

CHD7 regulates otic lineage differentiation and deafness gene expression in human inner ear organoids

Jing Nie

Indiana University School of Medicine <https://orcid.org/0000-0002-7141-1830>

Yoshitomo Ueda

Indiana University School of Medicine

Alexander Solivais

Indiana University School of Medicine

Eri Hashino (✉ ehashino@iu.edu)

Indiana University School of Medicine

Article

Keywords:

Posted Date: December 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1132148/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Communications on November 17th, 2022. See the published version at <https://doi.org/10.1038/s41467-022-34759-8>.

CHD7 regulates otic lineage differentiation and deafness gene expression in human inner ear organoids

Jing Nie¹, Yoshitomo Ueda¹, Alexander J. Solivais¹, & Eri Hashino^{1,2,*}

¹ Department of Otolaryngology-Head and Neck Surgery, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA.

² Stark Neurosciences Research Institute, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA.

*Corresponding Author

Abstract

Mutations in the chromatin remodeling enzyme CHD7 cause CHARGE syndrome, which affects multiple organs including the inner ear. We investigated how *CHD7* mutations affect otic development in human inner ear organoids. We found loss of CHD7 or its chromatin remodeling activity leads to complete absence of hair cells and supporting cells, which can be explained by dysregulation of key otic development-associated genes in mutant otic progenitors. Further analysis of the mutant otic progenitors suggested that CHD7 can regulate otic genes through a chromatin remodeling-independent mechanism. Results from transcriptome profiling of hair cells revealed disruption of deafness gene expression as a potential underlying mechanism of CHARGE-associated sensorineural hearing loss. Notably, co-differentiating *CHD7* knockout and wild-type cells in chimeric organoids partially rescued mutant phenotypes by restoring otherwise severely dysregulated otic genes. Taken together, our results suggest that *CHD7* plays a critical role in regulating human otic lineage differentiation and deafness gene expression.

Introduction

CHARGE syndrome is a congenital multi-organ disorder mainly caused by *de novo* mutations in the *CHD7* gene¹, which encodes an ATP-dependent chromatin remodeling protein that regulates target genes expression via changes in nucleosome accessibility². The most prevalent clinical features of CHARGE syndrome include malformation of the inner ear structures, accompanied by prelingual deafness and vestibular dysfunctions³⁻⁵. Previous mouse

model studies provided important insights into CHARGE syndrome inner ear phenotypes⁶⁻¹³, but it remains unclear how loss of *CHD7* affects the transcriptomes in key otic lineage cell types such as otic progenitors, hair cells, and supporting cells. We have previously established a pluripotent stem cell-derived inner ear organoid system capable of recapitulating inner ear development through sequential generation of non-neural ectoderm (NNE), otic-epibranchial progenitor domain (OEPD), otic placodes/pits, and otic vesicles. The otic progenitor cells in the otic vesicles go through self-guided differentiation and form mechanosensitive hair cells, supporting cells, and neurons forming synapses with the hair cells¹⁴⁻¹⁷. In this study, we generated multiple *CHD7* mutant human embryonic stem cell (hESC) lines including a complete knockout (KO) and a patient-specific missense mutant, and used these lines for studying mechanisms underlying inner ear phenotypes due to mutations in *CHD7*.

Results

***CHD7* is expressed at key otic developmental stages.**

We first accessed the gene expression profiles of *CHD7* during human inner ear organoid differentiation. To circumvent the low specificity issue of available *CHD7* antibodies (**Supplementary Fig. 1**), we tagged the endogenous *CHD7* gene with a 3×Flag tag in hESCs with CRISPR (**Supplementary Fig. 2**). Anti-Flag detection of *CHD7*-3×Flag eliminated nearly all non-specific binding (**Fig. 1a**), and showed *CHD7* expression in all early otic developmental stages, including NNE, OEPD, otic pits, and otic vesicles (**Fig. 1b–f**). In addition, it revealed *CHD7* expression in hair cells as well as weaker expression in supporting cells (**Fig. 1g**). These results revealed that *CHD7* is expressed in all key stages of inner ear development.

Loss of CHD7 or its chromatin remodeling function leads to failure of sensory epithelium formation.

To model phenotype manifestation of CHARGE syndrome, we next created mono- and bi-allelic *CHD7* KO hESC lines by targeting the first of the 38 coding exons of *CHD7* with CRISPR. Frameshift indel formation at this early coding region leads to nonsense-mediated mRNA decay (NMD)¹⁸, or if the mutated transcript escapes the NMD mechanism, it results in early truncation of the CHD7 protein prior to any functional domain, therefore creating a null deletion. We chose two clonal lines with frameshift indels in one, or both alleles, and designated them as *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC lines for further analysis (**Fig. 2a and Supplementary Fig. 3–4**). Western blotting confirmed the complete elimination of CHD7 protein in the *CHD7*^{KO/KO} mutant, and reduced protein levels in *CHD7*^{KO/+} (**Fig. 2b**). When differentiated into inner ear organoids, both mutant lines generated morphologically normal PAX2⁺ PAX8⁺ EPCAM⁺ otic vesicles at differentiation day 20 (d20) (**Fig. 2l, t**). Consistent with the normal morphology of stereocilia-bearing hair cells in heterozygous *Chd7* deficient mice¹⁹, at day 70 (d70), the *CHD7*^{KO/+} mutant organoids generated hair cells and supporting cells that are indistinguishable from the WT control. In addition, these mono-allelic KO hair cells exhibited stereocilia with a normal morphology (**Fig. 2m–o**). In contrast, neither hair cells nor supporting cells were observed in *CHD7*^{KO/KO} organoids (**Fig. 2u–w**) ($n = 233$ aggregates from 7 independent organoid cultures).

In addition to the complete knockouts, we also introduced a patient-specific missense mutation in *CHD7*. A heterozygous serine to phenylalanine substitution at the CHD7 residue 834 (p.S834F, c.2501C>T) was described by two independent clinical studies in three CHARGE patients and in a patient with idiopathic hypogonadotropic hypogonadism^{20, 21}. This missense

mutation occurs in a highly conserved sequence motif at one of the chromodomains of CHD7. Previous biochemical analysis demonstrated that this single amino acid substitution completely abolishes CHD7's ATPase activity. Consistent with this mutant protein's inability to hydrolyze ATP, its chromatin remodeling activity is also completely abolished². We used a CRISPR base editor²² and created mono- and bi-allelic *CHD7* S834F mutations in hESCs (**Fig. 2c and Supplementary Fig. 5**). When differentiated towards the otic lineage in inner ear organoids, the mono-allelic and bi-allelic *CHD7* S834F mutants essentially phenocopied the corresponding KO phenotypes at both d20 and d70. The *CHD7*^{S834F/+} mutant gave rise to morphologically normal otic vesicles, supporting cells, and stereocilia-bearing hair cells (**Fig. 2h–k**), while no hair cells or supporting cells were found in the *CHD7*^{S834F/S834F} mutant organoids ($n = 210$ aggregates from 5 independent organoid cultures), despite the presence of normal-looking otic vesicles (**Fig. 2p–s**). Collectively, these results demonstrated that hair cell and supporting cell derivation requires the ATP-dependent chromatin remodeling activities of CHD7.

Loss of CHD7 leads to dysregulation of otic development genes in the otic progenitors.

To investigate the mechanisms underlying the failure of hair cell and supporting cell generation in d70 *CHD7*^{KO/KO} organoids, we performed transcriptome profiling in the seemingly normal d20 *CHD7*^{KO/KO} otic progenitor cells using scRNA-seq. We have previously generated a *PAX2*-2a-nGFP (*PAX2*^{nG}) reporter hESC line to label the otic progenitor cells with nuclear GFP (Hashino lab, unpublished data), and all our four *CHD7* mutant lines were built on this *PAX2*^{nG} genetic background (**Supplementary Fig. 6**). To enrich otic progenitors, we dissociated d20 WT and *CHD7*^{KO/KO} organoids and FACS-isolated the *PAX2*^{nG+} cells (**Supplementary Fig. 7a**). scRNA-seq analysis suggested that the WT *PAX2*^{nG+} cells mainly consisted of otic progenitor

cells (89.6%). As neuroblast cells that delaminate from the otic vesicles and hindbrain neurons are also known to be *PAX2*-positive²³, these cell populations were also present in the WT dataset. In the *CHD7*^{KO/KO} samples, the otic progenitors made up a smaller percentage (39.7%), while two additional clusters of *PAX2*^{nG+} *EPCAM*⁻ non-epithelial cells made up more than half of all mutant cells, with one of these two clusters showing high levels of cell cycle marker gene expression. In addition to these abnormal *PAX2*⁺ non-epithelial cells, *CHD7* depletion also appeared to affect the neuroblast cells, as the mutant cells only made up 4.9% of all neuroblast cells (**Fig. 3a** and **Supplementary Fig. 7b–d**).

As it is the otic progenitor cell population that gives rise to hair cells and supporting cells, we focused on the otic progenitors for further analysis (**Fig. 3b**). Differential gene expression analysis revealed 323 upregulated genes and 129 downregulated genes in *CHD7*^{KO/KO} otic progenitors relative to WT control (fold change ≥ 2.0 , $P \leq 1 \times 10^{-10}$) (**Fig. 3d**). Among these differentially expressed (DE) genes, 15 deafness genes listed in the OtoSCOPE gene panel²⁴ were downregulated, including *TBX1*, *LMX1A*, and *SOX10* (**Fig. 3c**). Gene set enrichment (GSE) analysis of the downregulated genes using the iDEA pipeline²⁵ suggested that genes in many inner ear development-related Gene Ontology (GO) categories were disrupted, including *DLX5* and *SIX1* in the inner ear morphogenesis GO term (**Fig. 3c, g**). In addition, GSE analysis also revealed dysregulation of FGF and WNT signaling pathways, as well as enrichment of multiple gene sets closely related to the cellular functions of the otic progenitors, such as cell junction organization and extracellular matrix organization (**Fig. 3g**). To examine whether cell lineage identity is affected, we compared our dataset with a list of otic lineage-specific genes systematically identified by Hartman et al.²⁶. While 57.7% of these otic-specific genes remained largely unaffected, 38.5% of them, including the highest ranked otic-specific genes *FBXO2*,

COL9A2, and *OC90*, were significantly downregulated in the *CHD7*^{KO/KO} otic progenitors, suggesting the otic identity is partially impaired (**Fig. 3c, e**). Consistent with the aberrant otic gene expression profile, a large number of genes not normally found in the developing otic vesicle were significantly upregulated, including a cohort of HOX genes (e.g., *HOXB9*, *HOXA7*, and *HOXD3*) from the embryonic skeletal system development GO category (**Fig. 3c**). We also noticed upregulation of multiple cell cycle marker genes (e.g., *PCNA* and *MCM3*) (**Fig. 3c**), which could explain the extensive expansion of *PAX2*⁺ epithelial vesicle structures in ~d25 *CHD7*^{KO/KO} organoids (**Fig. 3f and Supplementary Fig. 6**). Taken together, these results suggest that the failure of hair cell and supporting cell derivation from the *CHD7* null mutant stemmed from a multitude of dysregulation at the gene and gene set levels, including downregulation of genes essential to hearing and inner ear development, dysregulation of components and regulators of signaling pathways, cell junction, and extracellular matrix, dysregulation of cell cycle control, as well as a partially drifted otic lineage identity.

A chromatin remodeling-independent function of CHD7 is responsible for regulating a subset of otic genes.

To determine differential gene expression in the *CHD7*^{S834F/+}, *CHD7*^{KO/+}, and *CHD7*^{S834F/S834F} mutants, as well as to confirm the *CHD7*^{KO/KO} scRNA-seq results, we next examined protein level expression of 7 key markers with all the four *CHD7* mutant lines. While SOX2 expression followed the gene dosage changes of *CHD7* and showed moderate dysregulation in the mono-allelic mutants and more severe dysregulation in the bi-allelic mutants (**Fig. 4a–g**), there were markers not following this trend. For sample, *COL9A2* was downregulated in the bi-allelic mutants to the same extent as the mono-allelic mutants, while the

non-otic HOXB9 protein was not upregulated in the two mono-allelic mutants, indicating a more faithful otic lineage identity. In addition, SOX10 showed an opposite direction of dysregulation between the mono- and bi-allelic mutants (**Fig. 4v–ab, aj–aw**). These results demonstrated a complex dysregulation pattern of *CHD7* downstream proteins, which did not simply follow the changes in *CHD7* gene dosage.

With regard to S834F and its corresponding mono- or bi-allelic KO mutant, while most of them showed comparable levels of dysregulation, there are several exceptions. For example, SIX1 and FBXO2 showed significant downregulation in *CHD7*^{KO/KO} compared to *CHD7*^{S834F/S834F}, and DLX5 showed significant downregulation in *CHD7*^{KO/+} compared to *CHD7*^{S834F/+} (**Fig. 4h–u, ac–ai**). Considering the complete abolishment of ATPase and chromatin remodeling activities in the S834F mutant², the differential downstream protein expression between S834F and its corresponding KO mutant revealed the presence of CHD7 function(s) beyond its ATPase and chromatin remodeling activities. Such chromatin remodeling-independent mechanism appears to be solely responsible for regulating key otic genes such as *FBXO2* (**Fig. 4ag–ai**).

Deafness genes were dysregulated in *CHD7* mutant hair cells.

To investigate how decreased *CHD7* expression affects the morphologically normal hair cells and supporting cells at the transcriptome level, we performed scRNA-seq in d70 WT and *CHD7*^{KO/+} organoids. We first labeled the hair cells with a highly specific *POU4F3*-2a-ntdTomato (*POU4F3*^{nT}) fluorescence reporter in the WT (Hashino lab, unpublished data) and *CHD7*^{KO/+} genetic backgrounds (**Supplementary Fig. 8**). We FACS-separated the *POU4F3*^{nT+} and *POU4F3*^{nT-} cells from micro-dissected WT and *CHD7*^{KO/+} d70 organoids (**Supplementary**

Fig. 9a) and performed scRNA-seq of these four groups of cells in four separate reactions. As hair cells only constitutes ~1–2% of all cells in inner ear organoids²⁷, this experimental design allowed an adequate number of hair cells to be collected for downstream analysis. Indeed, we obtained 9,884 hair cells (28.5%) and 6,273 supporting cells (18.1%) when analyzing the merged dataset (**Fig. 5a and Supplementary Fig. 9b–e**). Unsupervised cell clustering grouped hair cells into three clusters, one of them had immature hair cell gene expression profiles. The mature hair cells segregated into WT and *CHD7*^{KO/+} clusters, while the WT and *CHD7*^{KO/+} supporting cells were intermingled in one cluster (**Fig. 5b and Supplementary Fig. 9f–g**), which is consistent with the lower expression levels of *CHD7* in supporting cells and therefore a lesser extent of influence (**Fig. 1g and Supplementary Fig. 10a–b**). In contrast to the upregulation-oriented differential expression pattern in d20 otic progenitors (**Fig. 3d**), the majority of DE genes in d70 hair cells and supporting cells were downregulated (fold change ≥ 2.0 , $P \leq 1 \times 10^{-10}$), suggesting that *CHD7* shifted its predominant role from a transcriptional repressor in the otic progenitors to an activator in the sensory epithelium (**Fig. 5c–d**). GSE analysis of downregulated genes showed enrichment of hair cell differentiation and Notch signaling gene sets in hair cells, and WNT signaling and cell junction gene sets in the supporting cells (**Fig. 5e–h**). Notably, a number of deafness genes (e.g., *SIX1*, *USH1C*, and *STRC*) from the OtoSCOPE gene panel²⁴ were dysregulated in the *CHD7*^{KO/+} hair cells (**Fig. 5e**), providing potential explanations to the cause of hearing loss in individuals with CHARGE syndrome. Collectively, these d70 scRNA-seq results unveiled the dysregulated genes and gene sets in the *CHD7*^{KO/+} hair cells and supporting cells bearing normal morphological properties.

Co-differentiating WT cells with *CHD7* mutant cells in chimeric organoids partially rescued the *CHD7* mutant phenotypes.

When analyzing the d20 scRNA-seq data, we noticed that several signaling ligands, including those involved in BMP, FGF, Notch, TGF β , and WNT signaling pathways were downregulated in the *CHD7*^{KO/KO} otic progenitors (**Fig. 6a**). It is highly likely that dysregulation of signaling ligands also occurred in other cell types and in other developmental stages, and these abnormal levels of signaling cues likely contributed to the *CHD7* mutant phenotypes. To test this possibility, we established a chimeric organoid system by which normal levels of signaling ligands are supplied from WT cells to the neighboring mutant cells. To distinguish between WT and mutant cells, we first labeled WT cells with a cell membrane-bound tdTomato expressed under a ubiquitous pCA promoter at the AAVS1 locus (AAVS1^{mT}) (**Supplementary Fig. 11**). We aggregated a mixture of AAVS1^{mT}-labeled WT hESCs and unlabeled *CHD7*^{KO/KO} hESCs into chimeric organoids and differentiated them towards the otic lineage. As the organoids grew and differentiated, the initial single cells and small clusters expanded into larger clones, forming a mosaic WT-mutant tissue organization. Meanwhile, the *CHD7*^{KO/KO} clones received normal signaling inputs from neighboring otic and non-otic WT tissues throughout the developmental stages (**Fig. 6b**). Immunostaining of d20 chimeric organoids with a *CHD7* antibody confirmed the expected mosaic pattern of the mT⁺ *CHD7*⁺ WT tissues and the mT⁻ *CHD7*⁻ mutant tissues (**Fig. 6c, l**). Remarkably, under this chimeric culture condition, the otherwise severely dysregulated *FBXO2*, *SOX10*, *DLX5*, and *HOXB9* proteins were restored comparable to the WT level (**Fig. 6g–j, p–s**). As the *CHD7* mutant phenotypes most likely resulted from both aberrant extrinsic signaling inputs and dysregulated intrinsic gene expression networks, we did not anticipate this strategy to provide a full rescue for all affected genes. Consistent with this

expectation, we observed partial rescues of COL9A2, SOX2 and SIX1 (**Fig. 6d–f, m–o**). At d70, all derived hair cells were mT⁺ WT cells and no mT⁻ *CHD7* null mutant hair cells were observed in chimeric organoids ($n = 180$ aggregates in 12 independent chimeric organoid cultures) (**Fig. 6k**). The failure of *CHD7*^{KO/KO} hair cell generation likely resulted from the incomplete rescue of key otic genes such as COL9A2, SOX2 and SIX1 at earlier developmental stages (**Fig. 6d–f, m–o**). Collectively, these results demonstrated that the *CHD7* KO phenotypes can be partially rescued, and the drifted otic lineage identity can be partially restored in the otic progenitors in chimeric organoid cultures.

Discussion

In this study, we recapitulated pathogenesis of CHARGE syndrome with human inner ear organoids as a model system. Previous mouse studies did not allow for investigation of mature inner ear phenotypes associated with homozygous *Chd7* mutation, as these mice survive only up to embryonic day 10.5 (E10.5), by which time the otic lineage cells only developed to the otic progenitor stage. Conditional KO of *Chd7* in otic lineage cells has been used as an alternative approach, but none of the Cre recombinases used in these experiments were expressed before E8.5^{7, 8, 10-13}, making it difficult to assess the effect of *Chd7* deficiency initiated at earlier developmental stages. In the stem cell-derived inner ear organoid system, the survival and development of the otic lineage cells are not dependent on the proper development of vital organs such as the heart, making organoids an ideal platform to study embryonically lethal genes. Using this approach, we revealed that genetic ablation of *CHD7* or its chromatin remodeling activity leads to dysregulation of early otic lineage genes and a partially drifted otic lineage identity, resulting in a complete absence of hair cells and supporting cells.

By epigenetically altering the chromatin architecture to modulate the nucleosome accessibility, along with other less understood mechanism(s), *CHD7* exerts transcriptional control of hundreds of tissue-specific downstream genes²⁸⁻³¹. The combined effects of many of these dysregulated downstream genes likely resulted in the CHARGE inner ear phenotypes. However, one of the *CHD7* otic target genes, *SOX2*, appears to be a leading cause. We identified *SOX2* as one of the top differentially expressed genes following *CHD7* deletion. *SOX2* expression in otic progenitors is completely abolished in the bi-allelic *CHD7* KO or S834F mutants, and the corresponding mono-allelic *CHD7* mutations lead to reduced *SOX2* expression. Notably, loss of *Sox2* expression in mouse otic tissues results in failure of hair cell and supporting cell generation³², which phenocopies the absence of hair cell and supporting cells in *CHD7* null organoids. Moreover, reduced *SOX2* expression leads to truncated semicircular canals, shortened cochleae, and hearing impairment^{32, 33}, which are also the hallmark phenotypes in heterozygous *Chd7* mutant mice and individuals with CHARGE syndrome^{19, 34-36}. Thus, the downregulation of *SOX2* alone accounts for many of the *CHD7* mutant phenotypes. In addition to being a *CHD7* downstream gene, *Sox2* has been reported to physically interact with *Chd7* and co-occupy genomic binding sites to regulate common target genes³⁷. Therefore, the reduction of *SOX2* expression in *CHD7* mutants may further dysregulate genes downstream of the *CHD7*-*SOX2* complex.

Previous studies have showed that heterozygous *Chd7* deficient mice had stereocilia-bearing hair cells that are morphologically indistinguishable from the WT counterpart¹⁹, but it is unclear whether the gene expression profile or the function of the hair cells is affected in these mutant mice. Using single-cell transcriptome profiling, we have demonstrated that multiple deafness genes, including *SIX1*, *USH1C*, and *CLDN9*, were dysregulated in the mono-allelic

CHD7 KO organoid hair cells. These results suggest that the dysregulation of deafness genes in *CHD7*^{+/-} hair cells may be one of the underlying mechanisms of CHARGE syndrome-associated hearing loss and balance dysfunction, which are accompanied by middle and inner ear malformations and neurogenic defects as revealed by previous mouse model and human patient studies^{7, 19, 35, 38}.

CHD7 is known as an ATP-dependent chromatin remodeling enzyme³⁹. The S834F mutation has been shown to completely abolish *CHD7*'s ATPase and chromatin remodeling activities², raising the possibility that this patient-specific missense mutant protein represents a functional null. Indeed, we observed similar cell morphological phenotypes between the S834F mutant and its corresponding mono- or bi-allelic KO mutants, and in most cases, similar downstream gene expression levels between these two types of mutants. However, there are notable exceptions. For example, the highly specific otic lineage marker *FBXO2* was severely downregulated in the bi-allelic KO, but its expression was maintained to the WT level in the bi-allelic S834F mutant. These results imply the presence of additional *CHD7* protein function(s) beyond its ATP-dependent chromatin remodeling activities, and such function(s) play critical roles in regulating the expression of some of the key otic genes. Recently, a chromatin remodeling-independent function of *CHD7* was identified in mouse cardiovascular development, where *CHD7* binds to *WDR5*, a core component of a H3K4 methyltransferase complex, to recruit this histone-modifying enzyme complex to its target gene loci to regulate target gene expression⁴⁰. Future studies are needed to test if similar chromatin remodeling-independent functions of *CHD7* is responsible for regulating otic developmental genes such as *FBXO2*, *SIX1*, and *DLX5* in the inner ear.

Genetic chimeric animals can be generated through genomic integration of complicated

sets of gene circuits as seen in the *Drosophila* MARCM mosaic tissue generation system⁴¹ and in the mouse MADM system⁴². Chimeric animals can also be generated by simply mixing the WT and mutant mouse embryonic stem cells in the blastocysts at preimplantation stage⁴³. Inspired by these models, we have established a chimeric organoid culture system to analyze cell autonomous and non-autonomous actions by mixing mutant and fluorescently labeled WT embryonic stem cells. Through co-differentiating these cells in chimeric organoids, the expression levels of otherwise dysregulated genes in the *CHD7* KO otic progenitors were completely or partially restored. This can be explained by diffusible morphogens or cell surface signaling ligands provided from the developing WT cells to the neighboring co-developing mutant cells during the co-differentiation process. The complete rescue of *FBXO2*, *SOX10*, *DLX5*, *HOXB9* and the partial rescue of *COL9A2*, *SOX2*, *SIX1* genes indicate that paracrine signaling from WT cells helped the *CHD7* KO cells to develop more faithfully the otic lineage path, thereby restoring the expression of these otic genes to full or partial extents. It will be interesting to test if the paracrine signaling ligands supplementing approach can rescue *CHD7* phenotypes in other organs, or alleviate phenotypes of other genetic diseases.

In summary, we demonstrate a critical role of *CHD7* in regulating human otic lineage differentiation and deafness gene expression. Loss of *CHD7* or its ATP-dependent chromatin remodeling function resulted in a failure of hair cell and supporting cell generation in human inner ear organoids. The differential expression of a subset of *CHD7* downstream genes between S834F and its corresponding mono- or bi-allelic KO mutant suggest the presence of *CHD7* functions beyond its ATP-dependent chromatin remodeling activities. Notably, the aberrant otic lineage identity and the dysregulation of key otic genes in *CHD7* KO cells can be partially rescued by co-culturing and co-differentiating with WT cells in a mosaic manner, highlighting

the contribution of abnormal extrinsic signaling inputs in the *CHD7* mutant phenotypes. Our findings shed light on the molecular basis of inner ear phenotypes associated with CHARGE syndrome and reveal potential therapeutic target genes and pathways. The various human *CHD7* mutant lines established in this study allow for future drug screening and testing and for validating human genome-specific gene therapy approaches.

Acknowledgements

This work was supported by a National Institute of Health grant R01 DC015788 (E.H.), the Indiana University Health-Indiana University School of Medicine Strategic Research Initiative (E.H.), and Indiana Clinical and Translational Sciences Institute core pilot grant (E.H.). The authors would like to thank J. Harkin, A. Botros, and S. Moore for their technical assistance. We thank the following Indiana University School of Medicine core facilities for their support: Center for Medical Genomics, Flow Cytometry Resource Facility, and the Specimen Storage Facility.

Author Contributions

J.N. designed and led the study, performed experiments, analyzed data, and wrote the manuscript. Y.U. performed experiments and analyzed data. A.J.S. analyzed the data. E.H. conceived and designed the study, oversaw experiments, and wrote the manuscript. All authors read and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

References

1. Basson, M.A. & van Ravenswaaij-Arts, C. Functional Insights into Chromatin Remodelling from Studies on CHARGE Syndrome. *Trends Genet* **31**, 600-611 (2015).
2. Bouazoune, K. & Kingston, R.E. Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. *Proc Natl Acad Sci U S A* **109**, 19238-19243 (2012).
3. van Ravenswaaij-Arts, C.M., Hefner, M., Blake, K. & Martin, D.M. CHD7 Disorder, in *GeneReviews((R))*. (eds. M.P. Adam *et al.*) (Seattle (WA); 2020).
4. Hsu, P. *et al.* CHARGE syndrome: a review. *J Paediatr Child Health* **50**, 504-511 (2014).
5. Zentner, G.E., Layman, W.S., Martin, D.M. & Scacheri, P.C. Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *Am J Med Genet A* **152A**, 674-686 (2010).
6. Hurd, E.A. *et al.* Loss of Chd7 function in gene-trapped reporter mice is embryonic lethal and associated with severe defects in multiple developing tissues. *Mamm Genome* **18**, 94-104 (2007).
7. Hurd, E.A., Poucher, H.K., Cheng, K., Raphael, Y. & Martin, D.M. The ATP-dependent chromatin remodeling enzyme CHD7 regulates pro-neural gene expression and neurogenesis in the inner ear. *Development* **137**, 3139-3150 (2010).
8. Hurd, E.A., Micucci, J.A., Reamer, E.N. & Martin, D.M. Delayed fusion and altered gene expression contribute to semicircular canal defects in Chd7 deficient mice. *Mech Dev* **129**, 308-323 (2012).
9. Van Nostrand, J.L. *et al.* Inappropriate p53 activation during development induces features of CHARGE syndrome. *Nature* **514**, 228-232 (2014).
10. Micucci, J.A. *et al.* CHD7 and retinoic acid signaling cooperate to regulate neural stem cell and inner ear development in mouse models of CHARGE syndrome. *Hum Mol Genet* **23**, 434-448 (2014).
11. Yao, H. *et al.* CHD7 represses the retinoic acid synthesis enzyme ALDH1A3 during inner ear development. *JCI Insight* **3** (2018).
12. Balendran, V. *et al.* Chromatin remodeler CHD7 is critical for cochlear morphogenesis and neurosensory patterning. *Dev Biol* **477**, 11-21 (2021).
13. Duran Alonso, M.B. *et al.* Meis2 Is Required for Inner Ear Formation and Proper Morphogenesis of the Cochlea. *Front Cell Dev Biol* **9**, 679325 (2021).
14. Koehler, K.R., Mikosz, A.M., Molosh, A.I., Patel, D. & Hashino, E. Generation of inner ear sensory epithelia from pluripotent stem cells in 3D culture. *Nature* **500**, 217-221 (2013).
15. Koehler, K.R. *et al.* Generation of inner ear organoids containing functional hair cells from human pluripotent stem cells. *Nat Biotechnol* **35**, 583-589 (2017).

16. Liu, X.P., Koehler, K.R., Mikosz, A.M., Hashino, E. & Holt, J.R. Functional development of mechanosensitive hair cells in stem cell-derived organoids parallels native vestibular hair cells. *Nat Commun* **7**, 11508 (2016).
17. Koehler, K.R. & Hashino, E. 3D mouse embryonic stem cell culture for generating inner ear organoids. *Nat Protoc* **9**, 1229-1244 (2014).
18. Brogna, S. & Wen, J. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16**, 107-113 (2009).
19. Adams, M.E. *et al.* Defects in vestibular sensory epithelia and innervation in mice with loss of Chd7 function: implications for human CHARGE syndrome. *J Comp Neurol* **504**, 519-532 (2007).
20. Delahaye, A. *et al.* Familial CHARGE syndrome because of CHD7 mutation: clinical intra- and interfamilial variability. *Clin Genet* **72**, 112-121 (2007).
21. Kim, H.G. *et al.* Mutations in CHD7, encoding a chromatin-remodeling protein, cause idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am J Hum Genet* **83**, 511-519 (2008).
22. Zafra, M.P. *et al.* Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* **36**, 888-893 (2018).
23. Durruthy-Durruthy, R. *et al.* Reconstruction of the mouse otocyst and early neuroblast lineage at single-cell resolution. *Cell* **157**, 964-978 (2014).
24. Shearer, A.E. *et al.* Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc Natl Acad Sci U S A* **107**, 21104-21109 (2010).
25. Ma, Y. *et al.* Integrative differential expression and gene set enrichment analysis using summary statistics for scRNA-seq studies. *Nat Commun* **11**, 1585 (2020).
26. Hartman, B.H., Durruthy-Durruthy, R., Laske, R.D., Losorelli, S. & Heller, S. Identification and characterization of mouse otic sensory lineage genes. *Front Cell Neurosci* **9**, 79 (2015).
27. Tang, P.C. *et al.* Defective Tmprss3-Associated Hair Cell Degeneration in Inner Ear Organoids. *Stem Cell Reports* **13**, 147-162 (2019).
28. Narlikar, G.J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**, 490-503 (2013).
29. Chai, M. *et al.* Chromatin remodeler CHD7 regulates the stem cell identity of human neural progenitors. *Genes Dev* **32**, 165-180 (2018).
30. Schnetz, M.P. *et al.* Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. *Genome Res* **19**, 590-601 (2009).
31. He, D. *et al.* Chd7 cooperates with Sox10 and regulates the onset of CNS myelination and remyelination. *Nat Neurosci* **19**, 678-689 (2016).
32. Kiernan, A.E. *et al.* Sox2 is required for sensory organ development in the mammalian inner ear. *Nature* **434**, 1031-1035 (2005).
33. Hagstrom, S.A. *et al.* SOX2 mutation causes anophthalmia, hearing loss, and brain anomalies. *Am J Med Genet A* **138A**, 95-98 (2005).
34. Vissers, L.E. *et al.* Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* **36**, 955-957 (2004).

35. Choo, D.I., Tawfik, K.O., Martin, D.M. & Raphael, Y. Inner ear manifestations in CHARGE: Abnormalities, treatments, animal models, and progress toward treatments in auditory and vestibular structures. *Am J Med Genet C Semin Med Genet* **175**, 439-449 (2017).
36. Kiernan, A.E. *et al.* ENU mutagenesis reveals a highly mutable locus on mouse Chromosome 4 that affects ear morphogenesis. *Mamm Genome* **13**, 142-148 (2002).
37. Engelen, E. *et al.* Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* **43**, 607-611 (2011).
38. Green, G.E., Huq, F.S., Emery, S.B., Mukherji, S.K. & Martin, D.M. CHD7 mutations and CHARGE syndrome in semicircular canal dysplasia. *Otol Neurotol* **35**, 1466-1470 (2014).
39. Basson, M.A. & van Ravenswaaij-Arts, C. Functional Insights into Chromatin Remodelling from Studies on CHARGE Syndrome. *Trends Genet* **31**, 600-611 (2015).
40. Yan, S. *et al.* CHD7 regulates cardiovascular development through ATP-dependent and -independent activities. *Proc Natl Acad Sci U S A* **117**, 28847-28858 (2020).
41. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451-461 (1999).
42. Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479-492 (2005).
43. Rossant, J. & Spence, A. Chimeras and mosaics in mouse mutant analysis. *Trends Genet* **14**, 358-363 (1998).

Figure Legends

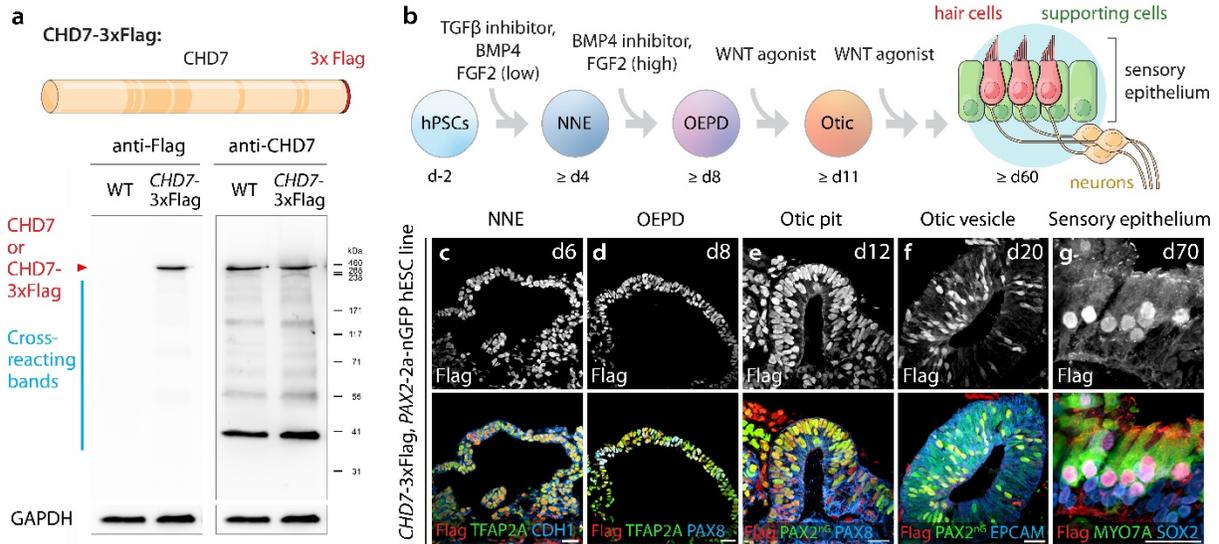


Figure 1. CHD7 is expressed at key otic development stages. **a**, Western blotting of WT (*PAX2*-2a-nGFP cell line, hereafter *PAX2*^{nG}) and *CHD7*-3 \times Flag hESCs using anti-Flag and anti-CHD7 antibodies. Calculated molecular weight of CHD7 and CHD7-3 \times Flag are 336 kDa and 339 kDa, respectively. **b**, Schematics of otic lineage differentiation during human inner ear organoid culture. **c–g**, Immunostaining at key otic development stages in *CHD7*-3 \times Flag *PAX2*^{nG} human inner ear organoids using an anti-Flag antibody, as well as antibodies against NNE markers TFAP2A and CDH1, OEPD markers TFAP2A and PAX8, otic placode/pit and otic vesicle markers *PAX2*^{nG}, *PAX8*, and *EPCAM*, and hair cell markers *MYO7A* and *SOX2* and SC marker *SOX2*. Scale bars, 25 μ m.

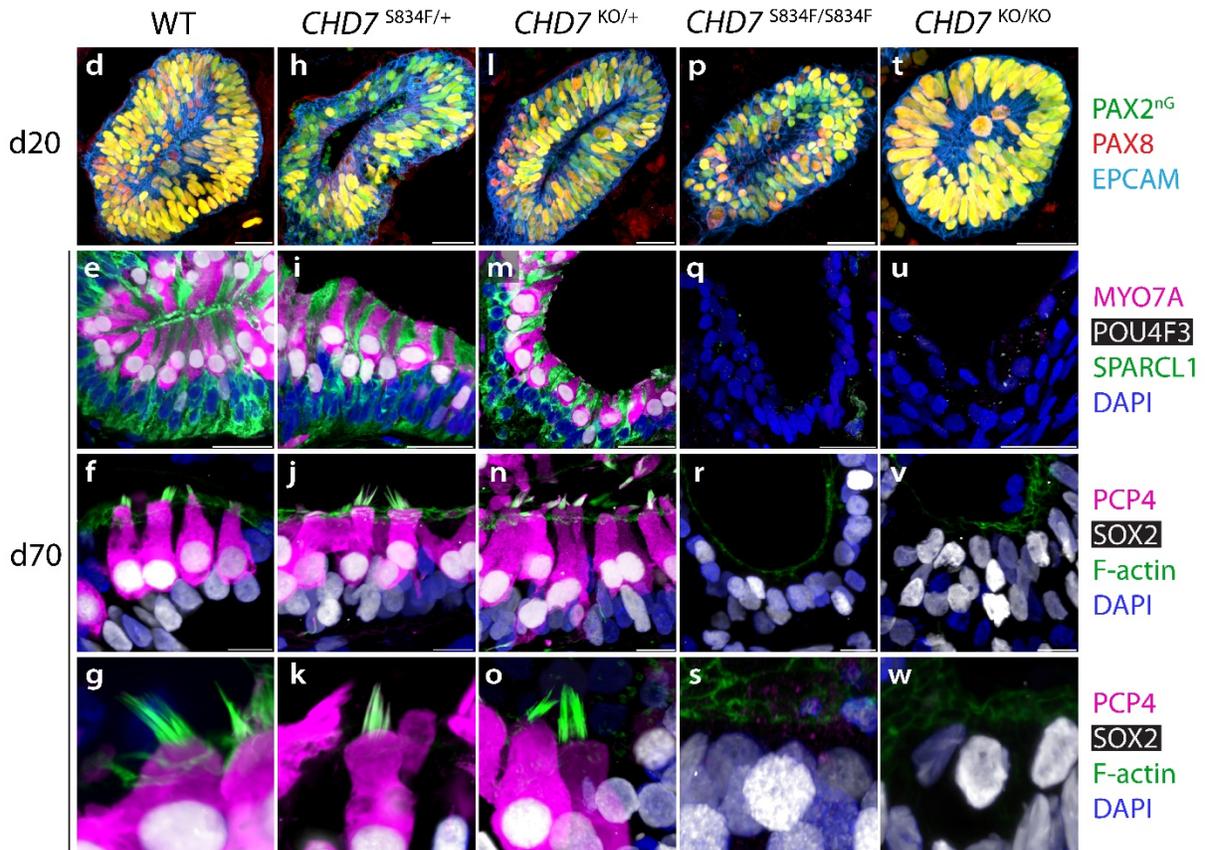
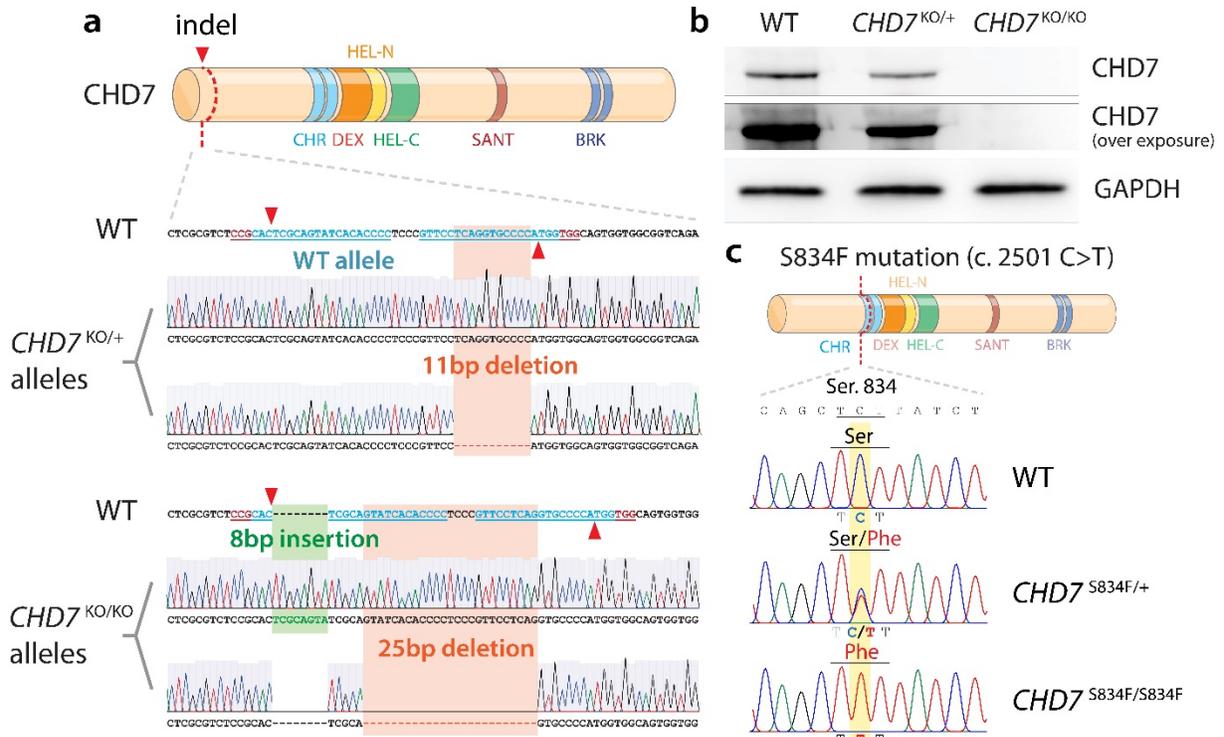


Figure 2. CHD7 and its ATP-dependent chromatin remodeling activities are required for sensory epithelium derivation. **a**, Sanger sequencing chromatograms of *CHD7*^{KO/+} and *CHD7*^{KO/KO} alleles cloned into TOPO vectors. **b**, Western blotting of WT (*PAX2*^{nG}), *CHD7*^{KO/+}, and *CHD7*^{KO/KO} hESCs using an anti-CHD7 antibody. **c**, Sanger sequencing chromatograms of WT (*PAX2*^{nG}), *CHD7*^{S834F/+}, and *CHD7*^{S834F/S834F} hESCs at the *CHD7* c.2501 locus. **d–w**, Immunostaining of d20 and d70 WT and *CHD7* mutant organoids. Antibodies highlight otic progenitors (*PAX2*^{nG}, *PAX8*, and *EPCAM*), hair cells (*MYO7A*, *POU4F3*, *PCP4*, *SOX2*, and F-actin for stereocilia of hair cells), and supporting cells (*SPARCL1* and *SOX2*). Scale bars, 25 μ m (top two rows of **d–w**), 10 μ m (third row of **d–w**), and 5 μ m (bottom row of **d–w**).

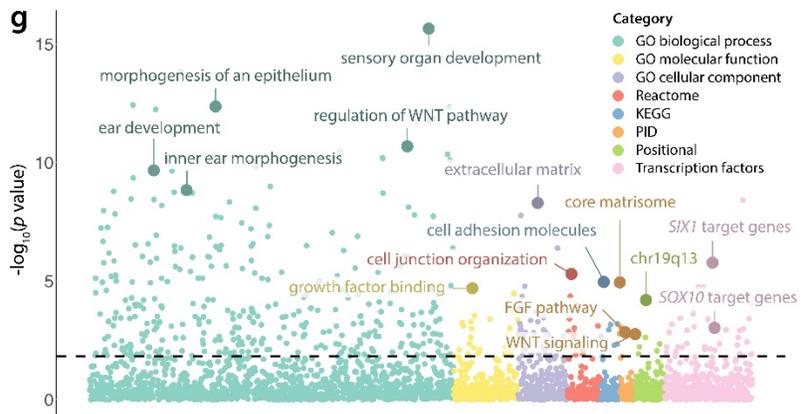
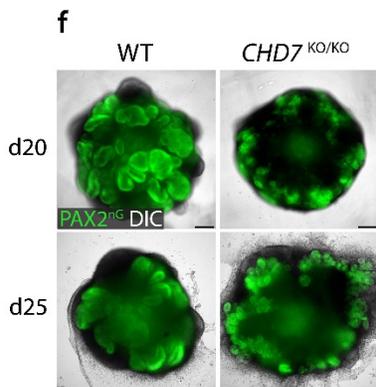
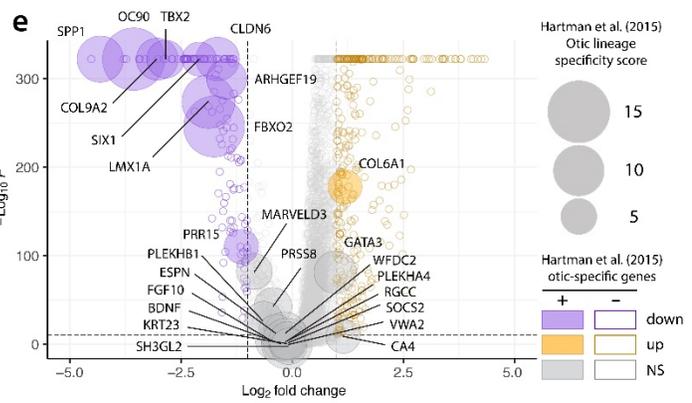
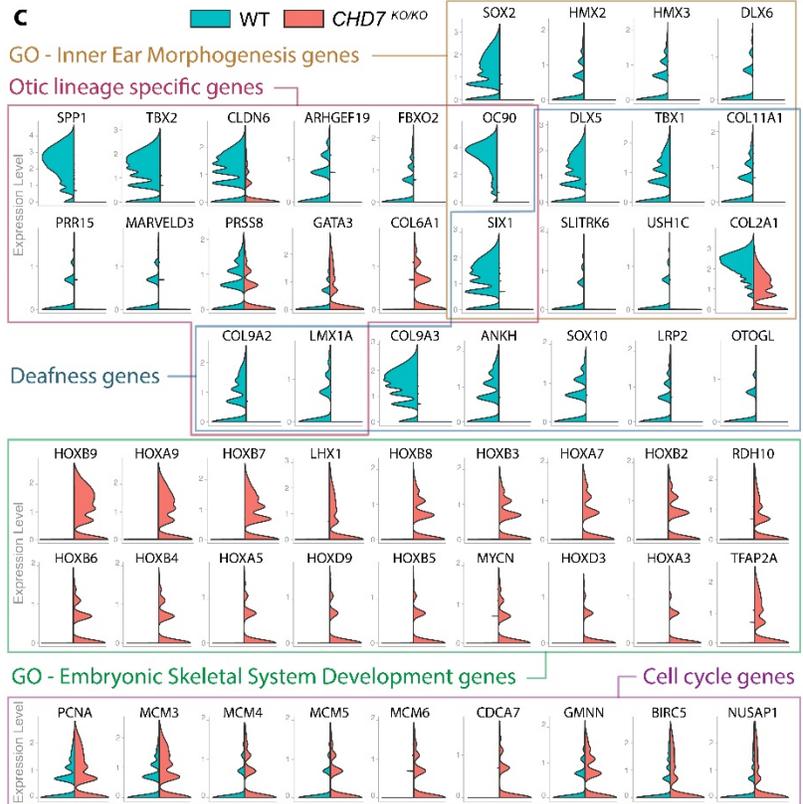
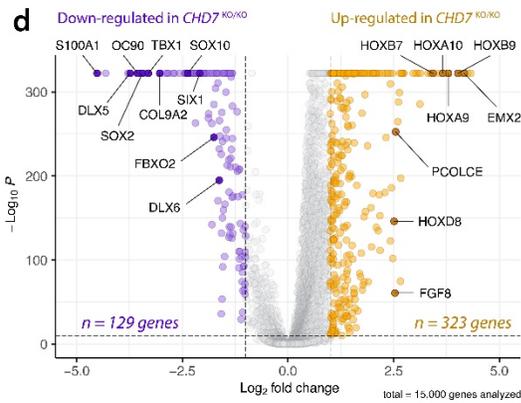
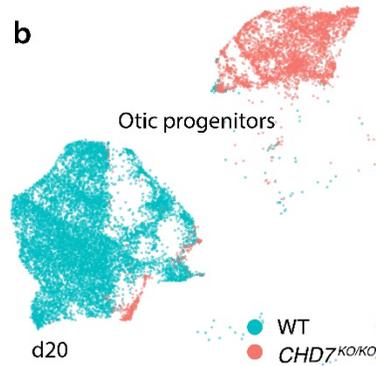
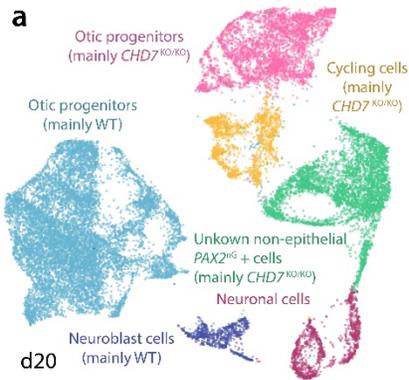


Figure 3. scRNA-seq of d20 WT and *CHD7*^{KO/KO} organoids revealed critical role of CHD7 in otic development. **a–b**, Uniform manifold approximation and projection (UMAP) plot and cluster annotation of FACS-isolated *PAX2*^{nG}-positive cells from d20 WT and *CHD7*^{KO/KO} organoids (**a**) and UMAP plot of the otic progenitor subset (**b**). Data represent 22,390 cells (**a**) and 14,894 cells (**b**). **c**, Split-violin plots of key genes dysregulated in *CHD7*^{KO/KO} otic progenitors. **d**, Volcano plot of differentially expressed genes in *CHD7*^{KO/KO} otic progenitors. **e**, Bubble plot of dysregulated and unaffected otic lineage-specific genes in *CHD7*^{KO/KO} otic progenitors. The areas of the bubbles represent the gene expression fold-change values of E10.5 otic vesicles versus non-otic tissues as reported by Hartman et al. (2015) **f**, Live imaging of d20 and d25 WT and *CHD7*^{KO/KO} organoids. **g**, Bubble plot of enriched gene sets from downregulated genes in d20 *CHD7*^{KO/KO} otic progenitors. Scale bars, 250 μ m.

Figure 4. Genes essential for otic development were dysregulated in *CHD7* mutant otic vesicles. **a–aw** left panel, UMAP plots of key dysregulated genes in d20 WT and *CHD7*^{KO/KO} otic progenitors. In each UMAP plot, the bottom left cluster consists mainly of WT cells (97.2%), and the top right cluster consists mainly of *CHD7*^{KO/KO} cells (98.4%). The color bars to the bottom right show the log-normalized expression scale. **a–aw** middle panel, Immunostaining of SOX2, SIX1, DLX5, COL9A2, FBXO2, SOX10, or HOXB9 along with otic progenitor markers PAX2^{nG} and EPCAM in WT, *CHD7*^{S834F/+}, *CHD7*^{KO/+}, *CHD7*^{S834F/S834F}, and *CHD7*^{KO/KO} organoids. Dotted lines mark the boundaries of otic vesicles. **a–aw** right panel, Violin plot quantifications of immunofluorescence signal intensities (arbitrary unit, A.U.) as shown in (**a–aw** middle panel). Dashed and dotted lines indicate the median and quartile values, respectively. *n* = 3 otic vesicles per genotype. All PAX2^{nG+} EPCAM⁺ cells from each otic vesicle were quantified. ****, *P* < 0.0001; ***, *P* < 0.001; ns, not significant. Significance was accessed by Kruskal-Wallis test followed by Dunn's multiple comparisons test. Scale bars, 25 μm.

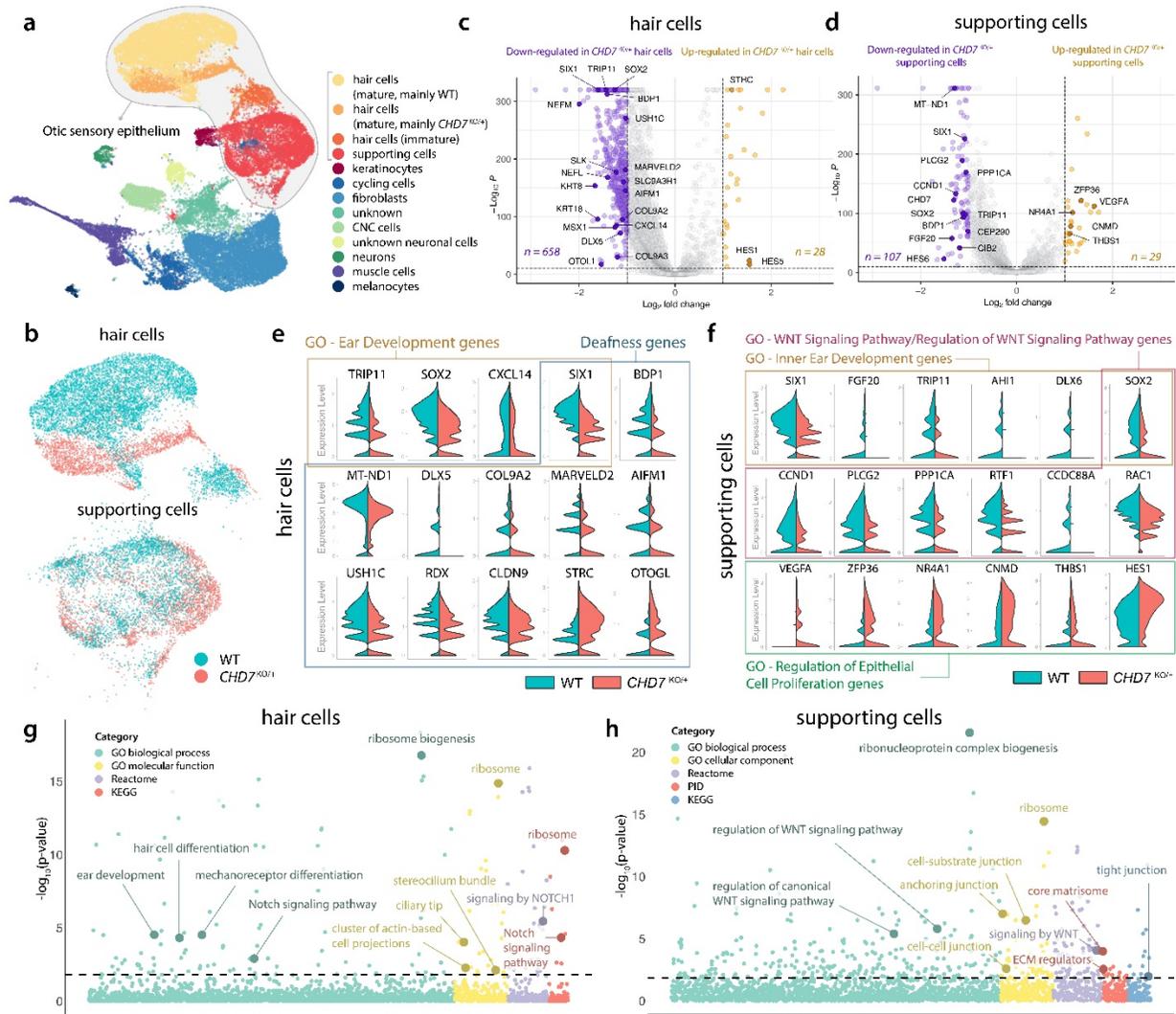


Figure 5. scRNA-seq of d70 WT and $CHD7^{KO/+}$ organoids revealed the importance of CHD7 in sensory epithelium development. **a–b**, d70 WT and $CHD7^{KO/+}$ organoids were micro-dissected and FACS-isolated into $POU4F3^{nT}$ -positive and -negative populations. UMAP plot and cluster annotation of merged datasets of these four groups of cells were shown in **(a)**, and UMAP plots of the hair cell and supporting cell subsets were shown in **(b)**. Data represent 34,703 total cells **(a)**, 9,884 hair cells and 6,273 supporting cells **(b)**. **c–d**, Volcano plots of differentially expressed genes in $CHD7^{KO/+}$ hair cells **(c)** and supporting cells **(d)**. **e–f**, Split-violin plots of key genes

dysregulated in *CHD7*^{KO/+} hair cells (**e**) and supporting cells (**f**). **g–h**, Bubble plot of enriched gene sets from downregulated genes in d70 *CHD7*^{KO/+} hair cells (**g**) and supporting cells (**h**).

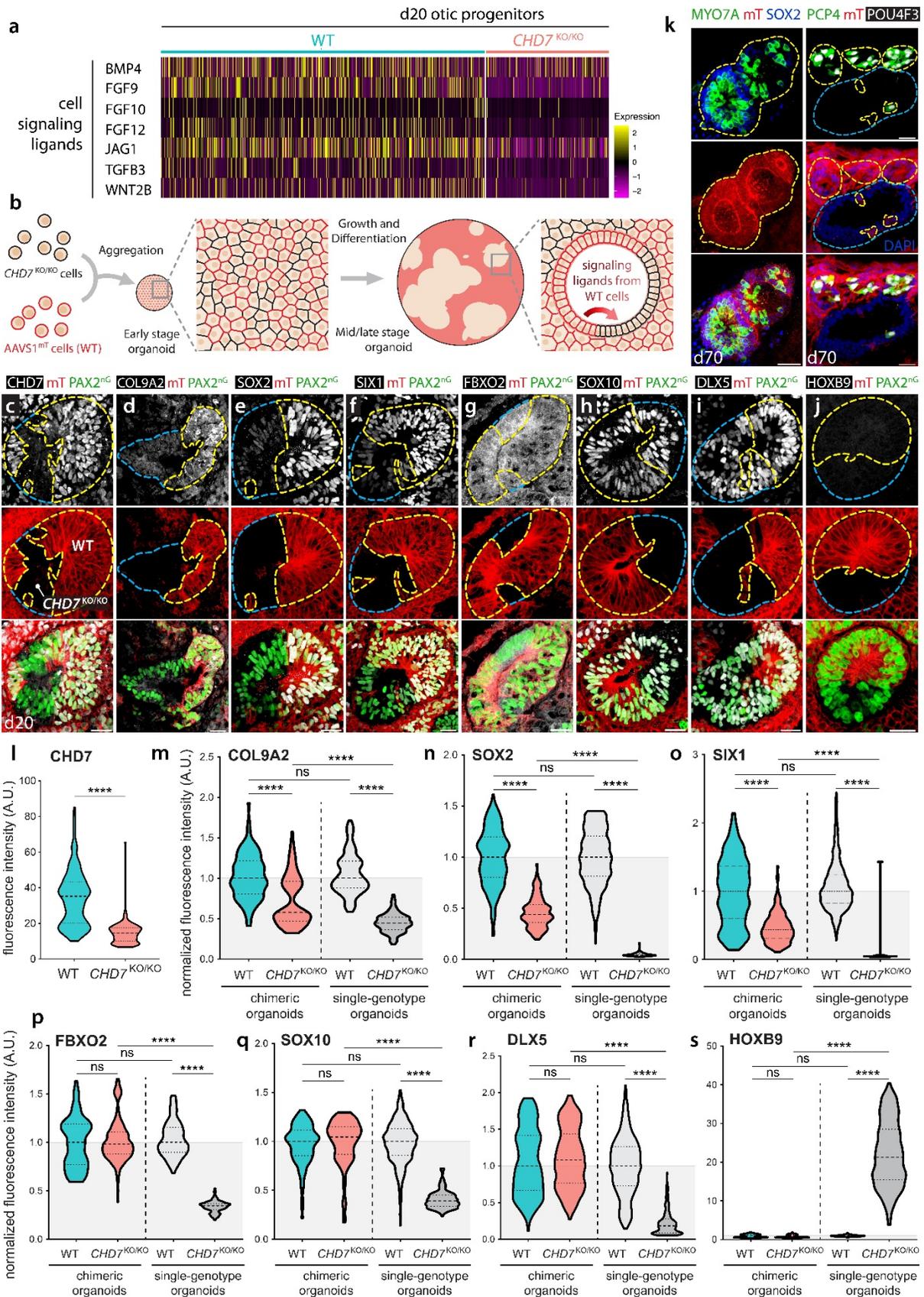
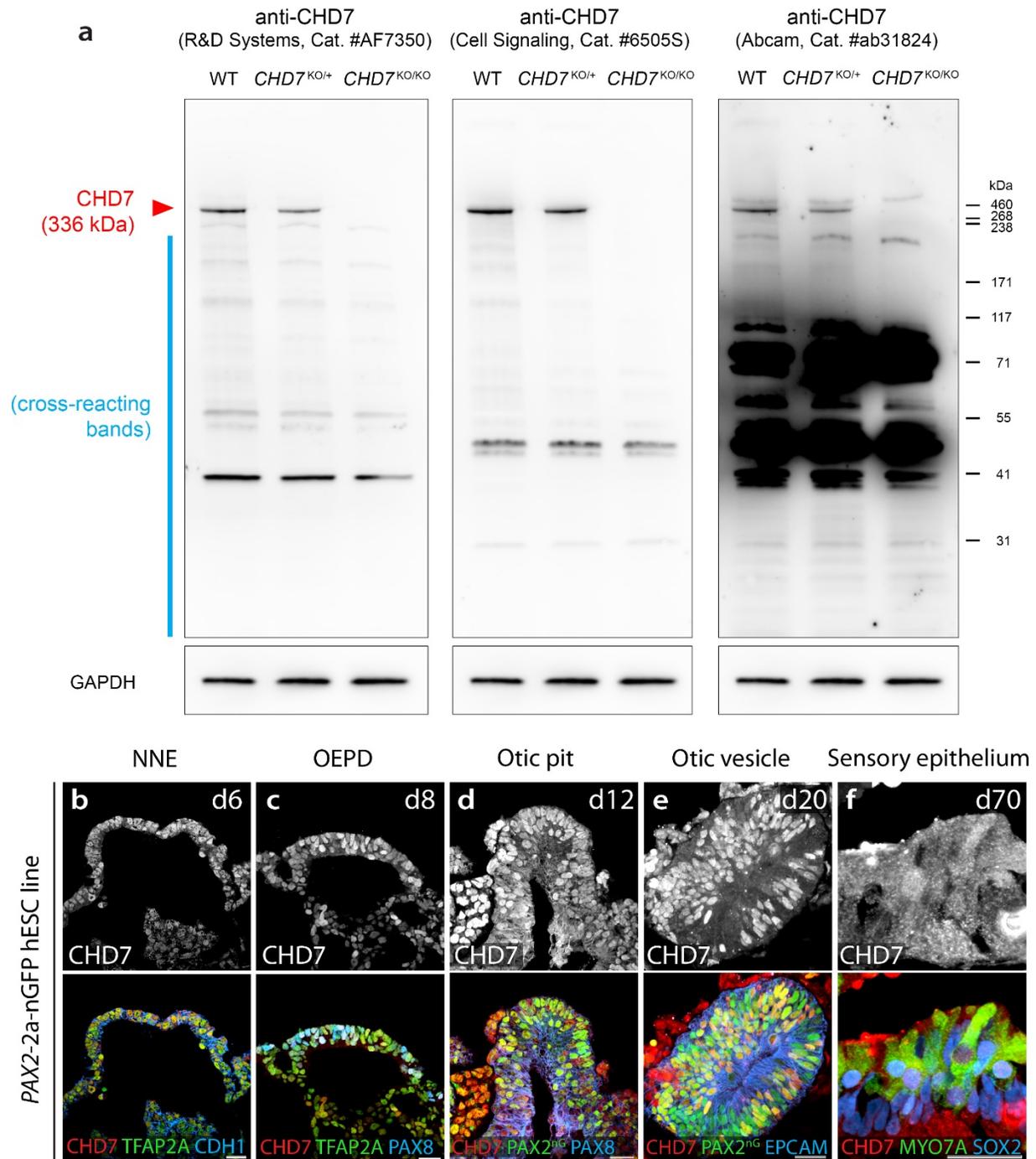
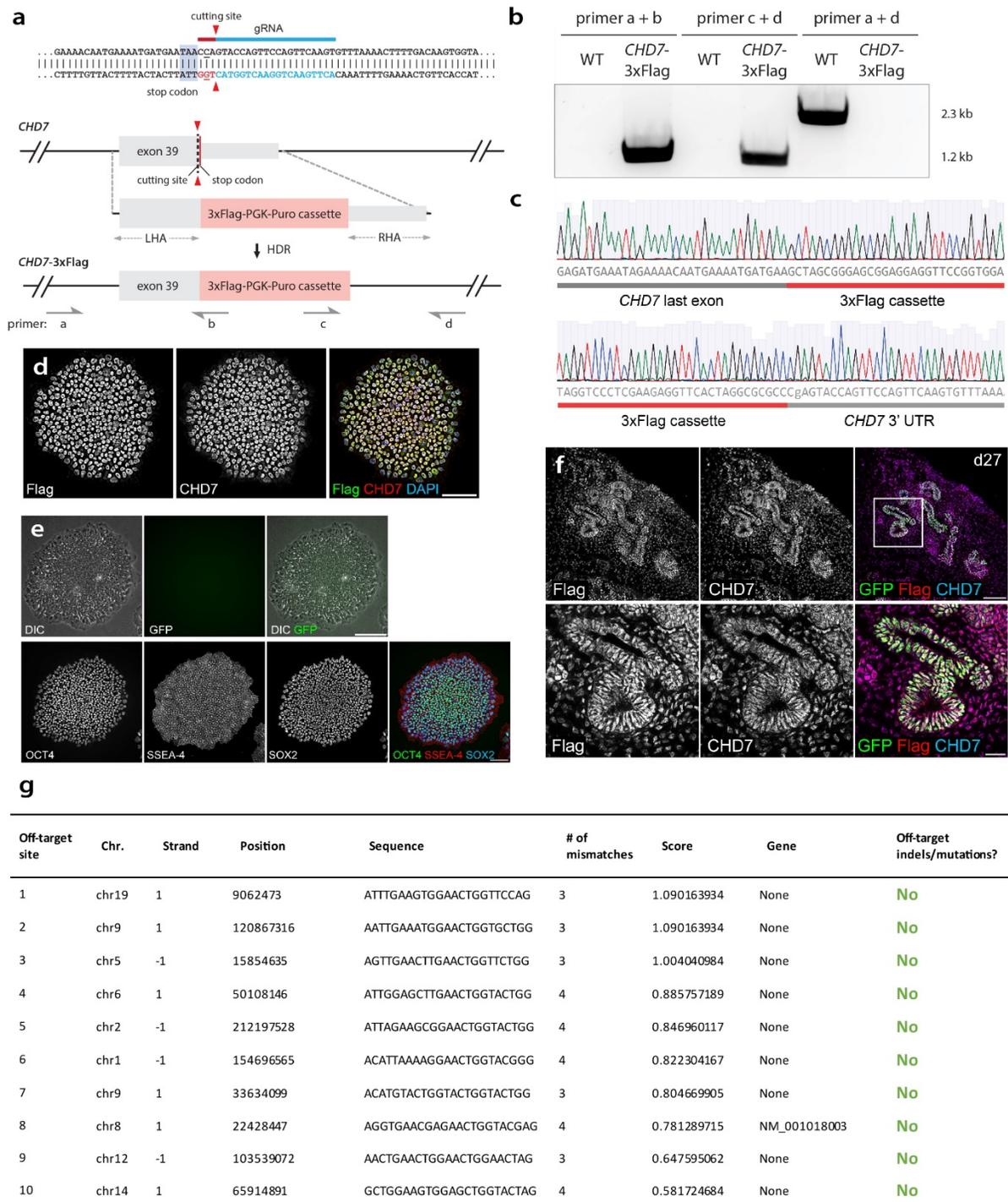


Figure 6. *CHD7*^{KO/KO} phenotypes in the otic progenitors can be partially rescued by co-differentiating with WT cells in chimeric organoids. **a**, Heatmap showing the scaled log-normalized expression of several cell signaling ligand genes in d20 WT and *CHD7*^{KO/KO} otic progenitors. **b**, WT-*CHD7*^{KO/KO} chimeric organoid culture strategy for supplying *CHD7*^{KO/KO} mutant cells with diffusible morphogens and cell surface signaling ligands from neighboring and nearby AAVS1^{mT}-labeled WT cells. **c–j**, Immunostaining of otic vesicles in d20 WT-*CHD7*^{KO/KO} chimeric organoids. Yellow dotted lines highlight the AAVS1^{mT}-labeled WT clones in an otic vesicle, and the blue dotted lines highlight the *CHD7*^{KO/KO} otic vesicle clones. **k**, hair cells in d70 WT-*CHD7*^{KO/KO} chimeric organoids labeled with MYO7A, PCP4, POU4F3, and SOX2. Yellow dotted lines highlight the AAVS1^{mT}-labeled WT clones in the sensory epithelia, and the blue dotted lines highlight the *CHD7*^{KO/KO} clones. **l–s**, Violin plot quantifications of immunofluorescence intensity data as shown in (**c–j**). WT and *CHD7*^{KO/KO} data from chimeric organoids were normalized to the chimeric organoid WT cell median values, and were shown in blue and red, respectively. Regular single-genotype WT and *CHD7*^{KO/KO} organoids data from (**Fig. 3**) were normalized to regular WT organoid median values, and were shown in light and dark gray, respectively. Dashed and dotted lines indicate the median and quartile values, respectively. $n = 3$ otic vesicles per genotype. All PAX2^{nG+} EPCAM⁺ WT and *CHD7*^{KO/KO} cells from each otic vesicle were quantified. ****, $P < 0.0001$; ns, not significant. Significance was accessed by Kruskal-Wallis test followed by Dunn's multiple comparisons test. Scale bars, 25 μm .



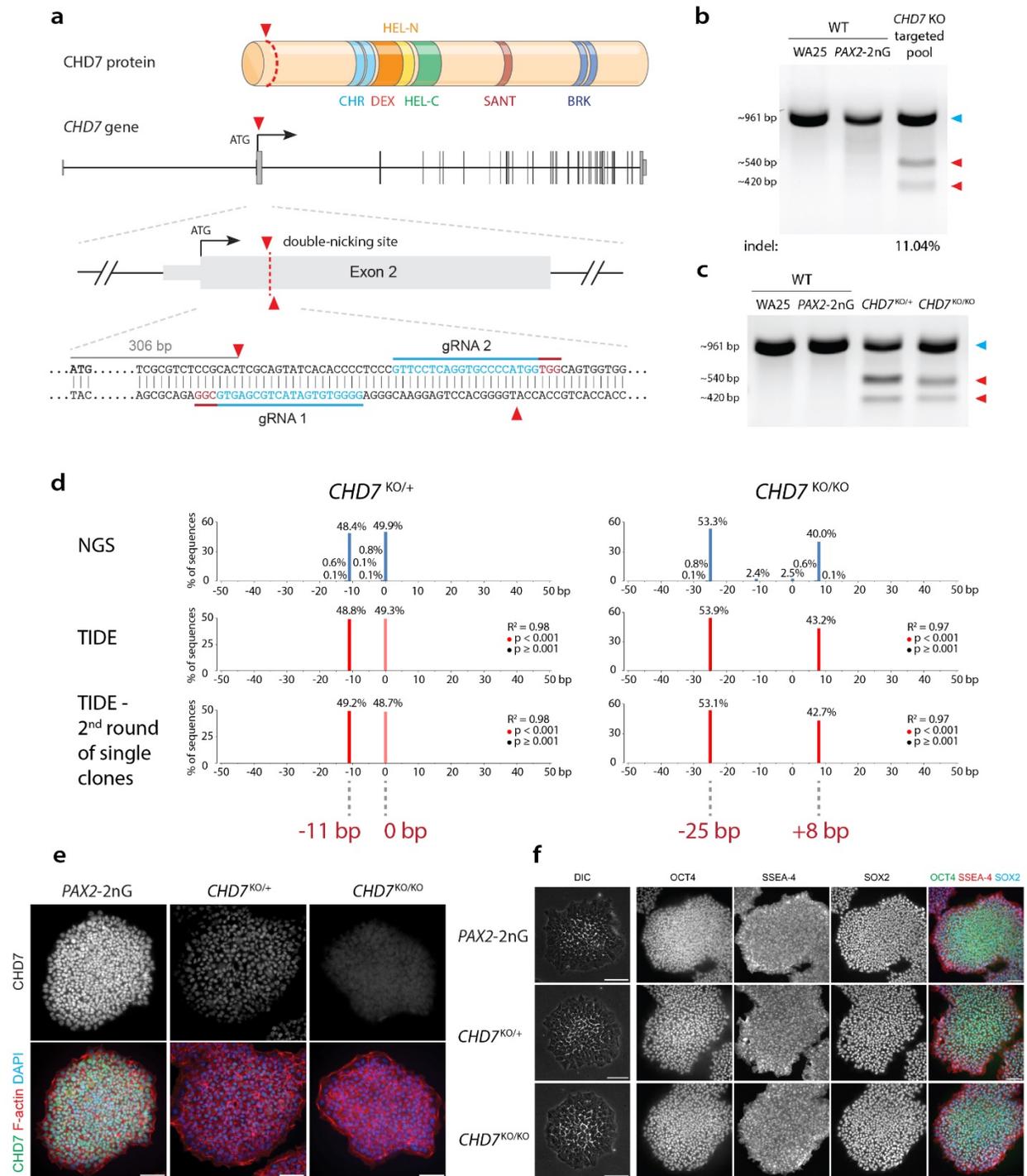
Supplementary Figure 1. Low specificity of widely used CHD7 antibodies. **a**, Western blot of WT, *CHD7*^{KO/+}, and *CHD7*^{KO/KO} hESCs showing high level of cross reactivities in three

commercially available anti-CHD7 antibodies (R&D Systems #AF7350, Cell Signaling #6505S, and Abcam #ab31824). **b–f**, Immunostaining at key otic development stages in *PAX2^{nG}* human inner ear organoids using an anti-CHD7 antibody (R&D Systems #AF7350), as well as antibodies against NNE markers TFAP2A and CDH1, OEPD markers TFAP2A and PAX8, otic placode/pit and otic vesicle markers *PAX2^{nG}*, PAX8, and EPCAM, and hair cell markers MYO7A and SOX2 and supporting cell marker SOX2. Scale bars, 25 μ m.



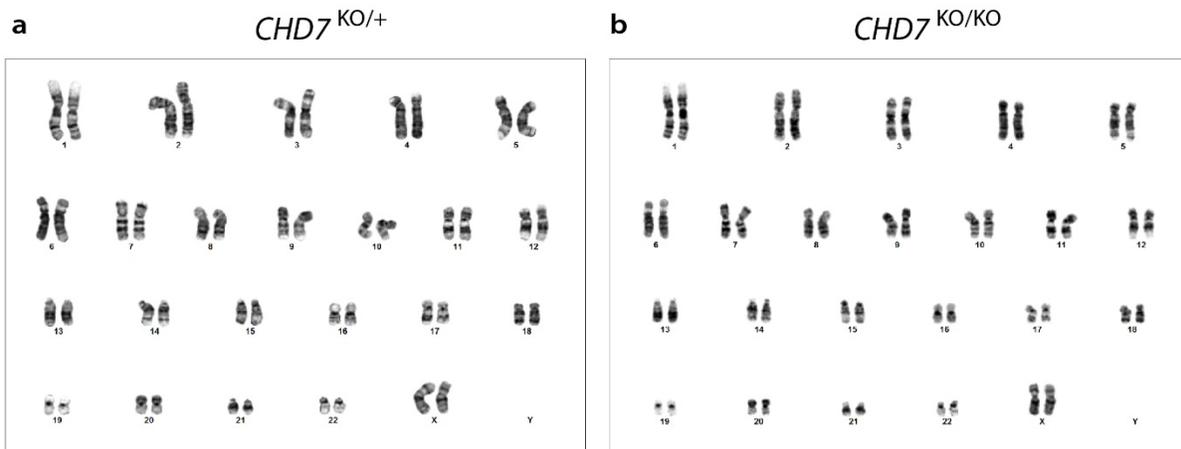
Supplementary Figure 2. Labeling of endogenous CHD7 with a 3×Flag tag. **a**, CRISPR knockin design of *CHD7*-3×Flag. **b**, PCR genotyping of WT and *CHD7*-3×Flag cell lines using

primers shown in **(a)** suggested successful bi-allelic 3×Flag knockin. The large PGK-Puro antibiotic selection cassette was difficult to amplify, so this PCR genotyping experiment did not use a long extension time setting to attempt to amplify the entire 3×Flag-PGK-Puro insertion. **c**, Sanger sequencing chromatograms of the left and right junction areas between the genomic DNA and the 3×Flag-PGK-Puro cassette insertion. **d**, Co-localization of anti-Flag and anti-CHD7 signals in *CHD7*-3×Flag hESCs. **e**, Normal hESC morphology and absence of *PAX2*-2a-nGFP expression, as well as normal pluripotency marker (OCT4, SSEA4, and SOX2) expression in *CHD7*-3×Flag hESCs. **f**, Co-localization of anti-Flag and anti-CHD7 signals in d27 *CHD7*-3×Flag organoids. **g**, No off-target mutations were detected at the top 10 predicted off-target sites of *CHD7*-3×Flag hESCs. Scale bars, 100 μm (**d**, **e**), 25 μm (**f**).



Supplementary Figure 3. Generation of *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC lines with CRISPR gene editing. **a**, CRISPR double nicking targeting strategy at the first coding exon of *CHD7*. **b**,

T7 endonuclease 1 (T7E1) assay detected ~11.04% indel formation from *CHD7* KO targeted population of hESCs. **c**, T7E1 assay detected indel formation from *CHD7*^{KO/+} and *CHD7*^{KO/KO} clonal hESC lines. **d**, Next generation sequencing (NGS) and Tracking of Indels by DEcomposition (TIDE) analysis revealed a WT allele and a 11 bp frameshift deletion allele in the *CHD7*^{KO/+} clonal hESC line, as well as a 25 bp frameshift deletion allele and an 8 bp frameshift insertion allele in the *CHD7*^{KO/KO} clonal hESC line. Second round of clonal cell line isolation followed by TIDE analysis confirmed that the *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC lines are homogeneous cell lines. **e**, Anti-CHD7 immunostaining in *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC. Note that the CHD7 antibody (R&D Systems, #AF7350) is not highly specific based on western blot analysis (**Supplementary Fig. 1a**), thus resulted in moderate levels of background noise. **f**, *CHD7*^{KO/+} and *CHD7*^{KO/KO} lines showed normal hESC morphology and normal pluripotency marker (OCT4, SSEA4, and SOX2) expression. Scale bars, 50 μ m.



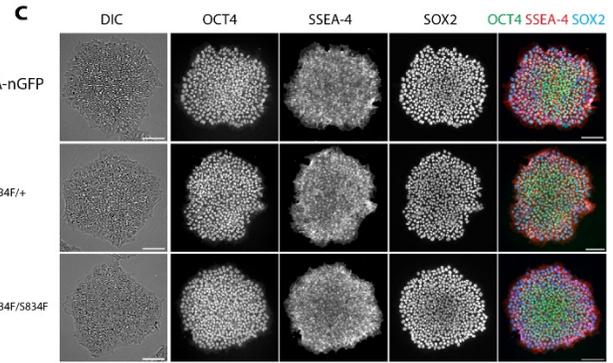
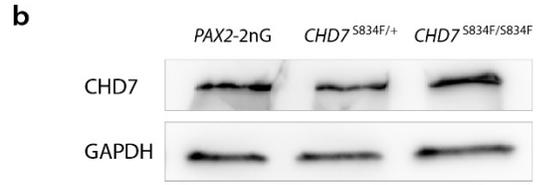
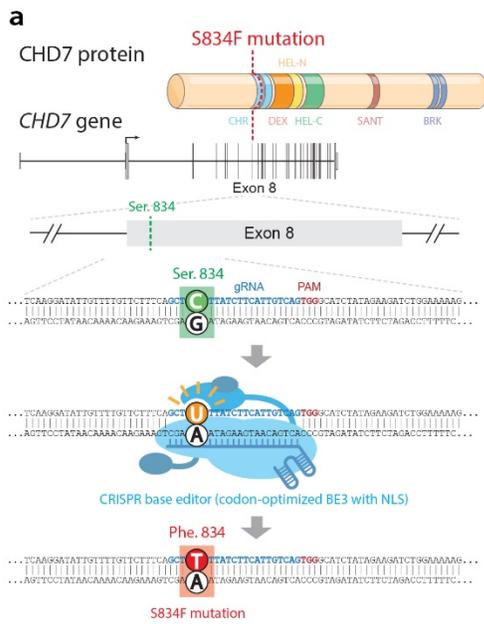
c

Off-target site	chr.	strand	position	sequence	# mismatches	score	gene	off-target indels/mutations?	
								<i>CHD7</i> KO/+	<i>CHD7</i> KO/KO
gRNA1-1	chr20	1	20337865	AGGGGGTCATACTGCGAGTGGAG	3	2.606310014	None	?	?
gRNA1-2	chr20	1	20338204	AGGGGGTCATACTGCGAGTGGAG	3	2.606310014	None	?	?
gRNA1-3	chr20	1	20336883	GAGGGGGTCATACTGCGAGTGGAG	3	2.543507363	None	?	?
gRNA1-4	chr11	1	129245133	GGGCTGCGAGACTGCGAGTGGAG	3	1.599975502	None	No	No
gRNA1-5	chr5	-1	146780153	GGCTGGTGACACTGCGAGTGCCAG	4	1.258804523	None	No	No
gRNA1-6	chr11	1	112829584	TGGGAGTGATACTGGGAGTGTGG	3	0.844444444	None	No	No
gRNA1-7	chr18	-1	54960416	TGTGTGTAATACTGCGTGTGTAG	4	0.611927885	None	No	No
gRNA1-8	chr7	-1	42610888	TGGGAGAGGTACTGCGAGTGTGG	4	0.58761959	None	No	No
gRNA1-9	chr6	-1	157639441	GGTGTGTGATGCTGCGAGGGGAG	3	0.577631667	None	No	No
gRNA1-10	chr9	1	131144867	AGGGAGTGAGACTGCCGAGAGAGG	4	0.485227553	None	No	No
gRNA2-1	chr15	-1	77305172	CTTCCTGAGGTGCCCATGGAGG	2	4.569366197	None	No	No
gRNA2-2	chr19	-1	14694078	CTGCCTCAGTGCCTCATGGTGG	3	2.48975671	NM_207390	No	No
gRNA2-3	chr19	1	3969396	GTTGCTCTGGTGCCCATGCCAG	3	1.397354497	None	No	No
gRNA2-4	chr8	-1	140655071	GCTGCTCAGGTTCCCATGGGGG	3	1.384888889	None	No	No
gRNA2-5	chr2	-1	121570106	GGGCCTCAGGTGCCCATGGTGG	3	1.365500444	None	No	No
gRNA2-6	chr1	1	206686559	ATTCCTCAGTTCCCATGGAGG	3	1.310427397	None	No	No
gRNA2-7	chr17	1	32058055	GATGCTCGGATGCCCATGGAAG	4	1.296862648	None	No	No
gRNA2-8	chr15	1	41199793	GTCCATCTGTTGCCCATGGTGG	4	1.258804523	None	No	No
gRNA2-9	chr10	1	72467883	GTTCTCAGATGCCCATGGAGG	2	1.059798592	None	No	No
gRNA2-10	chr22	-1	30178740	GCTCCCCAGGTGCCCATGACAG	3	0.902714689	None	No	No

Supplementary Figure 4. Validation of the established *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC lines.

a, *CHD7*^{KO/+} and *CHD7*^{KO/KO} lines showed normal karyotyping results. **b**, Off-target sequencing of the top 10 predicted off-target sites of both double-nicking gRNAs in *CHD7*^{KO/+} and *CHD7*^{KO/KO} hESC lines. The top 3 predicted off-target sites of gRNA #1 were in a stretch of highly repetitive genomic DNA sequence region, resulting in failure of PCR amplifications

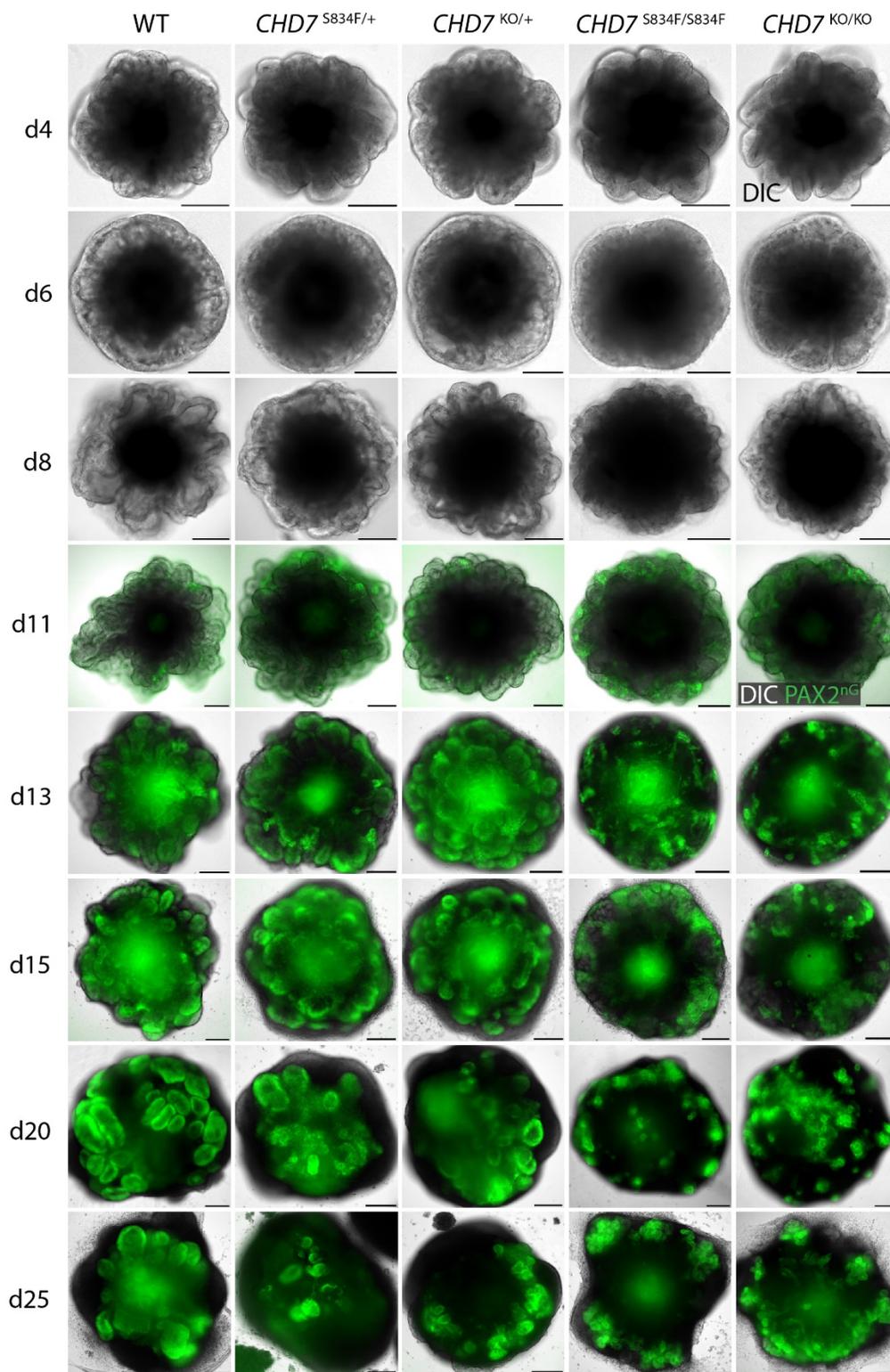
despite numerous attempts using various PCR and molecular cloning strategies. These three off-target sites were far away from any coding gene, and karyotyping analysis (**a**) detected no chromosomal translocation or other chromosomal abnormalities. Moreover, the Cas9n nickase is unlikely to induce indel formation from a single nicking on one DNA strand. Therefore, these three predicted sites are not expected to contain mutations. Even if mutation(s) exist, they are unlikely to affect *CHD7*^{KO/+} and *CHD7*^{KO/KO} phenotypes.



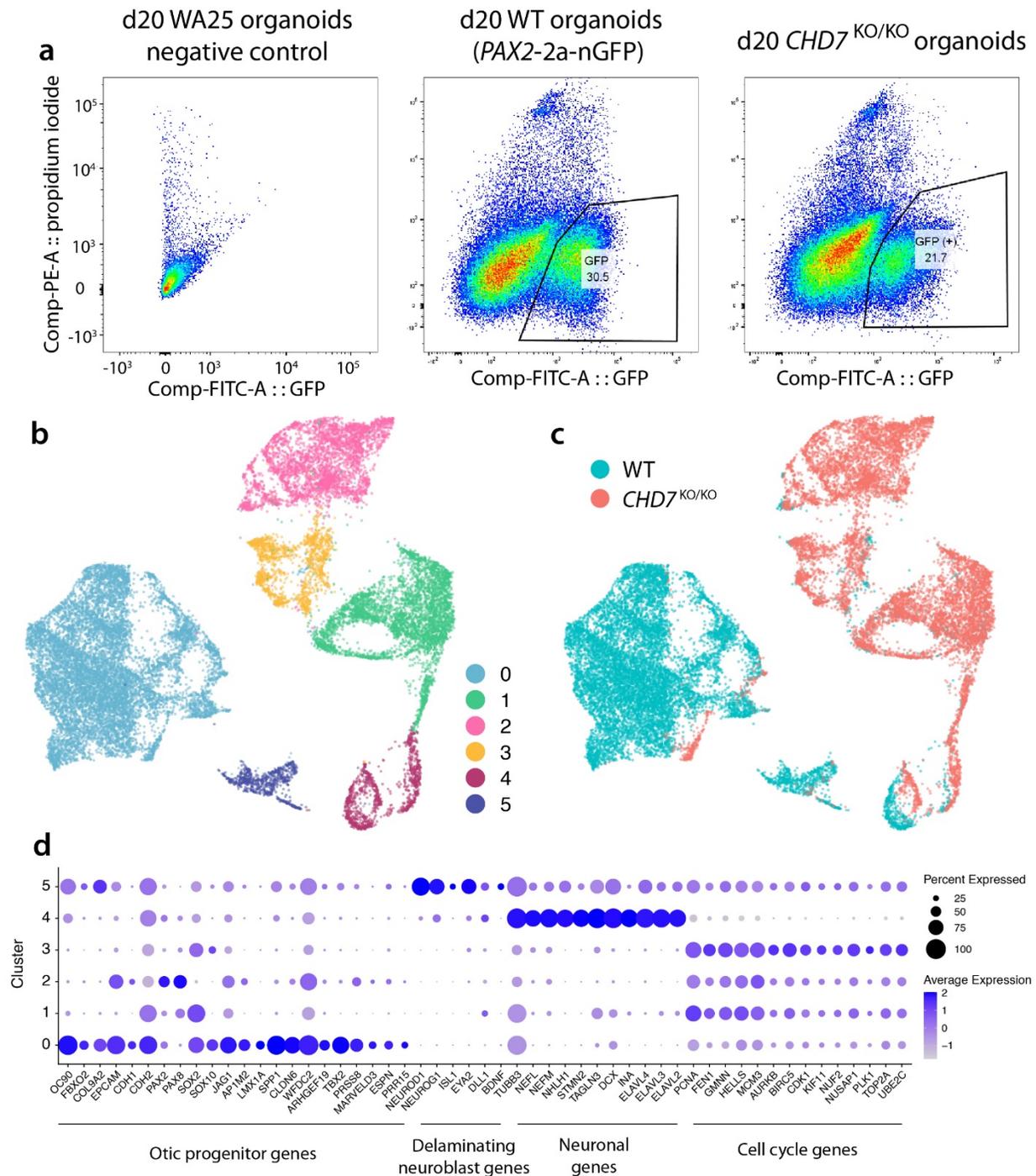
f

Off-target site	Chr.	Strand	Position	Sequence	# of mismatches	Score	Gene	Off-target mutations in the CHD7 ^{S834F/+} cell line?	Off-target mutations in the CHD7 ^{S834F/S834F} cell line?
1	chr3	-1	41046557	GTCCTTATCTCATTGTCAT	2	8.611956522	None	No	No
2	chrX	-1	94295223	ACTTTTATCTCATTGTCAG	2	5.722891566	None	No	No
3	chr7	1	38218640	GCACATATTTTCATTGTCAG	2	4.030448592	None	No	No
4	chrX	1	17453098	AATCTTTTCTCATTGTCAG	2	1.825175809	None	No	No
5	chrX	1	12425343	TCCCTTCTCATTGTCAG	3	1.799623347	None	No	No
6	chr7	1	84144288	GCTCATTTCTCATTGTCAG	3	1.626097551	None	No	No
7	chr10	-1	117770683	GCTAATATTTTCATTGTCAG	3	1.579455782	None	No	No
8	chr2	-1	159748105	CCGCCTACCTTCATTGTCAG	4	1.469717573	MARCH7 (ENSG00000136536)	No	No
9	chr12	-1	51981191	GCCCTTACCTTTATTGTCAG	3	1.442430047	ACVR1B (ENSG00000135503)	No	No
10	chr5	1	96962862	GCTTTTATCAACATTGTCAG	3	1.413674672	None	No	No

Supplementary Figure 5. Generation of *CHD7*^{S834F/+} and *CHD7*^{S834F/S834F} hESC lines with a CRISPR base editor. **a**, CRISPR base editing strategy for inducing the C to T mutation at the *CHD7* Ser. 834 locus. **b**, Western blot of WT (*PAX2*^{nG}), *CHD7*^{S834F/+}, and *CHD7*^{S834F/S834F} hESCs suggested that protein expression levels were not affected. **c**, *CHD7*^{S834F/+} and *CHD7*^{S834F/S834F} lines showed normal hESC morphology and normal pluripotency marker (OCT4, SSEA4, and SOX2) expression. **d–e**, *CHD7*^{S834F/+} and *CHD7*^{S834F/S834F} lines showed normal karyotyping results. **f**, No mutations were found from the top 10 predicted off-target sites in the *CHD7*^{S834F/+} and *CHD7*^{S834F/S834F} hESC lines. Scale bars, 50 μ m.

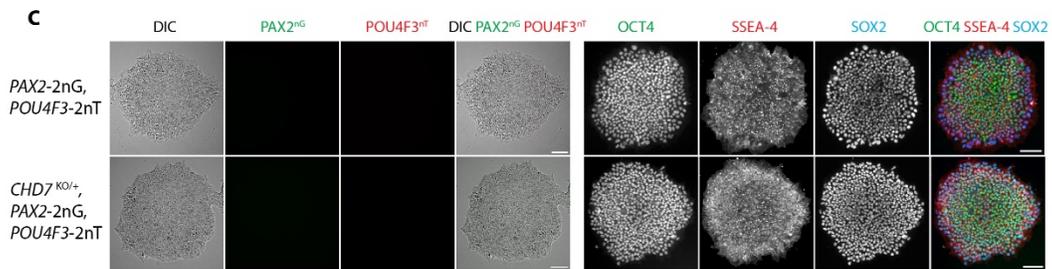
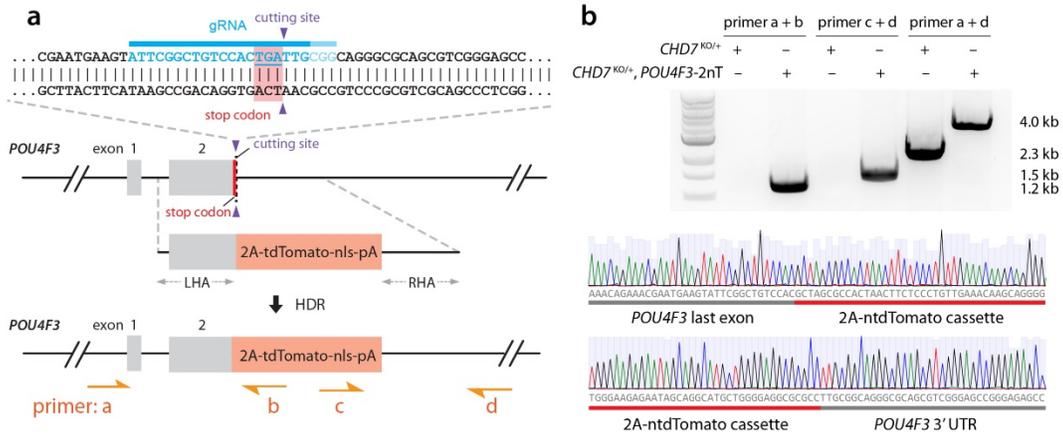


Supplementary Figure 6. Morphology and *PAX2*^{nG} fluorescence signals of WT (*PAX2*^{nG}), *CHD7*^{S834F/+}, *CHD7*^{KO/+}, *CHD7*^{S834F/S834F}, and *CHD7*^{KO/KO} inner ear organoids. D4–d8 organoids were DIC live imaged only, and d11–d25 organoids were live imaged in both DIC and GFP channels. Scale bars, 250 μ m.



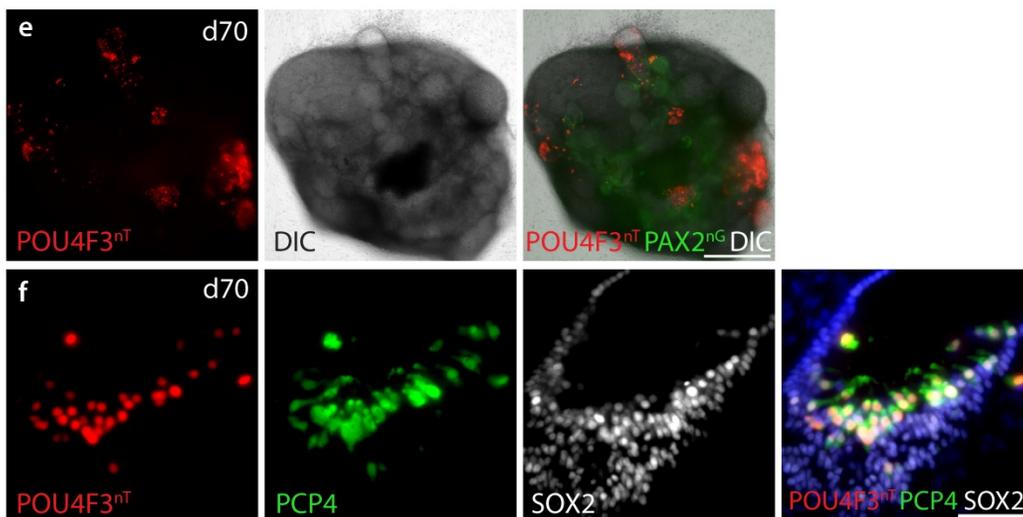
Supplementary Figure 7. scRNA-seq analysis of FACS-isolated *PAX2*^{nG} cells from d20 WT (*PAX2*^{nG}) and *CHD7*^{KO/KO} inner ear organoids. **a**, FACS isolation of *PAX2*^{nG}-positive cells from

d20 WT ($PAX2^{nG}$) and $CHD7^{KO/KO}$ inner ear organoids. **b–c**, UMAP plots of $PAX2^{nG}$ -positive cells from d20 WT ($PAX2^{nG}$) and $CHD7^{KO/KO}$ inner ear organoids plotted by cluster (**b**) or by genotype (**c**). **d**, Dot plot of otic progenitor genes, neuroblast genes, neuronal genes, and cell cycle genes from clusters shown in (**b**). Gene expression frequency was indicated by dot size and expression level was indicated by color intensity.



d

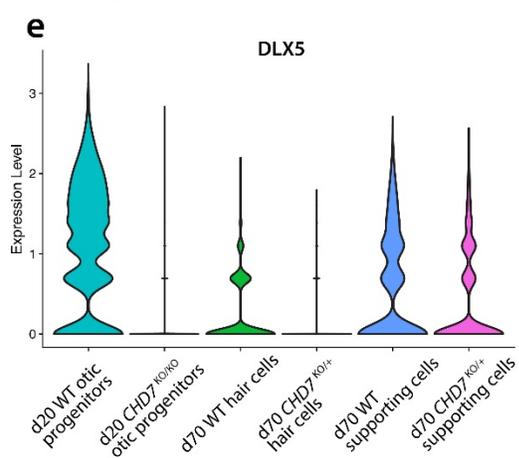
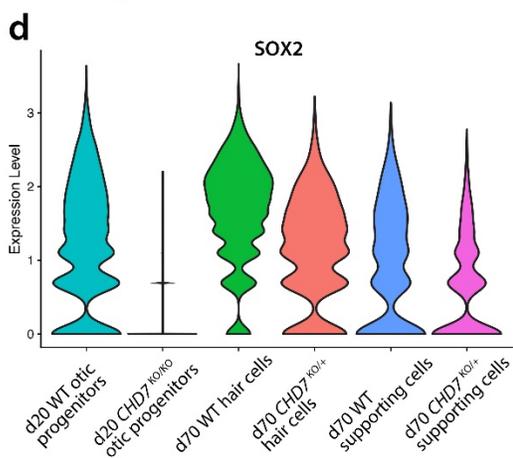
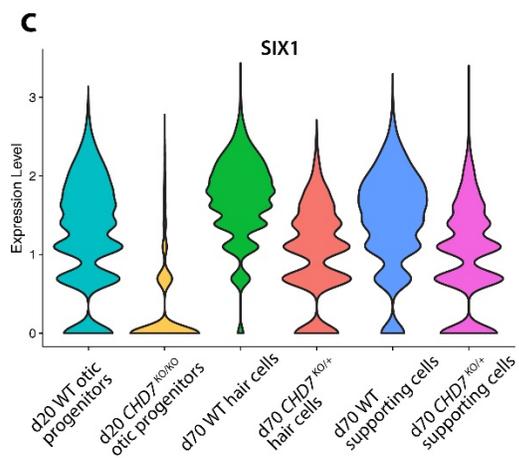
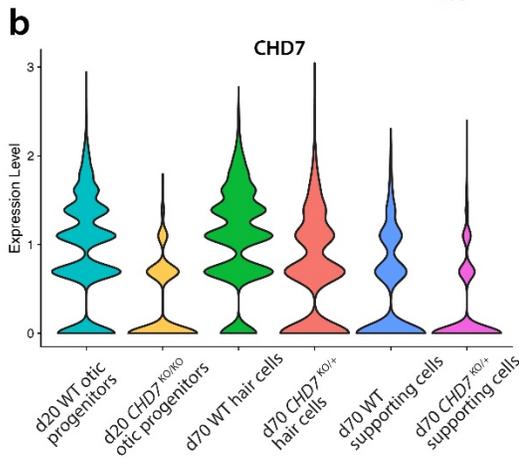
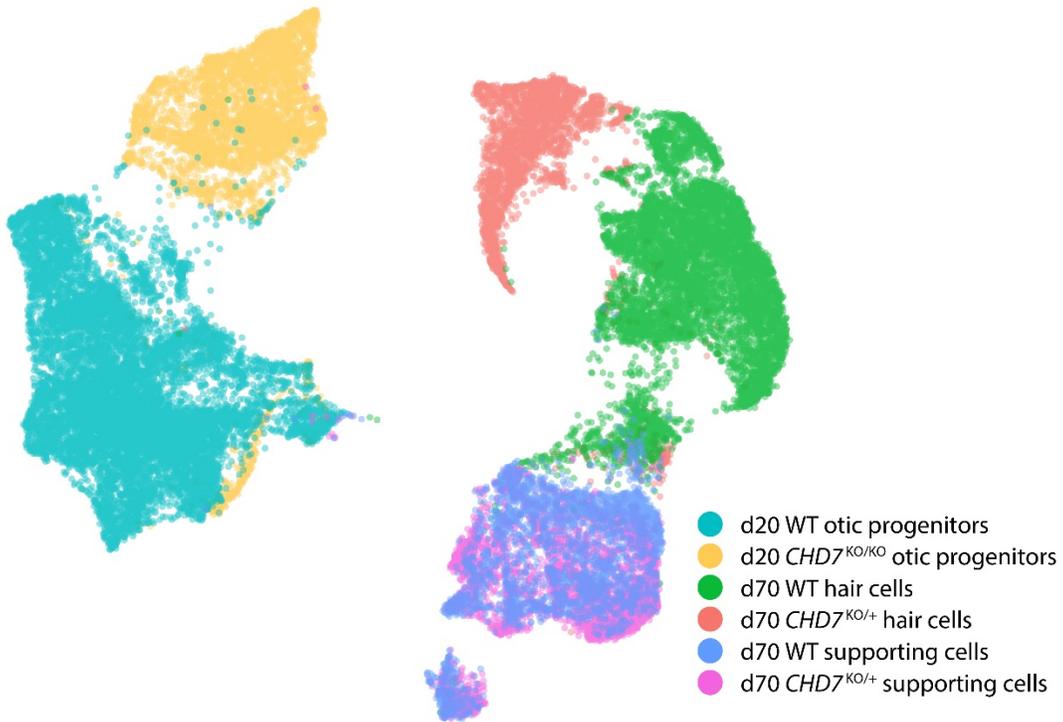
Off-target site	Chr.	Strand	Position	Sequence	# of mismatches	Score	Gene	Off-target mutations?
1	chr2	1	16391400	GTTCTGCTTCCACTGATTGGAG	3	1.632771	None	No
2	chr3	-1	188314251	AGTTGTCTGCCACTGATTGTAG	3	1.468072	None	No
3	chr12	-1	101614310	TCTCTGCTGCCACTGATTAAG	4	0.710795	None	No
4	chr4	1	53847222	TTTCTGCGTCCACTGATTAAG	4	0.710795	None	No
5	chr18	1	68394376	AGTTAGCTGCCACTGATTCAGG	4	0.697447	None	No
6	chr11	1	128463498	ATGCAGCAGTCCACTGATTACAG	4	0.675007	None	No
7	chr2	1	152741440	AATCTGCTGCCACTGAGTGGGG	3	0.65679	None	No
8	chr8	1	93202010	GGTCTGCTGCCCTGATTGTGG	4	0.581725	None	No
9	chr1	-1	194082508	TTTCTGCTGGCCACTGGTTGAAG	4	0.571588	None	No
10	chr9	-1	8278749	TTTCTTATGCCACTGATTGGAG	4	0.564015	None	No



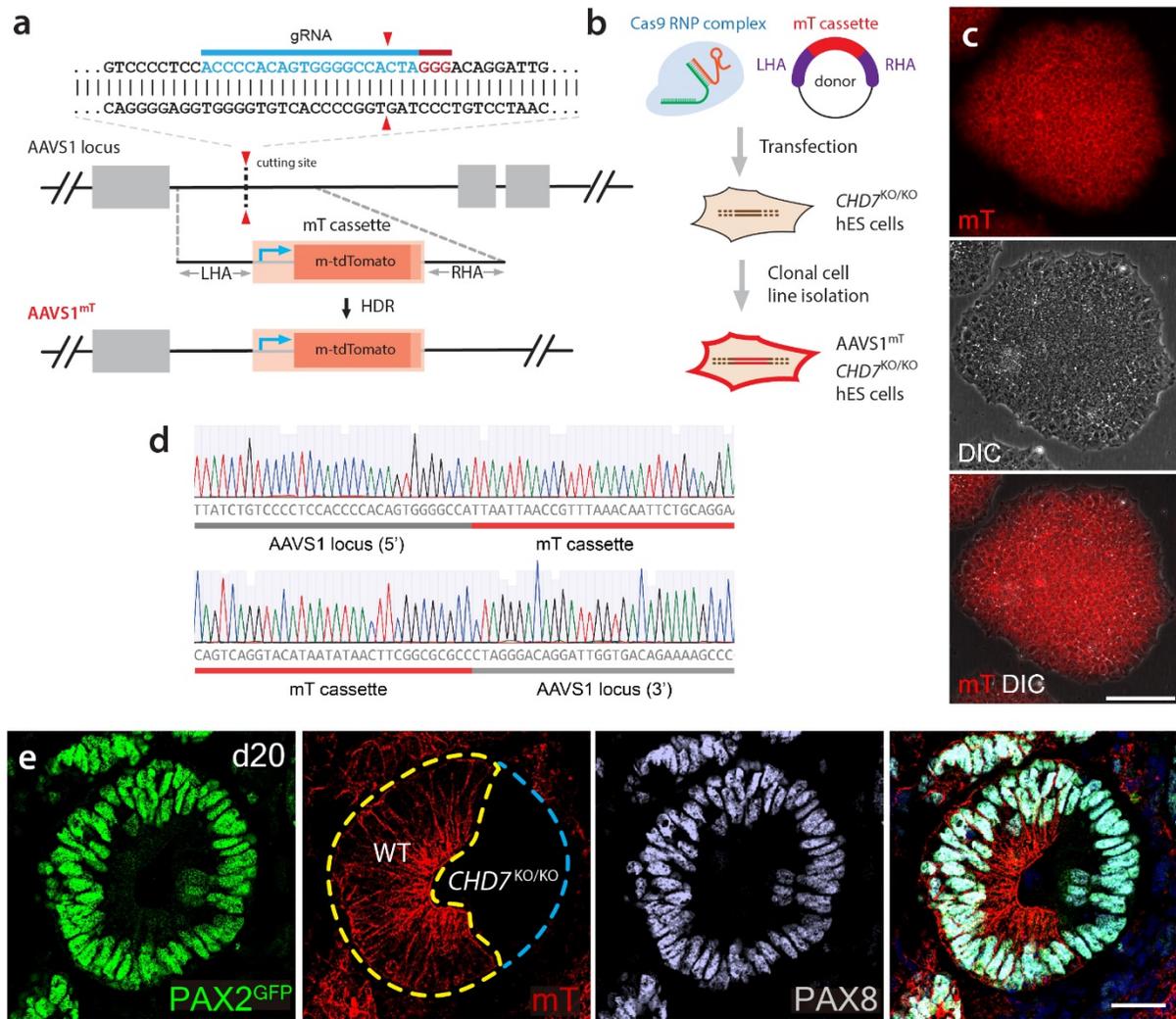
Supplementary Figure 8. Generation of *POU4F3*-2a-ntdTomato (*POU4F3^{nT}*) knockin hESC line on the *CHD7^{KO/+}* genetic background. **a**, CRISPR genome engineering design of 2a-ntdTomato cassette knockin to the *POU4F3* stop codon locus. **b**, Genotyping PCR and sequencing of the *CHD7^{KO/+} POU4F3^{nT}* hESC line using primers shown in **(a)** suggested successful bi-allelic 2a-ntdTomato knockin at the *POU4F3* stop codon locus. **c**, The *CHD7^{KO/+} POU4F3^{nT}* line showed normal hESC morphology, normal pluripotency marker expression (OCT4, SSEA4, and SOX2), and absence of *PAX2^{nG}* or *POU4F3^{nT}* fluorescence signals in hESCs. **d**, No off-target mutations were detected from the top 10 predicted off-target sites. **e**, Live imaging of d70 *CHD7^{KO/+} POU4F3^{nT}* organoids, showing *POU4F3^{nT}* fluorescence signals from *CHD7^{KO/+}* hair cells. **f**, Immunostaining of the hair cell markers PCP4, the hair cell and supporting cell marker SOX2, as well as *POU4F3^{nT}* fluorescence in d70 *CHD7^{KO/+} POU4F3^{nT}* organoids. Scale bars, 500 μm (**e**), 50 μm (**f**).

Supplementary Figure 9. scRNA-seq analysis of FACS-isolated *POU4F3*^{nT}-positive and -negative cells from micro-dissected d70 WT (*POU4F3*^{nT} *PAX2*^{nG}) and *CHD7*^{KO/+} (*CHD7*^{KO/+} *POU4F3*^{nT} *PAX2*^{nG}) inner ear organoids. **a**, FACS isolation of *POU4F3*^{nT}-positive and -negative cells from micro-dissected d70 WT and *CHD7*^{KO/+} inner ear organoids. **b–d**, UMAP plots of merged datasets of d70 WT and *CHD7*^{KO/+} inner ear organoids plotted by cluster (**b**), by genotype (**c**), or by FACS isolation (**d**). **e**, Dot plot of marker genes and highly expressed genes from each cluster shown in (**e**). Gene expression frequency was indicated by dot size and expression level was indicated by color intensity. **f**, Subsetting of the three hair cell clusters. **g**, Dot plot of early, intermediate, and late hair cell marker genes from each cluster shown in (**f**).

a d20 and d70 otic lineage cells



Supplementary Figure 10. scRNA-seq analysis of merged d20 otic progenitor dataset (WT and *CHD7*^{KO/KO}) and d70 hair cell and supporting cell datasets (WT and *CHD7*^{KO/+}). **a**, UMAP plot of the merged datasets. **b–e**, Violin plots showing the expression levels of *CHD7*, *SIX1*, *SOX2*, and *DLX5* in d20 otic progenitors (WT and *CHD7*^{KO/KO}) and d70 hair cells and supporting cells (WT and *CHD7*^{KO/+}).



Supplementary Figure 11. Generation of AAVS1-pCA-mtdTomato (AAVS1^{mT}) knockin hESC line on the *CHD7*^{KO/KO} genetic background. **a**, CRISPR genome engineering design of pCA-mtdTomato knockin to the AAVS1 locus. **b**, CRISPR knockin and clonal cell line isolation workflow. **c**, Membrane-bound tdTomato is expressed in AAVS1^{mT} hESCs. **d**, Sanger sequencing chromatograms of junction regions showing correct insertion of the pCA-mtdTomato cassette to the AAVS1 locus. **e**, In WT-*CHD7*^{KO/KO} chimeric organoids, AAVS1^{mT} labeled the WT clones while the *CHD7*^{KO/KO} were unlabeled. Scale bars, 50 μ m (c), 25 μ m (e).

Methods

hESC culture

Human ESCs (WA25 hESCs and genomically edited cell lines based on the WA25 background, passage 13–55) were cultured in Essential 8 Flex (E8f) medium (Thermo Fisher) supplemented with 100 µg/mL Normocin (Invivogen) (hereafter, E8fn medium) on truncated recombinant human Vitronectin-N (Thermo Fisher)-coated Nunclon Delta surface-treated 6-well plates (Thermo Fisher) according to an established protocol¹. At 60%–80% confluency or every 3 or 4 days, the cells were passaged at a split ratio of 1:100–1:10 into 2–4 wells of a 6-well plate using 0.5 mM Ethylenediaminetetraacetic acid (EDTA) in Dulbecco's phosphate-buffered saline (DPBS). 1X RevitaCell (Thermo Fisher) was supplemented in the E8fn medium for 1 d after passaging for increased viability. The WA25 hES cell line was acquired from the WiCell Research Institute. For additional validation and testing information refer to the cell line webpage: <https://www.wicell.org/home/stem-cell-lines/catalog-of-stem-cell-lines/wa25.cmsx>.

CRISPR genomic editing

To generate the *CHD7*-3×Flag cell line, a *CHD7*-3×Flag-PGK-Puro donor plasmid was constructed by cloning the following fragments into pUC19 plasmid backbone using Gibson Assembly²: left and right homology arms (LHA and RHA) homologous to 1 kb genomic DNA upstream and downstream of the *CHD7* stop codon locus, respectively (PCR amplified from WA25 genomic DNA), Glycine-Serine linker-3×Flag-STOP-bGH polyA (gBlock DNA, IDT), and a PGK-Puro cassette (Addgene #31938)³. Ribonuclease protein (RNP) complex⁴ targeting the *CHD7* stop codon locus was assembled by incubating high fidelity Cas9 protein (HiFi Cas9

nuclease 3NLS, IDT) with crRNA:tracrRNA duplex (*CHD7* crRNA: 5'–ACTTGAAGTGGAACTGGTAC –3', IDT). The RNP complex, the donor plasmid, as well as an electroporation enhancer (IDT) were transfected into the WA25-based *PAX2-2A-nGFP* reporter hES cell line (Hashino lab, unpublished data) with 4D Nucleofector (Lonza) using the P3 Primary Cell 4D-Nucleofector X kit (Lonza) and Program CB-150. After nucleofection, cells were plated in RevitaCell-supplemented E8fn medium for one day for improved cell survival rate, followed by treatment with 1 μ M Scr7 (Xcessbio) on the second day for enhanced HDR efficiency⁵. 0.25–0.5 μ g/mL puromycin (Thermo Fisher) selection was performed for four days starting from the third day post-nucleofection. Clonal cell lines were established by low-density seeding followed by isolation of hESC colonies after 6 days of expansion. Genotypes of cell lines were analyzed by PCR and Sanger sequencing. Cell lines with bi-allelically inserted 3 \times Flag at the *CHD7* stop codon locus were used for downstream validation and experiments.

CHD7 mono-allelic and bi-allelic knockout (KO) cell lines were generated with the double-nicking CRISPR strategy⁶ to minimize any potential off-target mutation. Two gRNAs (5'–GGGGTGTGATACTGCGAGTG –3' and 5'–GTTCTCAGGTGCCCCATGG –3', offset = 4 bp) targeting the first coding exon of *CHD7* were individually cloned into a gRNA expression plasmid pSPgRNA (Addgene #47108)⁷. The two gRNA plasmids, along with a double-nicking Cas9 expression plasmid hCas9_D10A (Addgene #41816)⁸ and a puromycin expression plasmid pPGKpuro (Addgene #11349)⁹ were transfected into the WA25-based *PAX2-2A-nGFP* reporter hES cell line with a 4D Nucleofector. After nucleofection, cells were plated in RevitaCell-supplemented E8fn medium for one day for improved cell survival. To enrich transfected cells, 0.5 μ g/mL puromycin selection was performed for two days starting from 48 h post-nucleofection. After confirming successful indel formation at the cell population level using the

T7 endonuclease 1 assay (T7E1, New England Biolabs), clonal cell lines were established by low-density seeding (1–3 cells/cm²) of accutase (Thermo Fisher)-dissociated single cells followed by isolation of hESC colonies after 5–7 d of expansion. T7E1 assay showed 26 out of 94 (28%) established clonal cell lines contained indels. Next generation sequencing (NGS) of amplicon DNA was performed on the 26 T7E1-positive cell lines at the *CHD7* first coding exon locus (Genome Engineering and iPSC Center, Washington University). Among these cell lines, 5 cell lines (19%) had frame-shift indels on both alleles (bi-allelic mutant cell lines, indel size ranging from -25 bp to +79 bp), 13 cell lines (50%) had frame-shift indels on one allele and wild-type (WT) sequence on the other allele (mono-allelic mutant cell lines, indel size ranging from -20 bp to +38 bp), and the remaining 8 cell lines contained in-frame indels on one or both alleles (31%). A bi-allelic cell line with a 25 bp deletion and an 8 bp insertion, as well as a mono-allelic cell line with a 11 bp deletion and a WT allele were selected for further validation and analysis. To test whether each of these two cell lines were homogenous cell lines or mixtures of cells with different genotypes, multiple second-round clonal cell lines were derived from each of these two parental lines. TIDE (Tracking of Indels by DEcomposition) analysis¹⁰ of the second-round clonal cell lines showed identical indel patterns as their parental lines, with ~50% frequencies for both of the two peaks, thus confirming that these two *CHD7* mutant cell lines were homogenous.

To generate mono-allelic and bi-allelic *CHD7* S834F missense point mutation hESC lines, a gRNA (5'–GCTCTTATCTTCATTGTCAG–3') targeting the *CHD7* p.S834 c.2501C locus was cloned into the gRNA expression plasmid pSPgRNA (Addgene #47108)⁷. An optimized CRISPR base editor expression cassette (BE-FNLS-2a-Puro) was sub-cloned from its original lentiviral vector (Addgene #110841)¹¹ into a pUC19 vector backbone. These two plasmids were

nucleofected into the WA25-based *PAX2-2A-nGFP* reporter hES cell line followed by 0.5 µg/mL puromycin selection on the following day. After a cell passaging on post-nucleofection day 7, genomic DNA from an aliquot of day 13 targeted population of cells was harvested for PCR and sequencing. The genotyping results demonstrated ~44% C to T conversion rate as analyzed by the EditR tool¹². A total of 48 clonal cell lines were established and screened with PCR and Sanger sequencing, in which 27 cell lines contained a total of 44 c.2501C>T p.S834F mutant alleles. After excluding cell lines with C>T conversions at c.2506C, cell lines with indels, and cell lines showing signs of heterogeneity, we obtained 7 bi-allelic and 5 mono-allelic c.2501C>T p.S834F mutant cell lines. One cell line from each of the two genotypes was used for downstream validation and experiments.

To generate the *POU4F3-2a-ntdTomato (POU4F3^{nT}) CHD7^{KO/+}* cell line, we constructed a pUC19-*POU4F3-2a-tdTomato-nls-bGHpA* donor plasmid via restriction enzyme digestion and T4 ligase-based subcloning and Gibson assembly. RNP complex targeting the *POU4F3* stop codon locus was assembled by incubating high fidelity Cas9 protein (HiFi Cas9 nuclease 3NLS v3, IDT) with a *POU4F3* sgRNA (5'–ATTCGGCTGTCCACTGATTG–3', Synthego). The RNP complex, the donor plasmid, a puromycin expression plasmid pPGKpuro (Addgene #11349)⁹, and an electroporation enhancer (IDT) were transfected into the established *CHD7^{KO/+}* parental hESC line by nucleofection. After nucleofection, cells were plated in RevitaCell-supplemented E8fn medium for one day, followed by treatment with 1 µM Scr7 on the second day for enhanced HDR efficiency⁵. 0.25 µg/mL puromycin selection for transfected cells was performed on the second and the third day after nucleofection. Clonal cell lines were established by low-density seeding followed by isolation of hESC colonies. Genotypes of cell lines were analyzed

by PCR and Sanger sequencing. Cell lines with bi-allelically inserted 2a-ntdTomato at the *POU4F3* stop codon locus were used for downstream validation and experiments.

To generate the AAVS1-mT (membrane-localized tdTomato) cell line, a AAVS1-pCA-mT donor plasmid was constructed by cloning 1 kb AAVS1 LHA and RHA (PCR amplified from WA25 genomic DNA) and pCA-mT (Addgene #17787)¹³ into pUC19 plasmid backbone using Gibson Assembly. The donor plasmid was nucleofected into WA25 hESCs along with a high fidelity Cas9 RNP complex (AAVS1 crRNA: 5'–ACCCACAGTGGGGCCACTA–3'), an electroporation enhancer, and a pPGKpuro plasmid. RevitaCell, 1 μM Scr7, and 0.5 μg/mL puromycin were used to treat the transfected population of cells, and low-density cell seeding was performed as described above. As the ubiquitous pCA promoter drives the expression of mT in all cell types at all developmental stages, including in hESCs, only hESC colonies emitting the mT fluorescence signals were isolated to establish clonal cell lines. PCR and Sanger sequencing were used to confirm successful mT knockin at the AAVS1 locus in these cell lines.

For all CRISPR genomically engineered hES cell lines, pluripotency was verified by immunohistochemistry of pluripotency markers (OCT4, SSEA4, and SOX2), and normal hESC colony morphologies were verified with a bright-field microscope. Top 10 predicted off-target sites (www.crispr.mit.edu and www.guidescan.com¹⁴) of each gRNA were PCR amplified (~1 kb) from the genomic DNA of established cell lines and were Sanger sequenced to test for off-target mutations. Karyotyping assays were performed at KaryoLogic, Inc.

Human inner ear organoid culture

Human inner ear organoids were derived from hESCs following our previous protocol¹⁵ with modifications. Briefly, to start differentiation, hESCs cultured on 6-well plates were washed three times with DPBS (Thermo Fisher) followed by dissociation with accutase (Thermo Fisher) for 8 min at 37°C. Dissociated cells were pelleted by centrifuging for 3 min at 100×g and were resuspended in E8fn medium containing 20 μM Y-27632 (Stemcell Technologies) to a final concentration of 35,000 cells/mL. 100 μL of cells were added to each well (3,500 cells per well) of Nunclon Sphera low-binding 96-well U-bottom plate(s) (Thermo Fisher) and were centrifuged to the bottom of the wells at 120×g for 5 min. After ≥4 h of incubation at 37°C 5% CO₂, 100 μL of E8fn were added to each well to decrease the concentration of Y-27632 to 10 μM. Following a 48 h incubation after cell seeding, the aggregates were transferred to fresh low-binding 96U plates in 180 μL of chemically defined medium (CDM) containing 2% Matrigel (Corning), 10 μM SB-431542 (Stemcell Technologies), 4 ng/mL FGF-2 (Stemcell Technologies), and 100 pg/mL BMP-4 (ReproCell) to initiate non-neural induction – that is, differentiation day 0 (d0). CDM contained a 1:1 mixture of F-12 Nutrient Mixture with GlutaMAX (Thermo Fisher) and Iscove's Modified Dulbecco's Medium with GlutaMAX (IMDM; Thermo Fisher), additionally supplemented with 0.5% Bovine Serum Albumin (BSA, Sigma), 1× Chemically Defined Lipid Concentrate (Thermo Fisher), 7 μg/mL Insulin (Sigma), 15 μg/mL Transferrin (Sigma), 450 μM Mono-Thioglycerol (Sigma), and 100 μg/mL Normocin (Invivogen). On day 4 of differentiation culture, 45 μL of CDM containing 250 ng/mL FGF-2 (50 ng/mL final concentration) and 1 μM LDN-193189 (200 nM final concentration; ReproCell) was added to the pre-existing 180 μL of media in each well. On day 8 of differentiation culture, 45 μL of CDM containing 18 μM CHIR-99021 (3 μM final concentration; Stemcell Technologies) was added to the pre-existing 225 μL of media in each well. On differentiation day 11, the aggregates were pooled together, washed with

DMEM:F12 with HEPES (Thermo Fisher), and resuspended in freshly prepared Organoid Maturation Medium (OMM) supplemented with 1% Matrigel and 3 μ M CHIR-99021. The OMM medium contains a 1:1 mixture of Advanced DMEM:F12 (Thermo Fisher) and Neurobasal Medium (Thermo Fisher) supplemented with 0.5 \times N2 Supplement (Thermo Fisher), 0.5 \times B27 without Vitamin A (Thermo Fisher), 1 \times GlutaMAX (Thermo Fisher), 0.1 mM β -Mercaptoethanol (Thermo Fisher), and 100 μ g/mL Normocin. Starting from day 11, the aggregates were cultured stationary on non-coated 100 mm dishes. On day 13 and day 15, medium was changed with OMM supplemented with 3 μ M CHIR-99021. Starting on day 18, the aggregates were cultured in OMM without additional supplements. OMM medium change were performed twice a week or when the color of the medium start to turn orange or slightly yellow. Aggregates can be gently washed off the dishes and transferred to new 100 mm dishes during medium change to get rid of the migrating cells growing adherently on the dishes, which may compete with the aggregates to consume the culture medium.

To start differentiation of the WT-*CHD7*^{KO/KO} chimeric cultures, WT (AAVS1^{mT}) and *CHD7*^{KO/KO} hESCs were separately dissociated and resuspended to a final concentration of 35,000 cells/mL. Equal volume of WT and *CHD7*^{KO/KO} hESCs were mixed in a fresh tube, followed by seeding of 100 μ L of the cell mixture to each well of low-binding 96U plates (3,500 cells per well). The rest of the culture procedures were the same as regular organoid cultures.

Organoid dissociation and scRNA-seq

For dissociation of d20 aggregates, 30 aggregates with *PAX2*-2a-nGFP signal (as well as 10 WA25 control aggregates dissociated in separate tubes and wells) were washed three times with

DPBS, three times with 1.1 mM EDTA, followed by resuspension in an accutase solution. 10–15 aggregates were transferred to each well of a Nunclon Sphera low-binding 6-well plate (Thermo Fisher) along with 3 mL of accutase. The plates were incubated at 37°C 5% CO₂ for 90 min with gentle trituration every 5–10 min with a wide-bore P1000 pipet tip. Dissociated cells were filtered through a 100 µm cell strainer and then a 40 µm cell strainer (Corning), and then centrifuged in 2 mL round-bottom tubes at 100 × g for 3 min. Cell pellets were resuspended in a DMEM:F12 solution (with HEPES, no phenol red; Thermo Fisher) supplemented with 10% FBS (Thermo Fisher) and 1:500 propidium iodide (Thermo Fisher) cell viability dye. GFP⁺ propidium iodide⁻ cells were sorted into a DMEM:F12 (with HEPES, no phenol red) solution supplemented with 10% FBS on a SORP Aria FACS machine (BD Biosciences) for 1 h at Indiana University Flow Cytometry Resource Facility, using dissociated cells from d20 WA25 aggregates as negative control for gating.

To enrich hair cells and supporting cells from d70 WT (*POU4F3*^{nT} *PAX2*^{nG}) and *CHD7*^{KO/+} (*CHD7*^{KO/+} *POU4F3*^{nT} *PAX2*^{nG}) organoids, tissues containing vesicle structures harboring the *POU4F3*^{nT}-positive hair cells were micro-dissected from the rest of the d70 aggregates with fine tweezers (Dumont) under a fluorescence stereomicroscope. The dissected d70 organoid tissues were dissociated and FACS sorted in a similar way as d20 organoids, with the exception that no viability dye was used. tdTomato⁺ and tdTomato⁻ cell populations were collected in separate tubes for separate downstream scRNA-seq reactions.

Sorted single cells were captured, lysed, and cDNA libraries were generated using a Chromium Controller and Single Cell 3' Reagent Kits V3 (10X Genomics) following the manufacturer's instructions. cDNA library quality was verified using a bioanalyzer (Agilent Technologies),

followed by sequencing using the NovaSeq 6000 sequencing system (Illumina) at Indiana University School of Medicine Center for Medical Genomics.

scRNA-seq data analysis

Illumina's Real Time Analysis software was used to generate a BCL file, which was subsequently de-multiplexed and converted to a FASTQ file by the bcl2fastq Conversion Software (Illumina). The Cell Ranger pipeline was used to process the FASTQ file as follows: De-multiplexed reads were mapped to the GRCh38/hg38 human reference genome with the STAR (Spliced Transcripts Alignment to a Reference) aligner, mapped reads were grouped by cell barcode, and single-cell gene expression was quantified using unique molecular identifiers (UMIs). The resulting filtered gene-barcode (count) matrix was used as input for downstream analysis.

Using the Seurat¹⁶ v4.0.3 R package, scRNA-seq datasets were loaded to R and converted to Seurat objects using Seurat functions `Read10X` and `CreateSeuratObject`, respectively. Low quality cells with extremely high or low number of detected UMIs and cells with high percentage of mitochondrial reads were filtered out from subsequent analysis. Datasets were merged across samples, followed by data normalization, scaling, and variable gene identification using the `SCTransform` function. Principal component analysis (PCA) was performed and the first 30 principal components were retained for downstream analysis. Clustering was performed with `FindNeighbors` and `FindClusters` functions, and cluster markers were identified using the `FindMarkers` function in Seurat.

Differential expression (DE) analysis was performed using the DESeq2 package¹⁷ in conjunction with zingeR in R, and DE genes were visualized on volcano plots using the EnhancedVolcano R package (<https://github.com/kevinblighe/EnhancedVolcano>). To generate the bubble plot for differentially expressed otic-specific genes, the gene expression fold-change values of E10.5 otic vesicles versus non-otic tissues reported by Hartman et al.¹⁸ were used to calculate the area of gene-correlated spots on the volcano plot.

Gene set enrichment (GSE) analysis was performed using the Integrative Differential expression and gene set Enrichment Analysis (iDEA) R package¹⁹. The DESeq2 differential expression analysis results were used as a summary statistics input for iDEA, and upregulated genes and downregulated genes were analyzed separately with iDEA. All gene sets used for GSEA, some of which were included with the iDEA package¹⁹, can be downloaded from the MSigDB database (<http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>).

Immunohistochemistry

Aggregates were fixed with 4% paraformaldehyde (PFA, Electron Microscopy Sciences) for 30 min at room temperature (RT) followed by graded treatment of 15% and 30% sucrose and embedding in tissue-freezing medium. Frozen tissue blocks were sectioned into 10–20 μm cryosections on a Leica CM-1860 cryostat. For fixation of hESCs, cells growing on the 6-well plates were fixed with 4% PFA for 15 min at RT on the plates. No sucrose treatments or cryosections were performed on fixed hESCs. For immunostaining of both the aggregates and the hESCs, a 10% horse serum (Vector Laboratories) in 0.1% Triton X100 1 \times PBS solution was used for blocking, and a 3% horse serum in 0.1% Triton X100 1 \times PBS solution was used for

primary and secondary antibody incubations. Alexa Fluor conjugated anti-mouse (IgG1, IgG2a, and IgG2b), rabbit, sheep, or goat secondary antibodies (Thermo Fisher) were used for primary antibody detection. ProLong Gold antifade reagent with DAPI (Thermo Fisher) was used to mount the samples and visualize cellular nuclei.

Microscopy images of hES cells and sectioned aggregates were captured on a Nikon A1R-HD25 confocal microscope or a Leica DMI8 inverted microscope. Primary antibodies used in this study were listed in Extended Data Table 1. All anti-CHD7 immunostaining assays in this study used the CHD7 antibody from R&D Systems (#AF7350).

Western blot

hESCs were lysed in of RIPA buffer (Thermo Fisher) supplemented with 1× Halt protease inhibitor cocktail (Thermo Fisher) and 5 mM EDTA for 15 min on ice. After cell lysis, the sample were mixed with Laemmli sample buffer (Bio-Rad) and DTT (Thermo Fisher) to final concentrations of 1× and 25 mM, respectively. Samples were heated at 95°C for 10 min, and then centrifuged at 14,000×g for 15 min to pellet the cell debris. Supernatants were loaded onto a 4–15% Mini-PROTEAN TGX precast gel (Bio-Rad) and were subject to electrophoresis in 1× running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) (Bio-Rad) at 200 V for 35 min. After electrophoresis, samples were transferred to a PVDF membrane (Bio-Rad) using a Trans-Blot Turbo system (Bio-Rad) at 1.3 A, up to 25 V for 10 min. The PVDF membrane was briefly immersed in a wash buffer (0.05% Tween-20 (Sigma) in PBS), and then in a blocking buffer (0.05 g/mL blotting-grade blocker (Bio-Rad) and 0.05% Tween-20 in PBS) for 30 min shaking at RT. Primary antibodies and HRP-conjugated secondary antibodies were diluted in the

blocking buffer, and antibody incubation was performed rocking overnight at 4°C for primary antibodies and rocking for 1–2 h at RT for secondary antibodies. Three times of 10 min washing in the wash buffer was performed following primary and secondary antibody incubations. Bands were detected with an ECL substrate (Bio-Rad) and the membrane was imaged with a ChemiDoc imager (Bio-Rad). For HRP-conjugated primary antibodies, the secondary antibody incubation step and the subsequent washing steps were omitted. For blotting with a different primary antibody on the same membrane, stripping with a Restore Plus stripping buffer (Thermo Fisher) was performed to remove prior antibodies and chemiluminescent substrates. Primary antibodies used in this study were listed in Extended Data Table 1. Unless otherwise noted, all anti-CHD7 western blotting assays in this study used the CHD7 antibody from R&D Systems (#AF7350).

Representative data and reproductivity

Prior to harvesting organoid samples for scRNA-seq or immunostaining, d19–d20 aggregates were pre-screened based on the epithelial *PAX2*^{nG} fluorescence signals, and d60–d70 aggregates with the *POU4F3*^{nT} reporter knockin were pre-screened based on the hair cell-specific *POU4F3*^{nT} fluorescence signals. Low quality organoids with few *PAX2*^{nG}-positive epithelial vesicle structures or few *POU4F3*^{nT}-positive hair cells were not used from subsequent scRNA-seq or immunostaining experiments.

Unless stated otherwise, all immunostaining of organoids shown in this article are representative of a minimum of three aggregates from a minimum of three separate experiments.

Statistical information

For samples to be used for comparison of immunofluorescence intensities, all samples were processed for immunostaining at the same time using the same tubes of diluted primary or secondary antibody mixtures, and samples were blocked, incubated, and washed for the same durations of time. These stained samples were imaged with the Nikon A1R-HD25 confocal microscope using the same image capture settings, including the same laser power, the same pinhole size, the same HV gain and offsets, etc. The exported TIFF images from each fluorescence channel were merged but were not adjusted in any other way prior to fluorescence intensity measurement in the Fiji software. The circle click tool from the ROI 1-click tool sets in Fiji was used to manually measure fluorescence intensities of nuclear proteins such as SOX2, HOXB9, and DLX5. The polygon selection tool from Fiji was used for whole cell fluorescence intensity manual measurements for cell body-localized proteins COL9A2 and FBXO2. To visualize and locate the nuclei and the cell membranes of otic progenitors during measurement, PAX2^{nG} and EPCAM channels were merged with the channel of interest. The Fiji software records intensity data from each channel separately, and only the data from the channel of interest were used for downstream analysis.

Graphical plots and statistical analysis of measured fluorescence intensity data were performed in GraphPad Prism 9. All sample were subject to Anderson-Darling, D'Agostino & Pearson, Shapiro-Wilk, and Kolmogorov-Smirnov normality testing. Datasets containing sample(s) that did not pass the normality test were analyzed by Kruskal-Wallis test followed by Dunn's multiple comparisons test. Data collection and analysis were not performed blind to the conditions of the experiments. Violin plots display the full distribution of individual data points.

References

1. Beers, J. *et al.* Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nat Protoc* **7**, 2029-2040 (2012).
2. Gibson, D.G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345 (2009).
3. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**, 731-734 (2011).
4. Kim, S., Kim, D., Cho, S.W., Kim, J. & Kim, J.S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res* **24**, 1012-1019 (2014).
5. Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538-542 (2015).
6. Ran, F.A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).
7. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**, 973-976 (2013).
8. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
9. Tucker, K.L. *et al.* Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev* **10**, 1008-1020 (1996).
10. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014).
11. Zafra, M.P. *et al.* Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* **36**, 888-893 (2018).
12. Kluesner, M.G. *et al.* EditR: A Method to Quantify Base Editing from Sanger Sequencing. *CRISPR J* **1**, 239-250 (2018).
13. Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593-605 (2007).
14. Perez, A.R. *et al.* GuideScan software for improved single and paired CRISPR guide RNA design. *Nat Biotechnol* **35**, 347-349 (2017).
15. Koehler, K.R. *et al.* Generation of inner ear organoids containing functional hair cells from human pluripotent stem cells. *Nat Biotechnol* **35**, 583-589 (2017).
16. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
17. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
18. Hartman, B.H., Durruthy-Durruthy, R., Laske, R.D., Losorelli, S. & Heller, S. Identification and characterization of mouse otic sensory lineage genes. *Front Cell Neurosci* **9**, 79 (2015).
19. Ma, Y. *et al.* Integrative differential expression and gene set enrichment analysis using summary statistics for scRNA-seq studies. *Nat Commun* **11**, 1585 (2020).