



This is a repository copy of *Helios is a key transcriptional regulator of outer hair cell maturation*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/146863/>

Version: Accepted Version

Article:

Chessum, L., Matern, M.S., Kelly, M.C. et al. (20 more authors) (2018) Helios is a key transcriptional regulator of outer hair cell maturation. *Nature*, 563 (7733). pp. 696-700. ISSN 0028-0836

<https://doi.org/10.1038/s41586-018-0728-4>

© 2018 Springer Nature Limited. This is an author-produced version of a paper subsequently published in *Nature*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Published in final edited form as:

Nature. 2018 November 21; 563(7733): 696–700. doi:10.1038/s41586-018-0728-4.

***Ikzf2*/helios is a key transcriptional regulator of outer hair cell maturation**

Lauren Chessum^{#1}, Maggie S Matern^{#2}, Michael C Kelly³, Stuart L Johnson⁴, Yoko Ogawa², Beatrice Milon², Mark McMurray², Elizabeth C Driver³, Andrew Parker¹, Yang Song⁵, Gemma Codner¹, Christopher T Esapa¹, Jack Prescott¹, Graham Trent², Sara Wells⁶, Abigail K Dragich⁷, Gregory I. Frolenkov⁷, Matthew W Kelley³, Walter Marcotti⁴, Steve DM Brown¹, Ran Elkon^{8,9}, Michael R Bowl^{1,12,*}, and Ronna Hertzano^{2,5,10,12,*}

¹Mammalian Genetics Unit, MRC Harwell Institute, Harwell Campus, Oxfordshire, OX11 0RD, UK

²Department of Otorhinolaryngology Head and Neck Surgery, University of Maryland School of Medicine, Baltimore, MD 21201, USA

³National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD 20892, USA

⁴Department of Biomedical Science, University of Sheffield, Sheffield, S10 2TN, UK

⁵Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD 21201, USA

⁶Mary Lyon Centre, MRC Harwell Institute, Harwell Campus, Oxfordshire, OX11 0RD, UK

⁷Department of Physiology, College of Medicine, University of Kentucky, Lexington, Kentucky 40536-0298, USA

⁸Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

⁹Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

¹⁰Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, MD 21201, USA

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence and requests for materials should be addressed to M.R.B. (m.bowl@har.mrc.ac.uk) or R.H.

(RHertzano@som.umaryland.edu).

¹²These authors jointly supervised this work

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/>).

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. analyzed 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. and M.R.B. and R.H. wrote the manuscript.

Author information Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing financial interests.

These authors contributed equally to this work.

Abstract

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

The mature mammalian cochlea contains two distinct types of sensory cells, named inner and outer hair cells (HCs), each of which are highly specialized and, in humans, do not regenerate once damaged or lost¹. Progressive loss of these cells, particularly the outer HCs (OHCs), underlies much of the aetiology of age-related hearing loss – a worldwide epidemic^{2,3}. While these two cell types were first described by Retzius in the 1800's, the mechanisms underlying the specification of their common progenitor cells to functional inner versus outer HCs remain poorly understood. Additionally, attempts to direct stem cells towards HC fates have, to date, resulted only in the formation of immature HC-like cells that lack many of the markers of mature inner or outer HCs⁴. 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 ultimately for 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 prestin-CreER^{T2} mice, which can be induced to express Cre-recombinase specifically in OHCs, with the RiboTag mouse model, allowing for OHC-specific ribosome immunoprecipitation (IP)^{5,6}. Cochlear ducts from *RiboTag^{HA/+};prestin^{CreERT2/+}* mice were collected at five postnatal time points (postnatal day (P) 8, P14, P28, 6 weeks (wk), and 10 wk), and actively translated OHC transcripts were enriched for by ribosome IP, followed by RNA sequencing (RNA-seq) of all IP and paired input RNA (Extended Data Fig. 1a-b, Supplementary Table 1). We calculated an OHC enrichment factor (EF) based on the IP/input log₂ fold change (LFC) for each gene at each time point (Supplementary Table 2). Reassuringly, known postnatal HC- and OHC-expressed genes such as *Pou4f3*, *Gfi1*, *Strc*, *Ocm* and *Slc26a5* generally had high EFs across time points (EF>1), while prominent IHC marker genes such as *Otof*, *Atp2a3* and *Slc17a8* were generally depleted from the IP samples (EF<-1). Additionally, 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 depletion of IHC markers identified by a published adult

mouse OHC and IHC transcriptomic dataset (Liu et al.)⁷, and overall classified the OHC-enriched transcripts into three clusters (Extended Data Fig. 1d-f, Supplementary Table 3). Intersecting genes whose transcripts were enriched in OHCs in our most mature RiboTag OHC data point (10wk, EF>0.5) with the Liu et al. dataset 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 (TF) binding motifs^{8,9}. A TF binding motif prediction analysis of the 100 OHC marker genes identified multiple enriched motifs in the 20 kb regions centered around the transcription start site, the top five of which correspond to the TFs HNF4A, MZF1, POU3F2, IKZF2/helios and RFX310. Of these TFs, only IKZF2/helios: was included in the list of 100 OHC marker genes; was found to be markedly enriched in OHCs at all time points (Fig. 1b-c); and showed a ~4-fold enrichment in OHCs compared to IHCs in the Liu et al. dataset (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 *Ikzf2*/helios in regulating the OHC transcriptome from early postnatal to adult stages.

A recent phenotype-driven 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 TF (Fig. 1g, Extended Data Fig. 2b-d)¹¹. A combination of *in silico* mutation analyses, structural 3D modelling, immunolabeling 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 Fig. 2e and 3). We further investigated the functional role of *Ikzf2* in hearing by assessing Auditory Brainstem Response (ABR) thresholds in wild-type and *cello* mice across several time points. Results show that *Ikzf2^{cello/cello}* mice have progressive deterioration of hearing function starting as early as P16 (>60 dB 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 HC stereocilia bundles in the *cello* mice appear normal up to 1-month of age, after which the OHCs bundles, and later the IHCs bundles, begin to degenerate (Extended Data Fig. 4d, 5a-d, Supplementary Tables 5-6). These data indicate that the hearing impairment in *cello* mice precedes the loss of HC bundles, and suggest that the helios mutation instead leads to a functional OHC deficit. Furthermore, by utilizing a second *Ikzf2* mutant allele (*Ikzf2^{del1890}*), which leads to an in-frame deletion of the 3rd coding exon, we confirm *Ikzf2* as the causative gene underlying the auditory dysfunction in the *cello* mutants. At 1-month of age, *Ikzf2^{cello/del1890}* compound heterozygotes display elevated ABR thresholds (up to 40 dB SPL) compared to heterozygotes and wild-type mice (Extended Data Fig. 5e-f), confirming *Ikzf2^{cello}* as the causative allele in the *cello* mutant.

To explore the effect of the *cello* mutation on OHC physiology we investigated the basolateral properties of OHCs. We found that the 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 *Ikzf2^{cello/cello}* OHCs. The resting membrane potential (V_m) of OHCs is also similar between genotypes (*Ikzf2^{cello/+}*: -68 ± 2 mV; *Ikzf2^{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^{12–14}. However, *Ikzf2^{cello/cello}* OHCs show significantly reduced movement compared to *Ikzf2^{cello/+}* control OHCs (Fig. 2e), even when the values are normalized to their reduced surface area (Fig.2f). We also found that young adult *Ikzf2^{cello/cello}* mice have significantly reduced 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 cochleae of P8 *Ikzf2^{cello/cello}* and their wild-type 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 TFs selected from Li et al., 2016 (Fig.2i)¹⁵, suggesting that the observed OHC gene dysregulation results from disruption of a specific transcriptional cascade. Interestingly, by P16 the transcript levels of *Car7*, *Ocm*, and *Slc26a5*, but not *Ppp1r17*, are similar to wild-type littermate controls, suggesting the possibility that other factors may be compensating for the functional loss of *Ikzf2* by this time point (Extended Data Fig.6i).

To further characterize the transcriptional cascade downstream of *Ikzf2*/helios, we performed *in vivo* Anc80L65 adeno-associated virus (AAV) gene delivery of a myc-tagged *Ikzf2* or GFP (from hereon Anc80-*Ikzf2* and Anc80-*eGFP*) to neonatal inner ears of *Myo15^{Cre/+};ROSA26^{CAG-tdTomato}* mice, sorted the cochlear HCs at P8, and measured resultant changes in gene expression using single cell RNA-seq (scRNA-seq) (Fig.3a, Extended Data Fig.7)^{16,17}. The HCs from Anc80-*Ikzf2* injected inner ears separated into two distinct sets of clusters, containing both IHCs and OHCs. One set of IHCs and OHCs completely overlapped with the HCs from the Anc80-*eGFP* control injected ears (Fig.3b, bottom clusters), while the other set clustered separately (Fig.3b, top clusters). Separation of the two sets of clusters showed a clear correlation with Anc80-*Ikzf2* transgene expression (Fig.3b), where HCs in the bottom clusters had a lower expression of Anc80-*Ikzf2*, and the HCs in the top clusters had a higher expression of Anc80-*Ikzf2* (hereon defined as Anc80-*Ikzf2* low (-) and high (+), respectively). Because the HCs defined as Anc80-*Ikzf2* (-) clustered together with the HCs transduced with Anc80-*eGFP*, these two groups of HCs were merged and named Anc80-*Ikzf2* (-) IHCs and OHCs for all downstream analyses (Fig. 3b-c).

While overexpression of *Ikzf2* in IHCs and OHCs did not change the expression of HC 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 HC populations, including *Otof*, *Rprm*, *Atp2a3*, and *Fgf8* (Fig. 3e, Extended Data Fig. 8, Supplementary Tables 8, 9, 10). Interestingly, some of the genes that are downregulated in both Anc80-*Ikzf2* transduced IHCs and OHCs are genes that are normally expressed in both cell types in early

postnatal development, and that later become IHC-specific (e.g., *Pvalb* and *Otof*, Supplementary Table 10)^{18,19}. This suggests that helios overexpression in OHCs results in an accelerated downregulation of these genes. Additionally, helios overexpression in IHCs results in the upregulation of genes that are normally enriched in OHCs, such as *Ocm*, *Pde6d*, *Ldhd* and *Lbh* (Fig.3f, Extended Data Fig.8). Overall, these data suggest that during normal OHC development, helios likely functions to both decrease the expression of early pan-HC markers, such as *Otof*, in the maturing OHCs, as well as to upregulate OHC marker genes. A correlation analysis further validates the role of *Ikzf2* in regulating OHC-related gene expression (Extended Data Fig.8, 9, Supplementary Table 11). The effect of *Ikzf2* transduction on IHC gene expression was further validated by immunolabeling or *in situ* hybridization for OTOF, VGLUT3, OCM, prestin and *Fcrlb* (Fig.4, Extended Data Fig. 10a-b). Further 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 (non-linear) capacitance in IHCs (Extended Data Fig.10d-e), which is an electrical “signature” of prestin-dependent OHC electromotility^{20,21}. These data indicate that Anc80-*Ikzf2* transduced IHCs start to acquire the major function of normal OHCs.

In conclusion, our study demonstrates that *Ikzf2*/helios is necessary for hearing and is a critical regulator of gene expression in the maturing postnatal OHC. In particular, our results suggest that *Ikzf2*/helios functions to suppress IHC and early pan-HC gene expression in OHCs, as well as upregulate canonical OHC marker genes. It further shows that *Ikzf2*/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 – supportive of the notion that additional OHC-expressed transcription factors are involved in postnatal OHC development. This is the first study to demonstrate functional shifts in postnatal HC molecular identities via viral gene delivery, and suggests that delivery of combinations of TFs may lead to successful regeneration of functional OHCs in the deafened cochlea.

Methods

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^{5,6,22}, and were generously 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)26Sor^{tm14}(CAG-tdTomato)Hze/J* mice (stock #007914, referred to as *ROSA26^{CAG-tdTomato}*) were procured from the Jackson Laboratory (Bar Harbor, ME). Prestin^{CreERT2} specificity was determined by crossing *prestin^{CreERT2/CreERT2}* mice to

ROSA26^{CAG-tdTomato} mice, and 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 mice²³. Recombination was induced by tamoxifen injection (3 mg/40 g body weight in mice younger than 21 days, 9 mg/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 (Frederick, MD) or the University of Maryland School of Medicine Veterinary Resources (Baltimore, MD). Resulting neonates were injected with Anc80L65 virus between P1 – P3, and dissected for later analyses between P8 and 8wk. For the Anc80L65 transduced IHC scRNA-seq experiment, *Myo15^{Cre/Cre}* mice were crossed to *ROSA26^{CAG-tdTomato}* mice, resulting offspring were injected with Anc80L65 virus between P1-P3, and the cochlear epithelium was collected at P8. Additionally, a number of 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 non-linear (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 licenced 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 Screen¹¹. In this screen, ENU-mutagenized C57BL/6J males were mated with wild-type ‘sighted C3H’ (C3H.Pde6b+) females²⁴. The resulting ♂ males were crossed with C3H.Pde6b+ females to produce G₂ females, all of which were screened for the *Cadherin23^{ahl}* allele²³. *Cadherin23^{+/+}* 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 in Sanz et al., 2009⁶. Briefly, 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, 1X protease inhibitor cocktail, 200 U/mL RNaseOUT, 100 µg/ml cycloheximide, 1 mg/ml heparin). Homogenates were spun

down (10,000 rpm for 10 minutes at 4°C) to remove particulates. 40 µl of homogenate was reserved for total RNA isolation (input control), and the remaining homogenate was incubated with 5 µg HA antibody (BioLegend) at 4°C under gentle rotation for 4 – 6 hours. 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 following 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 cycloheximide) at 4°C for 10 minutes, rotating. 350 µl of buffer RLT from the RNeasy Plus Micro kit (Qiagen) was then added to the beads or reserved input sample, and vortexed for 30 seconds to 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 recommended by Sanz et al. This method yielded an average of 10.9 ng of IP RNA (average concentration = 0.68 ng/µl) and 185.6 ng of input RNA (average concentration = 10.9 ng/µl) 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/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/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 v2.0.825, and HTSeq was used to quantify the number of reads aligning to predicted coding regions²⁶. 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 DESeq²⁷. 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 10th percentile of the distribution of the expression levels was applied (i.e., all expression values below the 10th percentile were set to the 10th percentile value). The OHC enrichment factors (EF) were calculated for each gene and time point by comparing the RiboTag IP samples to matched input samples, and are defined as the Log₂ ratio of expression levels between the IP and input samples. Inspection of these EFs 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). The RiboTag OHC RNA-seq and P8 *cello* cochlea RNA-seq data have been submitted to the Gene Expression Omnibus database (GEO accession numbers GSE116703 and GSE116702), and are additionally available for viewing through the gEAR Portal (<https://umgear.org/>).

Gene expression analyses

Genes with a changed level of expression in OHC IP samples at any time point relative to P8 were subjected to a clustering analysis using the CLICK algorithm, implemented in the EXPANDER package^{28,29}. Gene Ontology (GO) enrichment analysis was carried out using the EXPANDER implemented tool TANGO²⁸. The adult mouse IHC and OHC transcriptomic dataset used for comparisons was generated by Liu et al., 2014 and can be accessed through the Gene Expression Omnibus database (GEO accession number GSE111348)⁷. The expanded motif prediction analysis was performed using iRegulon¹⁰ through the Cytoscape visualization tool³⁰. The analysis was performed on the putative regulatory region of 20 kb centered around the TSS using default settings.

Immunohistochemistry

For cochlear sections, mice were culled by cervical dislocation and inner ears fixed in 4% paraformaldehyde (PFA) overnight at 4°C then decalcified in 4% ethylenediaminetetraacetic acid (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 minutes at room temperature (RT) then labelled with primary antibodies for 3 hours at RT. To enable detection, samples were incubated with fluorophore-coupled secondary antibodies for 2 hours at RT then stained with DAPI (1:2500, Thermo Fisher) for 5 minutes. Sections were transferred to WillCo glass bottom dishes (Intracel) and visualized free-floating in PBS using a Zeiss 700 inverted confocal microscope (10x – 40x magnification). Primary antibodies: goat anti-Helios M-20 (1:400, Santa Cruz Biotechnology); 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 minutes at 4°C. Post-fixation, ears were fine dissected to expose the sensory epithelium then immediately permeabilized in 0.2% Triton-X for 10 minutes and

blocked with 10% donkey serum (Sigma) for 1 hour at RT. 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 hour at RT. Samples were washed with DAPI (1:2500, Thermo Fisher) for 60 seconds to stain nuclei then mounted onto slides with SlowFade® Gold (Life Technologies) and visualized using a Zeiss LSM 710 fluorescence confocal microscope and 63x 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, employing 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 in order 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 BACC2 Genomic DNA Extraction Kit (GE Healthcare) and sequenced by the Oxford Genomics Centre (Wellcome Trust Centre for Human Genetics, Oxford, UK) 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 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)^{31–33}. Structural 3D representations of wild-type and H517Q helios ZnF6 were predicted with RaptorX³⁴, 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

(generously provided by Dr. Chris Esapa), to yield N-terminally tagged *Ikzf2*⁺ and *Ikzf2*^{cello} helios. See Supplementary Table 13 for cloning and mutagenesis oligonucleotide sequences.

Constructs were subsequently employed for subcellular localization studies using male *Cercopithecus aethiops* SV40 transformed kidney cells (Cos-7) cells (generously provided by Dr. Chris Esapa) that had been seeded onto 22 x 22 mm glass coverslips in six-well plates, at a volume of 1x10⁵ cells per well. After 24 hours (or when at 50 – 60% confluency), cells were transiently transfected with 1 µg DNA of *Ikzf2*⁺-Myc or *Ikzf2*^{cello}-Myc helios construct using JetPEI® DNA Transfection Reagent (Polyplus Transfection). At 24 hours post-transfection, cells were fixed in 4% PFA for 10 minutes and permeabilised with 1% Triton-X for 15 minutes at RT. After blocking in 10% donkey serum (Sigma) for 1 hour at RT, 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 hour at RT. Cells were washed with DAPI (1:2500, Thermo Fisher) for 60 seconds. Coverslips were mounted onto slides with SlowFade® Gold (Life Technologies) and cells were visualized using a Zeiss LSM 710 multiphoton fluorescence confocal microscope and 63x oil magnification.

Constructs were also utilized for co-immunoprecipitation studies using *Homo sapiens* embryonic kidney cells (HEK293T) cells (generously provided by Dr. Chris Esapa) that had been seeded directly onto six-well plates at a volume of 5x10⁵ cells per well. Cells were transiently co-transfected 24 hours 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 1x RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 50 mM Tris pH 7.5 in milliQ water) at 48 hours post-transfection, then incubated with Protein G Sepharose® Beads (Sigma) for 2 hours 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 (1:5000, Developmental Studies Hybridoma Bank) and custom-made rabbit anti-GFP (1:1000, 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 (1:15000, LI-COR) and goat anti-rabbit IRDye 800CW secondary antibodies (1:15000, LI-COR) and imaged using the Odyssey® CLx Infrared Imaging System (LI-COR). For quantification, band intensities were determined using the Image Studio Lite Ver 5.2

software and used to calculate the relative ratio of Co-IP to IP signal. Cos-7 and HEK293T cells were grown at 37°C under 5% carbon dioxide (CO₂) conditions in Dulbecco's Modified Eagle Medium (Invitrogen) containing 10% heat-inactivated foetal bovine serum (FBS) (Invitrogen) and 1X penicillin/streptomycin (Invitrogen).

Auditory brainstem response (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 function³⁵. Mice were anaesthetized by intraperitoneal injection of ketamine (100 mg/ml at 10% v/v) and xylazine (20 mg/ml at 5% v/v) administered at the rate of 0.1 ml/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, Alachua, FL, USA) in conjunction with either BioSig RP (version 4.4.11) or BioSig RZ (version 5.7.1) software. The TDT system click ABR stimuli comprised clicks of 0.1 ms broadband noise spanning ~2-48 kHz, presented at a rate of 21.1/sec 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 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 blind to genotype. Animals were recovered using 0.1 ml of anaesthetic reversal agent atipamezole (Antisedan™, 5 mg/ml 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 in Mianné et al., 2016 (see Supplementary Table 13 for single guide RNA (sgRNA) sequences, donor oligonucleotide sequences and genotyping primers)³⁶. 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 *AflIII* in order to express sgRNAs under the T7 promoter.

The p_1.1_sgRNA plasmids were linearized with *XbaI*, phenol-chloroform purified and the products 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 per Codner et al., 2018, for inclusion in the microinjection mix³⁷.

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, the four *Ikzf2* sgRNAs, each at 50 ng/μl and the *Ikzf2* flox lssDNA donor at 50 ng/μl prepared in microinjection buffer. Injected embryos were re-implanted in pseudo-pregnant CD1 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 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 *cello* mutation in order to generate *Ikzf2^{cello/del890}* compound heterozygotes for complementation testing.

Scanning electron microscopy

Mice were culled 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 hours at 4°C. Following decalcification in 4.3% EDTA, cochleae were dissected to expose the organ of Corti, and subjected to ‘OTO’ processing (1 hour incubation in 1% osmium tetroxide (TAAB Laboratories Equipment Ltd.), 30 minute incubation in 1% thiocarbonylhydrazide (Sigma), 1 hour incubation in 1% osmium tetroxide), before dehydration in increasing concentrations of ethanol (25%, 40%, 60%, 80%, 95%, 2 x 100%) at 4°C. Samples were critical point dried with liquid CO₂ using an Emitech K850 (EM Technologies Ltd), 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 *cello* 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 (50X, without L-Glutamine) and MEM vitamins solution (100X) were added from concentrates (Fisher Scientific, UK). The pH was adjusted to 7.5 (osmolality ~308 mmol kg⁻¹). The dissected cochleae were transferred to a microscope chamber, immobilized as previously described³⁸ and continuously perfused with a peristaltic pump using the above extracellular solution. The organs of Corti were viewed using an upright microscope (Nikon FN1, Japan) with Nomarski optics (x60 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 disc³⁸. 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 MgCl_2 , 1 EGTA-CsOH, 5 Na_2ATP , 0.3 Na_2GTP , 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 Ltd, UK) amplifier. Patch pipettes were made from soda glass capillaries (Harvard Apparatus Ltd, UK) and had a typical resistance in extracellular solution of 2-3 $\text{M}\Omega$. In order to reduce the electrode capacitance, patch electrodes were coated with surf wax (Mr Zoggs SexWax, USA). Potassium current recordings were performed at RT (22 – 24°C) and the intracellular solution contained (in mM): 131 KCl, 3 MgCl_2 , 1 EGTA-KOH, 5 Na_2ATP , 5 Hepes-KOH, 10 Na_2 -phosphocreatine (pH 7.3; osmolality ~ 296 mmol kg^{-1}). Data acquisition was controlled by pClamp software (version 10) using Digidata 1440A boards (Molecular Devices, USA). 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, USA). 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 (75x magnification) with a Flash 4.0 SCCD camera (Hamamatsu, Japan). 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 nm (290 pixels = 10 μm).

Non-linear (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 (Cat #21083027, Gibco/ThermoFisher, USA) containing the following inorganic salts (in mM): NaCl (137), KCl (5.4), CaCl_2 (1.26), MgCl_2 (1.0), Na_2HPO_4 (1.0), KH_2PO_4 (0.44), MgSO_4 (0.81) and placed into the 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, Japan), equipped with a high numerical aperture (NA) objective (100x, 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 CoCl_2 , 10 mM CsCl, and 0.1 mM Nifedipine (all from Sigma, USA), while the intrapipette solution contained (in mM): CsCl (140), MgCl_2 (2.5), Na_2ATP (2.5), EGTA (1.0), HEPES (5). 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, USA), controlled by jClamp

software (SciSoft, USA). Membrane capacitance was measured during the voltage ramp with a dual sinusoidal, FFT-based method³⁹. 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 prestin⁴⁰, as follows:

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

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

where, 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 C_{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}).

Distortion Product Oto-Acoustic Emissions (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, Alachua, FL, USA). 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 f1 and f2 at a ratio of f2/f1 = 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 at 10% v/v), xylazine (20 mg/ml at 5% v/v) and acepromazine (2 mg/ml at 8% v/v) administered at a rate of 0.1 ml/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 f1 and f2 tones were presented continuously and a fast-Fourier transform was performed on the averaged response of 356 epochs (each ~21 ms). The level of the 2f1-f2 DPOAE response was recorded and the noise floor calculated by averaging the four frequency bins either side of the 2f1-f2 frequency.

NanoString validation

Cochlear RNA extracted from biological triplicates of *Ikzf2^{cxello/cello}*, *Ikzf2^{cxello/+}* 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 analyzed 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 polyadenylation (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^{16,17} were produced by the Gene Transfer Vector Core, Grousbeck Gene Therapy Center at the Massachusetts Eye and Ear Infirmary (Boston, MA) (<http://vector.meei.harvard.edu/>).

Inner ear gene delivery

For in vivo HC transductions, mice were injected with Anc80L65 AAVs between P1 to P3 via the posterior semicircular canal following the injection method described in Isgrig et al., 2017⁴¹. Briefly, 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 Nanoliter 2010 microinjection system (World Precision Instruments) equipped with a loaded glass needle was used to inject 700 nl of 1.13E+13GC/ml Anc80-*Ikzf2* or 500 nl of 4.85E+12GC/ml Anc80-*eGFP*. Injections into the inner ear were performed in 50 nl increments over the course of 2 minutes. 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.

Fluorescence activated cell sorting (FACS)

For the scRNA-seq analysis of Anc80-*Ikzf2* transduced HCs, 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 harvested 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 intensity¹⁷. Cochlear tissues were then dissociated for fluorescence activated cell sorting (FACS) following the method described in Elkon et al., 2015⁸. Briefly, 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 mm 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 with

assistance from Dr. X. Fan at the University of Maryland Marlene and Stewart Greenebaum Comprehensive Cancer Center Flow Cytometry Shared Service.

Single cell RNA-seq (scRNA-seq)

tdTomato positive sorted HCs were pelleted once (300 g at 4°C) and resuspended in a minimal remaining volume (~30 µl). HC-enriched single cell suspensions were then used as input on the 10X Genomics Chromium platform with 3' Single Cell v2 chemistry (10x Genomics). Following capture and library preparation, single cell RNA-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 10X Genomics cellranger (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 HC 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 Fig.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 HCs were excluded based on co-expression of a contaminating cell type. Secondary analyses, including shared nearest neighbor (SNN) clustering, tSNE embedding, and differential expression testing (using either Wilcoxon Ranked Sum for marker gene identification or MAST for pairwise comparison between control inner and outer HCs) were performed in R with Seurat (version 2.1.0)^{42,43}. 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)^{44,45}. These analyses utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>). scRNA-seq data have been submitted to the Gene Expression Omnibus database (GEO accession number GSE120462), and are additionally available for viewing through the gEAR Portal (<https://umgear.org/>).

Immunohistochemistry of AAV-injected cochleae

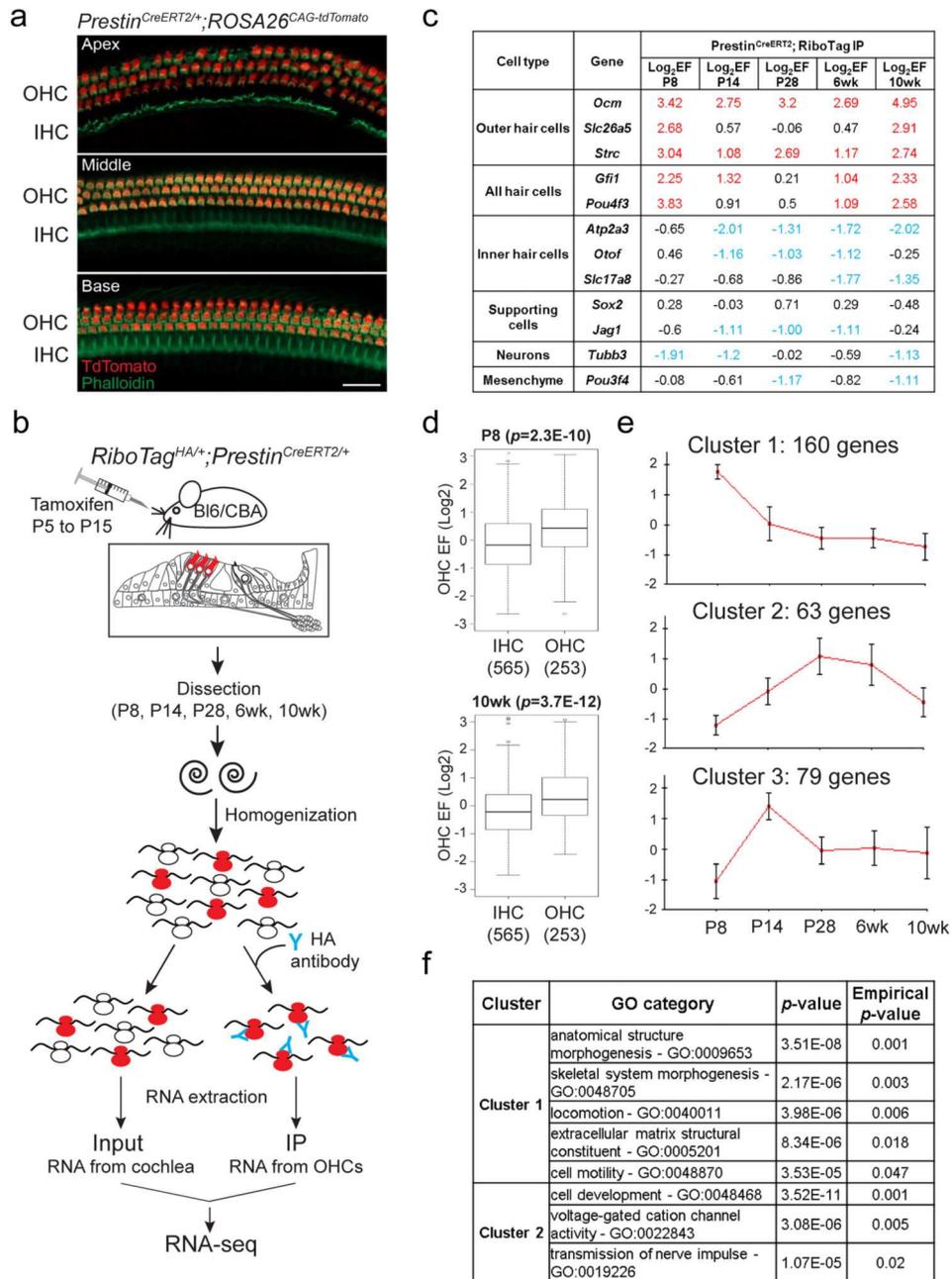
Mouse inner ears injected with either *Anc80-Ikzf2* or *Anc80-eGFP* were between P8 and 8 weeks, fixed in 4% PFA in PBS overnight at 4°C, and decalcified in a solution of 5% EDTA in RNA^{later} (Invitrogen). Decalcified ears were processed by sucrose gradient and embedded in OCT compound (Tissue-Tek) for cryosectioning, or fine dissected for whole-mount immunohistochemistry. 10 µm sections on positively charged glass slides were used for *in situ* hybridization (ISH) and section immunohistochemistry. For whole-mount immunolabeling at 6-8 weeks, HC 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 (1:1000, 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). Dr. Rebecca Seal generously donated the guinea pig anti-Vglut3 antibody used in this study (1:5000). Corresponding Alexa Fluor® 488 and 546 (1:800, Invitrogen) were used for secondary detection, Alexa Fluor® 488 Phalloidin (1:1000, Invitrogen) was used to mark F-actin, and 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI, 1:20,000, Thermo Fisher) was used to mark cell nuclei. Images were acquired using a Nikon Eclipse E600 microscope (Nikon, Tokyo, Japan) 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 63x oil magnification. Images were processed using Infinity Capture and Infinity Analyze software (Lumenera, Ottawa, ON), and ImageJ software.

RNA *in situ* hybridization (ISH)

ISH was performed as described in Geng et al., 201646. Briefly, slides were re-fixed in 4% PFA, and then treated with 2 ug/ml Proteinase-K for 10 minutes. 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). Following 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. Slide were then incubated in BM purple AP substrate precipitating solution (Roche) to localize bound anti-digoxigenin antibody.

Extended Data



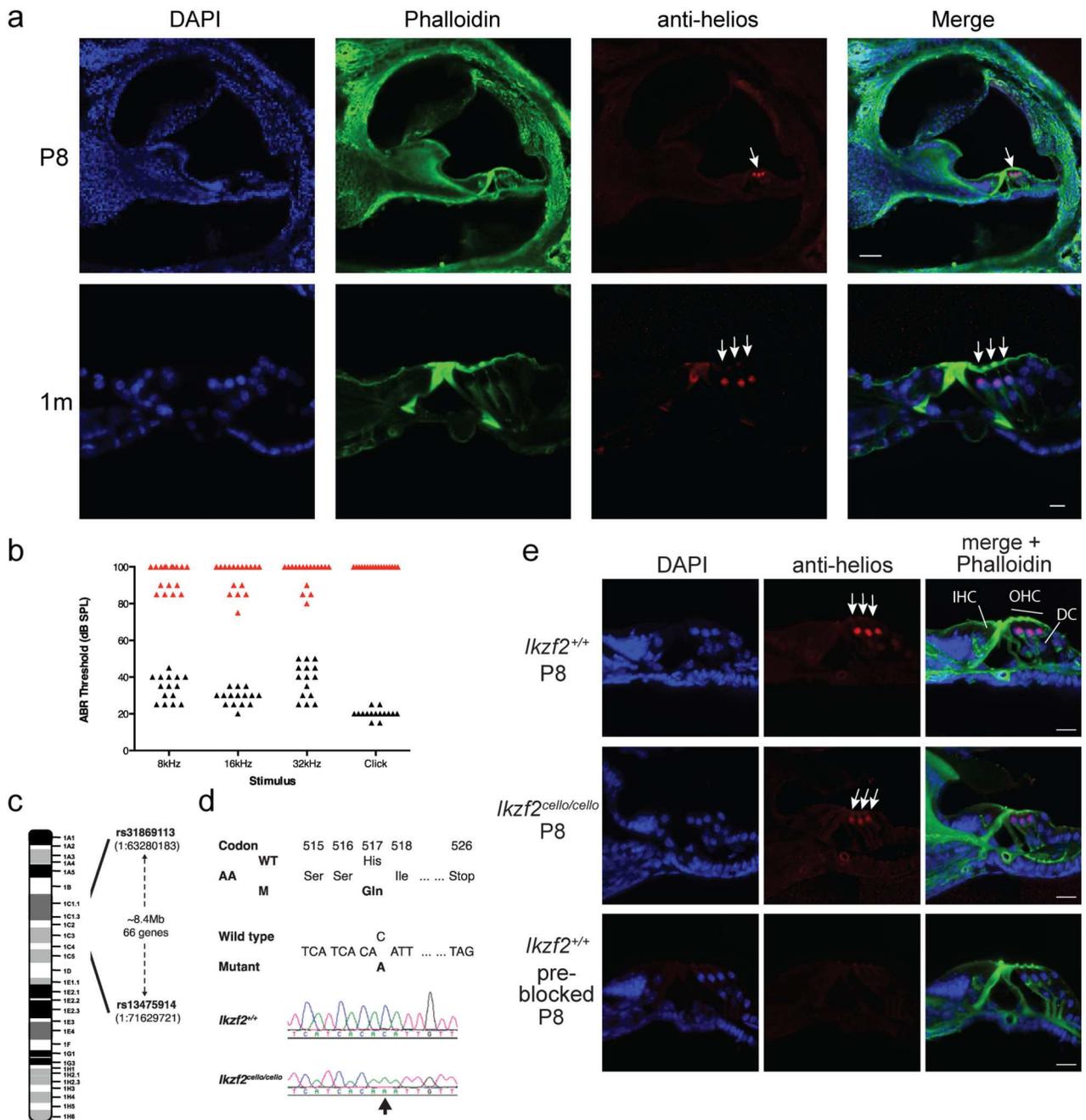
Extended Data Figure 1. RiboTag immunoprecipitation enriches for known OHC-expressed transcripts.

- (a) Representative *prestin*^{CreERT2/+}; *ROSA26*^{CAG-tdTomato} cochlear whole-mount. *Prestin*-*CreERT2*-driven tdTomato expression is OHC-specific at P21 (n=1). Scale=20 μm.
- (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 (enrichment factor (EF) = log₂(IP/input)).

(d) Genes at least 2-fold enriched in IHCs (n = 565 genes) or OHCs (n = 253 genes) in the dataset of Liu et al. are significantly depleted and enriched, respectively, by the RiboTag OHC immunoprecipitation (two-sided Wilcoxon's test). This was true for all time points examined. Black line represents median EF, box demarcates 1st and 3rd quartiles, whiskers demarcate 1st and 3rd quartiles $\pm 1.5 \times \text{IQR}$ values, dots represent single outliers.

(e) Clustering of genes differentially expressed across OHC postnatal development (error bars = SD). Prior to clustering, expression levels were standardized to mean=0 and SD=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.

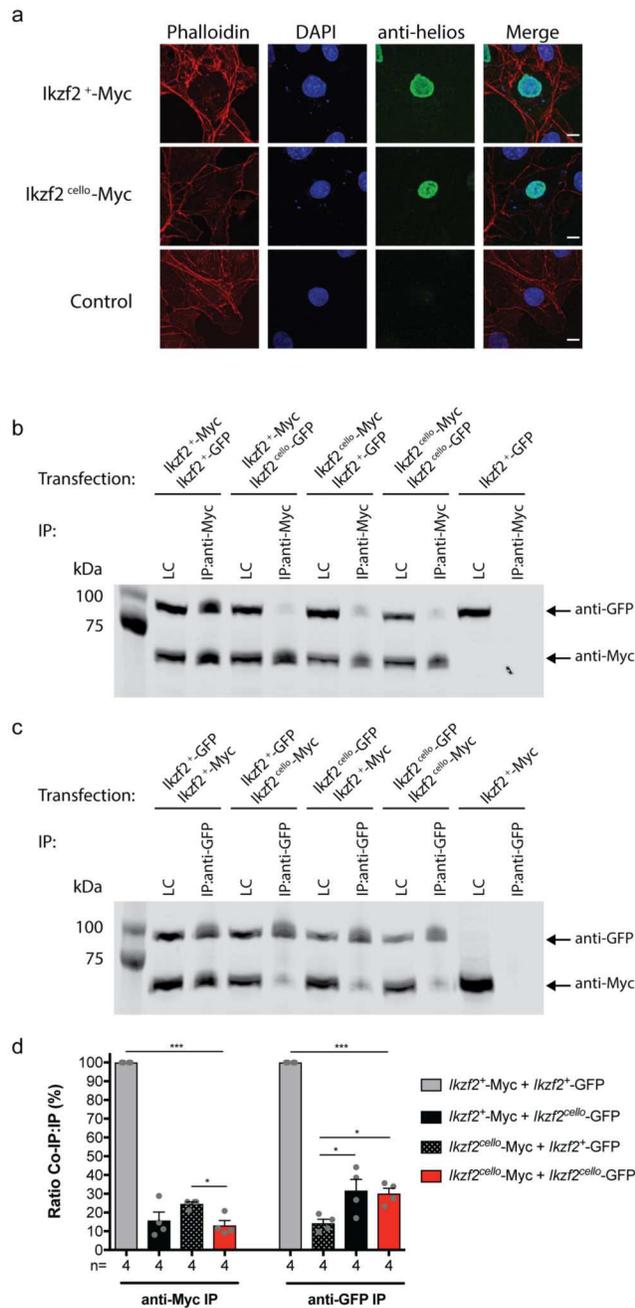


Extended Data Figure 2. Auditory phenotyping, SNP mapping and whole-genome sequencing of mouse pedigree MPC173, subsequently named *cello*.

(a) Specific expression of helios can be seen in the nuclei of wild-type P8 OHCs (white arrow, n=3 biologically independent samples, scale=50 μ m), and is maintained in wild-type OHCs at 1-month (white arrows, n=3 biologically independent samples, scale=10 μ m).

(b) Auditory brainstem response phenotyping of pedigree MPC173 at 9-months of age identified 17 biologically independent animals with elevated hearing thresholds (red triangles) compared to their normal hearing colony mates (n=15 biologically independent animals, black triangles).

- (c) The mutation mapped to an 8.4 Mb region on Chromosome 1 between SNPs 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 (WT) sequence CAC, encoding a histidine (His), to the mutant (M) sequence CAA, encoding a glutamine (Gln). 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=20 μ m.



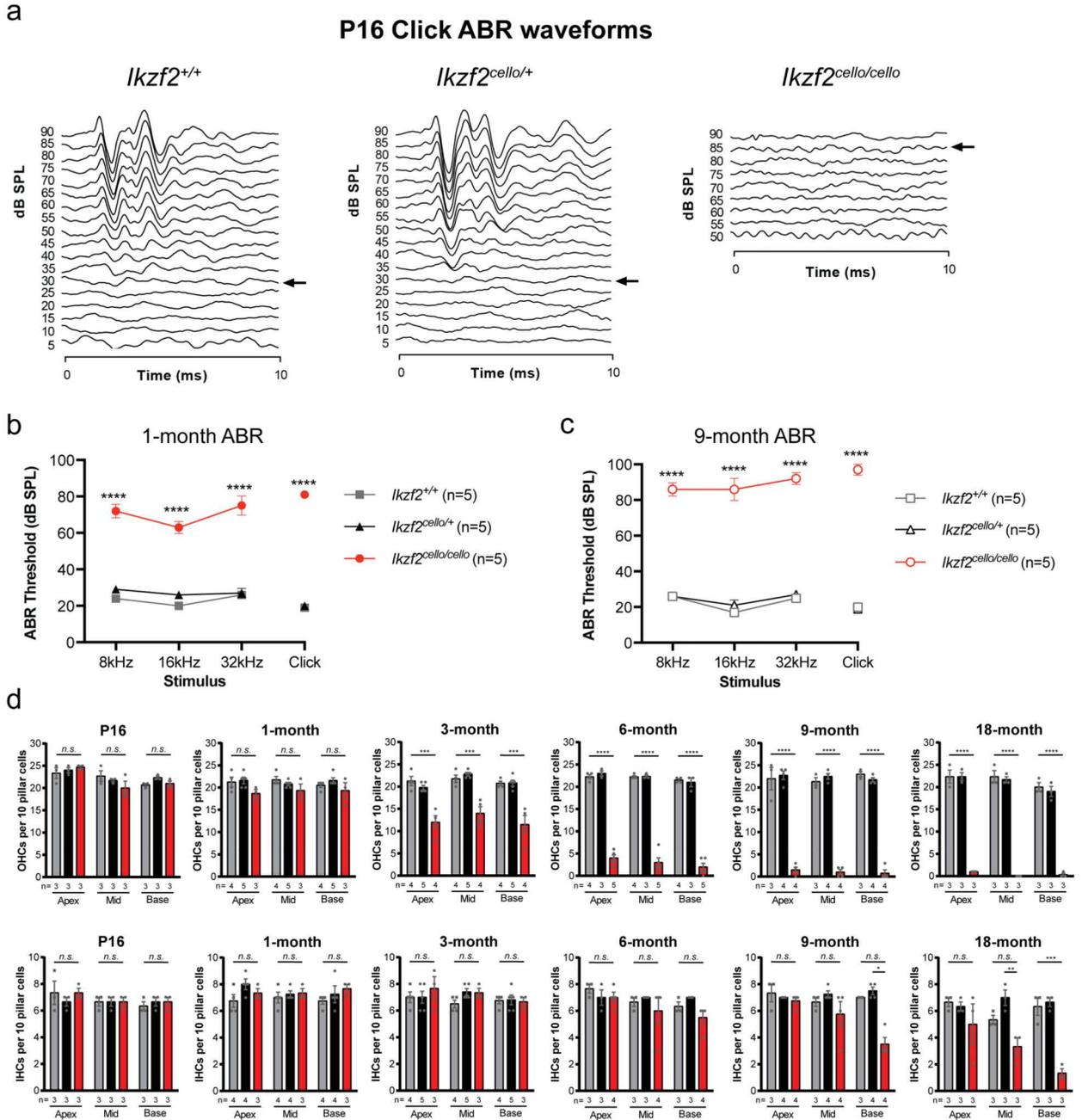
Extended Data Figure 3. The *Ikzf2^{cello}* mutation disrupts helios homodimerization.

(a) Cos-7 cells transfected with *Ikzf2^{+/Myc}* or *Ikzf2^{cello-Myc}*. Nuclear localization is unaffected by the *Ikzf2^{cello}* mutation (n=2 biologically independent experiments). Scale=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. kDa = kilodaltons, 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*^{cxello} helios. kDa = kilodaltons, LC = cell lysate loading control.

(d) Quantification of Co-IP western blots. Band intensities were determined and used to calculate the relative ratio of Co-IP to IP signal. Data shown are averaged percentages \pm s.e.m. (n=4 biologically independent experiments). Myc IP *p*-values: *Ikzf2*⁺-Myc + *Ikzf2*⁺-GFP vs *Ikzf2*⁺-Myc + *Ikzf2*^{cxello}-GFP <0.0001, *Ikzf2*⁺-Myc + *Ikzf2*⁺-GFP vs *Ikzf2*^{cxello}-Myc + *Ikzf2*⁺-GFP <0.0001, *Ikzf2*⁺-Myc + *Ikzf2*⁺-GFP vs *Ikzf2*^{cxello}-Myc + *Ikzf2*^{cxello}-GFP <0.0001, *Ikzf2*⁺-Myc + *Ikzf2*^{cxello}-GFP vs *Ikzf2*^{cxello}-Myc + *Ikzf2*⁺-GFP = 0.1488, *Ikzf2*⁺-Myc + *Ikzf2*^{cxello}-GFP vs *Ikzf2*^{cxello}-Myc + *Ikzf2*^{cxello}-GFP = 0.9020, *Ikzf2*^{cxello}-Myc + *Ikzf2*⁺-GFP vs *Ikzf2*^{cxello}-Myc + *Ikzf2*^{cxello}-GFP = 0.0476. GFP IP *p*-values: *Ikzf2*⁺-GFP + *Ikzf2*⁺-Myc vs *Ikzf2*⁺-GFP + *Ikzf2*^{cxello}-Myc <0.0001, *Ikzf2*⁺-GFP + *Ikzf2*⁺-Myc vs *Ikzf2*^{cxello}-GFP + *Ikzf2*⁺-Myc <0.0001, *Ikzf2*⁺-GFP + *Ikzf2*⁺-Myc vs *Ikzf2*^{cxello}-GFP + *Ikzf2*^{cxello}-Myc <0.0001, *Ikzf2*⁺-GFP + *Ikzf2*^{cxello}-Myc vs *Ikzf2*^{cxello}-GFP + *Ikzf2*⁺-Myc = 0.0202, *Ikzf2*⁺-GFP + *Ikzf2*^{cxello}-Myc vs *Ikzf2*^{cxello}-GFP + *Ikzf2*^{cxello}-Myc = 0.0346, *Ikzf2*^{cxello}-GFP + *Ikzf2*⁺-Myc vs *Ikzf2*^{cxello}-GFP + *Ikzf2*^{cxello}-Myc = 0.9894. Significance was assessed by one-way ANOVA with Tukey post-hoc test. See Supplementary Fig.1 for source images.



Extended Data Figure 4. Auditory function and HC bundle survival in *cello* mice.

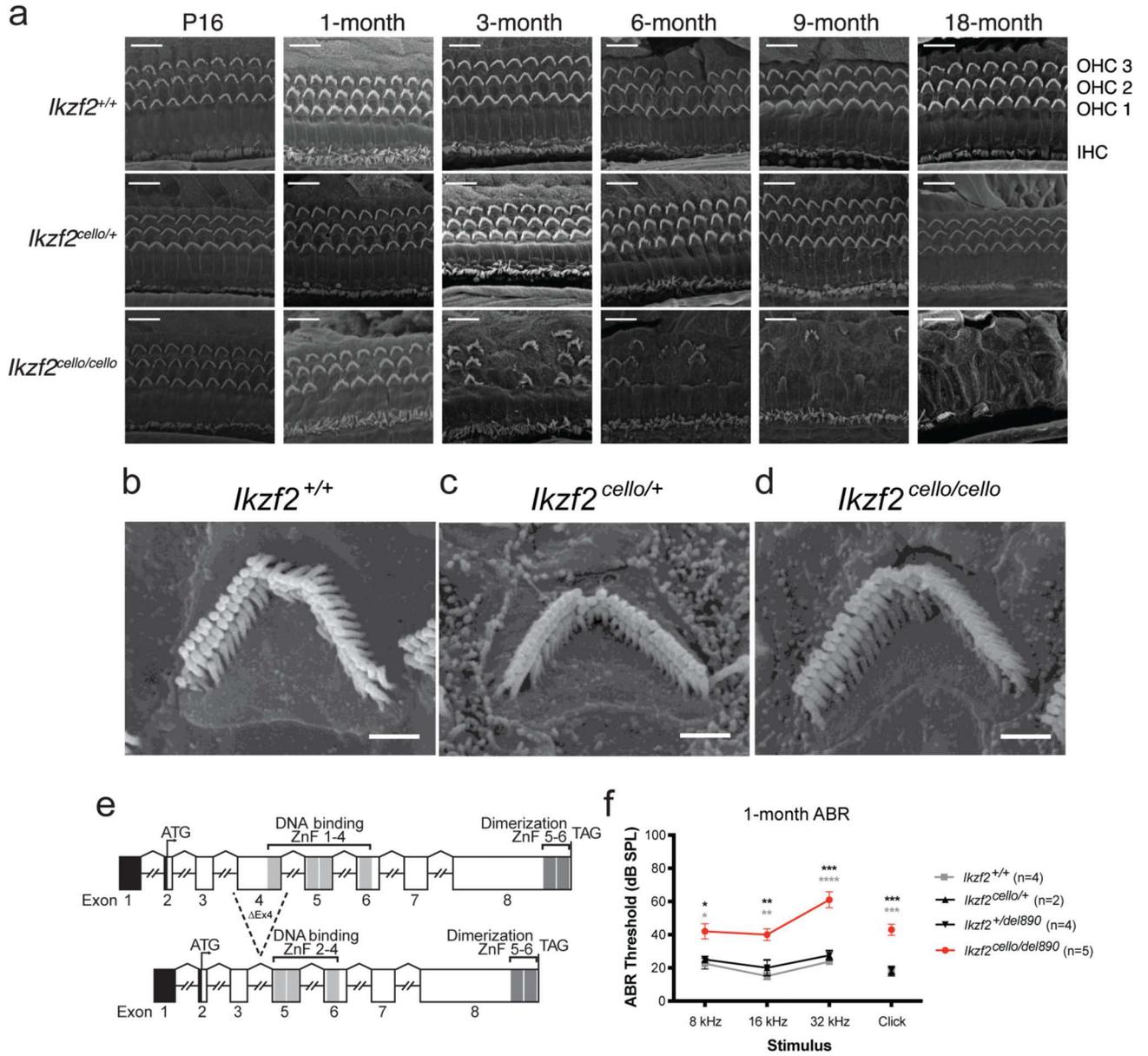
(a) Representative click ABR waveforms for *Ikzf2*^{+/+}, *Ikzf2*^{cello/+} and *Ikzf2*^{cello/cello} littermates at P16 (n=4 biologically independent animals per genotype).

(b-c) Averaged ABR thresholds for *cello* mice at 1-month of age (b, n=5 biologically independent animals per genotype) and 9-months of age (c, n=5 biologically independent animals per genotype). Age-matched *Ikzf2*^{+/+} and *Ikzf2*^{cello/+} controls display thresholds within the expected range (15 – 30 dB SPL) at all time-points tested. Data shown are averaged thresholds ± s.e.m. 1-month *Ikzf2*^{cello/cello} vs 1-month *Ikzf2*^{+/+} (b) *p*-values: 8

kHz < 0.0001, 16 kHz < 0.0001, 32 kHz < 0.0001, Click < 0.0001. 1-month *Ikzf2^{cello/cello}* vs 1-month *Ikzf2^{cello/+}* (b) *p*-values: 8 kHz < 0.0001, 16 kHz < 0.0001, 32 kHz < 0.0001, Click < 0.0001. 9-month *Ikzf2^{cello/cello}* vs 9-month *Ikzf2^{+/+}* (b) *p*-values: 8 kHz < 0.0001, 16 kHz < 0.0001, 32 kHz < 0.0001, Click < 0.0001. 9-month *Ikzf2^{cello/cello}* vs 9-month *Ikzf2^{cello/+}* (b) *p*-values: 8 kHz < 0.0001, 16 kHz < 0.0001, 32 kHz < 0.0001, Click < 0.0001. Significance was assessed by one-way ANOVA with Tukey post-hoc test.

(d) OHC and IHC bundle counts for *cello* mice from P16 to 18-months of age. Data shown are averaged number of HC bundles adjacent to ten pillar cells \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 bundle counts: P16 *Ikzf2^{+/+}* n=3, P16 *Ikzf2^{cello/+}* n=3, P16 *Ikzf2^{cello/cello}* n=3; 1-month *Ikzf2^{+/+}* n=4, 1-month *Ikzf2^{cello/+}* n=5, 1-month *Ikzf2^{cello/cello}* n=3; 3-month *Ikzf2^{+/+}* n=4, 3-month *Ikzf2^{cello/+}* n=5, 3-month *Ikzf2^{cello/cello}* n=4; 6-month *Ikzf2^{+/+}* n=4, 6-month *Ikzf2^{cello/+}* n=3, 6-month *Ikzf2^{cello/cello}* n=5; 9-month *Ikzf2^{+/+}* n=3, 9-month *Ikzf2^{cello/+}* n=4, 9-month *Ikzf2^{cello/cello}* n=4; 18-month *Ikzf2^{+/+}* n=3, 18-month *Ikzf2^{cello/+}* n=3, 18-month *Ikzf2^{cello/cello}* n=3. Number of biologically independent samples for IHC bundle counts: P16 *Ikzf2^{+/+}* n=3, P16 *Ikzf2^{cello/+}* n=3, P16 *Ikzf2^{cello/cello}* n=3; 1-month *Ikzf2^{+/+}* n=4, 1-month *Ikzf2^{cello/+}* n=4, 1-month *Ikzf2^{cello/cello}* n=3; 3-month *Ikzf2^{+/+}* n=4, 3-month *Ikzf2^{cello/+}* n=5, 3-month *Ikzf2^{cello/cello}* n=3; 6-month *Ikzf2^{+/+}* n=3, 6-month *Ikzf2^{cello/+}* n=3, 6-month *Ikzf2^{cello/cello}* n=4; 9-month *Ikzf2^{+/+}* n=3, 9-month *Ikzf2^{cello/+}* n=4, 9-month *Ikzf2^{cello/cello}* n=4; 18-month *Ikzf2^{+/+}* n=3, 18-month *Ikzf2^{cello/+}* n=3, 18-month *Ikzf2^{cello/cello}* n=3.

See also Supplementary Table 5 and 6.



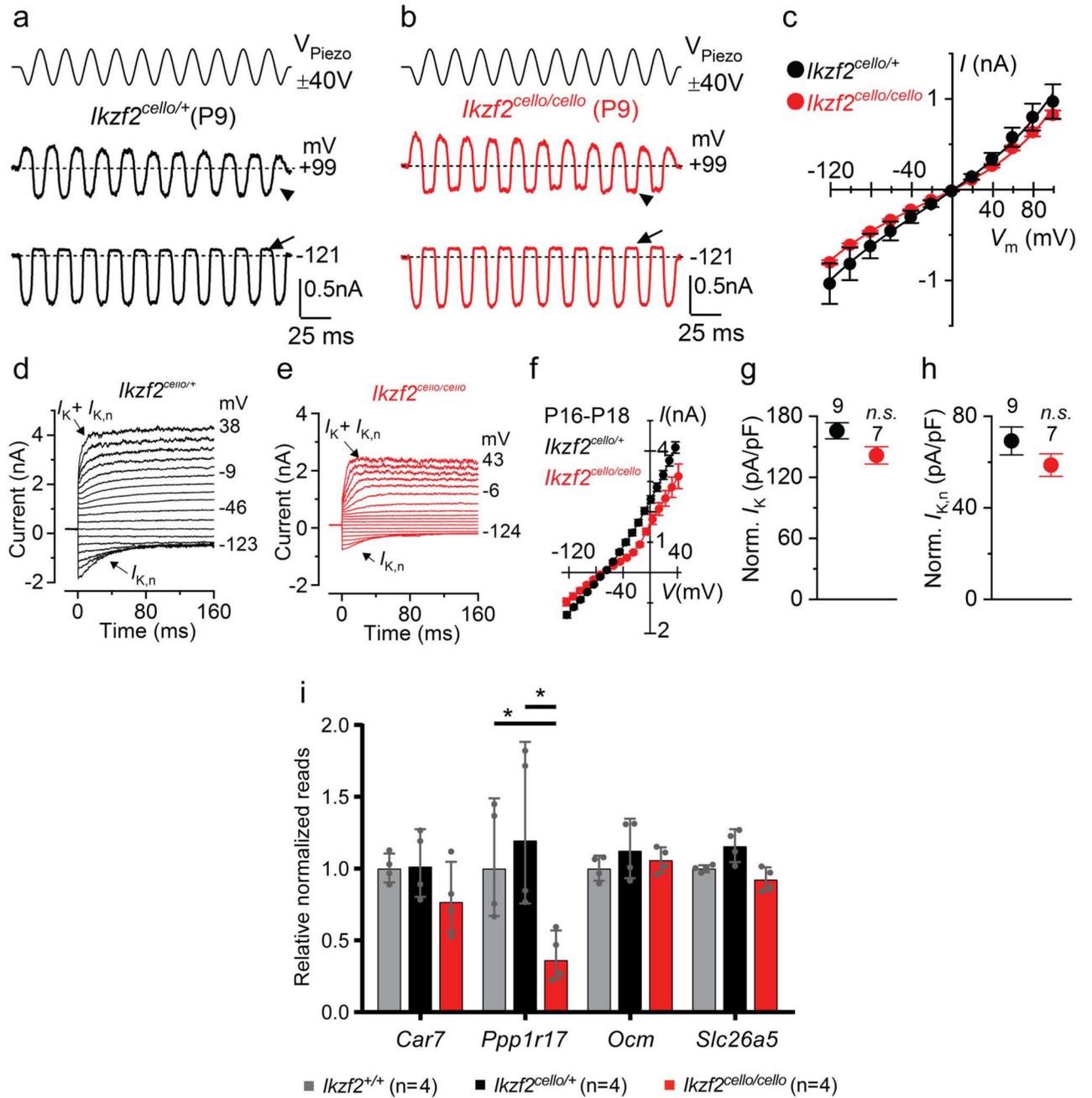
Extended Data Figure 5. Scanning electron microscopy of *cello* mice and auditory function of *Ikzf2*^{cello/del890} compound heterozygotes.

(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 = 10 μ m. Number of biologically independent samples: P16 *Ikzf2*^{+/+} n=4, P16 *Ikzf2*^{cello/+} n=3, P16 *Ikzf2*^{cello/cello} n=3; 1-month *Ikzf2*^{+/+} n=4, 1-month *Ikzf2*^{cello/+} n=5, 1-month *Ikzf2*^{cello/cello} n=3; 3-month *Ikzf2*^{+/+} n=4, 3-month *Ikzf2*^{cello/+} n=5, 3-month *Ikzf2*^{cello/cello} n=4; 6-month *Ikzf2*^{+/+} n=4, 6-month *Ikzf2*^{cello/+} n=5, 6-month *Ikzf2*^{cello/cello} n=4; 9-month *Ikzf2*^{+/+} n=3, 9-month *Ikzf2*^{cello/+} n=4, 9-month *Ikzf2*^{cello/cello} n=4; 18-month *Ikzf2*^{+/+} n=3, 18-month *Ikzf2*^{cello/+} n=3, 18-month *Ikzf2*^{cello/cello} n=3.

(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 = 1 μ m. Number of biologically independent samples: *Ikzf2*^{+/+} n=3, *Ikzf2*^{*cello*+} n=3, *Ikzf2*^{*cello/cello*} n=3.

(e) The genomic/domain structure of the *Ikzf2*^{*del890*} allele/protein. 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 surrounding intronic sequence.

(f) Averaged ABR thresholds for *Ikzf2*^{*cello/del890*} compound heterozygotes at 1-month of age, showing significantly elevated thresholds (40 dB SPL) at all frequencies tested compared to *Ikzf2*^{+/+}, *Ikzf2*^{*cello*+} and *Ikzf2*^{*del890*+} control colony mates. Data shown are averaged thresholds \pm s.e.m. Number of biologically independent samples: *Ikzf2*^{+/+} n=4, *Ikzf2*^{*cello*+} n=2, *Ikzf2*^{+/del890} n=4, *Ikzf2*^{*cello/del890*} n=5. *Ikzf2*^{*cello/del890*} vs *Ikzf2*^{+/+} p-values: 8 kHz = 0.011, 16 kHz = 0.002, 32 kHz <0.0001, Click = 0.0001; *Ikzf2*^{*cello/del890*} vs *Ikzf2*^{*cello*+} p-values: 8 kHz = 0.078, 16 kHz = 0.034, 32 kHz = 0.001, Click = 0.001; *Ikzf2*^{*cello/del890*} vs *Ikzf2*^{+/del890} p-values: 8 kHz = 0.025, 16 kHz = 0.009, 32 kHz = 0.0002, Click = 0.0002. Significance was assessed by one-way ANOVA with Tukey post-hoc test.



Extended Data Figure 6. The MET current is 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)⁴⁷. At hyperpolarised membrane potentials (-121 mV), saturating excitatory bundle stimulation (i.e., towards the taller stereocilia) elicited a large inward MET current from both *Ikzf2^{cello/+}* and *Ikzf2^{cello/cello}* OHCs, while inhibitory bundle stimulation (i.e. 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 (i.e., disappearance of the resting current) during inhibitory bundle displacements, arrowheads indicate the larger resting MET current at +99 mV compared to -121 mV.

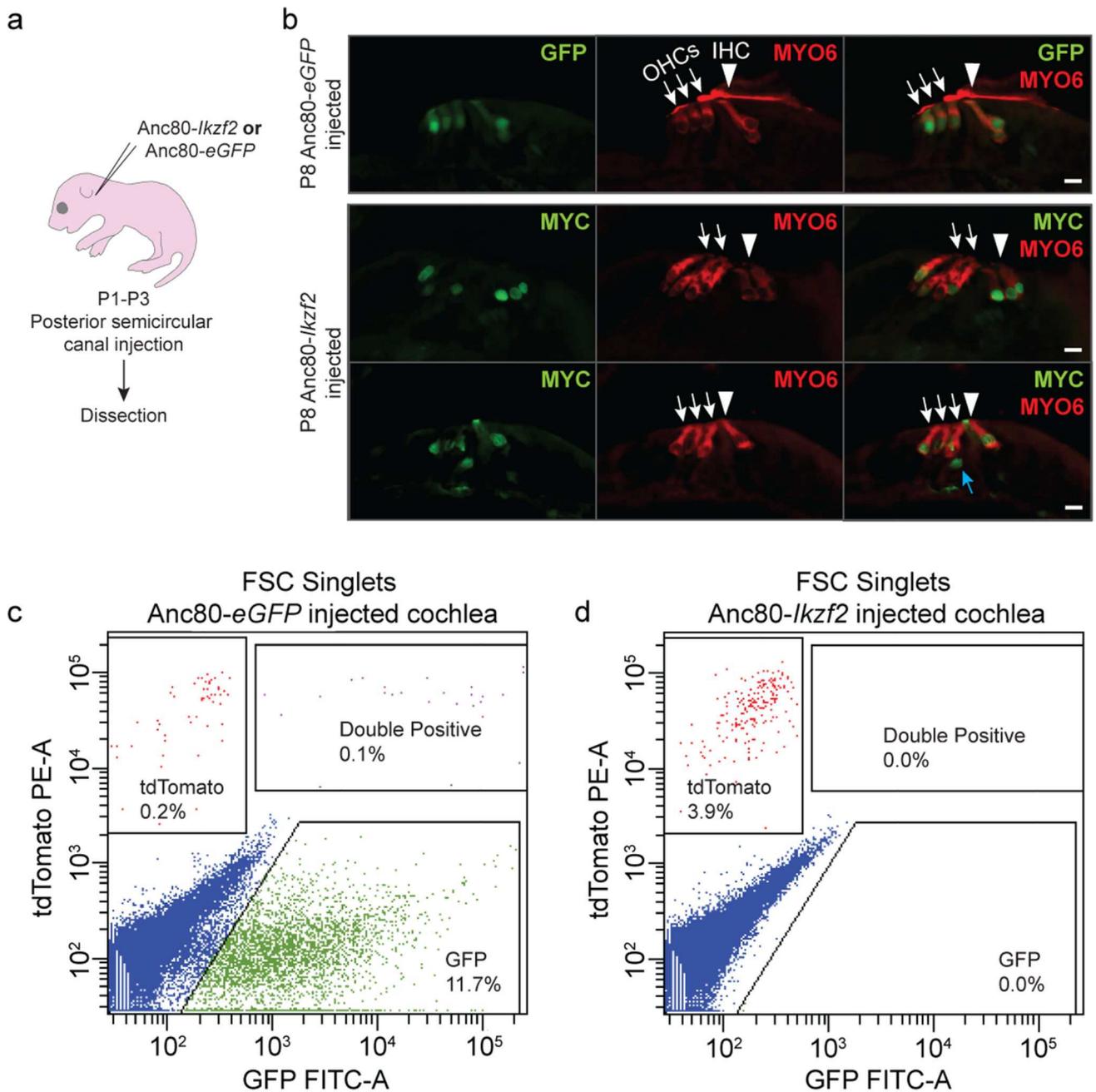
(c) Peak-to-peak current-voltage curves obtained from *Ikzf2^{cello/+}* (n=10 biologically independent samples) and *Ikzf2^{cello/cello}* (n=8 biologically independent samples) OHCs at P9. The maximal MET current and the resting open probability of the MET channel were found to be similar between the two genotypes. Data shown are mean values \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 : Marcotti and Kros, 1999), 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 shown are mean values \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. Data shown are mean values \pm s.e.m. *n.s.* non-significant, two-sided Welch' s t-test.

(i) NanoString validations of genes downregulated in P8 *Ikzf2^{cello/cello}* cochleae at P16. Data shown are mean normalized reads relative to wild-type \pm SD (n = 4 biologically independent samples per genotype). *Ppp17r1* in *Ikzf2^{cello/cello}* vs *Ikzf2^{+/+}* p -value = 0.038, *Ppp17r1* in *Ikzf2^{cello/cello}* vs *Ikzf2^{cello/+}* p -value = 0.037. Significance was assessed by two-sided Welch' s t-test.



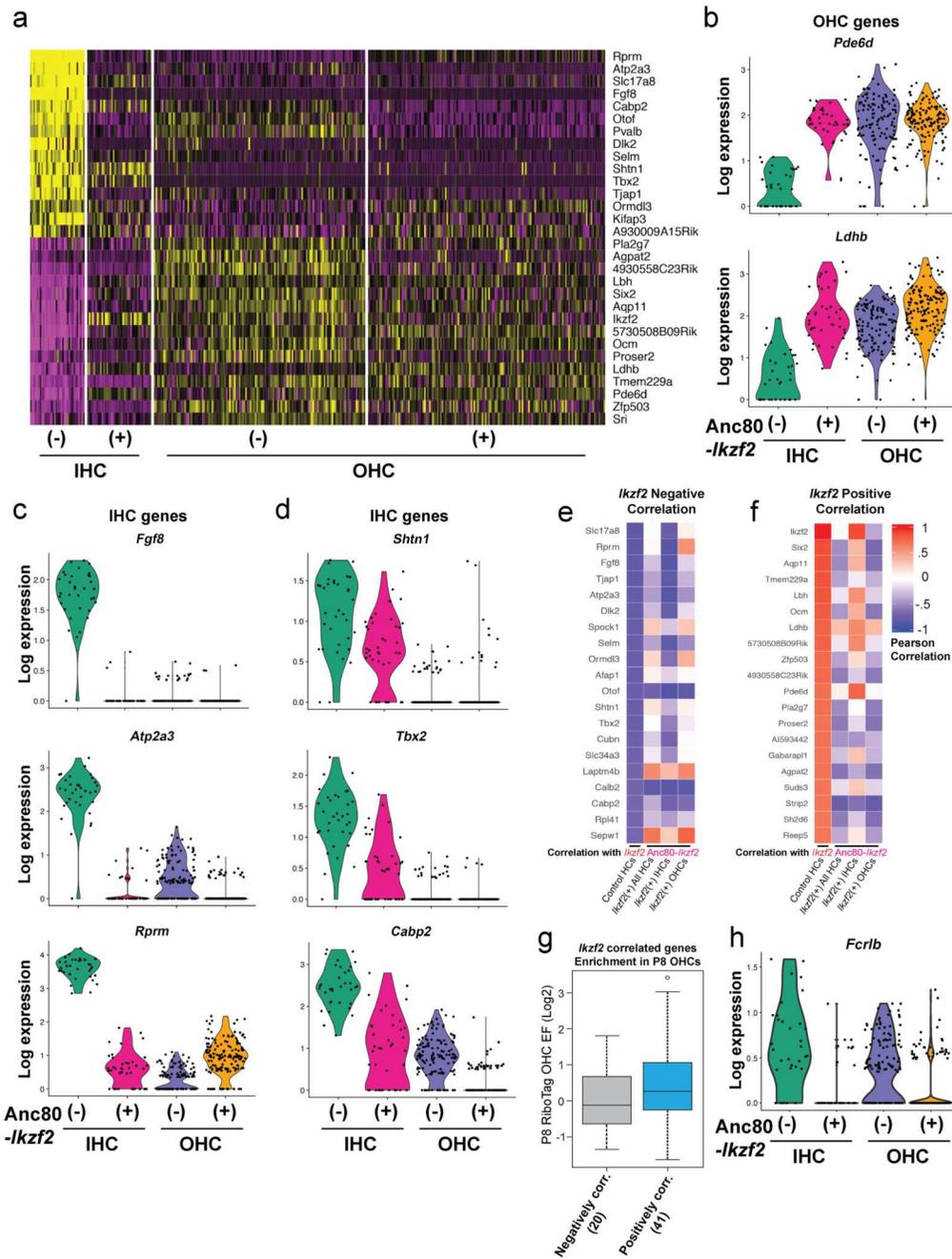
Extended Data Figure 7. Transduction of cochlear HCs using Anc80L65 and HC enrichment by flow cytometry.

(a) Schematic representation of inner ear viral gene delivery via the posterior semicircular canal of CD-1 mice for HC marker immunolabeling.

(b) Immunolabeling for GFP in the *Anc80-eGFP* injected, and MYC in the *Anc80-Ikzf2* injected ears, showed mainly HC 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 and white arrowheads indicate IHCs. Scale = 10 μ m.

(c-d) Fluorescence activated cell sorting (FACS) 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+GFP double positive cells were collected. For the *Anc80-Ikzf2* transduced cochlear sample, HCs were gated based on tdTomato single positive expression and collected.



Extended Data Figure 8. Transcriptional conversion of Anc80-Ikzf2 transduced IHCs.

(a) Heatmap for the top 30 differently expressed genes between all HCs 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* = 2.03E-12, *Ldhb* = 3.74E-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* = 3.30E-14, *Atp2a3* = 2.46E-13, *Rprm* = 2.27E-13 (Kruskal-Wallis test followed by post-hoc pairwise Wilcoxon Ranked Sum test adjusted for multiple comparisons).

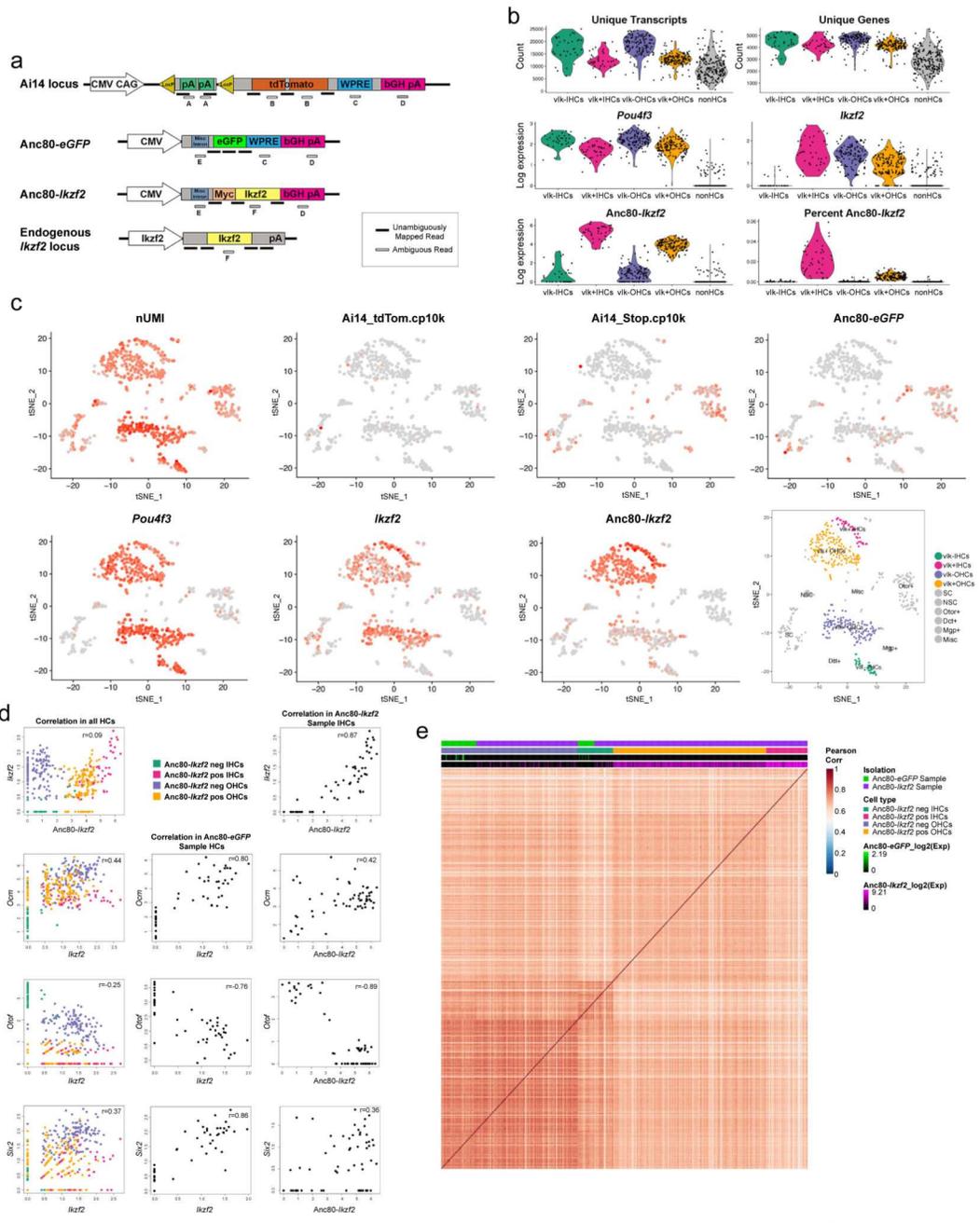
(d) IHC enriched genes that show only moderately reduced expression in Anc80-*Ikzf2*(+) IHCs. Anc80-*Ikzf2*(-) IHC (n=34) vs. Anc80-*Ikzf2*(+) IHC (n=40) FDR: *Shtn1* = 8.59E-05, *Tbx2* = 3.88E-08, *Cabp2* = 1.40E-10 (Kruskal-Wallis test followed by post-hoc pairwise Wilcoxon Ranked Sum test adjusted for multiple comparisons).

(e-f) Top 20 genes negatively (e) or positively (f) correlated with *Ikzf2* expression in control HCs, shown alongside corresponding correlations of gene expression within all Anc80-*Ikzf2* transduced HCs, Anc80-*Ikzf2* transduced IHCs, or Anc80-*Ikzf2* transduced OHCs. See also Extended Data Figure 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, BG, 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 BG (n=13,103) ($p = 0.025$, two-sided Wilcoxon's test). Black line represents median enrichment factor (EF, log₂ fold change), box demarcates 1st and 3rd quartiles, whiskers demarcate 1st and 3rd quartile $\pm 1.5 \times \text{IQR}$ values, dots represent single outliers.

(h) One of the most differentially expressed genes we observed in our scRNA-seq experiment was *Fcrlb*, a gene which encodes an Fc receptor like protein, and whose expression in the ear has not been previously described. *Fcrlb* is significantly downregulated in Anc80-*Ikzf2*(+) HCs. Anc80-*Ikzf2*(-) IHC (n=34) vs. Anc80-*Ikzf2*(+) IHC (n=40) FDR= 4.89E-06. Anc80-*Ikzf2*(-) OHC (n=132) vs. Anc80-*Ikzf2*(+) OHC (n=148) FDR= 6.88E-08 (Kruskal-Wallis test followed by post-hoc pairwise Wilcoxon Ranked Sum test adjusted for multiple comparisons).

See also Supplementary Tables 8-11.



Extended Data Figure 9. Single cell RNA-seq allows for high-resolution discrimination of cell types and their transcriptional changes due to overexpression of *Izkf2*/helios.

(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 (i.e. ambiguous reads).

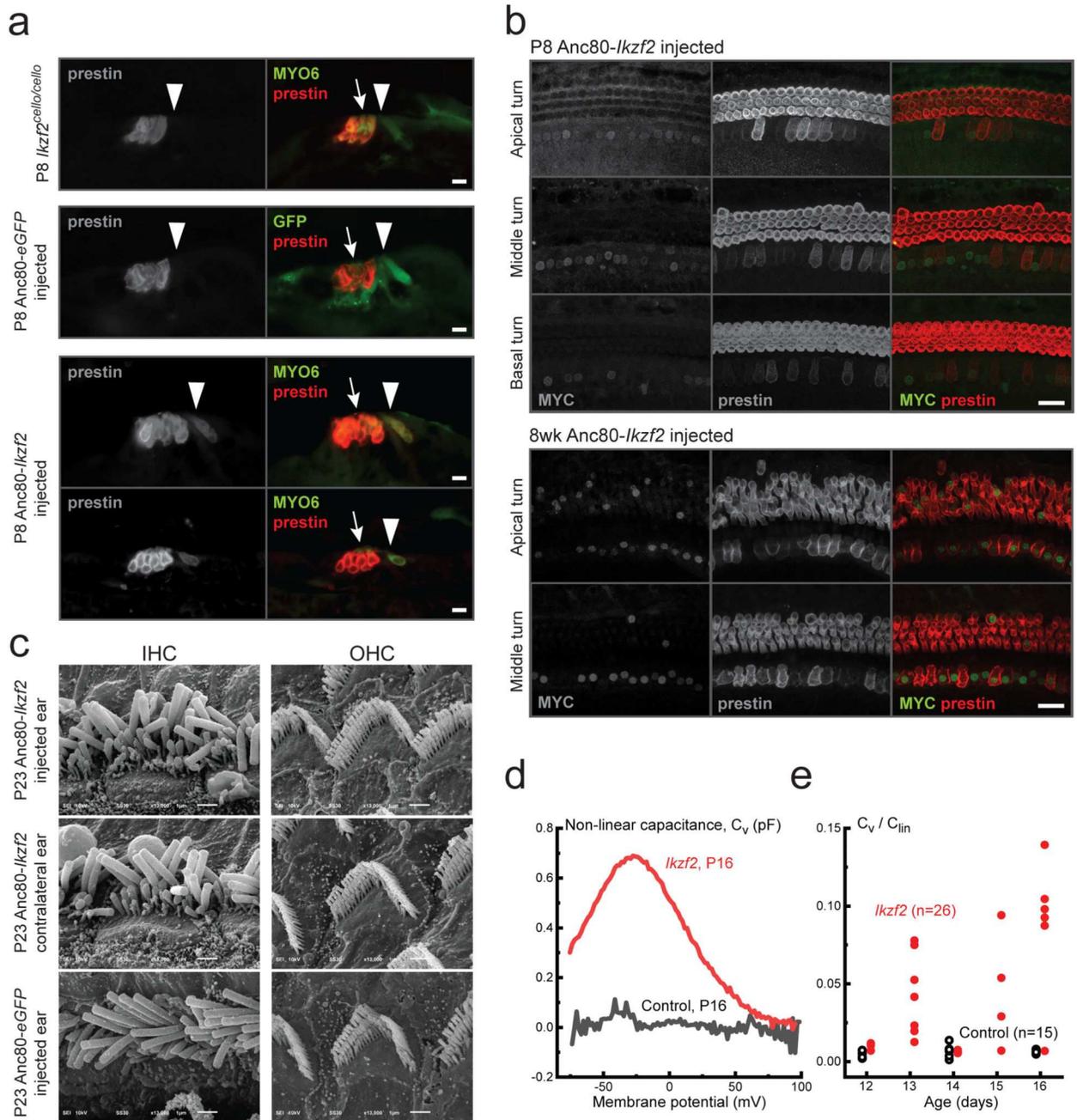
(b) Violin plots of the overall scRNA-seq detection metrics, including number of unique molecules detected in each of the major cell type cluster identified (low *Anc80-Izkf2* expressing IHCs: vIk- IHCs n=34; low *Anc80-Izkf2* expressing OHCs: vIk- OHCs n=132;

high *Anc80-Ikzf2* expressing IHCs: vIk+ IHCs n=40; high *Anc80-Ikzf2* expressing OHCs: vIk+ OHCs n=140; and non-HCs: NonHCs n=219).

(c) FeaturePlots with red showing higher expression across all profiled cells, including cells identified as non-HCs. Expression from loci captured with custom annotation shown to support cluster identification. A final labeled tSNE 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, and other clusters defined by the highest differentially expressed marker gene).

(d) Pearson correlation scatter plots for selected genes within all profiled HCs, HCs from the *Anc80-eGFP* sample, or IHCs from the *Anc80-Ikzf2* sample.

(e) A Pearson correlation heatmap of all HCs detected showing overall transcriptional similarities between the non-transduced IHCs and OHCs, along with the *Anc80-Ikzf2* transduced IHCs and OHCs.



Extended Data Figure 10. Helios 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). Additionally, 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=10 μ m

(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 = 20 μ m.

(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=1 μ m. Number of biologically independent samples (P16-P23): Anc80-*Ikzf2* injected cochlea n=8, Anc80-*Ikzf2* contralateral cochlea n=6, Anc80-*GFP* injected cochlea n=3.

(d) Representative traces of the voltage-dependent (non-linear) component of the membrane capacitance (an electrical “signature” of electromotility) in the IHCs of Anc80-*Ikzf2* injected mouse (red) and its non-injected littermate (black). Mice were injected with Anc80-*Ikzf2* at P2 and recorded at P16.

(e) Normalized maximal non-linear 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, the total number of cells is indicated in parentheses. Since 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 non-linear capacitance while the other IHCs do. In the IHCs with maximal non-linear capacitance of more than 0.25 pF (due to presumable *Ikzf2* expression), the parameters of the Boltzmann fit were as following (Mean \pm SEM): $Q_{max} = 0.10\pm 0.02$ pC; $V_{pk} = -31\pm 1$ mV; $z = 0.91\pm 0.02$; $C_{lin} = 11.7\pm 1.2$ pF; $C_{sa} = 0.14\pm 0.07$ pF (n=12). For information on the fitting procedure, see methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank L. Vizor, J. Sanderson and W. Chien for technical help and Z. Ahmed for helpful 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.

References

1. Ehret G. Development of absolute auditory thresholds in the house mouse (*Mus musculus*). *J Am Audiol Soc.* 1976; 1:179–84. [PubMed: 956003]
2. Bielefeld EC, Tanaka C, Chen G, Henderson D. Age-related hearing loss: is it a preventable condition? *Hear Res.* 2010; 264:98–107. [PubMed: 19735708]
3. World Health Organization. Deafness and hearing loss fact sheet. WHO Media centre. 2018. Available at: <http://www.who.int/mediacentre/factsheets/fs300/en/>
4. Mittal R, et al. Recent Advancements in the Regeneration of Auditory Hair Cells and Hearing Restoration. *Front Mol Neurosci.* 2017; 10:236. [PubMed: 28824370]
5. Fang J, et al. Outer hair cell-specific prestin-CreER T2 knockin mouse lines. *Genesis.* 2012; 50:124–131. [PubMed: 21954035]
6. Sanz E, et al. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *PNAS.* 2009; 106:13939–13944. [PubMed: 19666516]
7. Liu H, et al. Characterization of Transcriptomes of Cochlear Inner and Outer Hair Cells. *J Neurosci.* 2014; 34:11085–11095. [PubMed: 25122905]
8. Elkon R, et al. RFX transcription factors are essential for hearing in mice. *Nat Commun.* 2015; 6

9. 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.* 2011; 7:e1002309. [PubMed: 21980309]
10. Janky R, et al. iRegulon : From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. *PLOS Comput Biol.* 2014; 10:e1003731. [PubMed: 25058159]
11. Potter PK, et al. Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nat Commun.* 2016; 7
12. Ashmore JF. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J Physiol.* 1987; 388:323–47. [PubMed: 3656195]
13. Brownell WE, Bader CR, Bertrand D, de Ribaupierre Y. Evoked Mechanical Responses of Isolated Cochlear Outer Hair Cells. *Science* (80-.). 1985; 227:194–196.
14. Marcotti W, Kros CJ. Developmental expression of the potassium current $I_{K,n}$ contributes to maturation of mouse outer hair cells. *J Physiol.* 1999; 520(Pt 3):653–60. [PubMed: 10545133]
15. Li Y, et al. Transcription Factors Expressed in Mouse Cochlear Inner and Outer Hair Cells. *PLoS One.* 2016; 11:e0151291. [PubMed: 26974322]
16. Zinn E, et al. In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *CellReports.* 2015; 12:1056–1068.
17. Landegger LD, et al. A synthetic AAV vector enables safe and efficient gene transfer to the mammalian inner ear. *Nat Biotechnol.* 2017; 35:280–284. [PubMed: 28165475]
18. Roux I, et al. Otoferlin, Defective in a Human Deafness Form, Is Essential for Exocytosis at the Auditory Ribbon Synapse. *Cell.* 2006; 127:277–289. [PubMed: 17055430]
19. Simmons DD, Tong B, Schrader AD, Hawkes AJ. Oncomodulin Identifies Different Hair Cell Types in the Mammalian Inner Ear. *J Comp Neurol.* 2010; 518:3785–3802. [PubMed: 20653034]
20. Santos-Sacchi J. Reversible inhibition of voltage-dependent outer hair cell motility and capacitance. *J Neurosci.* 1991; 11:3096–110. [PubMed: 1941076]
21. Zheng J, et al. Prestin is the motor protein of cochlear outer hair cells. *Nature.* 2000; 405:149–155. [PubMed: 10821263]
22. 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 U S A.* 2011; 108:5825–30. [PubMed: 21436032]
23. Johnson KR, Zheng QY, Noben-Trauth K. Strain background effects and genetic modifiers of hearing in mice. *Brain Res.* 2006; 1091:79–88. [PubMed: 16579977]
24. Hoelter SM, et al. Sighted C3H mice—a tool for analysing the influence of vision on mouse behaviour? *Front Biosci.* 2008; 13:5810–23. [PubMed: 18508624]
25. Kim D, et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013; 14:R36. [PubMed: 23618408]
26. Anders S, Pyl PT, Huber W. Genome analysis HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015; 31:166–169. [PubMed: 25260700]
27. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010; 11:R106. [PubMed: 20979621]
28. Ulitsky I, et al. Expander: from expression microarrays to networks and functions. *Nat Protoc.* 2010; 5:303–322. [PubMed: 20134430]
29. Sharan R, Maron-Katz A, Shamir R. CLICK and EXPANDER: a system for clustering and visualizing gene expression data. *Bioinformatics.* 2003; 19:1787–99. [PubMed: 14512350]
30. Shannon P, et al. Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 2003; 12:2498–2504.
31. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010; 7:248–9. [PubMed: 20354512]
32. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One.* 2012; 7:e46688. [PubMed: 23056405]
33. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009; 4:1073–1081. [PubMed: 19561590]

34. Källberg M, et al. Template-based protein structure modeling using the RaptorX web server. *Nat Protoc.* 2012; 7:1511–22. [PubMed: 22814390]
35. Hardisty-Hughes RE, Parker A, Brown SDM. A hearing and vestibular phenotyping pipeline to identify mouse mutants with hearing impairment. *Nat Protoc.* 2010; 5:177–190. [PubMed: 20057387]
36. Mianné J, et al. Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. *Genome Med.* 2016; 8:16. [PubMed: 26876963]
37. Codner GF, et al. Application of long single-stranded DNA donors in genome editing: generation and validation of mouse mutants. *BMC Biol.* 2018; 16
38. Corns LF, Johnson SL, Kros CJ, Marcotti W. Calcium entry into stereocilia drives adaptation of the mechano-electrical transducer current of mammalian cochlear hair cells. *Proc Natl Acad Sci.* 2014; 111:14918–14923. [PubMed: 25228765]
39. Santos-Sacchi J. Determination of cell capacitance using the exact empirical solution of partial differential Y/partial differential Cm and its phase angle. *Biophys J.* 2004; 87:714–27. [PubMed: 15240504]
40. Santos-Sacchi J, Navarrete E. Voltage-dependent changes in specific membrane capacitance caused by prestin, the outer hair cell lateral membrane motor. *Pflügers Arch.* 2002; 444:99–106. [PubMed: 11976921]
41. Isgrig K, et al. Gene Therapy Restores Balance and Auditory Functions in a Mouse Model of Usher Syndrome. *Mol Ther.* 2017; 25:780–791. [PubMed: 28254438]
42. Finak G, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 2015; 16:278. [PubMed: 26653891]
43. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015; 33:495–502. [PubMed: 25867923]
44. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* Springer; 2009.
45. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics.* 2010; 11:367. [PubMed: 20598126]
46. Geng R, et al. Comprehensive Expression of Wnt Signaling Pathway Genes during Development and Maturation of the Mouse Cochlea. *PLoS One.* 2016; 11:e0148339. [PubMed: 26859490]
47. Corns LF, Johnson SL, Kros CJ, Marcotti W. Calcium entry into stereocilia drives adaptation of the mechano-electrical transducer current of mammalian cochlear hair cells. *Proc Natl Acad Sci.* 2014; 111:14918–14923. [PubMed: 25228765]

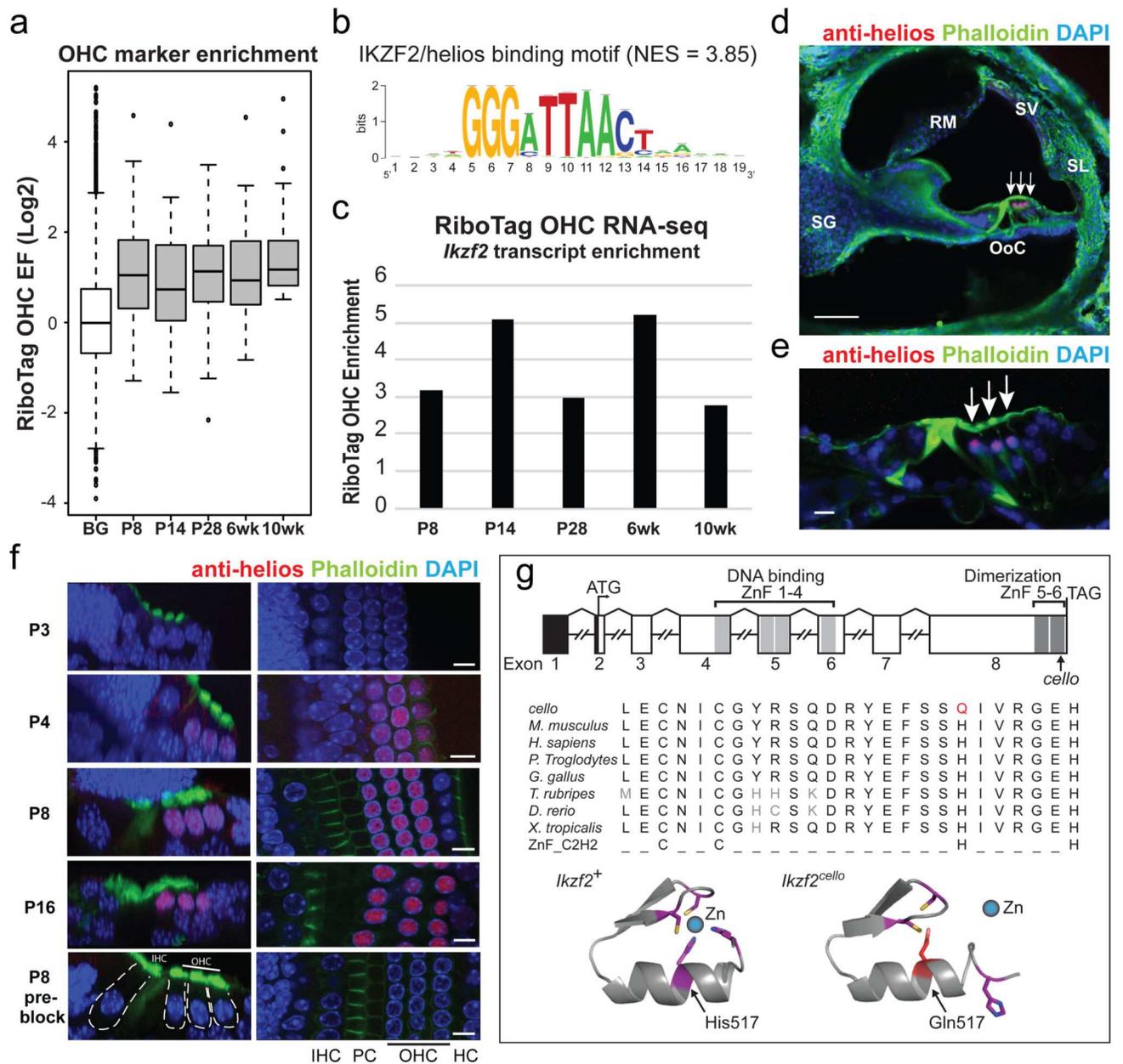


Figure 1. Helios is a candidate regulator of OHC genes.

(a) The 100 OHC marker genes (n=100) are enriched in OHCs at all RiboTag OHC dataset time points compared to expression of all other genes detected (background, BG) (n=13,044). *p*-values: P8 = 1.73E-17, P14 = 6.55E-12, P28 = 1.60E-18, 6wk = 7.79E-18, 10wk = 1.43E-33 (two-sided Wilcoxon's test). Black center line represents median enrichment factor (EF, log₂ fold change), box demarcates 1st and 3rd quartiles, whiskers demarcate 1st and 3rd quartile ± 1.5×IQR values, dots represent single outliers.

(b) Transcription factor binding motif analysis using the 100 highly confident OHC marker genes identifies the binding signature for IKZF2/helios as significantly overrepresented.

NES = normalized enrichment score. NES 3.0 corresponds to a false discovery rate of 3-9% (see Janky et al., 2014).

(c) *Ikzf2* transcript enrichment in OHCs as measured by the 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=50 μ m.

(e) Helios expression is maintained in wild-type OHCs at 1-month (white arrows, n=3 biologically independent samples). Scale=10 μ m.

(f) Helios is detected in wild-type OHCs from P4 and is maintained in mature P16 OHCs (P3 n=2, P4 n=4, P8 n=4, P16 n=4 biologically independent samples). Loss of labelling when the anti-helios antibody is 'pre-blocked' with its immunizing peptide confirms specificity (n=5 biologically independent samples). Scale=10 μ m.

(g) The genomic/domain structure of *Ikzf2*/helios. Black = 5' untranslated region, light grey = N-terminal DNA-binding domain, dark grey = C-terminal dimerization domain. The *Ikzf2^{cello}* mutation lies in ZnF6. Further alignment of the helios ZnF6 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 *Ikzf2⁺* ZnF6 and mutant *Ikzf2^{cello}* ZnF6 illustrates the requirement of residue His517 for zinc-coordination, which is not possible when residue Gln517 is substituted. HC, Hensen's cells; IHC, inner hair cells; OoC, organ of Corti; OHC, outer hair cells; PC, pillar cells; RM, Reissner's membrane; SG, spiral ganglion; SL, spiral ligament; SV, stria vascularis.

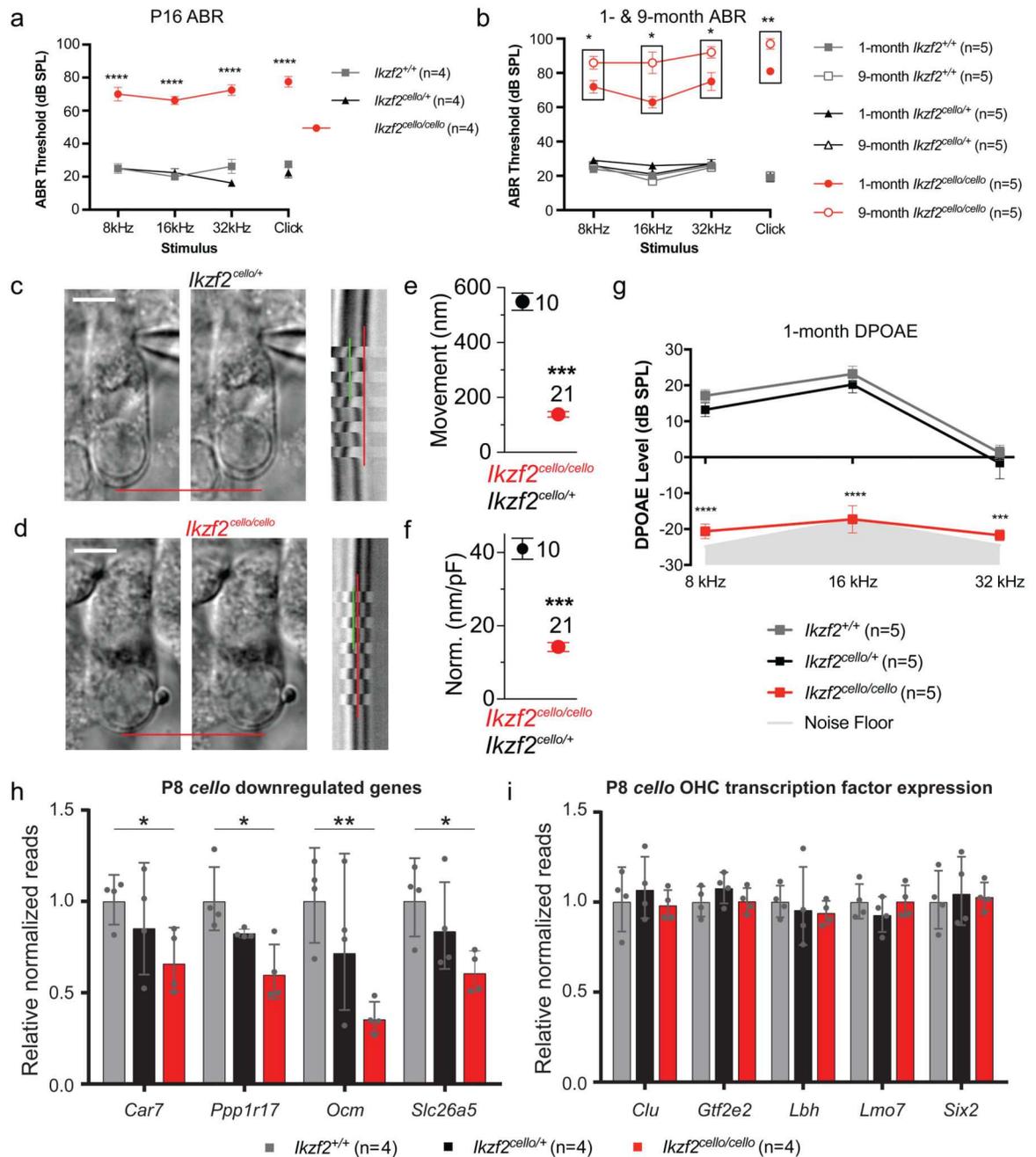


Figure 2. *Ikzf2/helios* is required for hearing and OHC electromotility.

(a-b) Averaged ABR thresholds for *cello* mice at P16 (a, n=4 biologically independent animals per genotype) and 1- and 9-months of age (b, n=5 biologically independent animals per genotype for each time point). Age-matched *Ikzf2*^{+/+} and *Ikzf2*^{cello/+} controls display thresholds within the expected range (15 – 30 dB SPL) at all time-points tested. Data shown are averaged thresholds ± s.e.m. P16 *Ikzf2*^{cello/cello} vs *Ikzf2*^{+/+} (a) *p*-values: 8 kHz <0.0001, 16 kHz <0.0001, 32 kHz <0.0001, Click <0.0001. P16 *Ikzf2*^{cello/cello} vs *Ikzf2*^{cello/+} (a) *p*-values: 8 kHz <0.0001, 16 kHz <0.0001, 32 kHz <0.0001, Click <0.0001. 1-month

Ikzf2^{cello/cello} vs 9-month *Ikzf2^{cello/cello}* (b) *p*-values: 8 kHz = 0.0284, 16 kHz = 0.0166, 32 kHz = 0.0303, Click = 0.0042. Significance was assessed by one-way ANOVA with Tukey post-hoc test (a) or two-sided Welch' s t-test (b). See also Extended Data Figure 4.

(c-d) Images showing a patch pipette attached to an OHC from control *Ikzf2^{cello/+}* (c) and mutant *Ikzf2^{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 from step from –64 mV to +56 mV, highlighting the shorting of the cells. Scale=5 μ m. Also shown are time-based z-stack projections (right), where red lines indicate the resting position of the basal membrane and the green lines indicate the movement. *Ikzf2^{cello/+}* n = 10 and *Ikzf2^{cello/cello}* n = 21 z-stack projections (one set per OHC) from 5 biologically independent animals per genotype.

(e-f) Average movement was significantly reduced in *Ikzf2^{cello/cello}* OHCs compared to *Ikzf2^{cello/+}* at P16–P18 (e), even after normalization to respective membrane capacitance (f) (for this set of recordings, *Ikzf2^{cello/+}*: 13.6 \pm 0.4 pF; *Ikzf2^{cello/cello}*: 10.0 \pm 0.3 pF). Data shown are averaged movement \pm s.e.m. *Ikzf2^{cello/+}* n = 10 and *Ikzf2^{cello/cello}* n = 21 OHCs from 5 biologically independent animals per genotype. *p*-value <0.0001, two-sided Welch' s t-test.

(g) Averaged DPOAE responses for *cello* mice at 1-month of age (n=5 biologically independent animals per genotype). Data shown are averaged thresholds \pm s.e.m.

Ikzf2^{cello/cello} vs *Ikzf2^{+/+}* *p*-values: 8 kHz <0.0001, 16 kHz <0.0001, 32 kHz = 0.0004. P16 *Ikzf2^{cello/cello}* vs *Ikzf2^{cello/+}* *p*-values: 8 kHz <0.0001, 16 kHz <0.0001, 32 kHz = 0.0012.

Significance was assessed by one-way ANOVA with Tukey post-hoc test.

(h-i) NanoString validations of genes downregulated in *Ikzf2^{cello/cello}* cochleae at P8 (h) and results showing no change in expression of other OHC TFs (i). Data shown are mean normalized reads relative to wild-type \pm SD (n=4 biologically independent samples per genotype). *Ikzf2^{cello/cello}* vs *Ikzf2^{+/+}* *p*-values: *Car7* = 0.028, *Ppp17r1* = 0.006, *Ocm* = 0.017, *Slc26a5* = 0.017 (two-sided Welch' s t-test).

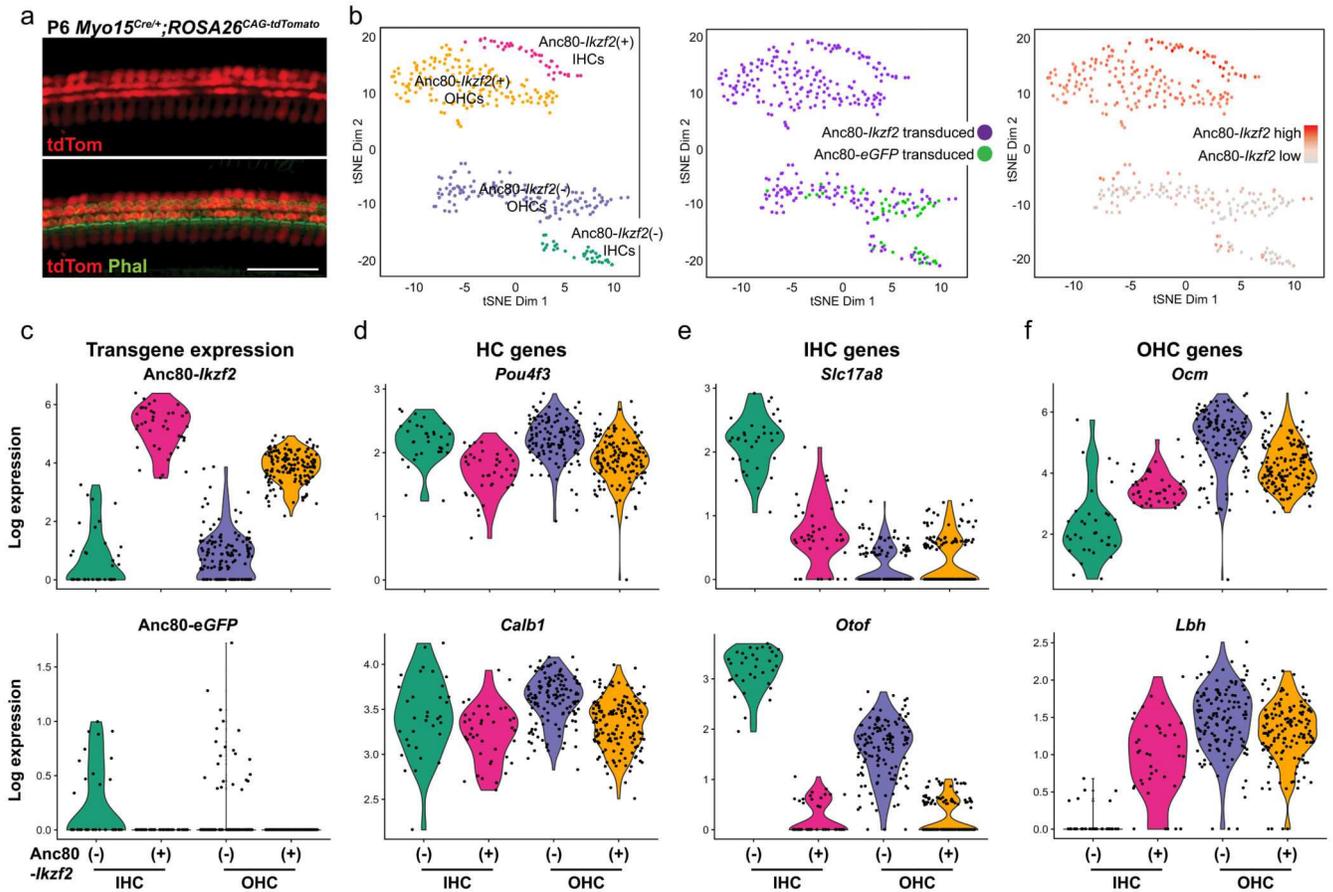


Figure 3. Partial transcriptional conversion of Anc80-*Ikzf2* transduced IHCs identified by scRNA-seq.

(a) Representative *Myo15^{Cre/+}; ROSA26^{CAG-tdTomato}* cochlear whole-mount. *Myo15-Cre*-driven tdTomato expression is HC specific at P6 (n=3 biologically independent samples with similar results). Scale=50 μ m.

(b) tSNE plots of all cochlear HCs profiled by scRNA-seq, including the cluster to which each cell was assigned, the experimental origin of each cell (Anc80-*Ikzf2* or Anc80-*eGFP* injected cochlea), and the relative transcript abundance of Anc80-*Ikzf2* measured in each cell.

(c) Anc80-*Ikzf2* is highly expressed in the Anc80-*Ikzf2*(+) IHCs and OHCs, whereas Anc80-*eGFP* expression is only seen in the cells assigned to the Anc80-*Ikzf2*(-) IHC and OHC clusters. Dots represent the expression values of individual cells, with width of violins summarizing overall relative distribution of expression.

(d) Canonical HC markers are highly expressed in all HC clusters, and not notably changed as a result of Anc80-*Ikzf2* expression.

(e) 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: *Slc17a8* = 2.25E-12, *Otof* = 6.76E-14. Significance was assessed by 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-Ikzf2*(+) IHCs. *Anc80-Ikzf2*(-) IHC (n=34) vs. *Anc80-Ikzf2*(+) IHC (n=40) FDR: *Ocm* = 3.65E-08, *Lbh* = 1.81E-10. Significance was assessed by Kruskal-Wallis test followed by post-hoc pairwise Wilcoxon Ranked Sum test adjusted for multiple comparisons. See also Extended Data Figures 8 and 9.

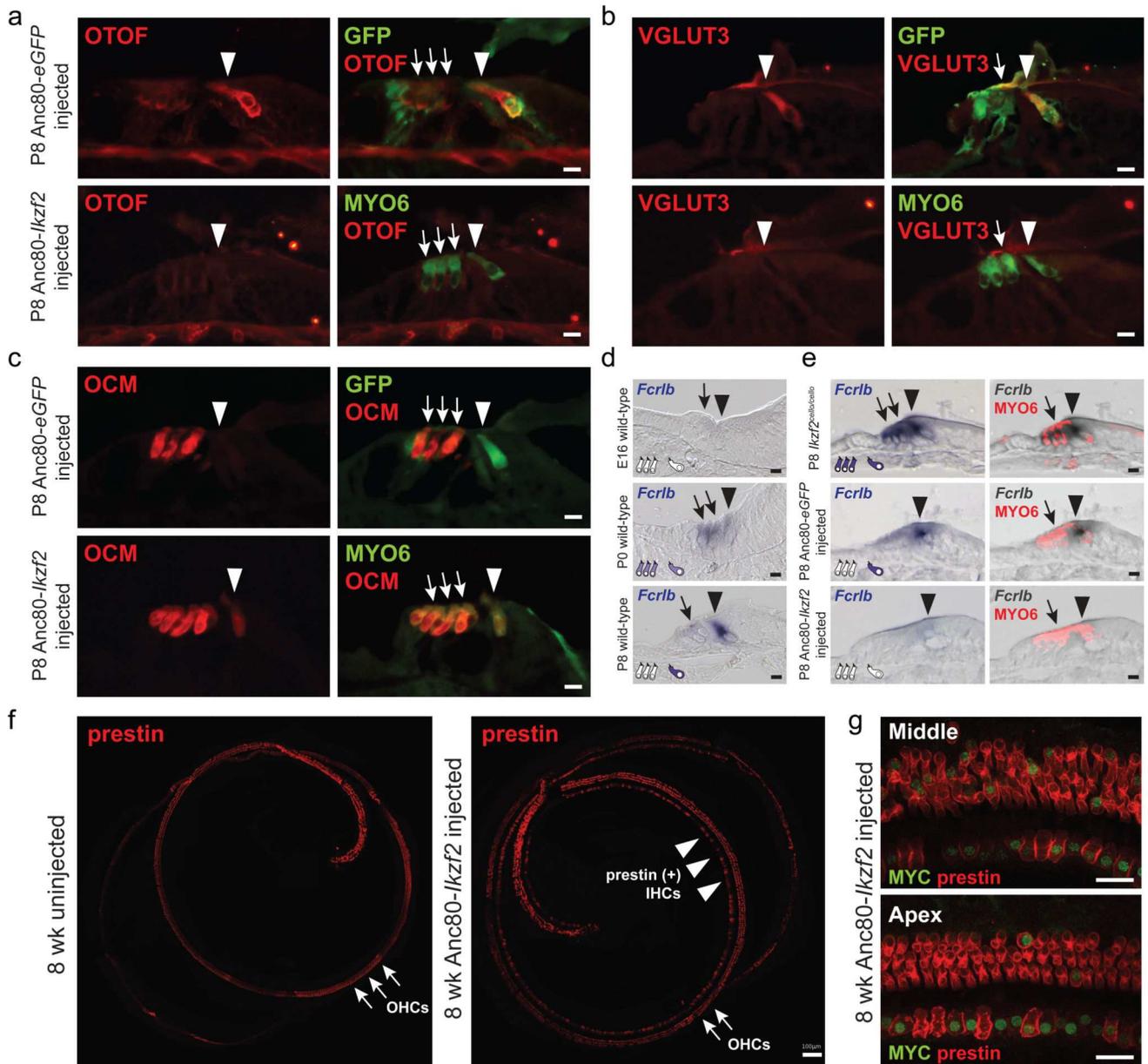


Figure 4. Helios overexpression modulates expression of HC markers.

(a-b) IHC markers OTOF and VGLUT3 are downregulated in Anc80-*Ikzf2* transduced IHCs (n=3 biologically independent samples). Arrows = OHCs, arrowheads = IHCs. Scale=10 μ M.

(c) The OHC marker OCM is expressed in Anc80-*Ikzf2* transduced IHCs (n=3 biologically independent samples per condition). Arrows = OHCs, arrowheads = IHCs. Scale=10 μ m.

(d) *Fcrlb* expression during wild-type mouse inner ear development as detected by *in situ* hybridization. While at E16, *Fcrlb* expression is not detected in the inner ear, by P0 it is detected in both IHCs and OHCs and by P8, *Fcrlb* expression is largely restricted to the IHCs (n=3 biologically independent samples per time point). Scale=10 μ m.

(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 HCs (n=3 biologically independent samples per condition). Scale=10 μ m.

(f-g) Expression of prestin can be seen in Anc80-*Ikzf2* transduced IHCs up to 8-weeks of age (n=3 biologically independent samples at 6-8 weeks) (f, scale=100 μ m), and overlaps with Myc staining (g, scale=20 μ m). See also Extended Data Figure 10.