Rare-variant association analysis of age-related hearing loss reveals known and new hearing loss genes.

Diana M. Cornejo-Sanchez, PhD, Guangyou Li, MSc, Tabassum Fabiha, Ran Wang, MSc, Anushree Acharya, MSc, Jenna L Everard, Magda K Kadlubowska, BS, Yin Huang, PhD, Isabelle Schrauwen, PhD, Gao T. Wang, PhD, Andrew T. DeWan, PhD, Suzanne M. Leal, PhD.

The supplementary information contains 10 figures (supplied in pdf format) and 6 tables (supplied in xlsx format). The description of each table or figure is supplied in their title.

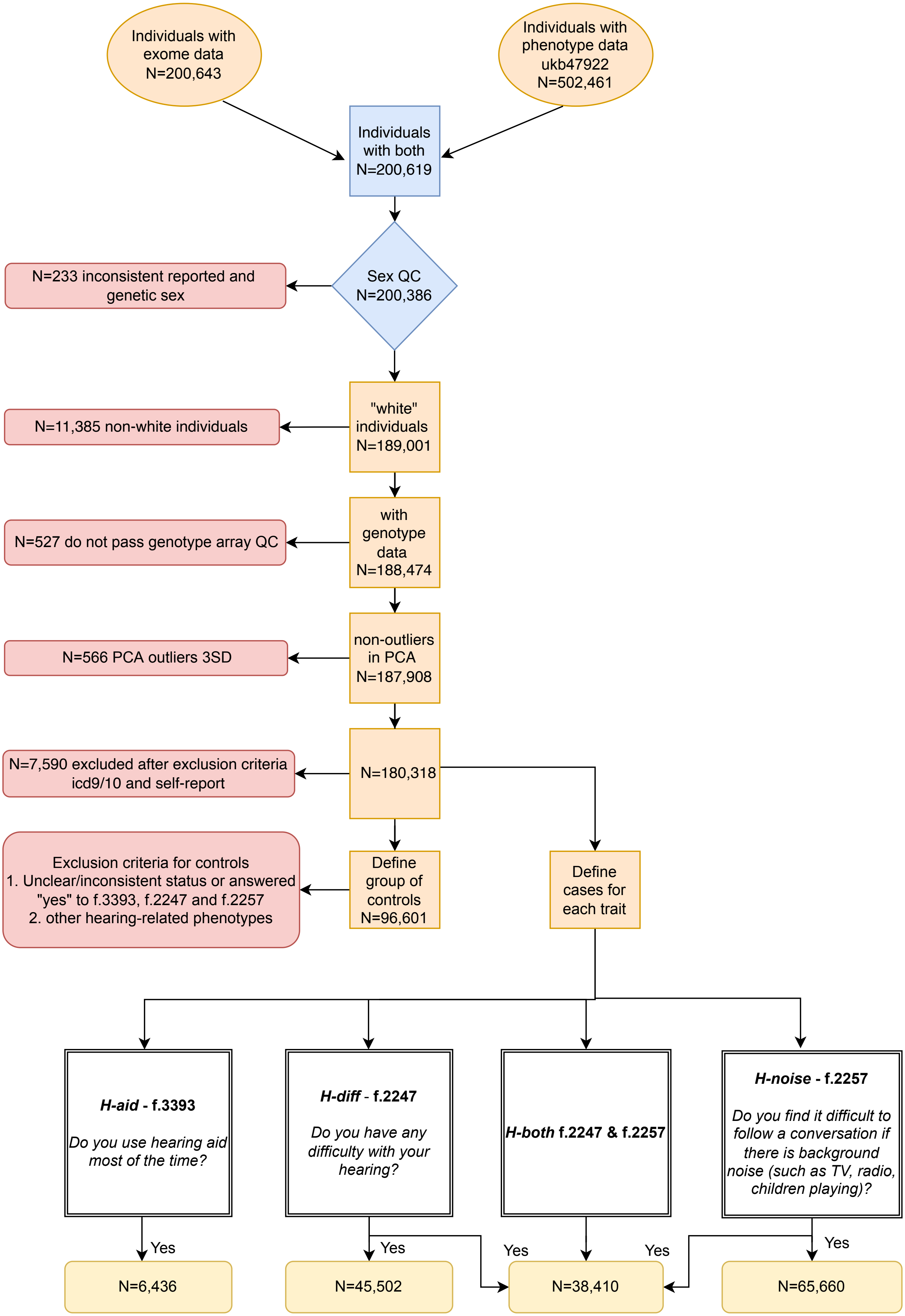

Figure S1. Phenotype definition of age-related hearing traits for individuals with exome sequence data for the mega-sample.

H-aid: hearing aid, *H-diff*: hearing difficulty, *H-both*: hearing difficulty and hearing difficulty with background noise, *H-noise*: hearing difficulty with background noise, QC: quality control, PCA: principal component analysis, SD: standard deviation

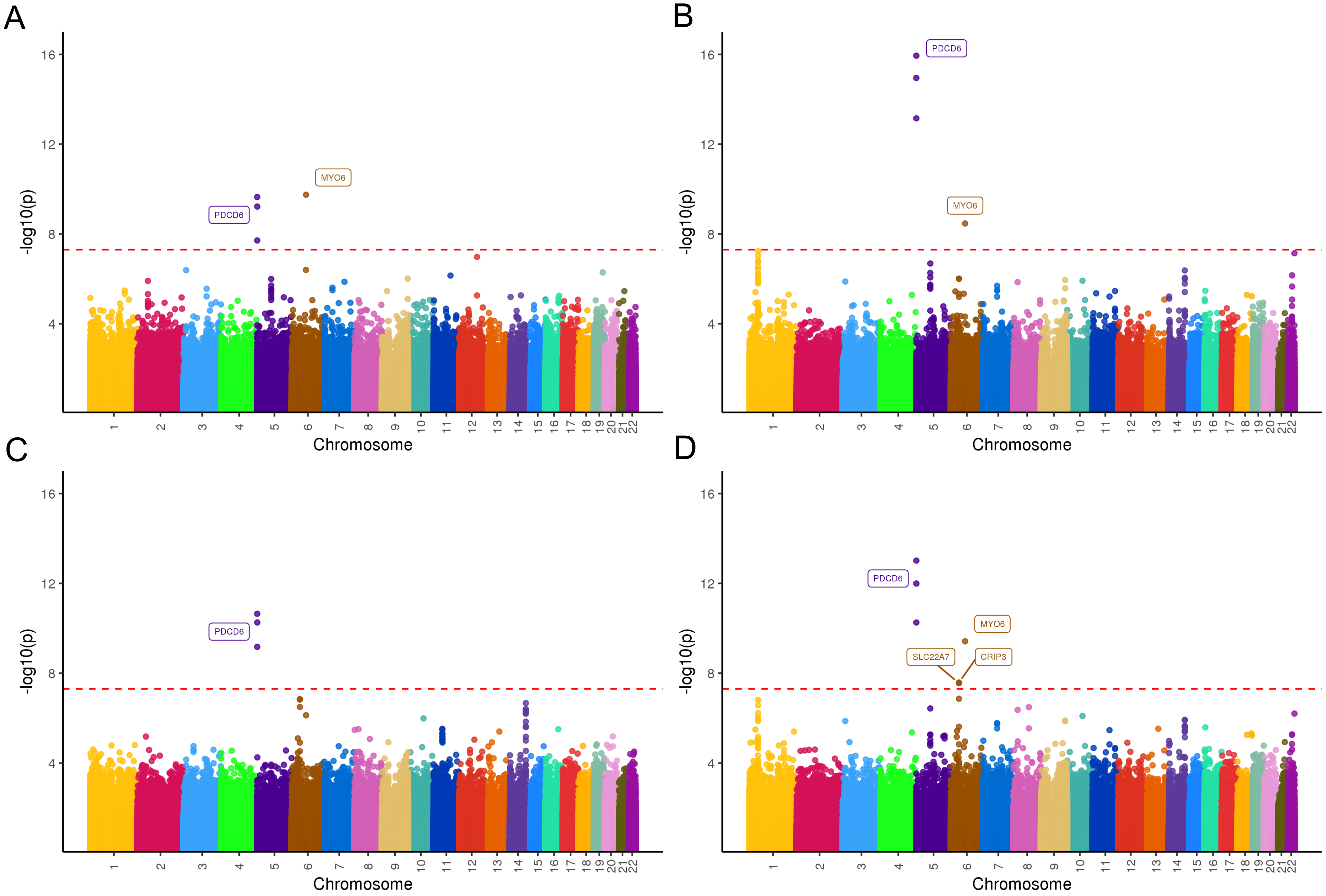

Figure S2 Single variant association analysis Manhattan plots for the discovery sample. Results are shown for the analysis of A) *H-aid*, B) *H-diff*, C) *H-noise* and D) *H-both*. The threshold for genome-wide significance ($p < 5.0 \times 10^{-8}$) is indicated by a red dotted line. Loci that reached genome-wide significance are annotated with the gene label.

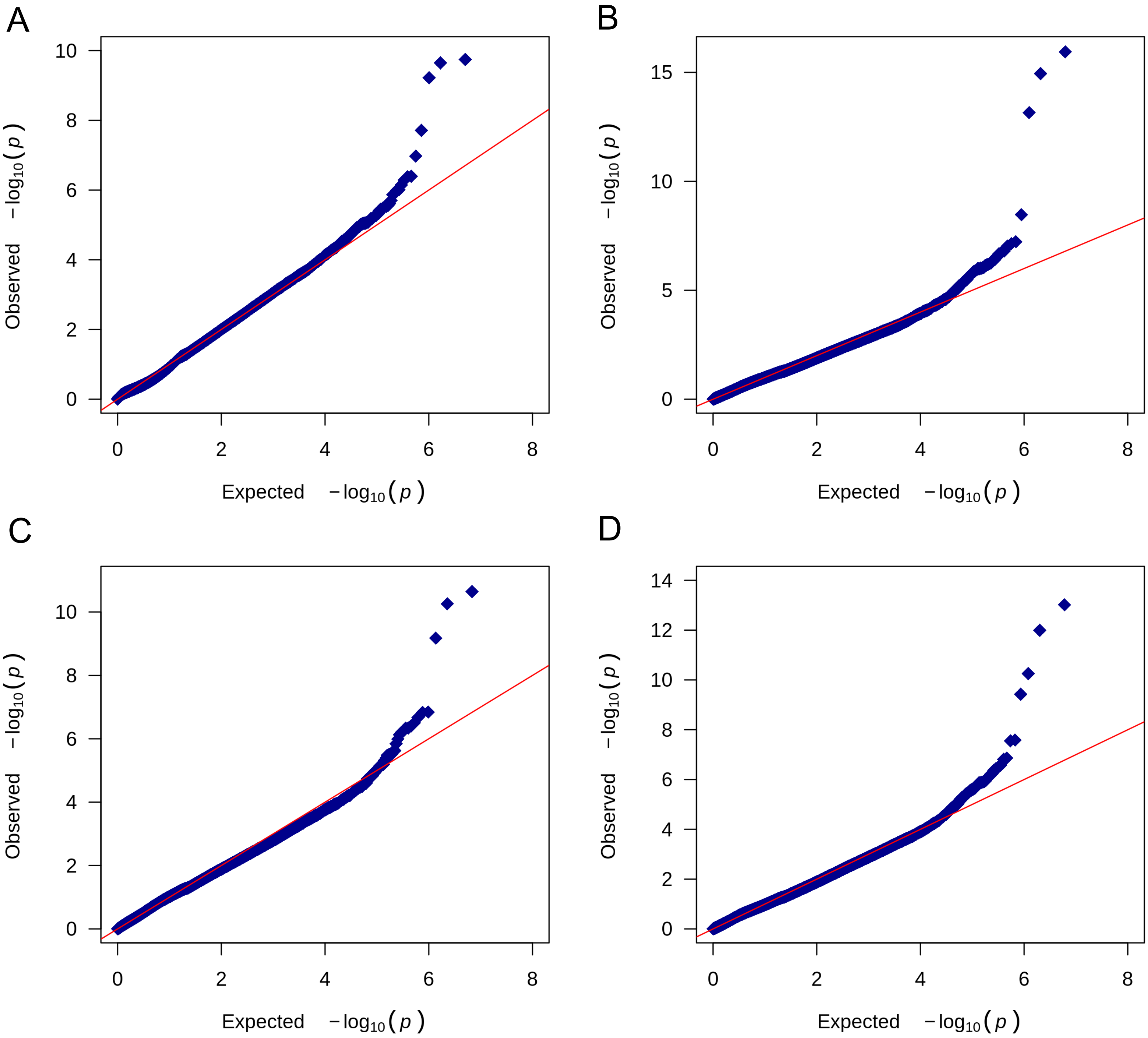

Figure S3 Single variant association analysis Quantile-Quantile (Q-Q) plots for the discovery sample. Results are shown for the analysis of A) *H-aid* (λ=0.90), B) *H-diff* (λ=1.05), C) *H-noise* (λ=1.05) and D) *H-both* (λ=1.08).

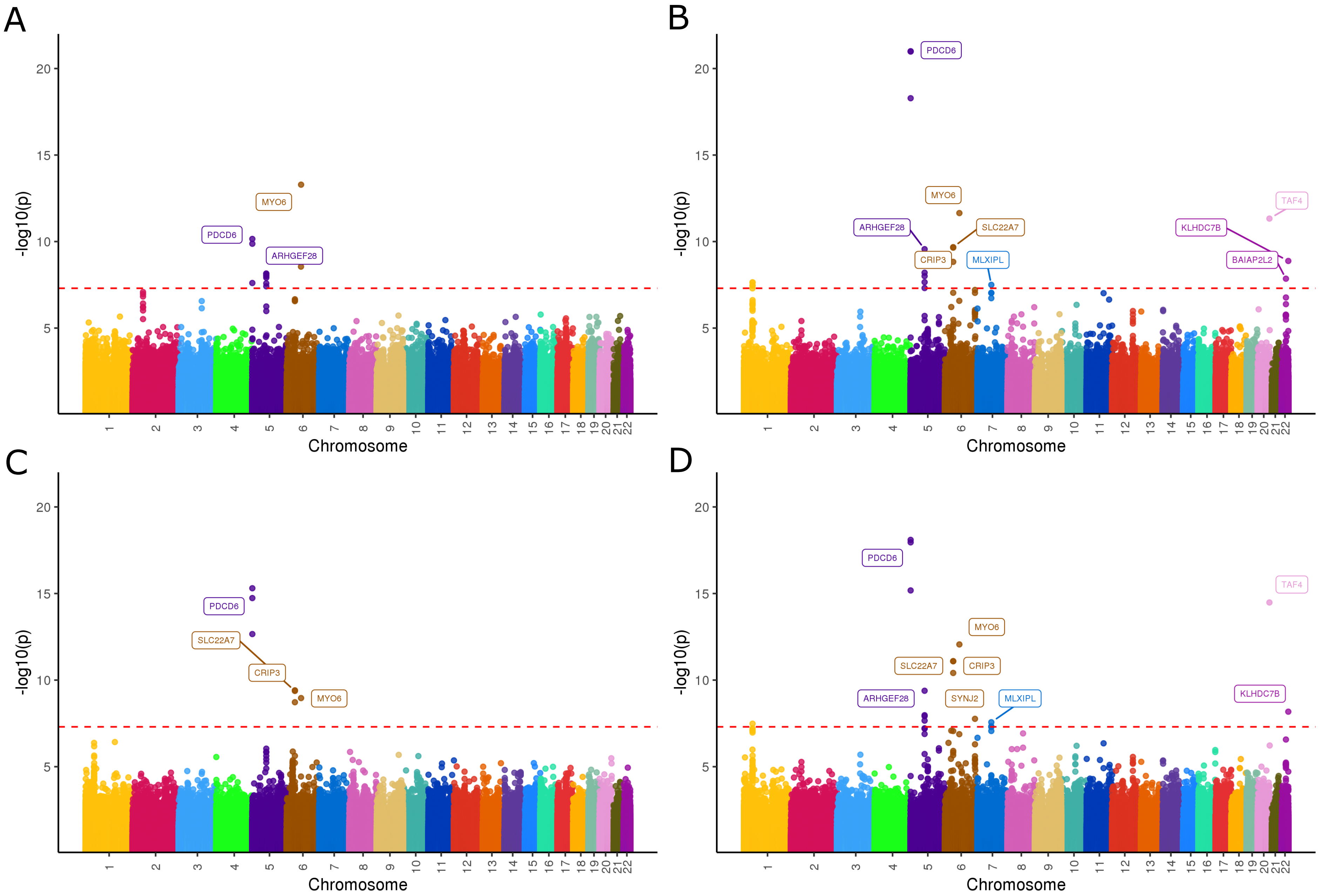

Figure S4 Single variant association analysis Manhattan plots for the mega-sample. Results are shown for the analysis of A) *H-aid*, B) *H-diff,* C) *H-noise* and D) *H-both*. The threshold for genome-wide significance (p < 5.0 x 10-8) is indicated by a red dotted line. Loci that reached genome-wide significance are annotated with the gene label.

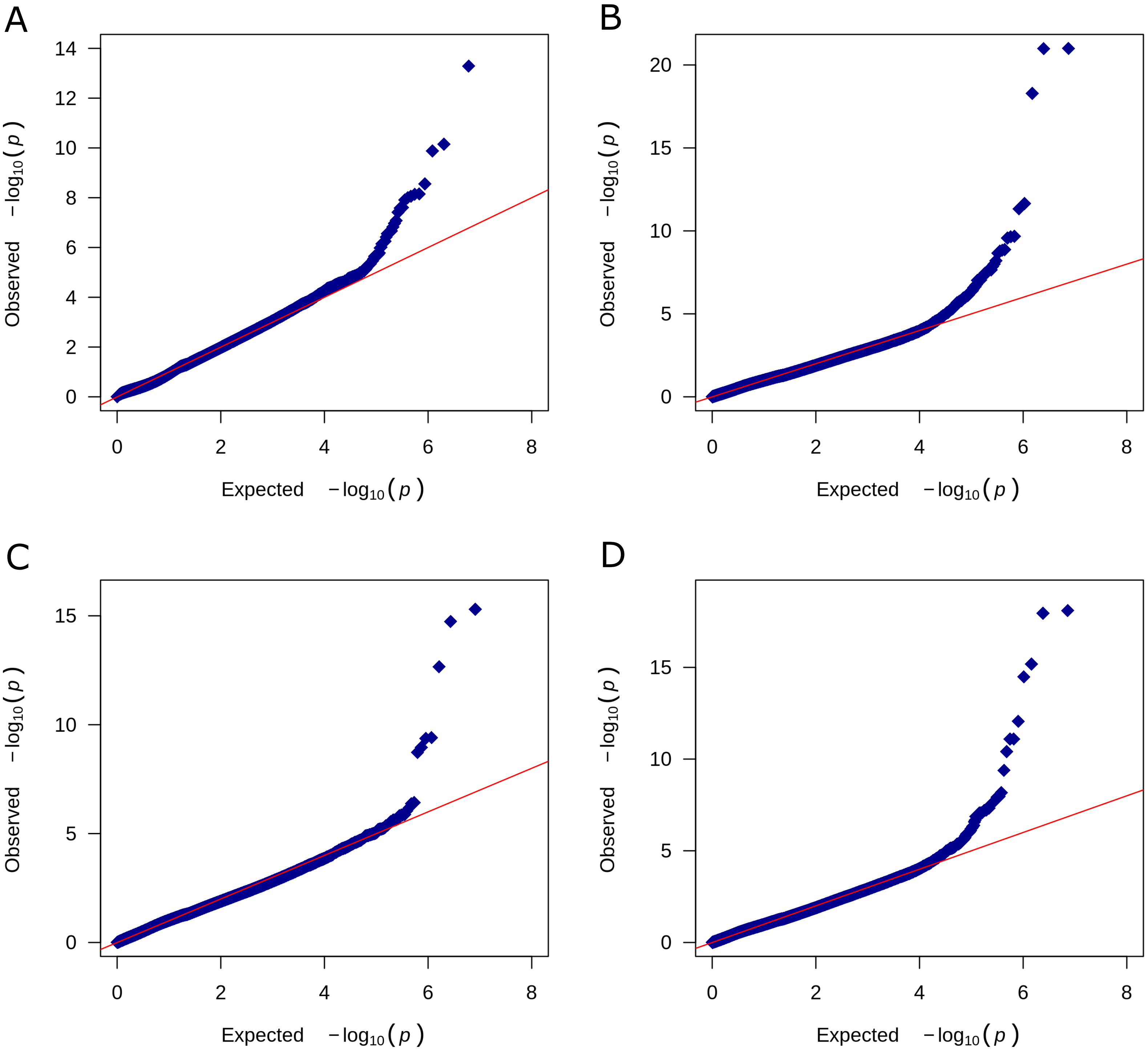

Figure S5 Single variant association analysis Quantile-Quantile (Q-Q) plots for the mega-sample. Results are shown for the analysis of A) *H-aid* (λ=0.94), B) *H-diff* (λ=1.04), C) *H-noise* (λ=1.05) and D) *H-both* (λ=1.07).

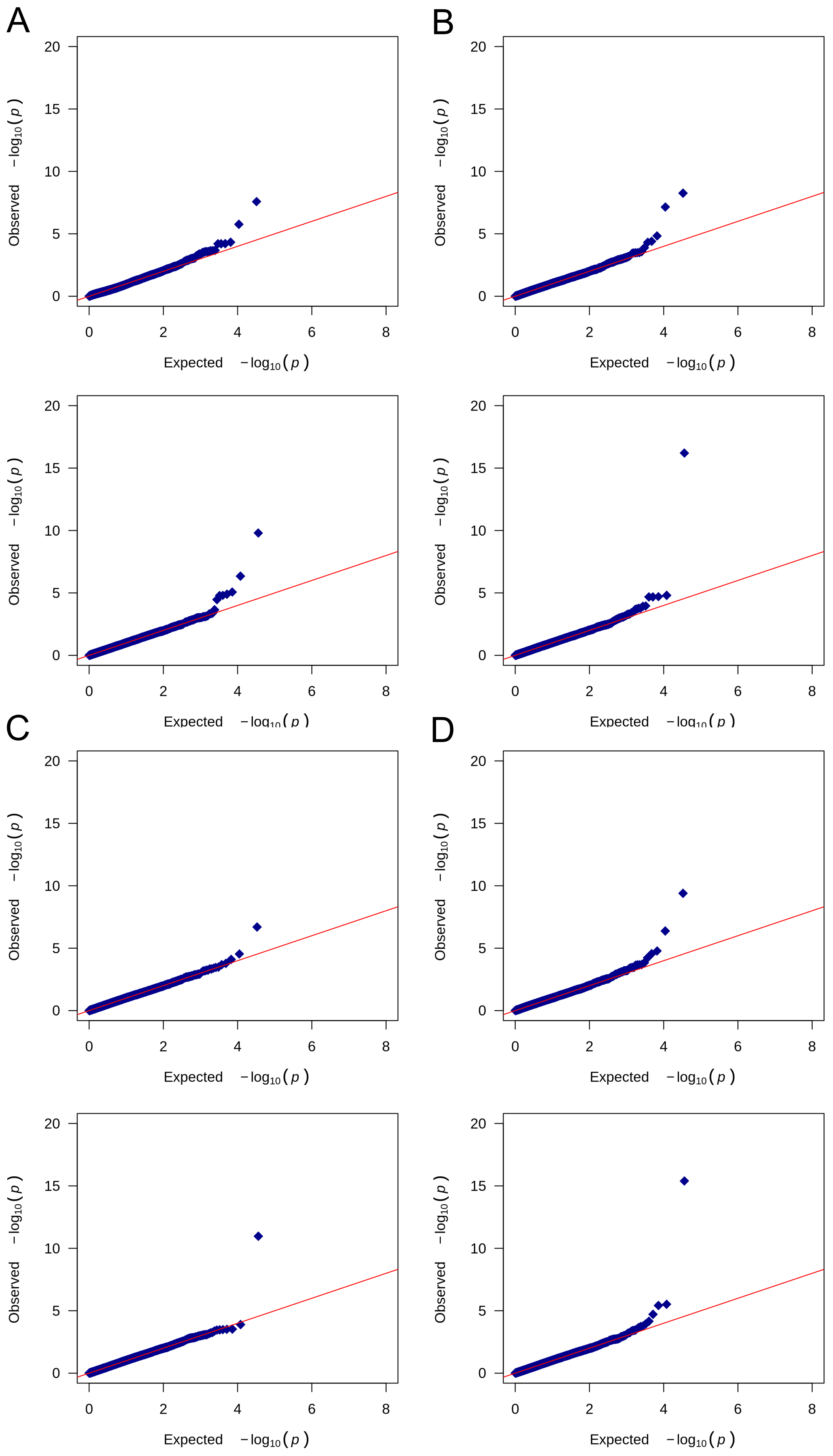

Figure S6 Rare-variant aggregate association analysis Quantile-Quantile (Q-Q) plots for the discovery sample. Results are shown for the analysis of A) *H-aid* (top λ=0.95; bottom λ=0.99), B) *H-diff* (top λ=1.09; bottom λ=1.04), C) *H-noise* (top λ=1.03; bottom λ=1.00) and D) *H-both* (top λ=1.09; bottom λ=1.02). For each panel the top half displays the result for the analysis of predicted loss of function (pLoF) variants and bottom half for missense and pLoF variants.

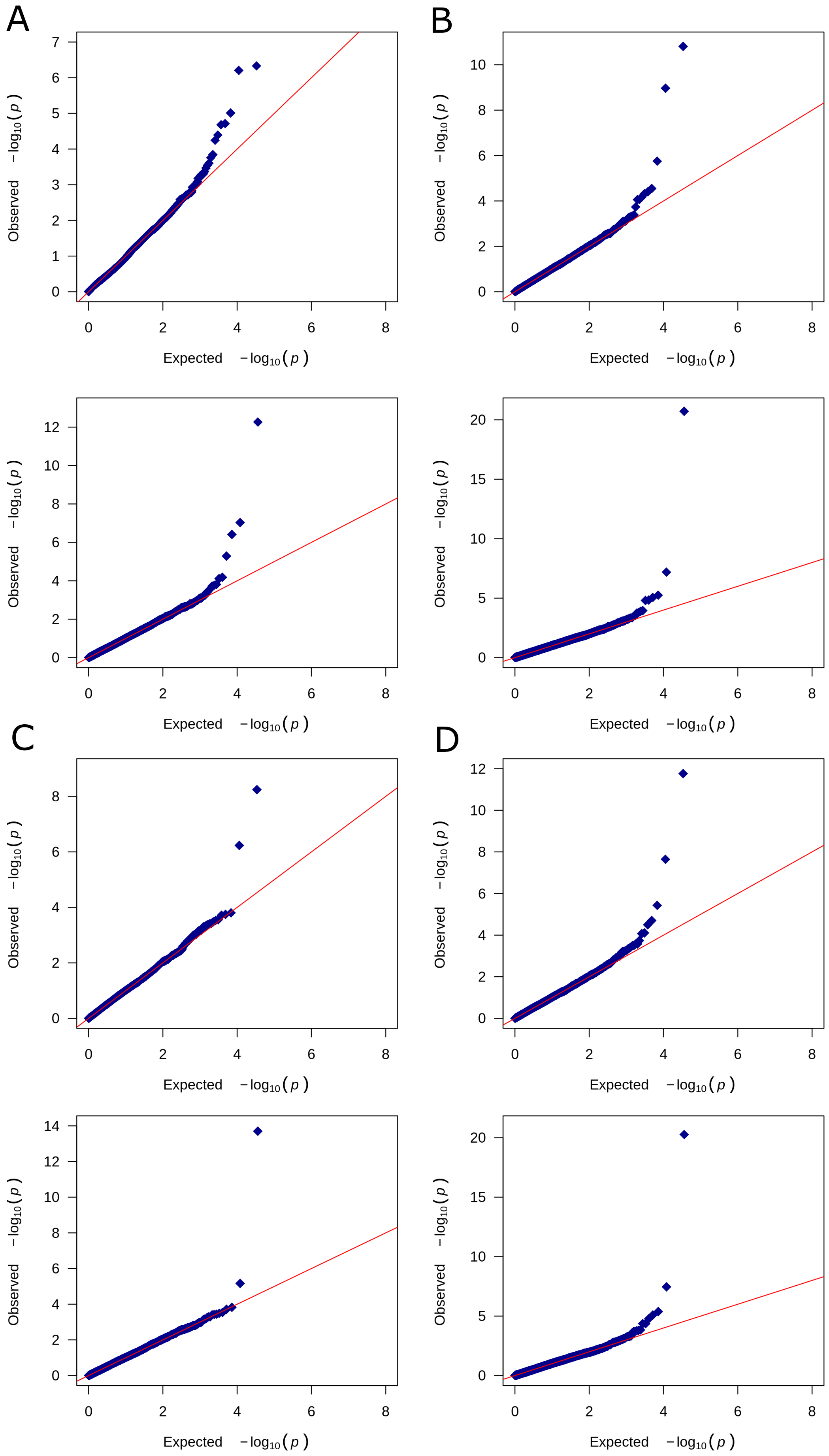

Figure S7 Rare-variant aggregate association analysis Quantile-Quantile (Q-Q) plots for the mega-sample. Results are shown for the analysis of A) *H-aid* (top λ=0.98; bottom λ=1.01), B) *H-diff* (top λ=1.04; bottom λ=1.04), C) *H-noise* (top λ=1.04; bottom λ=1.03) and D) *H-both* (top λ=1.06; bottom λ=1.03). For each panel the top half displays the result for the analysis of predicted loss of function (pLoF) variants and bottom half for missense and pLoF variants.

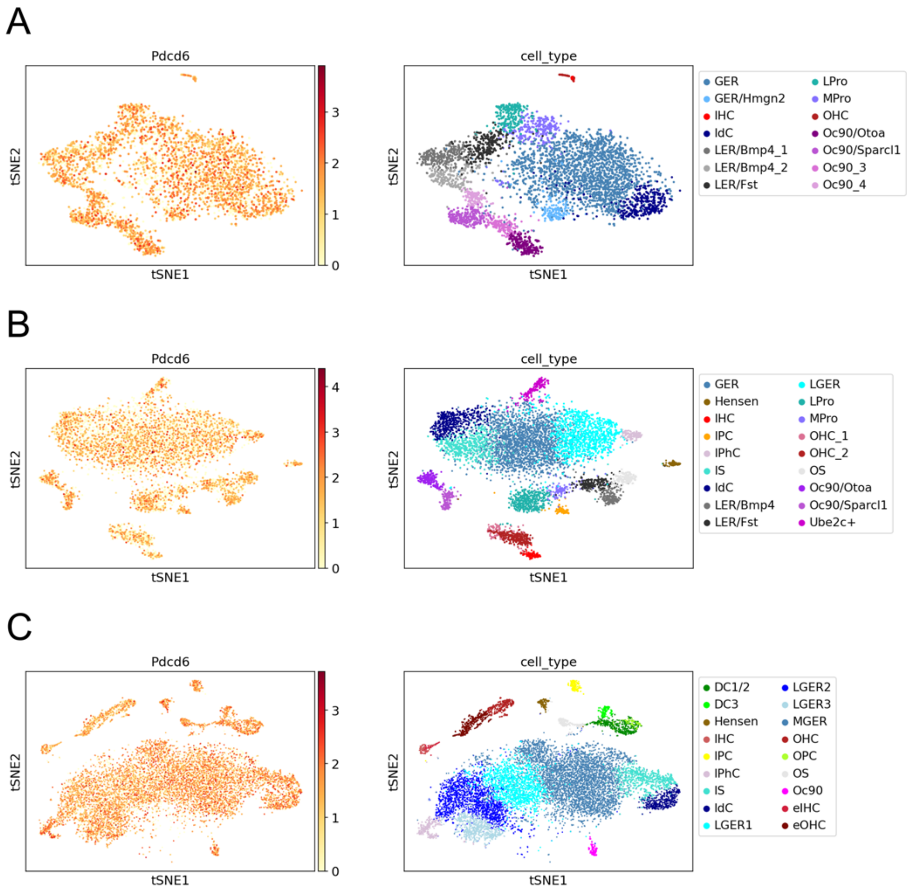

Figure S8 Single cell expression of *Pdcd6* in the cochlear floor epithelia during mouse development. *Pdcd6* expression is presented both in terms of its expression levels (left panel; red = high expression, yellow = low expression) and its location in the cochlear floor epithelium (right panel; colors indicate the cell type) at three developmental stages: **A**. E14, **B**. E16, and **C**. P1. The scale bar represents gene expression based on log transformed and normalized expression data. *Pdcd6* is widely expressed in the cochlear floor epithelia during various stages of development. Data is from a single cell RNA sequencing dataset of the cochlear floor epithelia of CD1 mice(1) and were obtained and plotted using the gEAR analysis suite (gene Expression Analysis Resource)(2).

Abbreviations:

Developing Supporting Cells: GER: Greater Epithelial Range; GER/Hmgn2: Greater Epithelial Range expressing Hmgn2; IdC: Interdental Cells; LER/Bmp4_1: Lesser Epithelial Ridge Cells Expressing Bmp4; LER/Bmp4_2: Lesser Epithelial Ridge Cells Expressing Bmp4; LER/Fst: Lesser Epithelial Ridge Cells Expressing Fst; Oc90: Cells expressing Oc90; Oc90/Otoa: Cells expressing Oc90 and Otoa; Oc90/Sparcl1: Cells expressing Oc90 and Sparcl1; Oc90_3: Cells expressing Oc90; Oc90_4: Cells expressing Oc90; OS: Outer Sulcus Cells; IS: Inner Sulcus Cells; IPhC: Inner Phalangeal Cells; IPC: Inner Pillar Cells; OPC: Outer Pillar Cells; Hensen: Hensen Cells; DC: Dieter Cells; DC1/2: Dieter Cells from rows 1 and 2; DC3: Dieter Cells from row 3; BM: Basilar Membrane Cells; Glia: Glial Cells; *Developing Prosensory Cells:* LPro: Lateral Prosensory Cells; MPro: Medial Prosensory Cells. *Developing Greater Epithelial Ridge Cells*: LGER1: Lateral Great Epithelial Ridge Cells, group 1; LGER2: Lateral Great Epithelial Ridge Cells, group 2; LGER3: Lateral Great Epithelial Ridge Cells, group 3; MGER: Medial Greater Epithelial Ridge Cells; MLGER: Medial Lateral Greater Epithelial Ridge Cells; *Developing Sensory Cells:* IHC: Inner Hair Cells; OHC: Outer Hair Cells; OHC_1: More mature developing outer hair cells; OHC_2: Less mature developing outer hair cells. eIHC: Less mature developing inner hair cells; eOHC: Less mature developing outer hair cells; *Uncategorized*: Ube2c+: Unannotated cells expressing Ube2c+.

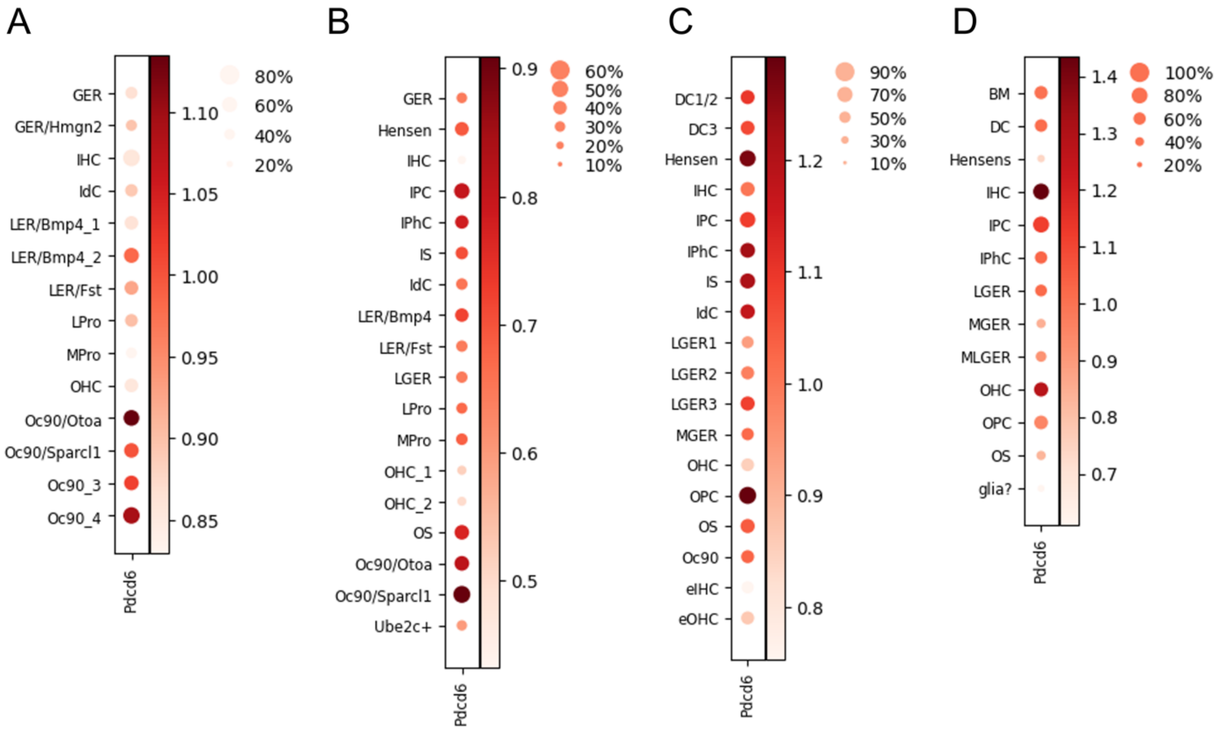

Figure S9 Expression of *Pdcd6* in the cochlear floor epithelia during mouse development per cell type/cluster. Log-transformed and normalized expression levels of *Pdcd6* are indicated for each cell type/cluster in the cochlear floor sensory epithelia of developing mice, with the dot size indicating the percentage of cells expressing *Pdcd6* and the color indicating the intensity of this expression (red = high expression, white = low expression). Expression is shown at four developmental stages: **A.** E14, **B.** E16, **C.** P1, **D.** P7. These plots illustrate that both the expression intensity of *Pdcd6* and the percentage of cells expressing *Pdcd6* generally increase as development progresses. There is an especially prominent upregulation of *Pdcd6* in the Inner Hair Cells (IHC) from E16 through P7. A less prominent upregulation also occurs in the Outer Hair Cells (OHC). Data were obtained from gEAR and plots were generated using its analysis suite(2).

Abbreviations:

Developing Supporting Cells: GER: Greater Epithelial Range; GER/Hmgn2: Greater Epithelial Range expressing Hmgn2; IdC: Interdental Cells; LER/Bmp4_1: Lesser Epithelial Ridge Cells Expressing Bmp4; LER/Bmp4_2: Lesser Epithelial Ridge Cells Expressing Bmp4; LER/Fst: Lesser Epithelial Ridge Cells Expressing Fst; Oc90: Cells expressing Oc90; Oc90/Otoa: Cells expressing Oc90 and Otoa; Oc90/Sparcl1: Cells expressing Oc90 and Sparcl1; Oc90_3: Cells expressing Oc90; Oc90_4: Cells expressing Oc90; OS: Outer Sulcus Cells; IS: Inner Sulcus Cells; IPhC: Inner Phalangeal Cells; IPC: Inner Pillar Cells; OPC: Outer Pillar Cells; Hensen: Hensen Cells; DC: Dieter Cells; DC1/2: Dieter Cells from rows 1 and 2; DC3: Dieter Cells from row 3; BM: Basilar Membrane Cells; Glia: Glial Cells; *Developing Prosensory Cells:* LPro: Lateral Prosensory Cells; MPro: Medial Prosensory Cells. *Developing Greater Epithelial Ridge Cells*: LGER1: Lateral Great Epithelial Ridge Cells, group 1; LGER2: Lateral Great Epithelial Ridge Cells, group 2; LGER3: Lateral Great Epithelial Ridge Cells, group 3; MGER: Medial Greater Epithelial Ridge Cells; MLGER: Medial Lateral Greater Epithelial Ridge Cells; *Developing Sensory Cells:* IHC: Inner Hair Cells; OHC: Outer Hair Cells; OHC_1: More mature developing outer hair cells; OHC_2: Less mature developing outer hair cells. eIHC: Less mature developing inner hair cells; eOHC: Less mature developing outer hair cells; *Uncategorized*: Ube2c+: Unannotated cells expressing Ube2c+.

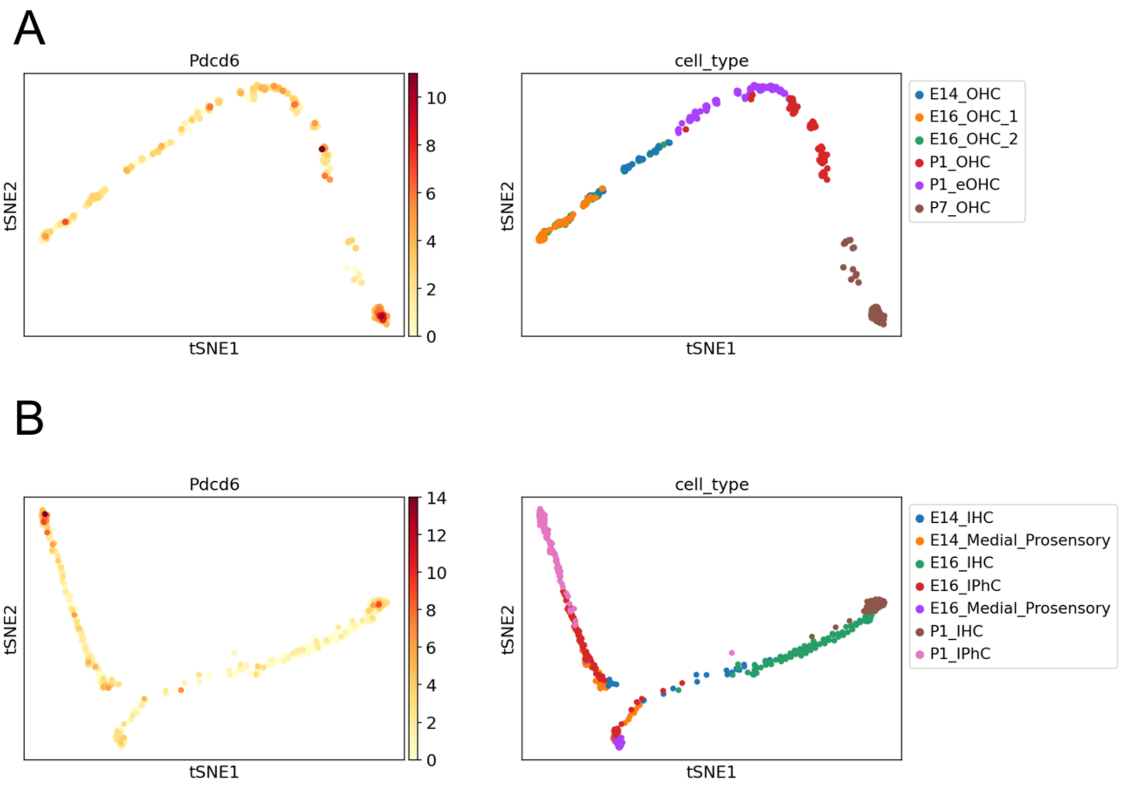

Figure S10 *Pdcd6* expression trajectories in Outer Hair Cells (OHC) and Inner Hair Cells (IHC). *Pdcd6* expression levels from a single cell dataset of the cochlear floor sensory epithelia of the developing mice, presented for OHC and IHC (log-transformed and normalized)(1). The scale bar represents gene expression ranging from low-yellow to high-red based expression data. **A.** *Pdcd6* raw expression values as determined by differential expression analysis (left) compared with OHC trajectories during E14-P7. The highest expression values occur in the mature OHC during P7. **B.** *Pdcd6* raw expression values as determined by differential expression analysis (left) compared with IHC, IPhC, and medial prosensory trajectories during E14-P1. Expression of *Pdcd6* is greatest in P1 in both IHC and IPhC, exhibiting low expression levels everywhere else. Data were obtained and generated with the gEAR analysis suite (2)

Abbreviations IHC: E14_IHC: Early developing inner hair cells E14; E14_Medial_Prosensory: Developing medial prosensory cells E14; E16_IHC: Developing inner hair cells E16; E16_IPhC: Developing inner phalangeal cells E16; E16_Medial_Prosensory: Developing medial prosensory cells E16; P1_IHC: Developing inner hair cells P1; P1_IPhC: Developing inner phalangeal cells P1. Abbreviations OHC: E14_OHC: Early developing outer hair cells E14; E16_OHC_1: More mature developing outer hair cells E16; E16_OHC_2: Less mature developing outer hair cells E16; P1_OHC: More mature developing outer hair cells P1: P1_eOHC: Less mature developing outer hair cells P1; P7_OHC: Developing outer hair cells P7.

References

1. Kolla L, Kelly MC, Mann ZF, Anaya-Rocha A, Ellis K, Lemons A, et al. Characterization of the development of the mouse cochlear epithelium at the single cell level. Nat Commun. 2020;11(1):2389.
2. Orvis J, Gottfried B, Kancherla J, Adkins RS, Song Y, Dror AA, et al. gEAR: Gene Expression Analysis Resource portal for community-driven, multi-omic data exploration. Nat Methods. 2021;18(8):843–4.