $n = 132$; high *Anc80-Ikzf2* expressing IHCs: *vIk*⁺ IHCs $n = 40$; high *Anc80-Ikzf2* expressing OHCs: *vIk*⁺ OHCs $n = 140$; and non-HCs: $n = 219$). **c**, FeaturePlots with red showing higher expression across all

profiled cells, including cells identified as non-hair cells. Expression from loci captured with custom annotation shown to support cluster identification. A final labelled t-SNE plot shows all cells profiled clustered by predicted cell type. Misc, cells from all miscellaneous clusters with fewer than 5 cells; NSC, non-sensory epithelial cell; SC, organ of Corti supporting cell. Other clusters are defined by the highest differentially expressed marker gene. **d**, Pearson correlation scatter plots for selected genes within all profiled hair cells, hair cells from the *Anc80-eGFP* sample, or IHCs from the *Anc80-Ikzf2* sample. **e**, A Pearson correlation heat map of all hair cells detected showing overall transcriptional similarities between the non-transduced IHCs and OHCs, along with the *Anc80-Ikzf2*-transduced IHCs and OHCs.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | *Ikzf2* overexpression induces prestin expression and electromotility in IHCs but does not affect hair bundle morphology. **a**, The OHC electromotility protein prestin is expressed in the OHCs of *Ikzf2^{cello/cello}* mutants ($n = 6$ biologically independent samples). In addition, the pattern of prestin expression is not affected by Anc80-eGFP transduction, but is induced in Anc80-*Ikzf2*-transduced IHCs ($n = 3$ biologically independent samples per condition). Scale bars, 10 μm . **b**, Expression of prestin can be seen in Anc80-*Ikzf2*-transduced IHCs as early as P8 and up to 8 weeks of age, and overlaps with MYC staining ($n = 6$ biologically independent samples at P8, $n = 3$ biologically independent samples at 6–8 weeks). Scale bars, 20 μm . **c**, Scanning electron micrographs of IHC and OHC stereocilia bundles of Anc80-*Ikzf2*- and Anc80-eGFP-injected mice at P23 showing expected bundle patterning. Images are from the mid-basal region of the cochlear spiral. Scale bars, 1 μm . Number of biologically independent samples (P16–P23): Anc80-*Ikzf2*-injected cochlea $n = 8$, Anc80-*Ikzf2* contralateral cochlea $n = 6$, Anc80-eGFP-injected cochlea $n = 3$. **d**, Representative

traces of the voltage-dependent (nonlinear) component of the membrane capacitance (an electrical signature of electromotility) in the IHCs of Anc80-*Ikzf2*-injected mouse (red) and its non-injected littermate (grey). Mice were injected with Anc80-*Ikzf2* at P2 and recorded at P16. **e**, Normalized maximal nonlinear capacitance in all recorded IHCs of mice injected with Anc80-*Ikzf2* at P2 (red) at different ages after injection and their non-injected littermates (black). Each symbol represents one biologically independent cell, and the total number of cells is indicated in parentheses. Because Anc80-*Ikzf2* transduction is not 100% efficient in the apical turn of the cochlea at the time points tested, some IHCs of Anc80-*Ikzf2*-injected mice do not show prominent nonlinear capacitance, whereas the other IHCs do. In the IHCs with maximal nonlinear capacitance of more than 0.25 pF (due to presumable *Ikzf2* expression), the parameters of the Boltzmann fit were as following (mean \pm s.e.m.): $Q_{\max} = 0.10 \pm 0.02$ pC; $V_{\text{pk}} = -31 \pm 1$ mV; $z = 0.91 \pm 0.02$; $C_{\text{lin}} = 11.7 \pm 1.2$ pF; $\Delta C_{\text{sa}} = 0.14 \pm 0.07$ pF ($n = 12$). For information on the fitting procedure, see Methods.

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

Structural 3D representations of wild-type and H517Q helios ZnF6 were predicted with RaptorX and visualized using pyMOL software (version 1.7). Electrophysiology data acquisition was controlled by pClamp software (version 10) using Digidata 1440A boards (Molecular Devices, USA). Cell body movement was tracked using Fiji software. Non-linear (voltage-dependent) capacitance of IHCs in *Anc80-lkzf2* injected mice and their non-injected littermates was recorded using jClamp software (SciSoft, USA). Auditory brainstem responses were collected, amplified and averaged using TDT System 3 in conjunction with either BioSig RP (version 4.4.11) or BioSig RZ (version 5.7.1) software. Distortion product oto-acoustic emissions tests were performed with the TDT RZ6 System 3 hardware and BioSig RZ (version 5.7.1) software. Images were captured and processed using Lumenera Infinity Capture and Infinity Analyze software.

Data analysis

For RNA-seq analysis, reads were aligned using TopHat v2.0.8, and HTSeq was used to quantify the number of reads aligning to predicted coding regions. Significant differential gene expression between samples was assessed using DEseq. The R lowess function was used to remove the transcript length bias of enrichment factors in the RiboTag dataset. Genes with a changed level of expression in the RiboTag IP samples at any time point relative to P8 were subjected to a clustering analysis using the CLICK algorithm, implemented in the EXPANDER package. GO enrichment analysis was done using the EXPANDER implemented tool TANGO. The expanded motif prediction analysis was performed using through the Cytoscape visualization tool. Genomic sequence data were assessed for variation using DNASTAR Lasergene software (version 12.0.0). Nanostring data were analyzed using nSolver 4.0 software. Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotyping version 2 (PolyPhen-2), and Protein Variation Effect Analyser (PROVEAN) were used to predict the functional effect of the cello mutation in silico. scRNA-seq reads were aligned using the 10X Genomics cellranger package (version 2.0.2). Secondary analyses were performed in R with Seurat (version 2.1.0). Additional plots were generated by NMF (version 0.20.6) and ggplot2 (version 2.2.1). Composite images were made using ImageJ software.

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 RiboTag OHC RNA-seq, P8 cello cochlea RNA-seq, and P8 *Anc80-Ikzf2* and *Anc80-eGFP* injected cochlea scRNA-seq data have been submitted to the Gene Expression Omnibus database (GEO Accession # GSE116703, GSE116702, and GSE120462), and are additionally available for viewing through 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

All RNA-seq analyses were performed in at least biological duplicates to ensure that statistical analyses could be accurately performed. No sample size calculation was performed for the RNA-seq experiments, and this is standard in the field. Representative images (n=1) were obtained to demonstrate the recombination pattern of the well established *prestinCreERT2*. All other analyses were performed in at least biological triplicate. All attempts at replication were successful.

The cello pedigree was identified from the phenotype-driven ENU-mutagenesis Harwell Aging Screen, wherein large cohorts of mice (n>50-100) were screened for phenotypes of interest (including hearing loss). Utilizing the data obtained from the cello cohort, we used GraphPad StatMate to carry out power calculations to determine sample size in an unpaired t-test using the standard deviation of the measured hearing thresholds, a significance level of p=0.01 (two-tailed), and a power of 95%. As such, the effect size was estimated using real data corresponding to the cello hearing loss phenotype compared with the hearing thresholds of wild-type and heterozygote littermates. This calculation determined that a cohort size of 4 mice/genotype would be sufficient to detect a 40 dB SPL threshold difference.

Data exclusions

No data were excluded from analysis.

Replication

Representative images (n=1) were obtained to demonstrate the recombination pattern of the well established *prestinCreERT2* models. All other analyses reported were replicated in at least biological duplicates, with similar results.

Randomization

Cello mutant mice and littermates were allocated into groups based on genotype. Mice for the viral transduction experiments were allocated into *Ikzf2*, GFP and non-injected groups randomly.

Blinding

For Distortion Product Otoacoustic Emissions (DPOAEs) and electrophysiology studies the operator was blind to the genotype of the animal being tested. For Auditory Brainstem Response tests, mice were housed in cages containing mixtures of genotypes, and mice from these cages were tested in a randomized order. Moreover, for the P16 ABRs the recorded primary data (i.e. the ABR waveform traces) were viewed and re-scored by a second operator blind to genotype.

Reporting for specific materials, systems and methods

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

Mouse anti-HA.11 monoclonal antibody (clone 16B12, purified, BioLegend, Cat# 901502, lot# B220767, 5ug/mL), Goat anti-Helios M-20 (Santa Cruz Biotechnology, Cat# sc-9864, lot# E0709, 1:400-1:600), Mouse anti-β-Actin (Abcam, Cat# ab6276, 1:500), Mouse anti-cMyc 9E10 (Developmental Studies Hybridoma Bank, Cat# 9E10, 1ug & 1:5000), Rabbit anti-GFP (custom-made CUK-1819 MGU-GFP-FL, 1-2ug & 1:1000), Mouse 12G10 anti-α-Tubulin (Developmental Studies Hybridoma Bank, Cat# 12G10, 1:10,000), Goat anti-prestin N-20 (Santa Cruz Biotechnology, Cat# sc-22692, lot# D3015, 1:200), Goat anti-Oncomodulin N-19 (Santa Cruz Biotechnology, Cat# sc-7446, lot# L0814, 1:100), Rabbit anti-MyosinVI (Proteus BioSciences, Cat# 25-6791, 1:1000), Rabbit anti-GFP (Life Technologies, Cat# A-11122, lot# 1925070, 1:100), Mouse anti-cMyc 9E10 (Santa Cruz Biotechnology, Cat# sc-40, lot# B0116, 1:100), Mouse anti-Otoferlin (Abcam, Cat# ab53233, lot# GR3186324-2, 1:100), Guinea pig anti-Vglut3 (Donated by Rebecca Seal, PhD, University of Pittsburgh, 1:5000), Sheep anti-Digoxigenin-AP (Sigma-Aldrich, Cat# 11093274910, lot# 14608125, 1:100), Alexa Fluor® 568 donkey anti-goat (Thermo Fisher, Cat# A-11057, lot# 1485187 & 1711491, 1:200), Alexa Fluor® 488 donkey anti-mouse (Thermo Fisher, Cat# A-21202, lot# 164464 & 1820538, 1:200-1:800), Alexa Fluor® 488 donkey anti-rabbit (Thermo Fisher, Cat# A-21206, lot# 1910751, 1:800), Alexa Fluor® 546 goat anti-mouse (Thermo Fisher, Cat# A-11030, lot# 1661231, 1:800), Alexa Fluor® 568 goat anti-guinea pig (Thermo Fisher, Cat# A-11075, lot# 949212, 1:800), Alexa Fluor® 546 donkey anti-rabbit (Thermo Fisher, Cat# A-10042, lot# 1606268, 1:800), Goat anti-mouse IRDye 680RD (LI-COR Biosciences, Cat# 926-68070, 1:15000), Goat anti-rabbit IRDye 800CW (LI-COR Biosciences, Cat# 926-32211, lot# C80718-15, 1:15000)

Validation

Mouse anti-HA.11 monoclonal antibody, from BioLegend – This antibody can be used to detect fusion proteins with an HA-tag. From the manufacturer: “The extreme specificity of the antibody allows unambiguous identification and quantitative analysis of the tagged protein. The HA.11 antibody recognizes HA epitopes located in the middle of protein sequences as well as at the N- or C-terminus.”

Goat anti-Helios M-20, from Santa Cruz Biotechnology – “Helios (M-20) is an affinity purified goat polyclonal antibody raised against a peptide mapping at the N-terminus of Helios of mouse origin.” This antibody recognizes nuclear Helios in mouse cochlear OHCs. Loss of labelling when the anti-helios antibody is ‘pre-blocked’ confirms specificity (see Figure 1).

Mouse anti-β-Actin, from Abcam – “Ab6276 was shown to specifically react with beta actin in wild-type HAP1 cells. No band was observed when beta actin knockout samples were used. Wild-type and beta actin knockout samples were subjected to SDS-PAGE.”

Mouse anti-cMyc 9E10, from Developmental Studies Hybridoma Bank - "9E10 was deposited to the DSHB by Bishop, J.M. (DSHB Hybridoma Product 9E10). Positive Tested Species Reactivity: Epitope tag/Fusion protein." Uniprot ID: P01106, Entrez Gene ID: 4609, Antibody Registry ID: AB_2266850. Initial publication: Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. Evan GI, Lewis GK, Ramsay G, Bishop JM. Molecular and cellular biology 1985 Dec.

Rabbit anti-GFP CUK-1819 MGU-GFP-FL, custom-made - This antibody was raised in rabbit immunized with a His fusion protein containing full-length GFP and the anti-sera was affinity purified using the antigen coupled to SulfoLink (Pierce, Rockford, IL, USA). It was validated at MRC Harwell Institute using lysates from HEK293T cells transiently transfected with a GFP-tagged construct. Western blot showed the antibody only bound protein from transfected cell lysates and not untransfected control cell lysates. Furthermore, immunoprecipitation of transfected cell lysates showed enrichment of GFP-tagged fusion proteins.

Mouse anti-α-Tubulin, from Developmental Studies Hybridoma Bank – “12G10 anti-alpha-tubulin was deposited to the DSHB by Frankel, J. / Nelsen, E.M. (DSHB Hybridoma Product 12G10 anti-alpha-tubulin). Positive Tested Species Reactivity: Algae, C. elegans, Cephalopod, Dictyostelium, Drosophila, Fish, Hermit Crab, Human, Medusa, Mouse, Protozoa, Saccharomyces, Tetrahymena, and Xenopus.” Uniprot ID: P41351, Entrez Gene ID: 7838142, Antibody Registry ID: AB_1157911. Initial publication: Polyglycylation domain of beta-tubulin maintains axonemal architecture and affects cytokinesis in Tetrahymena. Thazhath R, Liu C, and Gaertig J. Nature cell biology 2002 Mar.

Goat anti-prestin N-20, from Santa Cruz Biotechnology – “Prestin (N-20) is an affinity purified goat polyclonal antibody raised against a peptide mapping at the N-terminus of Prestin of human origin. Prestin (N-20) is recommended for detection of Prestin of mouse, rat and human origin by Western Blotting (starting dilution 1:200, dilution range 1:100-1:1000), immunofluorescence (starting dilution 1:50, dilution range 1:50-1:500) and solid phase ELISA (starting dilution 1:30, dilution range 1:30-1:3000).” Publications: Gao, J., et al. 2007. Prestin-based outer hair cell electromotility in knockin mice does not appear to adjust the operating point of a cilia-based amplifier. Proc. Natl. Acad. Sci. USA 104: 12542-12547; McGuire, R.M., et al. 2010. Cysteine mutagenesis reveals transmembrane residues associated with charge translocation in prestin. J. Biol. Chem. 285: 3103-3113.

Goat anti-Oncomodulin N-19, from Santa Cruz Biotechnology – “OCM (N-19) is an affinity purified goat polyclonal antibody raised against a peptide mapping at the N-terminus of OCM of human origin. OCM (N-19) is recommended for detection of OCM and OCML of mouse, rat and human origin by Western Blotting (starting dilution 1:200, dilution range 1:100-1:1000), immunoprecipitation [1-2 µg per 100-500 µg of total protein (1 ml of cell lysate)], immunofluorescence (starting dilution 1:50, dilution range 1:50-1:500) and solid phase ELISA (starting dilution 1:30, dilution range 1:30-1:3000).” Publications: Devarajan, K., et al. 2005. Circadian and light regulation of Oxytocin and parvalbumin protein levels in the ciliated ependymal layer of the third ventricle in the C57 mouse. Neuroscience 134: 539-547; Csillik, B., et al. 2009. Upregulated expression of oncomodulin, the β

isoform of parvalbumin, in perikarya and axons in the diencephalon of parvalbumin knockout mice. *Neuroscience* 165: 749-757. Rabbit anti-MyosinVI, from Proteus BioSciences – “Rabbits were immunized with amino acids 1049-1254 from the tail region of porcine myosin-VI.” This antibody is affinity purified, and is reported by the manufacturer to have broad species cross-reactivity. Publications: Hasson, T. and M. S. Mooseker (1994). *Journal of Cell Biology* 127:425-440; Hasson, T. et al. (1997), *Journal of Cell Biology* 137:1287-1307.

Rabbit anti-GFP, from Life Technologies – “The anti-GFP rabbit polyclonal antibody is raised against GFP isolated directly from *Aequorea victoria*.” This antibody is purified by ion-exchange chromatography. Lot# 1925070 (used in this study) was certified by Life Technologies using ELISA. Published species reactivity: Chimpanzee, Dog, Fish, Fruit fly, Hamster, Human, Mouse, Non-human primate, Pig, Rat, Sheep, Zebrafish.

Mouse anti-cMyc 9E10, from Santa Cruz Biotechnology – “c-Myc (9E10) is a mouse monoclonal antibody raised against an epitope corresponding to amino acids 408-439 within the C-terminal domain of c-Myc of human origin. Recommended for detection of c-Myc p67 and c-Myc tagged fusion proteins of mouse, rat, human, monkey, feline and canine origin by WB, IP, IF, IHC(P), FCM and ELISA; non cross-reactive with N-Myc or L-Myc proteins. Widely used in combination with eukaryotic expression vectors encoding proteins with c-Myc (amino acids 408-439) epitope tag.” Publications: Tsuga, H., et al. 1994. Sequestration of muscarinic acetylcholine receptor m2 subtypes. Facilitation by G protein-coupled receptor kinase (GRK2) and attenuation by a dominant-negative mutant of GRK2. *J. Biol. Chem.* 269: 32522-32527; Liu, Y., et al. 2017. miR-19a promotes colorectal cancer proliferation and migration by targeting TIA1. *Mol. Cancer* 16: 53.

Mouse anti-Otoferlin, from Abcam - From the manufacturer: “This antibody reacts specifically with human Otoferlin protein (220 kDa). Suitable for: IHC-Fr, Flow Cyt, IHC-FoFr, WB, IP, ICC/IF. Reacts with: Mouse, Rat, Human.” Publications: Richter KN et al. Glyoxal as an alternative fixative to formaldehyde in immunostaining and super-resolution microscopy. *EMBO J* 37:139-159 (2018); Meese S et al. Activity-Dependent Phosphorylation by CaMKId Alters the Ca²⁺-Affinity of the Multi-C2-Domain Protein Otoferlin. *Front Synaptic Neurosci* 9:13 (2017).

Guinea pig anti-Vglut3 – This antibody was generously donated by Rebecca Seal, PhD, of the University of Pittsburgh, and has been validated by the Seal laboratory. Publication: Seal RP, Akil O, Yi E, et al. Sensorineural Deafness and Seizures in Mice Lacking Vesicular Glutamate Transporter 3. *Neuron*. 2008;57(2):263-275. doi:10.1016/j.neuron.2007.11.032.

Sheep anti-Digoxigenin-AP, from Sigma-Aldrich – This antibody is used to detect Digoxigenin labelled RNA probes via a colorimetric reaction. From the manufacturer: “The polyclonal antibody from sheep is specific to digoxigenin and digoxin and shows no cross-reactivity with other steroids, such as human estrogens and androgens.”

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HEK293T and Cos7 cell lines were used in this study, and were provided by Dr. Chris Esapa, MRC Harwell Institute.
Authentication	None of the cell lines used were authenticated.
Mycoplasma contamination	The cell lines were tested and found to be free of myoplasma contamination.
Commonly misidentified lines (See ICLAC register)	The HEK293T cell line used in this study is a commonly misidentified cell line, found often to be misidentified HeLa cells. These cells were used for transient overexpression of tagged proteins and co-immunoprecipitation studies, therefore misidentification of HeLa cells as HEK293T cells would be unlikely to affect the results reported.

Animals and other organisms

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

Laboratory animals	RiboTag, prestinCreERT2 and Myo15Cre mouse models (maintained on a C57BL/6 background) were provided for this study by Dr. Mary Kay Lobo, Dr. Jian Zuo, and Drs. Christine Petit and Thomas Friedman, respectively. CBA/CaJ mice (stock #000654) and B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J mice (stock #007914, referred to as ROSA26CAG-tdTomato) were procured from the Jackson Laboratory. PrestinCreERT2 specificity was determined by crossing prestinCreERT2/CreERT2 mice to ROSA26CAG-tdTomato mice, and resulting offspring were dissected at P21 for whole-mount immunohistochemistry. Cochlear tissues from F1 RiboTagHA+/+;prestinCreERT2/+ offspring on a mixed CBA/C57BL/6 background were collected at the following ages: P8, P14, P28, 6 weeks, and 10 weeks. CD-1 timed-pregnant females were procured from Charles River (Frederick, MD). Resulting CD-1 neonates were injected with Ane80L65 virus between P1 – P3, and dissected for later analyses at either P8 or 8wk. The cello mutant mouse was originally identified from the MRC Harwell Institute phenotype-driven N-ethyl-N-nitrosourea (ENU) Ageing Screen (Potter et al., 2016). Auditory phenotyping was performed at 16 days, 3-, 6-, 9- and 12-months of age. Hair cell counts were performed at 16 days, 3-, 6-, 9- and 12-months of age. Electrophysiological recordings were made from OHCs of cello mice aged P9 – P18. For the cello RNA-seq and NanoString experiments, cochlear ducts from Izkf2+/+, Izkf2cello/+ and Izkf2cello/cello mice were dissected at P8. For the Ane80L65 transduced IHC scRNA-seq experiment, Myo15Cre/Cre mice were crossed to ROSA26CAG-tdTomato mice. Resulting offspring were injected with Ane80L65 virus between P1 – P3, and the cochlear epithelium was collected at P8. Both male and female animals were used for all experiments.
Wild animals	The study did not include wild animals.
Field-collected samples	The study did not include samples collected from the field.

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

Inner ears of neonatal Myo15Cre/+;ROSA26CAG-tdTomato mice were injected with Anc80-Ikzf2 (4 mice) or control Anc80-eGFP (2 mice) via the posterior semicircular canal. Cochlear tissues from both injected and uninjected ears were harvested at P8 and further dissected to reveal the sensory epithelium. The sensory epithelia from Anc80-eGFP and Anc80-Ikzf2 injected mice were pooled separately into 2 wells of a 48-well plate containing 0.5 mg/ml Thermolysin (Sigma). Tissues were incubated at 37°C for 20 minutes, after which the Thermolysin was removed and replaced with Accutase enzyme (MilliporeSigma). After a 3 minute incubation at 37°C, tissues were mechanically disrupted using a 23G blunt ended needle connected to a 1 ml syringe. This step was performed twice. After confirming tissue dissociation by direct visualization, the dissociation reaction was stopped by adding an equal volume of IMDM supplemented with 10% heat-inactivated FBS to the Accutase enzyme solution. Cells were passed through a 40 um cell strainer (BD) to remove cell clumps. tdTomato expressing HCs were sorted into ice cold tubes containing IMDM with 10% FBS on a BD FACSAria II (BD Biosciences) and processed for scRNA-seq. Flow cytometry analyses were performed at the University of Maryland Marlene and Stewart Greenebaum Comprehensive Cancer Center Flow Cytometry Shared Service.

Instrument

BD FACSAria II (BD Biosciences)

Software

FACSDiva Version 6.1.1

Cell population abundance

The population of interest (tdTomato positive hair cells) represented less than 1% of the total cell population. This low abundance has been well documented in the literature (Elkon et al., 2015, Matern et al., 2017). Therefore, post-sort analyses were not performed, as all cells were needed for downstream analysis. Hair cells and non-hair cells contained within the sorted population were later identified by single-cell RNA-seq.

Gating strategy

Cells were first gated by forward and side scatter to exclude doublets. For the Anc80-eGFP transduced cochlear sample, transduced cells were identified based on GFP expression, and hair cells were further identified by TdTomato expression. Tdtomato single positive, GFP single positive and TdTomato+GFP double positive cells were collected. For the Anc80-Ikzf2 transduced cochlear sample, hair cells were gated based on TdTomato single positive expression and collected.

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