



OPEN

Genome-wide association study suggests that variation at the *RCOR1* locus is associated with tinnitus in UK Biobank

Helena R. R. Wells^{1,2}, Fatin N. Zainul Abidin^{2,3}, Maxim B. Freidin¹, Frances M. K. Williams^{1,4}✉ & Sally J. Dawson^{2,4}✉

Tinnitus is a prevalent condition in which perception of sound occurs without an external stimulus. It is often associated with pre-existing hearing loss or noise-induced damage to the auditory system. In some individuals it occurs frequently or even continuously and leads to considerable distress and difficulty sleeping. There is little knowledge of the molecular mechanisms involved in tinnitus which has hindered the development of treatments. Evidence suggests that tinnitus has a heritable component although previous genetic studies have not established specific risk factors. From a total of 172,608 UK Biobank participants who answered questions on tinnitus we performed a case-control genome-wide association study for self-reported tinnitus. Final sample size used in association analysis was $N = 91,424$. Three variants in close proximity to the *RCOR1* gene reached genome wide significance: rs4906228 ($p = 1.7E-08$), rs4900545 ($p = 1.8E-08$) and 14:103042287_CT_C ($p = 3.50E-08$). *RCOR1* encodes REST Corepressor 1, a component of a co-repressor complex involved in repressing neuronal gene expression in non-neuronal cells. Eleven other independent genetic loci reached a suggestive significance threshold of $p < 1E-06$.

Tinnitus, often referred to as “ringing in the ears” has a reported prevalence of 10–15% in the adult population^{1–5} although the diagnosis and definition of tinnitus remains inconsistent. For most individuals, tinnitus is short lived and not unduly problematic but in some individuals it can be long-lasting and frequent leading to considerable distress and anxiety with 0.5% of individuals reporting it severely affects their ability to live a normal life³. Individual differences in the presentation, duration and frequency of these phantom sounds suggests that tinnitus is likely to be a heterogeneous condition representing a range of pathologies. It often manifests as secondary to hearing loss in both permanent hearing loss or temporary loss secondary to recreational or occupational noise exposure^{1–7}. Although most individuals with tinnitus also have a hearing loss, tinnitus can occur in isolation^{2,3}. The relative contributions of the central and peripheral auditory system to tinnitus remain uncertain but evidence suggests that acute temporary tinnitus after noise exposure is likely to reflect cochlear damage while long term chronic tinnitus has been shown to involve re-organisation or disturbance of neurons in the primary auditory cortex owing to a lack of signal from the cochlea^{1,8}. Lack of understanding of the mechanisms involved have meant that current therapies to address tinnitus symptoms are limited to masking by external sound sources or strategies designed to help with the anxiety and stress caused by tinnitus.

Genetic analysis using linkage or genome wide association studies (GWAS) can be a powerful tool to reveal underlying causes in heritable conditions. Evidence that tinnitus has a sizable genetic component has been mixed^{9,10}. There is little evidence to support familial segregation of tinnitus except in rare cases but a recent heritability estimate from more than 70,000 twins in the Swedish Twin Registry suggested that tinnitus heritability at 0.43¹¹. This rises to a relatively high rate of 0.68 when the trait is limited to bilateral tinnitus in men only¹¹, suggesting a higher genetic component in some. Another Swedish study using a large cohort of adoptees also estimated tinnitus to have a similar value of 0.38 for the heritability for tinnitus¹². Whether any genetic component to a tinnitus phenotype is specific or may instead represent a secondary phenotype due to increased

¹Department of Twin Research and Genetic Epidemiology, School of Life Course Sciences, King’s College London, London SE1 7EH, UK. ²UCL Ear Institute, University College London, London WC1X 8EE, UK. ³Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 7JE, UK. ⁴These authors jointly supervised this work: Frances M. K. Williams and Sally J. Dawson. ✉email: frances.williams@kcl.ac.uk; sally.dawson@ucl.ac.uk

Q. Do you get or have you had noises (such as ringing or buzzing) in your head or in one or both ears that lasts for more than 5 min at a time?					
Response	N	All %	% M	% F	Case control definition
Yes, now most or all of the time	10,719	6.37	8.29	4.76	Case
Yes, now a lot of the time	4110	2.44	2.88	2.08	Case
Yes, now some of the time	15,031	8.92	9.41	8.53	–
Yes, not now, but have in the past	18,868	11.21	11.13	11.28	–
Sub-total who answered Yes	48,728	28.95	31.71	26.65	–
No, never	119,600	71.05	68.29	73.36	Control

Table 1. Prevalence of tinnitus in UK Biobank participants. Percentage values correspond to the proportion of participants that selected one of these five responses (N = 168,328 participants). An additional 4280 participants selected either “Do not know” or “Prefer not to answer”. %M and %F is the percentage answers for males and females respectively. Case Control Definition designates if the participants who gave these responses were included in the GWAS as Cases or Controls. Entries for those included in the GWAS are shown in bold.

Q. How much do these noises worry, annoy or upset you when they are at their worst?		
Response	N	%
Severely	1714	3.55
Moderately	8037	16.64
Slightly	23,098	47.82
Not at all	15,455	31.99

Table 2. Self-reported effect of tinnitus on UKBB participants who reported tinnitus (n = 48,728). Percentage values correspond to the proportion of participants that selected one of these four responses. An additional 424 participants selected either “Do not know” or “Prefer not to answer”.

susceptibility to noise induced damage to the auditory system, remains to be clarified. Identification of the genetic variants involved in tinnitus would help reveal the nature of the mechanisms involved in generating tinnitus after hearing loss, a pre-requisite for development of treatments. Previous pilot genome-wide association studies and candidate gene studies for tinnitus^{10,13–18} have lacked sufficient power to establish specific genetic risk factors but the relatively high heritability demonstrates there is potential to use such approaches to reveal the underlying mechanisms^{9,10}. On this basis we performed a genome-wide association study using self-reported tinnitus available from 172,608 UK Biobank (UKBB)¹⁹ volunteers.

Results

Phenotype definition. In the UKBB 172,608 participants responded to the question regarding whether, and how often, they had experienced tinnitus symptoms (Table 1). The 48,728 who had responded yes to experiencing tinnitus answered an additional question regarding how much they were affected by their tinnitus (Table 2). Almost a third of the sample reported experiencing symptoms of tinnitus either at the time of data collection or in the past (Table 1). Of the subset that had experienced or currently experience symptoms, 3.55% reported that symptoms “severely worry, annoy or upset” them when they are at their worst, while a third of these participants’ reported that they are “not at all bothered” by the symptoms. For the genome-wide association study cases were assigned as participants who answered ‘yes’ tinnitus was present a lot, most, or all of the time (N = 14,829); just under 9% of respondents. We chose only these respondents as cases in order to limit our analysis to those individuals who reported frequent tinnitus to improve genetic power since there is evidence of higher heritability in more severe tinnitus subtypes¹¹. Controls were assigned to those who have never experienced tinnitus (N = 119,600), around 71% of the sample. Replies to the question of how much participants were affected by their tinnitus were not used for phenotype definition as there is a greater degree of subjectivity to this experience. For the GWAS, samples were then further selected based on ethnicity and additional quality control measures (see “Methods” for details) which resulted in a final sample size of N = 91,424 for association analysis.

Genome-wide association analysis. A linear mixed-effects model was used to test for association between 9,740,198 SNPs and tinnitus, using BOLT-LMM v.2.2²⁰. Three SNPs exceeded genome-wide significance ($P < 5E-08$) at the same locus on chromosome 14 (Figs. 1 and 2; Suppl. Table 1). An additional 88 SNPs were associated at a suggestive level of $P < 1E-06$ (Suppl. Table 1). Conditional and joint analysis using GCTA-COJO²¹ suggested they represent a single genome-wide significant locus and eleven independent suggestive loci (Table 3)²².

SNP heritability estimates for tinnitus calculated using BOLT-LMM gave $h^2g = 0.105$, $SE = 0.003$. The estimate was then recalculated to the liability scale, using a tinnitus prevalence derived from the cases and controls above

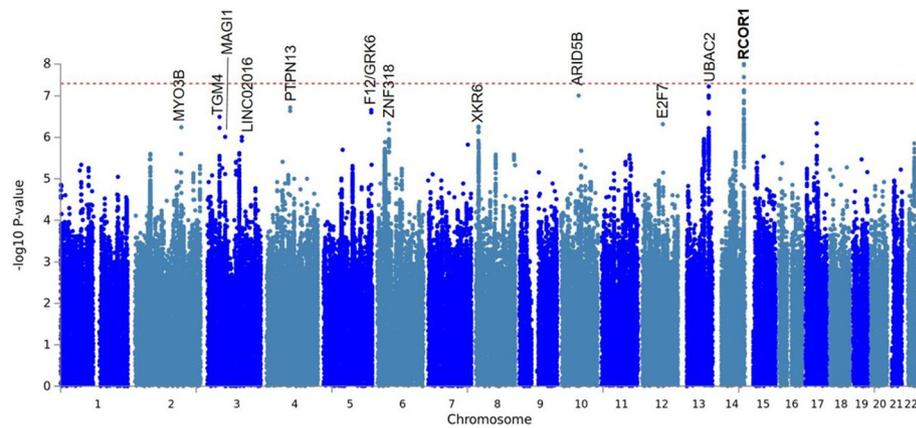


Figure 1. Manhattan plot for genetic association analysis with tinnitus symptoms. The red dotted line marks the genome-wide significance threshold of $P < 5E-08$. Gene loci with p values less than $p < 1 \times 10^{-6}$ are labelled. The genome wide significant locus is shown in bold.

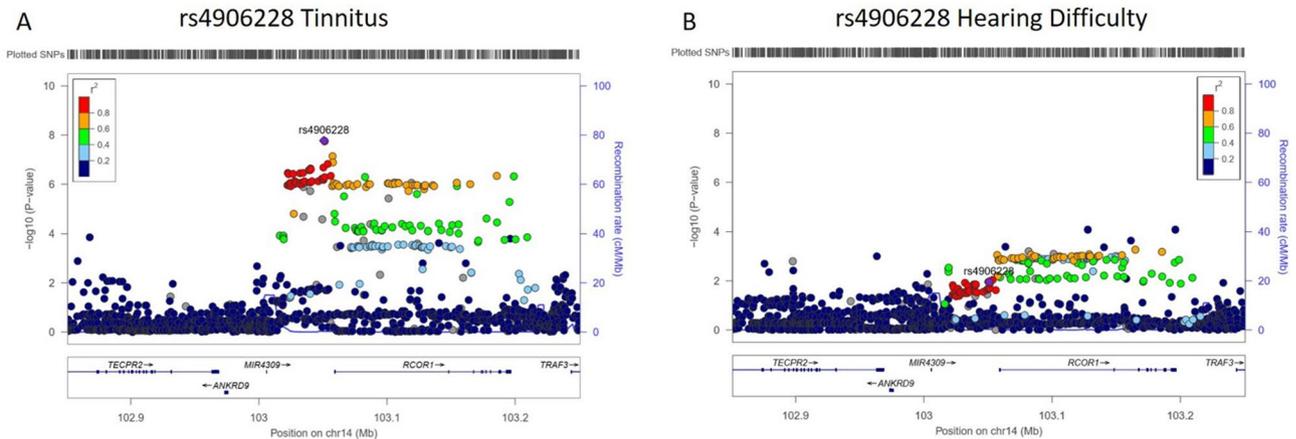


Figure 2. Locus Plots for rs4906228 association with tinnitus and hearing difficulty in UKBB. (A) Locus plot for lead SNP at genome-wide significance on Chr14, rs4906228 in genetic association analysis with frequency of tinnitus symptoms. Purple indicates lead independent SNP generated from GCTA-COJO conditional analysis. The colouring of remaining SNPs is based on the linkage disequilibrium (r^2) with the leads. Where LD information is not available, SNPs are coloured grey. The genes within the region are annotated and the direction of transcription is indicated by arrows. (B) Same plot is shown for association of rs4906228 with self-reported hearing difficulty as described previously²².

of 0.125 to give a heritability $h^2g = 0.30$. The LD score regression intercept for the tinnitus association analysis was 1.0003, suggestive of a negligible type 1 error rate inflation due to population stratification or confounding. The ratio $(\text{intercept}-1)/(\text{mean}(\chi^2)-1)$, was 0.0019, which represents the proportion of inflation in the χ^2 statistic attributable to alternative explanations than polygenicity.

The Variant Effect Predictor (VEP)²³ was used to map independent lead SNPs to the nearest protein coding genes, using the GRCh37 genomic reference. Summary statistics and gene annotation for the lead SNP at each of these regions are presented in Table 3. The genome-wide significant lead SNP rs4906228 lies upstream of the *RCOR1* gene which encodes REST Corepressor 1, a component of a transcriptional repressor complex which represses neuronal gene expression in non-neuronal cells²⁴. Having found a genome wide association in the two most severe tinnitus groups, we tested whether rs4906228 is also associated with tinnitus in the two least severe tinnitus groups (Table 1) and detected no association ($p = 0.29$). The rs4906228 SNP is multiallelic, incorporating a common A > C change ($\text{maf} = 0.24$) and an A > T change reported at the same position although this has only been reported at very low frequency (T: 7.16740e-05 from Ensembl). Table 3 shows the direction of association is that the reference allele A is associated with tinnitus. Of the 12 independent loci having $p < 1E-06$ in the GWAS, eight variants lie within gene introns and three more are positioned upstream or downstream of protein coding genes. Only variant rs72960531 does not lie within or proximal to a protein coding gene, this variant is within intron 3 of a long non-coding RNA (lncRNA), *LINC02016*.

Two of the 12 gene loci were also associated with self-reported hearing difficulty in our previous GWAS in UKBB; *ZNF318* and *ARID5B*²². In order to investigate whether tinnitus loci were associated secondary to a role in causing hearing loss we compared the Manhattan Plots and preformed regional heritability analysis for the

Chr	SNP	refA	Afreq	Beta	se	p	pJ	Gene	Distance	Context	CADD	GERP	PhyoP	Global	Max
14	rs4906228	A	0.24	0.08	0.014	1.70E-08	1.74E-08	RCOR1	8234	Upstream	4.159	- 1.60	0.5132	0.39	0.49
13	rs7336872	T	0.68	0.072	0.013	5.90E-08	5.94E-08	UBAC2	0	Intron	2.490	0.20	0.2213	0.84	1.00
10	rs118053011	C	0.93	- 0.127	0.024	9.80E-08	9.89E-08	ARID5B ^a	6901	Downstream	2.588	0.64	0.5602	0.98	1.00
4	rs113655471	T	0.97	- 0.188	0.036	1.90E-07	1.87E-07	PTPN13	0	Intron	2.531	- 0.79	0.2737	0.99	1.00
5	rs17876046	G	0.98	- 0.288	0.056	2.20E-07	2.18E-07	F12/GRK6	0	Intron	1.698	0.58	0.1065	0.99	1.00
3	rs1532898	A	0.71	- 0.069	0.013	3.20E-07	3.24E-07	TGM4	0	Intron	4.941	0.22	0.2722	0.57	0.88
6	rs553448379	T	0.38	- 0.063	0.013	4.60E-07	4.58E-07	ZNF318 ^a	10,571	Downstream	0.315	0.15	0.0325	n/a	n/a
12	rs7314493	A	0.69	- 0.067	0.013	4.80E-07	4.84E-07	E2F7	0	Intron	11.83	1.14	0.3912	0.76	0.93
8	rs4370496	G	0.73	- 0.07	0.014	5.50E-07	5.48E-07	XKR6	0	Intron	0.213	- 0.15	- 0.6709	0.80	0.99
2	2:171146084_CTT_C	CTT	0.99	- 0.283	0.057	5.70E-07	5.69E-07	MYO3B	0	Intron	n/a	n/a	n/a	0.99	1.00
3	rs557511691	C	0.64	0.065	0.013	9.70E-07	9.77E-07	MAG11	0	Intron	n/a	n/a	n/a	0.67	0.83
3	rs72960531	C	0.94	0.125	0.026	9.80E-07	9.80E-07	LINC02016 ^b	-	Intergenic	14.03	2.24	1.3932	0.95	1.00

Table 3. Summary statistics for lead SNPs that reached a significance threshold of $p < 1E-06$, ordered by p value. Bold denotes genome-wide significance at $p < 5E-08$. Chr, chromosome; SNP, lead SNP; bp, base position; refA, reference allele; beta, effect size from BOLT-LMM approximation to infinitesimal mixed model; se, standard error of the beta; p-value, non-infinitesimal mixed model association test p-value, pJ, joint p-value for variation at loci calculated with GCTA-COJO; Gene, denotes the closest protein coding gene to the variant; Distance, is the distance from the gene (b.p.); Context denotes the consequence of the variant identified by VEP. ^aDenotes genes also identified in previous UKBB GWAS of self-reported hearing²². ^brs72960531 is in an intergenic region without protein coding genes but is within an intron of the lincRNA *LINC02016*; Global, global effect allele frequency reported on Ensembl; Max, highest effect allele frequency reported in any group on Ensembl; n/a, data not available.

two traits (Figs. 3 and 4). No shared significant loci were identified in the regional heritability analysis and other than at the *ZNF318* and *ARID5B* loci there is little overlap between the two Manhattan Plots. Furthermore, examination of the individual locus plot for rs4906228 upstream of *RCOR1* in the hearing difficulty UKBB GWAS shows little evidence of an association with hearing loss (Fig. 2) suggesting this association is not purely a secondary effect of an association with hearing loss. Locus plots for each of the 11 suggestive loci are displayed in Supplementary Figure 1²².

Pathway and tissue enrichment analysis. In order to identify any properties or pathways common to the different tinnitus loci identified in the GWAS we undertook functional gene set enrichment analysis and annotation with genes mapped from SNPs associated at a suggestive level ($P < 1E-06$) (Supplementary Table 2). Although a number of processes and pathways involved in auditory function were enriched at nominal significance, including cytoskeletal protein binding and regulation of anion transport, none of these are significant after Bonferroni correction for multiple testing. One gene set enrichment finding which is significant after multiple testing correction is the presence of binding sites for the transcription factor hepatocyte nuclear factor 4 (HNF4) near six of the genes, *LMAN2*, *EXOSC7*, *ZDHHC3*, *SLC34A1*, *PDLIM7* and *F12* (V\$HNF4_DR1_Q3, Bonferroni corrected p value = 1.53E-03, Suppl. Table 2). None of the SNPs in these genes that are associated with tinnitus lie directly within HNF4 binding sites.

Genetic correlation analysis. We used the linkage disequilibrium score (LDSC) regression to quantify any shared genetic variance between the genome-wide common genetic variants that influence tinnitus with other traits. LD Hub is a centralised database of summary-level GWAS results and interface for LDSC regression analysis including SNP heritability and genetic correlation²⁵. We performed genetic correlation analysis between tinnitus and the 773 available traits on LD Hub. After removing 2 traits that were used to create the tinnitus phenotype used in this study, 61 traits were significantly correlated with tinnitus (Supplementary Table 3). Of these, “Hearing Difficulty/Yes” and “Hearing Difficulty/problems with background noise” were the most strongly correlated ($rg = 0.51$, $p = 5.36E-37$ and $rg = 0.43$, $p = 6.31E-26$ respectively) confirming a strong genetic component shared between tinnitus and hearing loss. Other highly correlated traits (see Suppl. Table 3) included several involving recent experience of pain ($rg = 0.40$, $p = 3.30E-18$ for “Pain type(s) experienced in last month: Neck or shoulder pain”) and traits evidencing low mood or depression ($rg = 0.32$, $p = 1.74E-16$ for “Fed-up feelings”).

GWAS in most severe tinnitus sub-group. We repeated our analysis this time with the only most severe tinnitus sub-group “Yes, now most or all of the time” against the control group. No genome wide significant associations were detected, 58 SNPs were found at the suggestive level ($p < 1E-06$) at eleven different gene loci (see Suppl. Fig. 2 and Supplementary Table 4). These included one SNP at the *RCOR1* locus 14:103042287_CT_C, although the lead SNP in the earlier GWAS rs4906228 did not quite reach suggestive level of significance ($p = 2.8E-06$).

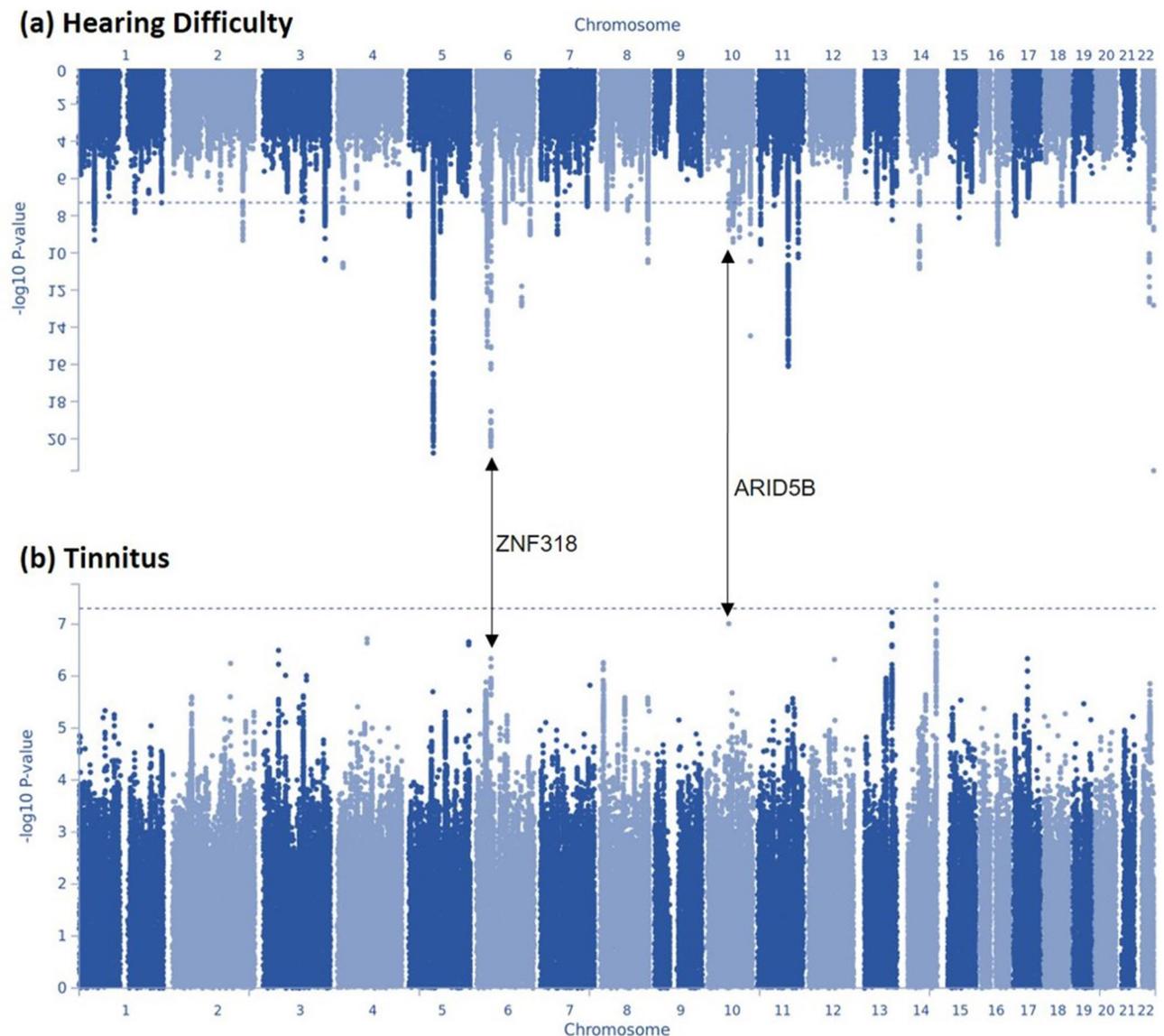


Figure 3. Manhattan plots displaying Hearing and Tinnitus GWAS results. (a) Displays GWAS results for the hearing difficulty phenotype defined by Wells et al.²² in the UK Biobank cohort, (b) displays GWAS results for the GWAS presented in this manuscript, with a tinnitus phenotype in the UK Biobank cohort. The Manhattan plots display the p values of all SNPs tested in discovery analyses. The threshold for genome wide significance ($p < 5 \times 10^{-8}$) is indicated by a dotted line. Loci common to both GWAS are annotated.

Discussion

We have investigated genetic risk factors for a self-reported tinnitus by performing a GWAS in the UKBB with a sample size over 90,000 individuals and identified 3 genome wide significant variant associations at the same locus. The lead SNP at this locus, rs4906228, is just over 8 kilobases upstream of the *RCOR1* gene in what Variant Effect Predictor²³ describes as a regulatory region/promoter flanking region. Examination of the locus plot around rs4906228 (Fig. 2) shows that the association with tinnitus, although strongest just upstream of the gene, spans the *RCOR1* coding region with 56 variants with $p < E-06$ over a 176 kilobase region along the length of the gene (Suppl. Table 1). *RCOR1* potential role in tinnitus is interesting: while not previously linked to hearing or tinnitus it is a co-factor of RE1 Silencing Transcription Factor (REST) with which it forms a transcriptional repressor complex known to downregulate the expression of neuronal genes in non-neuronal cells through histone de-acetylases (HDACs)²⁴. Dysregulation of the REST complex has been implicated in neurodegenerative disease including Alzheimer's Disease²⁵. A splicing mutation in *REST* has been reported as causing a progressive, non-syndromic, sensorineural hearing loss, DFNA27 in a North American family^{27,28}. Recently, Nakano et al.²⁷ established the mechanism underlying deafness in this family. The sensory receptor cells in the inner ear *hair cells* inactivate REST using a cell specific splicing mechanism thereby allowing neuronal gene expression in hair cells. The DFNA27 mutation disrupts this alternative splicing mechanism resulting in repression of neuronal gene expression in hair cells, leading to hair cell death and deafness²⁷. This suggests that there are distinct regulatory

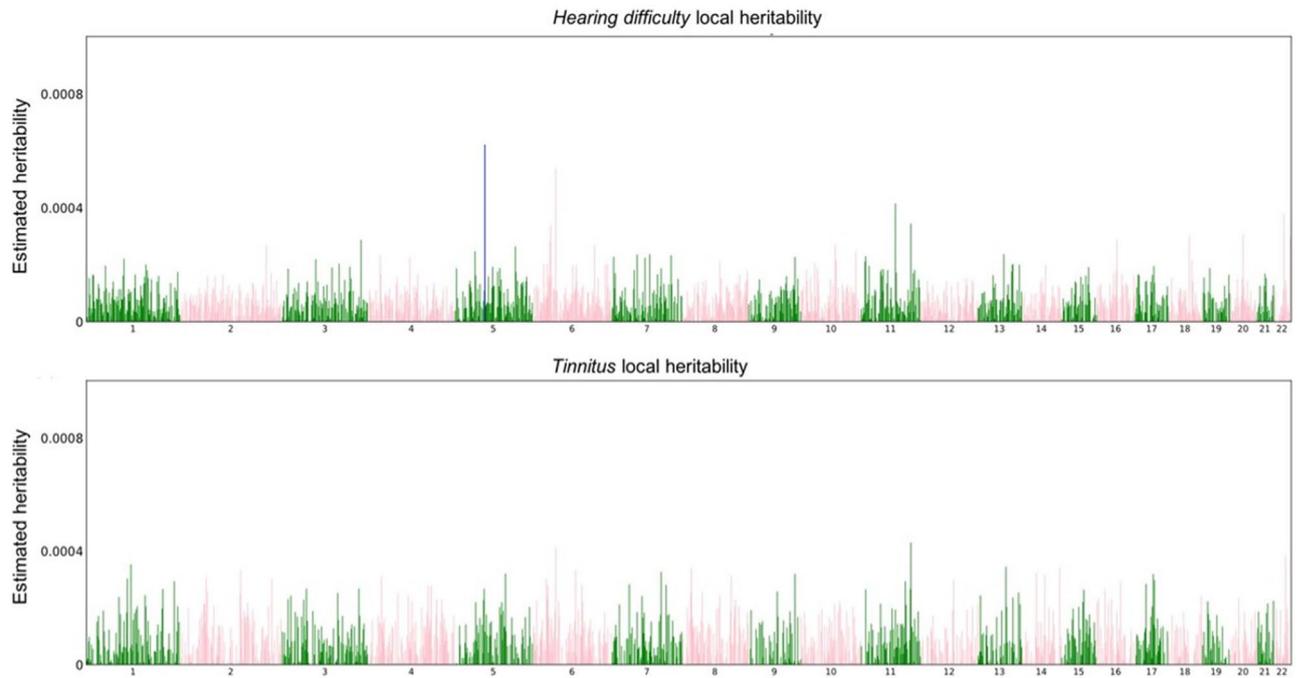


Figure 4. Manhattan-Style plots of regional heritability across the genome for hearing difficulty (top), and tinnitus (bottom). Estimated local SNP heritability for 1702 loci. Blue denotes an estimate of local $h^2_g = 0.0006$, $P < 1.157E-05$ on Chr 5 for hearing difficulty in the region between base positions 71240456–73759326. HESS total SNP heritability estimates were 0.1, SE = 0.005 for hearing difficulty and 0.0853, SE = 0.013 for tinnitus.

mechanisms and requirements for the RCOR1-REST repressor complex in both hair cells of the inner ear and neurons of the brain, both potential pathogenic sites for a tinnitus phenotype. We examined whether there was evidence that the association of *RCOR1* with tinnitus might be secondary to its effect on hearing by analysing any association with self-reported hearing difficulty in the UKBB cohort previous GWAS²². As Fig. 2 shows there are no associations at $P < 5E-04$ at this locus with hearing difficulty suggesting it is unlikely that the association with tinnitus is due to an increase in genetic risk of hearing loss.

Eleven other independent loci were associated with tinnitus at $P < 1E-6$, providing a candidate gene list for further analysis. The association with *UBAC2* is just below the genome-wide significance threshold ($p = 5.90E-08$) and lies in intron 2 of the gene. *UBAC2* encodes Ubiquitin-Associated Domain-Containing Protein 2 although there are a number of other genes that lie within this LD block including two G-protein coupled receptors *GPR18*, *GPR183*, and *FKSG9* which encodes Gasdermin-A, as well as 2 microRNAs (Suppl. Figure 1). Three of these 11 suggestive associations are in or close to genes previously linked to hearing. The 2 base deletion 2:171146084_CTT_C is within an intron of *MYO3B*, encoding myosin IIIb, one of two myosin III isoforms that are responsible for the organisation and elongation of hair cell stereocilia that are critical for the detection of sound. Mutations in *MYO3A* underlie a recessive form of human hearing loss, DFNB30²⁹. Furthermore, *Myo3a*^{-/-}*Myo3b*^{-/-} double knockout mice are profoundly deaf³⁰. Two suggestive associations are downstream of genes that were previously identified as being associated with self-reported hearing difficulty in the UKBB cohort²², *ARID5B* and *ZNF318* although the lead SNPs associated with the two traits are different in both cases (see Table 3 and Fig. 3). Both *ARID5B* and *ZNF318* encode transcription factors; AT-Rich Interaction Domain 5B and zinc finger protein 318. *ARID5B* has a role in regulating lipid metabolism, *ZNF318* is largely uncharacterized. It is therefore possible that the association of *MYO3B*, *ARID5B* and *ZNF318* with tinnitus is secondary to their role in hearing since tinnitus is usually manifested when there is a hearing loss present. However, it may be that the nature of the hearing loss caused by these genes variants creates a deficit which particularly potentiates the generation of tinnitus.

In order to investigate whether there are common functional properties or processes within genes associated with tinnitus that might pinpoint overlapping pathogenic mechanisms we performed gene set enrichment analysis. No enrichment of genes in specific molecular functions, biological processes, pathways or phenotypes was detected. However, HNF4 transcription factor binding sites were found to be significantly enriched within the gene set. HNF4 is a nuclear receptor and transcription factor, known to be critical in the regulation of hepatocyte development, nutrient transport and metabolism but it is also expressed outside the liver and thought to play a role in regulating gene expression relating to drug metabolism, lipid metabolism, cell proliferation, and inflammation³¹. Potential mechanisms connecting to tinnitus remain unknown but it is notable that *Hnf4a*^{tm1b(EUCOMM)Hmgu} knockout mice generated by the International Mouse Phenotyping Consortium are reported to have significantly reduced Acoustic Startle and Pre-pulse Inhibition response (PPI) (<https://www.mousephenotype.org/data/genes/MGI:109128>; accessed March 2020). Reduced PPI is often used as a surrogate indicator of tinnitus in animal behaviour experiments, although the validity of this has been questioned³². The

Current study										Clifford et al. Table 1				
SNP	CHR	BP	A1	A2	A1 freq	INFO	BETA	SE	P	A1 freq	INFO	BETA	SE	P
rs143424888	1	103456996	C	CACGTGATCT	0.508	0.973	-0.006	0.001351	9.00E-06	0.510	0.973	-0.023	0.004	8.50E-09
rs17249745	4	102547366	A	G	0.968	0.897	-0.018	0.003997	1.00E-05	0.967	0.897	-0.072	0.012	1.30E-09
rs553448379	6	43288656	T	TA	0.384	0.989	-0.007	0.001378	4.60E-07	0.386	0.989	-0.023	0.004	2.10E-08
rs11249981	8	10147398	C	T	0.449	0.990	-0.007	0.001346	1.20E-06	0.447	0.990	-0.022	0.004	4.60E-08
rs72815660	10	106614698	A	G	0.961	0.986	-0.014	0.003467	4.10E-05	0.960	0.986	-0.060	0.010	9.70E-09
rs11174489	12	62852271	G	A	0.626	0.995	-0.006	0.001384	1.10E-05	0.625	0.995	-0.024	0.004	3.80E-09

Table 4. A comparison of the summary statistics for six SNPs in the current study that were reported as genome wide significant in the study by Clifford et al. Bold denotes those SNPs that were replicated in the Clifford et al. study. Chr, chromosome; SNP, single nucleotide polymorphism; bp, base position; A1, allele A1; A2 allele A2; A1 freq, allele frequency of allele A1; Info, estimated imputation info score; beta, effect size from BOLT-LMM approximation to infinitesimal mixed model; se, standard error of the beta; p-value, non-infinitesimal mixed model association test p-value. NB. In Table 1 of Clifford et al. in the printed paper the effect sign is opposite for rs143424888, which appears to be an error; see their Supplementary eTable 5.

lack of gene enrichment within the tinnitus GWAS gene set may well indicate the heterogeneity of tinnitus with multiple pathogenic mechanisms implicated.

Our study has identified a number of interesting candidate genes for further investigation, most notably *RCOR1*, but these require further replication and validation. Very recently, during preparation of this manuscript, Clifford et al. 2020 published a similar analysis of tinnitus in the UK Biobank cohort but with little apparent overlap in the strongest genetic associations reported in the two studies³³. Clifford et al. reported genome-wide associations with 6 independent SNPs with 3 of these SNPs replicating in a second Million Veteran Program cohort. We therefore investigated the 6 genome-wide significant SNPs reported by Clifford et al. 2020 in our current analysis (Table 4). All 6 SNPs show suggestive level associations $p < 5E-05$ in the same direction as that reported by Clifford et al.³³. This suggests that there is some consistency between the two studies and that differences in the top hits are most likely due to differences in the phenotype definition of tinnitus used. Here, we used a case-control approach using only the two most severe categories of self-reported tinnitus, representing those individuals that have ongoing persistent tinnitus. Clifford et al. used a different approach using an ordinal definition of tinnitus frequency, scoring from 1 to 5 based on the severity of tinnitus which also includes people who have tinnitus sometimes or in the past. This again suggest some heterogeneity in the pathology of tinnitus with distinct pathologies and sub-types. Our top associations here might be those that are involved in a more severe and persistent subtype than those detected by Clifford et al. 2020. Further analysis in cohorts with more in-depth phenotyping which are not reliant on self-reported data are needed to refine these tinnitus subtypes and their distinct genetic susceptibilities. Our findings here are limited by lack of replication group with comparable power in which to confirm these associations. The availability of such cohorts is an important barrier to further progress in tinnitus research. This is particularly pertinent to tinnitus because validation and characterisation in mutant mouse models is compromised by the lack of an established biomarker for the presence of tinnitus in animals. Tinnitus as defined in this study, that is *frequent* tinnitus (“a lot, most or all of the time”) has a heritability h^2 of 0.30 which, although lower than some recent estimates in twin studies^{11,12} is still a sufficiently high genetic component to be tractable in future tinnitus GWAS if sufficiently large cohorts having the relevant phenotyping can be identified.

Methods

Participants and phenotype definition. The sample used for this study consisted of individuals who participated in the UKBB study. The UKBB is a national resource, initially set up to study lifestyle and genetic factors affecting ageing traits with the aim of understanding and improving healthy ageing at a population level. Over 500,000 volunteers attended 23 assessment centres across the UK between 2007 and 2013 where they donated samples for genotyping, completed lifestyle questionnaires and have standard measurements taken. All subjects provided informed consent to participate in UK Biobank. Ethical approval was granted by the North West Multi-Centre Ethics committee. All experiments were performed in accordance with these guidelines and regulations, under the UKBB ethics governance and framework, <https://www.ukbiobank.ac.uk/ethics/>. The UKBB resource is described extensively elsewhere¹⁹.

Two questions regarding tinnitus were included in the UKBB ‘Health and medical history’ questionnaire that participants completed on touchscreen monitors while attending an assessment centre. The first question relates to symptoms of tinnitus and the second relates to the severity/nuisance of said symptoms. The first question is “Do you get or have you had noises (such as ringing or buzzing) in your head or in one or both ears that lasts for more than 5 min at a time?” to which participants could select one of the following responses: “Yes, now most or all of the time”, “Yes, now a lot of the time”, “Yes, now some of the time”, “Yes, but not now, but have in the past”, “No, never”, “Do not know”, “Prefer not to answer”. Participants responded to the second question; “How much do these noises worry, annoy or upset you when they are at their worst?” with either “Severely”, “Moderately”, “Slightly”, “Not at all”, “Do not know”, “Prefer not to answer”. The work in this study used data collected during the main recruitment phase only.

Phenotype definition. To derive a phenotype for association analysis, study participants were categorized using a case–control design based on responses to the question “Do you get or have you had noises (such as ringing or buzzing in your head or in one or both ears that lasts for more than 5 min at a time?” Participants that responded either “Yes, now most or all of the time” or “Yes, now a lot of the time” were assigned ‘cases’ and those that responded ‘No’ were assigned controls, described in Table 1. Participants that selected any of the additional responses were not included in the analysis. The cohort used for association analysis consisted of UKBB participants with ‘White British’ ancestry. The UKBB sample classification ‘White British’ is derived from both principal component (PC) analysis and self-declared ethnicity³⁴. Samples with excess heterozygosity, excess relatedness and sex discrepancies were identified and removed prior to analysis, resulting in a sample size of $n = 91,424$.

Genotyping and imputation. Two arrays were used to genotype the ~500,000 UK Biobank samples; 50,000 samples were genotyped on the Affymetrix UK BiLEVE Axiom array and ~450,000 samples were genotyped on the Affymetrix UK Biobank Axiom array. The two arrays shared 95% of the >800,000 SNP genotype coverage. UKBB performed imputation centrally using the HRC reference panel and IMPUTE2³⁵. Further SNPs that did not feature on this panel were imputed with the UK 10K and 1000G panel. Analysis presented here was conducted with version 3 of the UKBB imputed data, following QC performed centrally by UKBB, 487,409 samples were imputed and available for analysis.

Association analysis. Genetic association analysis was performed using BOLT-LMM v2.2 that uses a linear mixed-effects model approach to test association between SNP dosages and the tinnitus phenotype. BOLT-LMM corrects for population stratification and within-sample relatedness. Additional adjustments were made for age, sex, UKBB genotyping platform and UKBB genetic principal components 1–10. Additional filters included the implementation of a minor allele frequency threshold of 0.01 and INFO score >0.7 and removal of samples with a genotype call of <98%.

Conditional and joint analysis. Conditional and joint SNP analysis was performed to identify independent signals within highly associated regions, using GCTA-COJO²¹. This analysis requires the linkage disequilibrium reference sample, which was obtained by random selection of 10,000 individuals from the UKBB cohort with White British ancestry. The reference sample size was selected to maximise power based on previous data simulations²¹. The distance assumed for complete linkage equilibrium was 10 Mb and a cut off value of $R^2 = 0.9$ was used to check for collinearity between the selected SNPs and those to be tested. Alleles with a frequency difference >0.2 between the reference sample and GWAS sample were excluded. Independent SNPs identified with GCTA-COJO were mapped to the nearest protein coding gene using variant effect predictor (VEP)²³, genome build GRCh37. VEP was used to establish whether the SNP was in an exonic, intronic or intergenic region, and also the functional consequence of the variant at that position.

Heritability estimates. BOLT-LMM was used to calculate heritability (h^2_g) for the tinnitus phenotype as $h^2_g = 0.105$, $SE = 0.003$. The estimate was then recalculated to the liability scale, using a case prevalence of 0.125. A region-based heritability estimation from summary statistics (HESS) was used on tinnitus and hearing difficulty trait from Wells et al.²² to partition genome-wide SNP heritability estimates into 1702 approximately independent loci. Partitioned heritability values for tinnitus, $h^2_g = 0.105$ and hearing difficulty, $h^2_g = 0.35$ were compared to investigate any loci sharing heritability percentage.

Pathway analysis. Pathway-based analysis in GWAS is being widely used to discover multi-gene functional associations. Many gene set enrichment tools have been developed to test the enrichment of the associated genes in pathways, and ToppGene Suite³⁶ is one of the sophisticated and easy to use tools suitable for this purpose. It comprises of many features such as pathway, Gene ontology, human and mouse phenotypes, diseases, drugs, etc. to be included in enrichment analysis and it reported the enrichment p-values at both uncorrected and corrected levels. Enrichment of features at nominal level is useful for smaller number of genes input and for exploratory analysis. We prepared the gene list by mapping the lead SNPs from the Tinnitus GWAS to the nearest genes within 100 kb of the gene transcription start/end site using Variant Effect Predictor as a positional mapping tool²³. These were then entered into the web-interface ToppGene Suite tool as the gene list. The resulting enrichment features and values were analysed and evaluated for any interesting pathway involved with Tinnitus GWAS genes.

LD score regression. The relationship between the test statistic and LD was studied (via univariate linkage disequilibrium score regression, LDSC), to calculate whether inflated test statistics are likely due to confounding bias, or the polygenic nature of the trait.

Genetic correlation analysis. LD Hub is a centralised database of summary-level GWAS results and interface for LDSC regression analysis including SNP heritability and genetic correlation²⁵. Here, we used LD Hub to analyse genetic correlations between tinnitus and all of the 771 other traits available from LD Hub. We set a multiple testing significance threshold of $p < 6.5E-5$ ($0.05/771$)²⁵.

References

1. Baguley, D., McFerran, D. & Hall, D. Tinnitus. *Lancet* **382**, 1600–1607. [https://doi.org/10.1016/S0140-6736\(13\)60142-7](https://doi.org/10.1016/S0140-6736(13)60142-7) (2013).
2. Bhatt, J. M., Lin, H. W. & Bhattacharyya, N. Prevalence, Severity, exposures, and treatment patterns of tinnitus in the United States. *JAMA Otolaryngol. Head Neck Surg.* **142**, 959–965. <https://doi.org/10.1001/jamaoto.2016.1700> (2016).
3. Davis, A. & Refaie, A. E. *Epidemiology of Tinnitus* 1–23 (Thomson Learning, 2000).
4. Folmer, R. L., McMillan, G. P., Austin, D. F. & Henry, J. A. Audiometric thresholds and prevalence of tinnitus among male veterans in the United States: Data from the National Health and Nutrition Examination Survey, 1999–2006. *J. Rehabil. Res. Dev.* **48**, 503–516. <https://doi.org/10.1682/jrrd.2010.07.0138> (2011).
5. Mahboubi, H., Oliaei, S., Kiumehr, S., Dwabe, S. & Djalilian, H. R. The prevalence and characteristics of tinnitus in the youth population of the United States. *Laryngoscope* **123**, 2001–2008. <https://doi.org/10.1002/lary.24015> (2013).
6. Dawes, P., Cruickshanks, K. J., Marsden, A., Moore, D. R. & Munro, K. J. Relationship between diet, tinnitus, and hearing difficulties. *Ear Hear* <https://doi.org/10.1097/AUD.0000000000000765> (2019).
7. Haider, H. F. *et al.* Biomarkers of presbycusis and tinnitus in a portuguese older population. *Front. Aging Neurosci.* **9**, 346. <https://doi.org/10.3389/fnagi.2017.00346> (2017).
8. Knipper, M., Van Dijk, P., Nunes, I., Ruttiger, L. & Zimmermann, U. Advances in the neurobiology of hearing disorders: Recent developments regarding the basis of tinnitus and hyperacusis. *Prog. Neurobiol.* **111**, 17–33. <https://doi.org/10.1016/j.pneurobio.2013.08.002> (2013).
9. Cederroth, C. R., Kahler, A. K., Sullivan, P. F. & Lopez-Escamez, J. A. Genetics of tinnitus: Time to biobank phantom sounds. *Front. Genet.* **8**, 110. <https://doi.org/10.3389/fgene.2017.00110> (2017).
10. Vona, B., Nanda, I., Shehata-Dieler, W. & Haaf, T. Genetics of tinnitus: Still in its infancy. *Front. Neurosci.* **11**, 236. <https://doi.org/10.3389/fnins.2017.00236> (2017).
11. Maas, I. L. *et al.* Genetic susceptibility to bilateral tinnitus in a Swedish twin cohort. *Genet. Med.* **19**, 1007–1012. <https://doi.org/10.1038/gim.2017.4> (2017).
12. Cederroth, C. R. *et al.* Association of genetic vs environmental factors in Swedish adoptees with clinically significant tinnitus. *JAMA Otolaryngol. Head Neck Surg.* **145**, 222–229. <https://doi.org/10.1001/jamaoto.2018.3852> (2019).
13. Coskunoglu, A. *et al.* Evidence of associations between brain-derived neurotrophic factor (BDNF) serum levels and gene polymorphisms with tinnitus. *Noise Health* **19**, 140–148. https://doi.org/10.4103/nah.NAH_74_16 (2017).
14. Doi, M. Y. *et al.* Association between polymorphism of interleukin-6 in the region-174G/C and tinnitus in the elderly with a history of occupational noise exposure. *Noise Health* **17**, 406–410. <https://doi.org/10.4103/1463-1741.169703> (2015).
15. Gilles, A., Van Camp, G., Van de Heyning, P. & Franssen, E. A pilot genome-wide association study identifies potential metabolic pathways involved in tinnitus. *Front. Neurosci.* **11**, 71. <https://doi.org/10.3389/fnins.2017.00071> (2017).
16. Marchiori, L. L. M., Dias, A. C. M., Goncalves, A. S., Poly-Frederico, R. C. & Doi, M. Y. Association between polymorphism of tumor necrosis factor alpha (tnfalpha) in the region -308 g/a with tinnitus in the elderly with a history of occupational noise exposure. *Noise Health* **20**, 37–41. https://doi.org/10.4103/nah.NAH_34_17 (2018).
17. Sand, P. G., Luettich, A., Kleinjung, T., Hajak, G. & Langguth, B. An examination of KCNE1 mutations and common variants in chronic tinnitus. *Genes* **1**, 23–37. <https://doi.org/10.3390/genes1010023> (2010).
18. Yuce, S. *et al.* Angiotensin-converting enzyme (ACE) I/D and alpha-adducin (ADD1) G460W gene polymorphisms in Turkish patients with severe chronic tinnitus. *J. Int. Adv. Otol.* **12**, 77–81. <https://doi.org/10.5152/iao.2016.1732> (2016).
19. Sudlow, C. *et al.* UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779. <https://doi.org/10.1371/journal.pmed.1001779> (2015).
20. Loh, P. R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290. <https://doi.org/10.1038/ng.3190> (2015).
21. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**(369–375), S361–363. <https://doi.org/10.1038/ng.2213> (2012).
22. Wells, H. R. R. *et al.* GWAS identifies 44 independent associated genomic loci for self-reported adult hearing difficulty in UK Biobank. *Am. J. Hum. Genet.* **105**, 788–802. <https://doi.org/10.1016/j.ajhg.2019.09.008> (2019).
23. McLaren, W. *et al.* The ensembl variant effect predictor. *Genome Biol.* **17**, 122. <https://doi.org/10.1186/s13059-016-0974-4> (2016).
24. Coulson, J. M. Transcriptional regulation: Cancer, neurons and the REST. *Curr. Biol.* **15**, R665–668. <https://doi.org/10.1016/j.cub.2005.08.032> (2005).
25. Zheng, J. *et al.* LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279. <https://doi.org/10.1093/bioinformatics/btw613> (2017).
26. Hwang, J. Y. & Zukin, R. S. REST, a master transcriptional regulator in neurodegenerative disease. *Curr. Opin. Neurobiol.* **48**, 193–200. <https://doi.org/10.1016/j.conb.2017.12.008> (2018).
27. Nakano, Y. *et al.* Defects in the alternative splicing-dependent regulation of REST cause deafness. *Cell* **174**, 536–548.e521. <https://doi.org/10.1016/j.cell.2018.06.004> (2018).
28. Peters, L. M. *et al.* A locus for autosomal dominant progressive non-syndromic hearing loss, DFNA27, is on chromosome 4q12–13.1. *Clin. Genet.* **73**, 367–372. <https://doi.org/10.1111/j.1399-0004.2008.00966.x> (2008).
29. Walsh, T. *et al.* From flies' eyes to our ears: Mutations in a human class III myosin cause progressive nonsyndromic hearing loss DFNB30. *Proc. Natl. Acad. Sci. USA* **99**, 7518–7523. <https://doi.org/10.1073/pnas.102091699> (2002).
30. Lelli, A. *et al.* Class III myosins shape the auditory hair bundles by limiting microvilli and stereocilia growth. *J. Cell Biol.* **212**, 231–244. <https://doi.org/10.1083/jcb.201509017> (2016).
31. Lu, H. Crosstalk of HNF4alpha with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids. *Acta Pharm. Sin. B* **6**, 393–408. <https://doi.org/10.1016/j.apsb.2016.07.003> (2016).
32. Galazyuk, A. & Hebert, S. Gap-prepulse inhibition of the acoustic startle reflex (GPIAS) for tinnitus assessment: Current status and future directions. *Front. Neurol.* **6**, 88. <https://doi.org/10.3389/fneur.2015.00088> (2015).
33. Clifford, R. E., Maihofer, A. X., Stein, M. B., Ryan, A. F. & Nievergelt, C. M. Novel risk loci in tinnitus and causal inference with neuropsychiatric disorders among adults of European ancestry. *JAMA Otolaryngol. Head Neck Surg.* <https://doi.org/10.1001/jamaoto.2020.2920> (2020).
34. Bycroft, C. *et al.* Genome-wide genetic data on ~ 500,000 UK Biobank participants. *bioRxiv* (2017).
35. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529. <https://doi.org/10.1371/journal.pgen.1000529> (2009).
36. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–311. <https://doi.org/10.1093/nar/gkp427> (2009).

Acknowledgements

The research was carried out using the UK Biobank Resource under application number 11516. HRRW was funded by a PhD Studentship Grant, S44, from Action on Hearing Loss. The study was also supported by funding from NIHR UCLH BRC Deafness and Hearing Problems Theme. We would like to thank all the participants of UK Biobank.

Author contributions

S.D. and F.W. conceived this study. H.W., F.Z.A. and M.F. conducted statistical and bioinformatics analyses. S.D. and H.W. drafted the manuscript, all authors contributed to, edited and revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-85871-6>.

Correspondence and requests for materials should be addressed to F.M.K.W. or S.J.D.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021