Doctoral program in Biomedicine (B11.56.1)

Genetic Contribution to Chronic Tinnitus in Patients with Meniere's Disease and Tinnitus Extreme Phenotype (MD-TEP)

Contribución genética al acúfeno crónico en pacientes de enfermedad de Meniere con fenotipo extremo de acúfeno (MD-TEP)

UNIVERSIDAD DE GRANADA

International PhD Thesis
Tesis Doctoral con Mencion International

Sana Amanat Ali
May, 2021

Granada

Editor: Universidad de Granada. Tesis Doctorales
Autor: Sana Amanat Ali
ISBN: 978-84-1306-904-3
URI: http://hdl.handle.net/10481/69097

Acknowledgements

It was not possible to carry out this thesis without the help and collaboration of many people. Thanks to my family especially my parents and sisters who supported my decision of moving abroad to pursue my PhD. Thanks for your constant support, I will never forget. Off course, I can't. This decision seems to have paid off now.

Decisions and coincidences that changed my life, Undergraduation in Computer science, Master in Computer science and then Doctorate in Biomedicine as a Bioinformatician.

I would like to take this opportunity to thanks Antonio for yout constant support, my supervisor. Thank you for trusting me and giving me the opportunity to be a part of this prestigious group. I was not very clear when I started but you have made sure that it was as straight as possible. Having little or no knowledge of exome sequencing, you taught me very well from basics to advance topics. It is time to thank all the colleagues with whom I have been lucky enough to work in GENyO. I would like to thanks Teresa Requena and Lidia Frejo for your kind support; though we spent very few months together but I will never forget. I am grateful to Álvaro Gallego for your help whenever i have questions and doubts. You helped me a lot to build my understanding about biological interpretation of data. I would like to thanks Pablo for his constant help. I would like to thanks Olga, you have been a great support for me in all administrative tasks related to my contract, even before my arrival. Thanks for churros con chocolate and funny moments we shared. Thank you very much Marisa, for making my stay exciting and cheerful. Thanks for the good and quality time we spent. I still remember few of Portuguese words, thanks for it. Thank you Estrella for the good time we spent, I will remember your name as Star. Alba, thanks for your help whenever I needed, you were always willing to help me. I would like to thanks all Master students for the discussion we had.

It is time to say a big thanks to everyone from ESIT. My ESIT colleagues and friends have been a great support for me throughout this journey. Thanks for the amazing trainings, workshops and courses. It was the best time. Thanks to Christopher for organizing my online stay at Karolinska Institute, and arranging amazing activities during this stay. I would like to extend my thanks to Meniere society, UK for hosting my stay.

Many thanks to Juanma, Patricia and all clinicians for your participation in this study. Thanks for providing us the samples and clinical data to made this study possible. I would like to thanks Patrick and Joseph for your amazing contribution in this study. I would like to say thanks to all the patients and their families for their support. I hope that together we can find the solution one day.

I don't want to forget anyone, Thanks to all those who are in my heart and silently supporting me. Thank to everyone even if I haven't mentioned the name.

Thanks to Ammi, Abu, Nida, Farah, Ghazala, Madeha, Farzana, Saeed, Moid, Zamna, Rehana, Mubarak, Ammar and Tayyaba. Thank a lot.

Contents

Funding 9
Resumen 10
Abstract 11
List of abbreviations 13
1. Introduction 14
1.1. Tinnitus: a complex trait 14
1.2. Inner ear and the peripheral auditory system 15
1.2.1. Anatomy of cochlea and its function 16
1.2.2. Anatomy of vestibular labyrinth and its function 17
1.2.3. The auditory central nervous system and tinnitus mechanism 18
1.3. Meniere disease 20
1.3.1. Vertigo 20
1.3.2. Sensorineural hearing loss 21
1.3.3. Tinnitus 21
1.3.4. Aural fullness 21
1.4. Extreme phenotype 21
1.5. Diagnosis and classification 24
1.6. Genetic variants 25
1.7. Pathogenicity of rare variants 27
1.8. Heritability and genetics contribution to tinnitus 29
2. Hypothesis 31
3. Objectives 32
3.1. Main objective 32
3.2. Specific objectives 32
4. Methods 33
4.1. Systematic review of extreme phenotype strategies 33
4.1.1. Search strategy 33
4.1.2. Research Question and Selection Criteria 33
4.1.3. Quality assessment of selected studies 34
4.1.4. Data Extraction and Synthesis 35
4.2. Subjects and definition of phenotype 35
4.3. Whole-exome sequencing 36
4.4. Bioinformatics analysis 36

4.5. Gene Ontology and gene-set enrichment analysis 39
4.6. Gene expression analysis using the Allen Brain Atlas 39
4.7. Statistical analysis 40
5. Results 41
5.1. Selection and Characteristics of EP Studies 41
5.2. Synthesized Findings of EP Studies 42
5.3. Rare missense variants in synaptic genes associated with tinnitus extreme phenotype in MD patients 42
5.3.1. Indel analysis 48
5.4. Replication of *ANK2*, *AKAP9* and *TSC2* in a Swedish tinnitus cohort with severe tinnitus 53
5.5. Rare missense variants in hearing loss genes associated with tinnitus extreme phenotype in Meniere disease 58
5.6. Gene ontology and Gene-set enrichment analysis in patients with tinnitus 58
5.7. *ANK2* gene expression profile in the mouse brain 60
5.8. Hypothesis free data-driven approach with tinnitus extreme phenotype in Meniere disease 63
5.8.1. Main genes found in the hypothesis-free driven analysis 64
5.8.2. Gene ontology and Gene-set enrichment analysis in patients with tinnitus 64
6. Discussion 67
6.1. Extreme phenotype strategies to uncover rare variants 67
6.1.1. Summary of main findings from systematic review 67
6.1.2. How to select EP in quantitative traits 67
6.1.3. Familial disorder as EP 68
6.1.4. An EP Strategy to Investigate the Genetic Contribution to Tinnitus 69
6.2. Synaptic genes involved in tinnitus 69
6.2.1. Association of *ANK2* and axon initial segment with tinnitus severity 70
6.2.2. Contribution of *AKAP9* and *TSC2* in tinnitus 73
6.2.3. Severe tinnitus and sex differences 75
6.2.4. Tinnitus, hearing loss and cognitive decline 75
6.3. Limitations 77
Conclusions 78
Conclusiones 79
References 80
Supplementary Material 90
Original Research Articles 128

Funding

Sana Amanat was supported by H2020 MSCA-ITN-2016–722046 as PhD student in the Biomedicine Program in Universidad de Granada. This study has been funded by European School of Interdisciplinary Tinnitus (ESIT), H2020 MSCA-ITN-2016–722046 (SA and JALE), the H2020-SC1-2019-848261 (JALE), and the GNP-182 GENDER-Net Co-Plus Fund (JALE). The project leading to these results has received funding from “la Caixa” Foundation (ID 100010434), under agreement LCF/PR/DE18/52010002 (JALE).

Resumen

Introducción: El acúfeno es la percepción de ruido en ausencia de una estimulación acústica externa que afecta a más del 15% de la población adulta. El acúfeno severo está presente en el 1% de la población general. Los estudios genéticos realizados en gemelos, adoptados y familias avalan una heredabilidad significativa. Varios estudios que seleccionaron genes candidatos no han mostrado hallazgos consistentes y la evidencia que respalda una contribución genética al acúfeno es débil y enfatiza la necesidad de seleccionar el fenotipo de acúfeno apropiado.

Objetivos: Identificar los genes implicados en el desarrollo y mantenimiento del acúfeno severo mediante la selección de individuos con fenotipo extremo para acúfeno mediante secuenciación del exoma y análisis de carga génica. La identificación de posibles vías y procesos biológicos se llevará a cabo mediante análisis de ontología genética y análisis de enriquecimiento de conjuntos de genes.

Métodos: Los exomas de 59 pacientes con enfermedad de Meniere (EM) y fenotipo extremo para acúfenos, y 32 pacientes con EM sin acúfeno grave de ascendencia española fueron secuenciados como control interno. Para la cohorte de replicación, se secuenciaron 97 pacientes suecos con acúfeno severo y se reclutó un tercer conjunto de datos de 701 pacientes con epilepsia genética generalizada de ascendencia europea. La frecuencia de alelos menores de las variantes se comparó con conjuntos de datos de referencia españoles y europeos no finlandeses. Se realizó un análisis de ontología genética y un enriquecimiento de conjuntos de genes para identificar las vías y los procesos biológicos.

Resultados: Encontramos un enriquecimiento de variantes raras sin sentido en 24 genes sinápticos en la cohorte española, siendo los más significativos *PRUNE2, AKAP9, SORBS1, ITGAX, ANK2, KIF20B* y *TSC2,* cuando se compararon con conjuntos de datos de referencia. Esta carga se replicó para el gen *ANK2* en la cohorte sueca con 97 individuos con acúfeno y en un subconjunto de 34 pacientes suecos con acúfeno severo para los genes *ANK2, AKAP9* y *TSC2*. Además, en el enfoque basado en datos sin hipótesis, los análisis de carga genética revelaron el gen *ADGRV1* y también confirmaron el exceso de variantes sin sentido para *ANK2* y *TSC2* en pacientes con acúfeno. Sin embargo, estas asociaciones no fueron significativas en la tercera cohorte de 701 individuos con epilepsia generalizada sin acúfeno. La ontología genética (GO) y los análisis de enriquecimiento de conjuntos de genes mostraron varias vías y procesos biológicos involucrados en el acúfeno severo, incluido el transporte a través de membranas y la unión de proteínas del citoesqueleto en las neuronas.

Conclusiones: En esta tesis, se han identificado los genes y las potenciales vías del acúfeno severo, lo cual sugiere que los genes sinápticos y la ramificación inicial del axón juegan un importante papel en el desarrollo y mantenimiento del acúfeno severo. Estos resultados podrían ser beneficiosos para futuros estudios genéticos del acúfeno.

Abstract

Introduction: Tinnitus is the perception of noise in the absence of an external acoustic stimulation affecting more than15% of adult population. Severe tinnitus is observed in 1% of the general population. The genetic studies conducted in twins, adoptees and families support a significant heritability. Several studies selecting candidate genes have shown no consistent findings and the evidence to support a genetic contribution to tinnitus is weak and emphasis the need to select an appropriate tinnitus phenotype.

Objectives: To identify the genes involved in the development and maintenance of severe tinnitus by selecting individuals with extreme phenotype for tinnitus using exome sequencing and gene burden analyses. The identification of potential pathways and biological processes will be carried out by gene ontology and gene-set enrichment analysis.

Methods: Exomes of 59 Meniere disease patients with extreme tinnitus and 32 MD patients without severe tinnitus of Spanish ancestry as internal control were sequenced. For replication cohort, 97 Swedish patients with severe tinnitus were sequenced and a third dataset of 701 patients with genetic generalized epilepsy of European ancestry were included. Minor allele frequency of variants was compared with Non-Finnish European and Spanish reference data sets. Gene ontology and gene-set enrichment was performed to identify the pathways and biological processes.

Results: We found an enrichment of rare missense variants in 24 synaptic genes in a Spanish cohort, the most significant being *PRUNE2, AKAP9, SORBS1, ITGAX, ANK2, KIF20B* and *TSC2*, when they were compared with reference datasets. This burden was replicated for *ANK2* gene in a Swedish cohort with 97 tinnitus individual, and in a subset of 34 Swedish patients with severe tinnitus for *ANK2*, *AKAP9* and *TSC2* genes. In addition, the hypothesis-free data driven approach, gene burden analyses revealed *ADGRV1* gene and also confirmed the excess of missense variants for *ANK2* and *TSC2* with tinnitus. However, these associations were not significant in the third cohort of 701 genetic generalized epilepsy individuals without tinnitus. Gene ontology (GO) and gene-set enrichment analyses showed several pathways and biological processes involved in severe tinnitus, including membrane trafficking and cytoskeletal protein binding in neurons.

Conclusions: In this Thesis, the potential genes and pathways for severe tinnitus have been identified, suggesting the role of synaptic genes and axon initial branching in the development and maintenance of severe tinnitus. These results could be beneficial for future genetic studies of tinnitus.

List of abbreviations

MD	Meniere disease
EP	Extreme phenotype
THI	Tinnitus handicap inventory
TRI	Tinnitus research initiative
SNHL	Sensorineural hearing loss
NGS	Next generation sequencing
TFI	Tinnitus functional index
MD-EP	Meniere disease- tinnitus extreme phenotype
MD-AEP	Meniere disease- tinnitus almost extreme phenotype
WES	Whole exome sequencing
SVA	Single variant analysis
GBA	Gene burden analysis
AIS	Axon initial segment
BMD	Body mass density
GOLD	Global initiative for chronic obstructive lung disease
MAF	Minor allele frequency
AAO-HNS	American academy of otolaryngology-head and neck surgery
SNV	Single nucleotide variants
CNV	Copy number variant
SIFT	Sorting Intolerant from tolerant
CADD	Combined annotation dependent depletion
PolyPhen-2	Polymorphism Phenotyping v2
FATHMM	Functional analysis through hidden Markov models
ACMG	American college of medical genetics
AMP	Association for molecular pathology
EF	Etiological fraction
MVP	Million veteran programs
PICOS	Population, Intervention, Comparison, Outcome, Study design
MeDiC	Meniere disease consortium
STOP	Swedish tinnitus outreach project
BAM	Binary alignment map
VCF	Variant calling format
GO	Gene ontology
SG	Synaptic genes

1. Introduction

1.1. Tinnitus: a complex trait

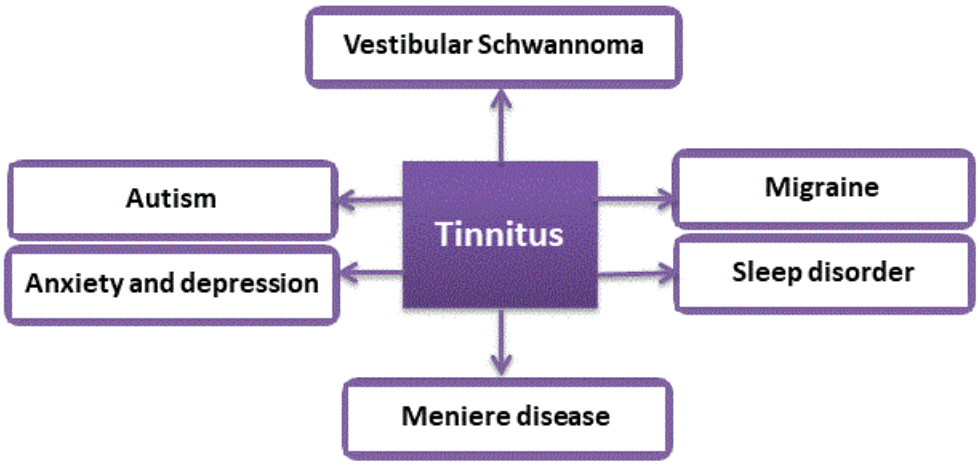
Tinnitus is the perception of auditory noise in the absence of an external acoustic stimulation. It can range from ringing, buzzing, hissing or beeping in the ear. It can be classified into two categories; subjective and objective tinnitus. Objective tinnitus is generated from an internal source such as turbulent flow of blood or spontaneous otoacoustic emissions. However, subjective tinnitus is originated due to abnormal spontaneous activity and only patient can hear it [1,2].

The most common form of tinnitus is subjective tinnitus affecting more than 15% of adult population[1]. Tinnitus can precisely characterized as pulsatile or non-pulsatile, intermittent or persistent, unilateral or bilateral, sporadic or familial[3]. Tinnitus can also described as a distressing disorder in 1% of the population affecting quality of life. Tinnitus in the light of complex disorder; can accompany and overlap with other several disorders including anxiety, sleep disorder, hyperacusis, hearing loss, headache or Meniere disease[4], (Figure 1). As a result, the heterogeneity nature of tinnitus and its wide range of characteristics have made this condition more complex.

To date, there is no consensus on the definition of tinnitus to classify it either as symptom or disorder. The diagnostic criteria of tinnitus have not been established so far due to the lack of operational or standardized definition. The current definitions do not differentiate patients with tinnitus and those experiencing tinnitus along with other comorbidities, functional disability or cognitive dysfunction. Recently, efforts have been made by a group of tinnitus experts under Tinnitus Research Initiative (TRI) consortium and proposed the definitions of tinnitus as a symptom and disorder.

Tinnitus as a symptom defined as the conscious awareness of a tonal or composite noise for which there is no identifiable corresponding external sound source. For tinnitus disorder the proposed definition is, when emotional distress, cognitive dysfunction, and/or autonomic arousal are associated with tinnitus; and leading to behavioural changes and functional disability.[5]

Figure 1: Tinnitus is a symptom observed in several disorders

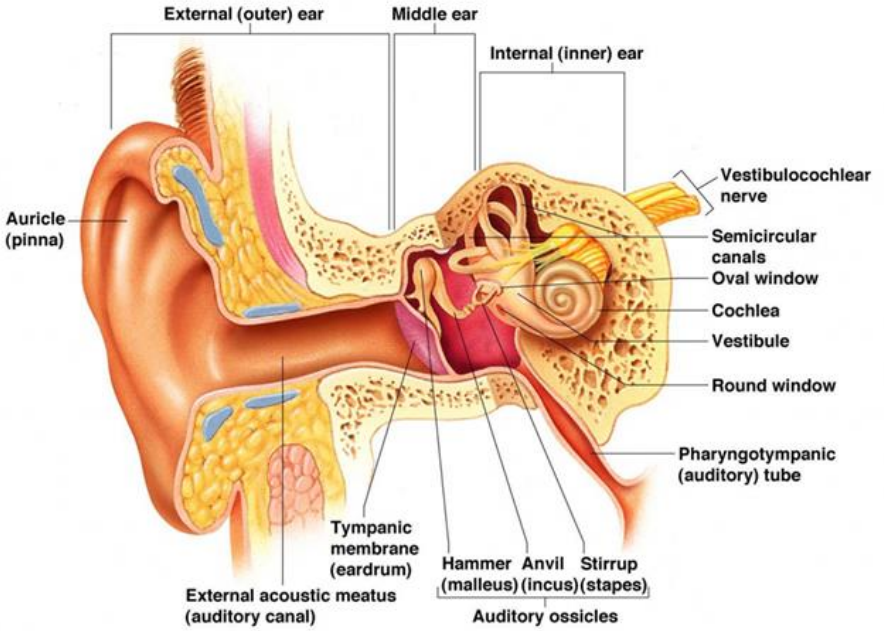


1.2. Inner ear and the peripheral auditory system

Human ear is responsible for hearing and balance functions; with the help of sensory organ which is located in the temporal bone of ear. The ear has three parts divided as outer, middle and inner ear. The external sound waves are particularly collected by the pinna; a part of outer ear. These sound waves then make the tympanic membrane vibrate and produce movements in the chain of three tiny bones, help in transferring these vibrations into the cochlea of the inner ear. The hair cells in the cochlea start to stimulate and generate electrical impulses which are then interpreted by the brain as sound (Figure 2). The detail of each part of the ear is as follows:

1. **Outer ear:** This part of the ear has pinna and external auditory canal; the tympanic membrane separates it from the middle ear.
2. **Middle ear:** composed of ossicular chain of three bones that includes malleus, incus and stapes
3. **Inner ear:** Morphology of inner ear is divided into cochlea or anterior labyrinth (organ of Corti) and posterior labyrinth (semi-circular canals and vestibular end organ)[6].

Figure 2: Anatomy of ear with detail of three basic parts outer, inner and middle ear



From John T. Hansen, David R. Lambert. Netter's Clinic Anatomy. Head & Neck. Elsevier 2014. p. 527-654. doi:10.1016/B978-84-458-1580-9.50008-4.

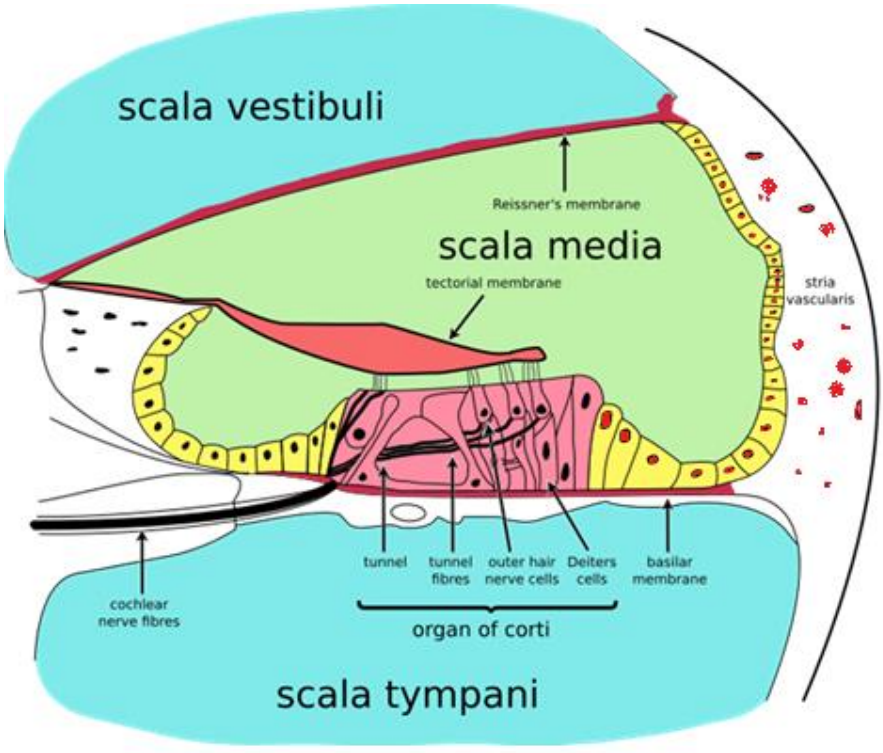
1.2.1. Anatomy of cochlea and its function

The cochlea is a spiral shaped tube with 2.5 turns which is responsible for the detection of sound and its analysis. Cochlea is located in bony labyrinth and has three primary parts as scala vestibule (vestibular duct), scala media (cochlear duct) and scala tympani (tympanic duct), see Figure 3. It is filled with two types of fluids i.e. endolymph and perilymph remains separated in their own chambers. The detail of cochlea sub parts is as follows:

1. **Scala vestibule:** It contains perilymph and reside above cochlea duct
2. **Scala media:** This chamber contains endolymph and organ of Corti. It is highly concentrated with K^+ and NA^+ ion
3. **Scala tympanic:** It is located below the scala media and also contains perilymph

The organ of Corti is a spiral like structure located in scala media and formed by two main types of cells i.e. hair cells and supporting cells (non-sensory). These hair cells are further divided into two categories inner and outer hair cell; responsible for sending and receiving neurotransmitter signals to code auditory information. This organ has very precise organization and pattern of stereocilia with one layer of inner hair and three layers of outer hair cells.

Figure 3: The detail of scala vestibule, scala media and scala tympani



(Under free license from Wikipedia common)

1.2.2. Anatomy of vestibular labyrinth and its function

The vestibular system plays a vital role in order to maintain the right balance. The structure of vestibular labyrinth is distributed into otolith organ (linear acceleration) and the semicircular canals (angular acceleration). The otoliths organs including sacculus (controls gravity and balance perception) and utricle (distinguishes the tilt degrees of a head) are focused in generating a respond to the linear acceleration and the position of a head, respectively. These structures contain macula, the sensory epithelium with hair cells which is covered by otolithic membrane. The otolithic membrane contains the calcium carbonate crystal which is called otoconia. The semicircular canals are composed of three main bony canals including anterior, posterior and the lateral semicircular canals, which are filled with endolymph. All these three canals have the dilation known as the ampulla which contains the sensory cells; responsible for collecting information about the angular acceleration of the head.

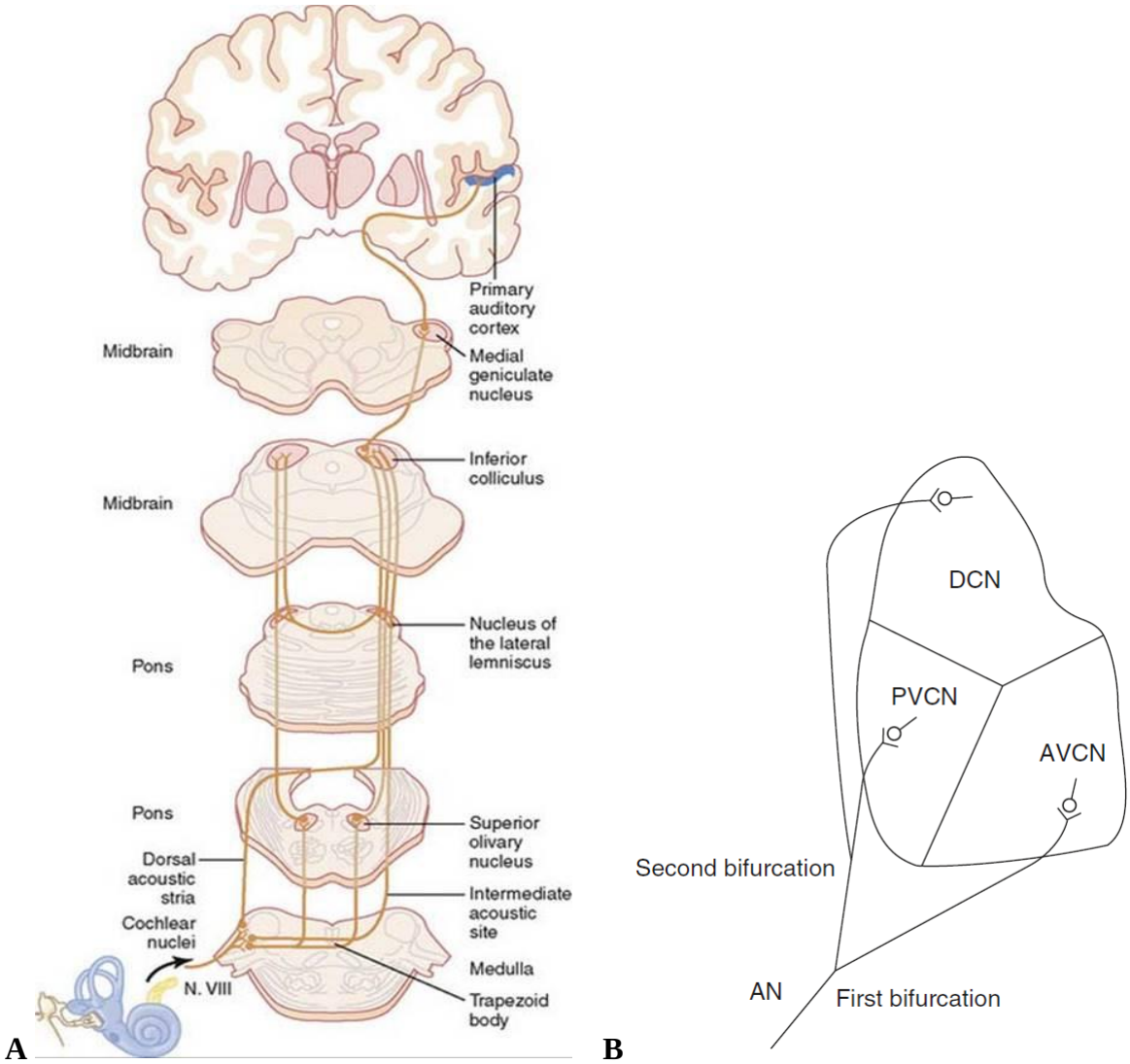
1.2.3. The auditory central nervous system and tinnitus mechanism

The auditory input is transmitted to the brain by either primary auditory pathway which is responsible to carry input from cochlea, or by the non-primary pathway known as reticular sensory pathway carrying sensory input. However, the input from peripheral auditory system reaches the central auditory nuclei by the auditory nerve. The auditory nerve then transmits the auditory input up in a form of series of nuclei to the primary auditory cortex. The series of nuclei includes:

- **Cochlear nuclei (Medulla)**
  - ❖ Anterior ventral cochlear nuclei: responsible for preserving and enhancing the precision of the temporal information in neural firing.
  - ❖ Posterior ventral cochlear nuclei: involves in complex stimulus analysis and in giving rise to dorsal auditory stream of brainstem.
  - ❖ Dorsal cochlear nuclei: It extracts the very complex patterns in the auditory stimulus and contains a range of cell types and several interneurons.
- **Superior olivary nuclei (Pons)**
- **Inferior colliculus (Midbrain)**
- **Medial geniculate nuclei (Midbrain)**

The auditory pathway is short and has 3-4 relays. The first relay occurs in cochlear nuclei in the brainstem which receive the Type I spiral ganglion axons/auditory nerve and decodes the frequency, intensity or the duration of signals. The second relay arises at superior olivary nuclei and the majority of auditory fibres synapse cross the midline and the third level neurons transmit the signals up to the inferior colliculus. These two relays determine the localisation of sound. The last relay occurs at the level of medial geniculate which consist of two types of cells i.e. ventral part and other include medial, dorsal. The thalamus then links to the auditory cortex and integrate the response[6,7], Figure 4.

Figure 4: The major auditory nervous pathways



A) Modified from Brodal A: The auditory system. In Neurological Anatomy in Relation to Clinical Medicine, 3rd ed. New York: Oxford University Press, 1981.

B) Schematic drawing of the cochlear nucleus to show the auditory nerve's connections with the three main divisions and the cochlear nucleus. *DCN* dorsal cochlear nucleus, *PVCN* posterior ventral cochlear nucleus, *AVCN* anterior ventral cochlear nucleus. (Reprinted from Møller, A.R., Sensory Systems: Anatomy and Physiology. 2003, Amsterdam: Academic Press, with permission from Elsevier.

The damage to ear, receptor organs, auditory nerve, or the nerve cells in nuclei of the auditory system can cause tinnitus. Since, tinnitus can occur in some individuals due to hair cells injuries, indicating the possibilities that some other factors are involve in tinnitus. The other possible mechanism of tinnitus generation could be the deprivation of input to the auditory nervous system.

The deprivation of input that occurs due to the certain parts of cochlea can cause tinnitus, providing inhibitory influence on neurons in auditory nervous system. In addition, the noise exposure can cause a change in the morphology of cochlear nucleus and can affect the

auditory nervous system. The abnormal input, overexposure to noise and hearing loss causing the deprivation of auditory input could lead towards the activation of neural plasticity. There is an evidence that the deprivation of auditory input, can result in unused parts being taken over by other systems e.g. unused auditory cortex by visual system[8]. The auditory input deprivation can cause two types of change at functional level: alteration in the balance between excitation and inhibition causing a gain increase in different parts of the auditory system; and it can also activate the neural plasticity. The high frequency sounds exhibit the strong inhibitory influence on neurons in cochlear nucleus as compared to low frequency sounds; indicating high frequency hearing loss can cause tinnitus by reducing normal inhibition occurrence[9,10]. Neural plasticity is considered an important factor in tinnitus generation, which alters the processing of sound and re-routing the signals in central nervous system. Plastic changes can stimulate the coherent firing of several neurons in the auditory pathway and can be a cause of tinnitus. In relation to tinnitus, the reduction in auditory input can activate the neural plasticity[11]. The neural plasticity is an essential part of the neurons in auditory system. Hearing loss can cause a change in the central auditory pathway i.e. tonotopic map reorganization. This change can increase in the firing rates of neurons in the primary auditory cortex; converting into an abnormal synchronous network that generates tinnitus [12].

1.3. Meniere disease

Meniere disease (MD) is a syndrome consisting of a spectrum of rare inner ear disorders characterized by the attacks of vertigo, sensorineural hearing loss (SNHL), tinnitus and sometime aural fullness. The sudden attacks of vertigo are associated with hearing loss and tinnitus. In early stages of disease, tinnitus tends to be more intense before and during the crises. The progression of disease is associated with the disappearance of vertigo, an increase in hearing loss thresholds; leading towards the permanence and an increase in the intensity of tinnitus in many patients [13,14].

1.3.1. Vertigo

Vertigo is one of the most common and initial symptom of MD as reported by most of the patients. It is a sensation of motion and rotation in the absence of real motion in the surroundings or environment, but the patient feels it. The duration of these attacks could last between 20 minutes to several hours; with a significant disturbing impact on health-related quality of life. The average number of vertigo attacks ranged between 6-11 per year and the severity of these attacks may change before turning into mild; but they are unpredictable [15].

1.3.2. Sensorineural hearing loss

SNHL in MD is a common symptom and it can get worse over time. It can be unilateral or bilateral with varying progression of development from weeks to years. The progression of SNHL seems to be slow in most patients, but it can develop from moderate to severe with early onset. To diagnose hearing loss and the progression of MD the serial pure tone audiograms are used; which can be used to differentiate MD from other inner ear disorders.

1.3.3. Tinnitus

Tinnitus is the perception of sound in the absence of an external sound in the environment. There are two types of tinnitus in patients with MD, according to the clinical presentation:

- The tinnitus reported in one ear during the episode of vertigo that usually is a low frequency noise of high loudness occurring immediately before the attack. This symptom is reported by the majority of patients during the attacks. Although the mechanism is unknown, it is hypothesized that it is related to acute changes in the composition of the endolymph.
- The tinnitus reported out of the attacks and this is probably related to the inner ear damage causing a lack of trophic support in the nerve fibres when patients develop a persistent SNHL.

Tinnitus is reported as a most troublesome symptom by many MD patients after several years; since its intensity can increases during vertigo attacks. However, it can occur together with several disorders making it heterogeneous and more complex.

1.3.4. Aural fullness

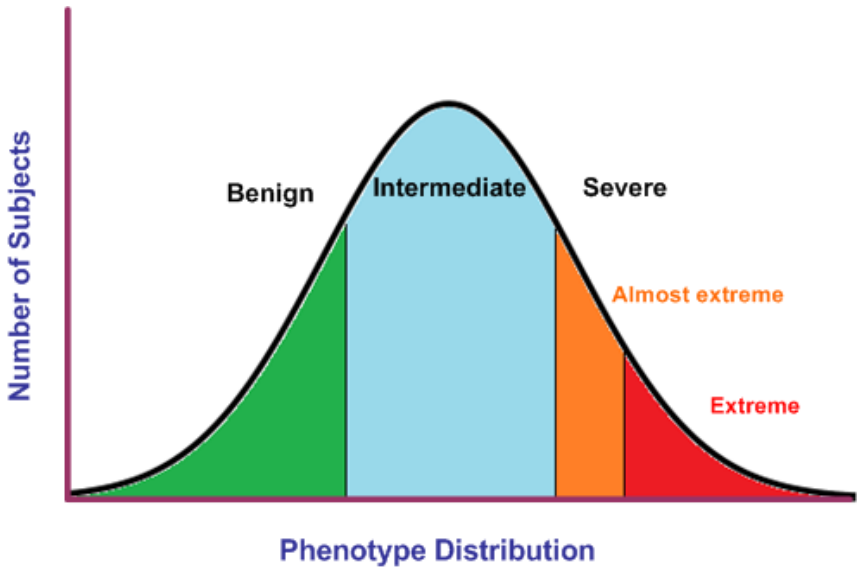
Aural fullness is a sensation of pressure, clog or blockage in the affected ear. Its intensity can fluctuate and become more intense during vertigo episodes. This symptom is explained by the accumulation of endolymph in the cochlear duct in the anterior labyrinth.

1.4. Extreme phenotype

A clinical phenotype is a set of observable signs, symptoms, and behavioral features associated with a human disorder. The phenotype can be described by several features or traits including categorical or quantitative characteristics of patients and disorder under consideration. In Mendelian genetics the variation in observable features of phenotype of a disorder is well known as expressivity and it can range from mild to severe phenotype[16,17,18]. The Phenotype variation in the quantitative traits can be represented by a bell-shaped graph

where mild and severe phenotypes are located at the tails of the distribution. However, the majority of the subjects show an intermediate phenotype (Figure 5).

Figure 5: Phenotypic variation in quantitative traits. Individuals' phenotypes can be classified as benign, intermediate, or severe according to general and disease-specific criteria. Extreme phenotypes are identified at the ends of the normal distribution (green, orange, and red areas).



The genetic architecture of human diseases allows a better understanding of the genetic variants that can influence the phenotype in complex diseases[19]. Next-Generation Sequencing (NGS) technology has been used to uncover missing heritability and to elucidate the genetic contribution to common and rare diseases with underlying heterogeneity. In particular, Whole-Exome Sequencing (WES) provides an opportunity to capture rare and ultra-rare alleles of protein-coding genes, which highly influence disease risk. In the last few years, several novel genes have been identified by utilizing WES for various neurological diseases, such as epileptic encephalopathies (*KCNQ2*, *STXBP1*, and *KCNB1)* and Parkinson's disease (*VPS13C*, *ARSB*, *PTPRH*, *GPATCH2L*, and *UHRF1BP1L*)[20,21,22].

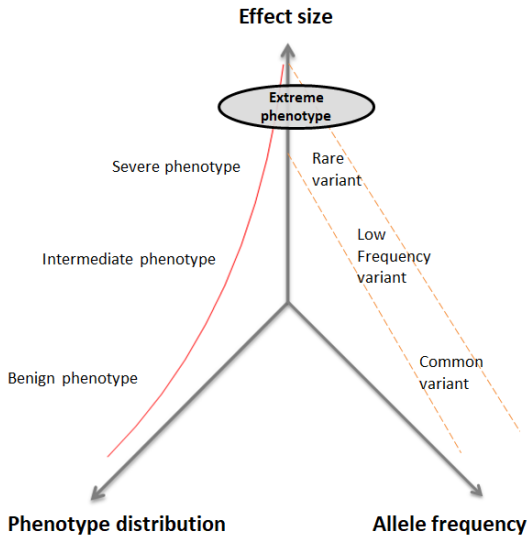
A significant increase in the prevalence of complex diseases has been reported the last decades such as bipolar disorder, coronary artery disease[23], type 2 diabetes, hypertension, obesity and cancer [24]. This increase could be related to environmental factors such as diet or lifestyle changes. However, the genetic contribution to complex conditions is still largely unknown, since the contribution of rare variations to heritability is still undetermined. There are several factors that limit the power of gene discovery approaches, such as phenotypic variance [25], the overlap of clinical features observed for similar conditions, minor allelic

frequency (MAF), the heterogeneous nature of loci, and the low effect size of potential risk alleles [26].

There is a well-established inverse relationship between the allelic frequency of a given variant and its effect size on the phenotype (Figure 6). The underlying hypothesis is that extreme phenotypes (EP) will occur in extreme cases with an excess of rare variants, with a moderate effect size on the phenotype in addition to the effect of the common variants for the trait of interest. The EP strategy aims to identify rare genetic variants causing a large effect on disease risk [27,28].

The EP study design includes the selection of individuals whose phenotypes are at the extreme ends of a disease phenotype distribution. These extreme subjects may be characterized by early or late age of onset, benign or severe forms of disease, family history, fast progression of symptoms, very high or very low scores on psychometric tests or extreme levels of a biomarker [29,30,31.] This strategy may identify rare genetic variants by sequencing a relatively small sample size and it can target novel candidate genes, since rare variants that contribute to a particular trait are enriched at the two extremes of a disease distribution [26]. The combination of EP with WES has successfully identified several rare variants and candidate genes for diabetic retinopathy [32], bipolar disorder [33], and cystic fibrosis [34] across diverse ethnic groups.

Figure 6: Distribution of genetic variants according to allelic frequency and effect size on the phenotype in quantitative traits. Individuals with extreme phenotypes will show a burden of rare variations with a moderate to large effect size (modified from Manolio et al., 2008 [35]).



1.5. Diagnosis and classification

The diagnosis of MD is based on the clinical symptoms reported by the patients during the episodes of vertigo. The heterogeneity nature of MD has made its diagnosis very challenging; since its symptoms may accompany other disorders as well including benign paroxysmal positional vertigo, otosclerosis, vestibular migraine, transient ischemic attacks in vertebrobasilar territory and others. The American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) developed the guidelines for diagnosis and therapy evaluation of MD in 1972 and revised them in 1985 and 1995; these criteria were further revised in 2015 in a consensus panel including representative of the Classification Committee of the Barany Society, the European Academy of Otology and Neurotology, the Korean Balance Society and the Japan Society for Equilibrium Research[36],Table 1.

Table 1: MD Diagnostic criteria according to the Barany Society guidelines (2015)

Symptoms	**Definite MD**	**Probable MD**
Vertigo	2 or more episodes of vertigo during 20 min to 12h	2 or more episodes of vertigo during 20 min to 24h
Hearing loss	Audiometrically documented low-to-medium frequency SNHL on an affected ear during/after one vertigo episode	-
Tinnitus/Aural fullness	Fluctuating aural symptoms	Fluctuating aural symptoms
Other	Not better explained by another vestibular disease	Not better explained by another vestibular disease

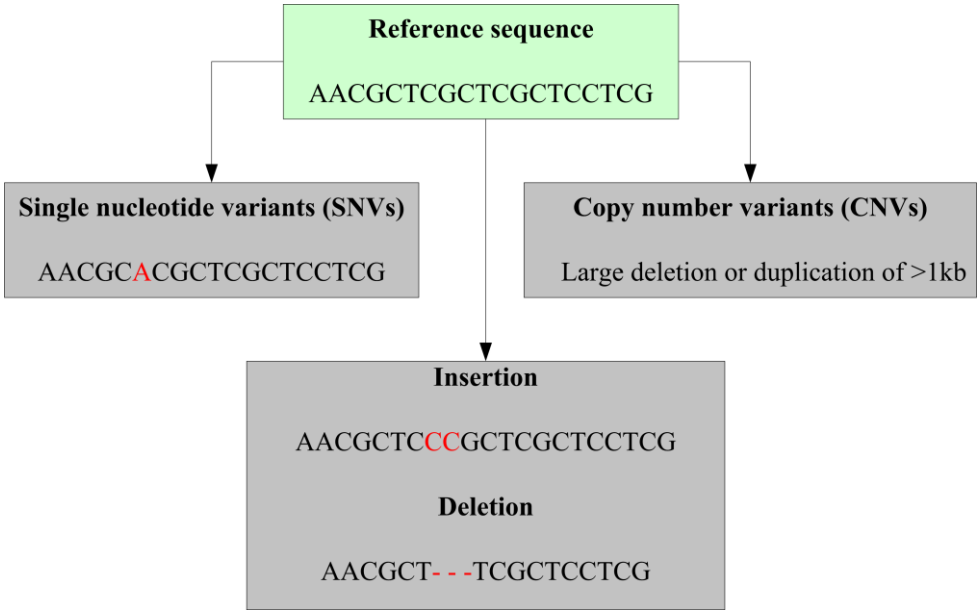
The symptoms of MD may show an overlap with other common or rare disorders making the diagnosis more complex. The overlapping disorders include the autoimmune inner ear disease, delayed hydrops, vestibular migraine and others.

1. Autoimmune inner ear disease is defined by the crisis of sudden to progressive bilateral sensorineural hearing loss.
2. Delayed hydrops is characterized by the suffering from longstanding unilateral profound sensorineural hearing loss.
3. Vestibular migraine is considered a common cause of episodic vertigo with overlapping symptoms with migraine.
4. Transient ischemic attacks of the brainstem, which are observed in the vertebrobasilar territory and may include tinnitus, vertigo, hearing loss and other transient neurological symptoms involving cranial nerves of the pons and medulla.

1.6. Genetic variants

Genetic variant is a well-known term that refers to the change that happens at DNA level as a result of errors occur during DNA replication. The irreversible change in the DNA is considered a primary source of genetic variation; i.e. mutation. It is endorsed that rare mutations located in the coding regions of the genome are more pathogenic and have a large effect in rare disease than common variants. There are several types of variants according to its effect in the nucleotide sequence and the reading frame (Figure 7).

Figure 7: Major types of genetic variants with an example



The detail of major types of genetic variants is as follows:

1. Single nucleotide variants (SNVs)

SNV is a change in the nucleotide sequence in the DNA as compared to the reference sequence on genome. SNVs with MAF lower than 5% are considered rare variants. The further division of SNVs based on functional affect is as follows:

1.1. Coding:

These variants usually affect the protein-coding regions and can cause a change in the resulting amino acid or maybe not.

- **Nonsynonymous**

  These variants change the amino acid of the resulting protein and can be further classified as missense and nonsense variants.

- **Missense:** cause a change in the amino acid
- **Nonsense:** tends to cause premature loss or gain of the stop codon

❖ **Synonymous**

In this type of variant, the resultant codon certainly codes the same amino acid with no change in protein production. These variants are also known as silent mutations or functionally neutral variants. However, some of the synonymous variants can be disease causative or can impact the total amount of the protein by changing the mean life of the mRNA product.

1.2. Non-coding

These variants occur in the non-coding regions of human genome and the functional significance of these variants is probably unknown. These variants include untranslated regions and non-coding RNA variants. Few of them may have regulatory effects in the coding regions by different mechanisms including alternate transcription start sites.

2. Insertion and deletion (INDELS)

Small insertion and deletion are well-known as INDELS with size range of 1 to 10,000 bp. INDELS occur frequently and most of the time are detected with SNVs. The effect on reading frame can subdivide the INDELS into the following categories as:

❖ **Frameshift:** resulting in a shift or change in the reading frame during the translation process and the resultant reading frame is no more divisible three.

3. Structural variants (SVs)

Large alterations on genome of more than 1000bp are termed as structural variants. These variants are not well studied but as equally important as indels with a significant contribution to the disease development. The major categories of structural variants are as follows:

❖ **Copy number variants (CNVs):** The well-studied type of SVs is CNVs, composed of large insertions, deletions or duplications responsible for a great proportion of variation in phenotypes.

❖ **Inversions and other SVs:** In this type of variation, the DNA is reversed with reference to the rest of human genome. These inversions are the root cause of several diseases including Angelman syndrome, Hunter syndrome and others. This type of SVs include translocations or segmental uniparental disomy and not very well-studied.

1.7. Pathogenicity of rare variants

In order to assess the pathogenicity of rare variants, several in silico predictive algorithms have been designed including Sorting Intolerant from Tolerant (SIFT), Combined Annotation Dependent Depletion (CADD), Polymorphism Phenotyping v2 (PolyPhen-2), Functional Analysis through Hidden Markov Models (FATHMM) and others.

In 2015, the American College of Medical Genetics (ACMG) and the Association for Molecular Pathology (AMP) published the guidelines to assess the evidence available to interpret the potential pathogenicity of rare variants for Mendelian disorders in the human genome and facilitate their profile as benign or pathogenic [37], Table 2.

Table 2: Pathogenicity criteria based on evidence according to ACGM guidelines

	Benign		**Pathogenic**			
Evidence	**Strong**	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very strong**
Population data	Minor allele frequency is too high for disorder OR observation in controls inconsistent with disease penetrance			Absent in population databases	Prevalence in affected statistically increased over controls	
Computational and predictive data		Multiple lines of computational evidence suggest no impact on gene /gene product	Multiple lines of computational evidence support a deleterious effect on the gene /gene product	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before. Protein length changing variant	Same amino acid change as an established pathogenic variant	Predicted null variant in a gene where loss-of-function is a known mechanism of disease

Functional data	Well-established functional studies show no deleterious effect		Missense in gene with low rate of benign missense variants and path. Missenses common	Mutational hot spot or well-studied functional domain without benign variation	Well-established functional studies show a deleterious effect	
Segregation data	Non-segregation with disease		Co-segregation with disease in multiple affected family members			
De novo data				*De novo* (without paternity & maternity confirmed)	*De novo* (paternity & maternity confirmed)	
Allelic data		Observed in *trans* with a dominant variant OR in *cis* with a pathogenic variant		For recessive disorders, detected in *trans* with a pathogenic variant		
Other database		Reputable source w/out shared data = benign	Reputable source = pathogenic			
Other data		Found in case with an alternate cause	Patient's phenotype or FH highly specific for gene			

In addition, further studies have proposed the use of the etiological fraction (EF)[38] to assess the pathogenicity of rare variants in gene burden tests; EF provides the quantitative estimates of rare variants being disease causative based on location, gene and variant classification (Table 3).

Table 3: Adaptation of ACMG/AMP guidelines for pathogenicity and etiological fraction

Current ACMG/AMP guidelines			
Pathogenicity			
Supporting	**Moderate**	**Strong**	**Very strong**
Missense in gene with low rate of benign, missense variants and pathogenic missenses common	Mutational hot spot or well-defined functional domain without benign variation		
Proposed adapted ACMG/AMP guidelines			
Pathogenicity			

Supporting	Moderate	Strong	Very strong
Non-truncating variant in gene or protein region with 0.8≤ EF <0.9	Non-truncating variant in gene or protein region with 0.9≤ EF <0.95	Non-truncating variant in gene or protein region with EF ≥0.95	

1.8. Heritability and genetics contribution to tinnitus

In literature, the evidence to support the genetic contribution to tinnitus is limited to few epidemiological and genetic studies, showing conflicting results. One of the tinnitus studies conducted on twins reported high concordance between monozygotic twins with bilateral tinnitus; indicating a genetic inheritance for severe bilateral tinnitus and a high heritability of 0.62 in young women with bilateral tinnitus. A study on Swedish adoptees was conducted to investigate the association of genetic or environmental factors with tinnitus. For this purpose the adoptees, biological parents and adoptive parents were recruited. The results showed that there is no association between the transmission of tinnitus and the shared environmental factors. However, a heritability of 0.32 was reported and suggested the association of tinnitus with genetic factors[39].

In addition, large population based familial studies have also been conducted. For the first large scale study, 198 families across Europe were recruited and the analysis revealed the heritability of 0.06–0.14 for siblings and 0.01–0.07 for parent-offspring; and the heritability of 0.11 in females and males, showed that the contribution of genetic factors in tinnitus is relatively low [3]. In Sweden, a large scale study on familial aggregation was conducted to determine the ratio of recurrence risk among siblings. This study has found that the recurrence risk ratio is significantly higher for severe tinnitus in females and highlighted the significance of tinnitus severity level and sex for future genetic studies[40].

Furthermore, a large scale GWAS on self-reported tinnitus individuals was performed to study the genetic association of tinnitus with neuropsychiatric disorders. The analysis showed 6 significant loci with small effect size, 27 genes and 6% of tinnitus estimated heritability. In addition, 3/6 loci were also replicated in Million Veteran Program (MVP) cohort. MVP cohort was not controlled for noise exposure but was known for tinnitus risk factors. The genetic correlation of tinnitus was also found with neuroticism, insomnia which was consistent with clinical and neuroimaging evidences[41].

There are several genotyping studies investigated the genetic contribution to tinnitus particularly chronic tinnitus. These studies have selected candidate genes for sequencing or genotyping and recruited a very small sample of patients with chronic tinnitus. However, none of the studies so far have investigated the genetic contribution of chronic tinnitus in MD patients. The existing researches on tinnitus have suggested the involvement of several candidate genes including *KCNE1*[42], *SLC6A4* [43], *KCNE3*[44], *GDNF,BDNF, KCTD12*[45,46], *GDNF*[47] and Metabolic pathway[48] but the evidence to support these findings is very weak[3].The main concerns include the selection of patients without an appropriate phenotype (severity level, age of onset, age group), the lack of replication in an independent cohort, and the selection bias of candidate genes, ignoring the multiple interaction of proteins in the complex biological processes.

The recommendations and suggestions to conduct genetic studies on tinnitus in humans are detailed below, according to [4]:

1. Deep phenotyping: an appropriate selection of patients with homogenous tinnitus phenotype and detailed clinical information to control biasness associated with comorbidities/disorders that have been previously associated with chronic tinnitus (i.e. hearing loss, depression, anxiety).
2. Well-defined inclusion and exclusion criteria i.e. age of onset, family history
3. Utilization of standardized tools to measure the severity of tinnitus i.e. Tinnitus Functional Index (TFI), Tinnitus Handicap Inventory (THI)
4. Replication in independent cohorts to validate the genetic association

2. Hypothesis

Tinnitus as a disorder is a complex trait observed in 1% of the general population. Epidemiological studies conducted in twins, adoptees and families support a significant heritability.

The hypothesis of this study is that patients with an extreme tinnitus phenotype (severe tinnitus) will have an enrichment of rare variants with functional effects in certain genes (i.e., synaptic genes). Further, the accumulation of rare and novel variants in these genes will target relevant genes and proteins which could play an important role in the molecular pathophysiology of severe tinnitus.

To demonstrate our hypothesis, we will use exome sequencing data from Spanish patients with MD and severe tinnitus, and compare them with datasets from MD without severe tinnitus (internal control), and Spanish and Non-Finnish European data (external controls).

3. Objectives

3.1. Main objective

The main objective of this Thesis is the identification of genes involved in the development and maintenance of severe tinnitus by selecting individuals with extreme phenotype for tinnitus and using exome sequencing and gene burden analyses.

3.2. Specific objectives

1. To assess the effectiveness of the extreme phenotype strategy in complex disorders to reveal a burden of rare variation in certain genes and define novel genes by a systematic review of genetic studies using extreme phenotype.

2. To identify the main genes associated with severe tinnitus in MD by selecting patients according to THI scores and performing exome sequencing and gene burden analysis.

3. To validate these genes by a replication study using sequencing data in an independent cohort of patients with severe tinnitus without MD.

4. To reveal the potential biological process and biochemical pathways involve in severe tinnitus by gene ontology and gene-set enrichment analyses.

4. Methods

4.1. Systematic review of extreme phenotype strategies

To achieve the first objective, a systematic review of genetic studies in complex diseases was performed and it followed Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines[49] and recommendations from the Human Genome Epidemiology Network (HuGENet) review handbook (https://www.cdc.gov/genomics/hugenet/default.htm).

4.1.1. Search strategy

Literature search for EP strategies was performed on 12 December 2019 using two bibliographic databases (PubMed and Embase). For EP strategies the keywords "phenotypic extreme", "extreme phenotype", "rare variant" and "genetics" were used to formulate the search string. The selected keywords could appear in the title, abstract, text word, author keywords, or MeSH Terms of the articles. The keyword string used for the literature search in PubMed was: ((((("phenotypic extreme"[Title/Abstract] OR "extreme phenotype")[Title/Abstract] AND ("rare variant"[Title/Abstract] OR "genetics")[Title/Abstract])) OR (("phenotypic extreme"[Text Word] OR "extreme phenotype")[Text Word] AND ("rare variant"[Text Word] OR "genetics")[Text Word])) OR (("phenotypic extreme" OR "extreme phenotype") AND ("rare variant" OR "genetics") [MeSH Terms]); that for Embase was: ('phenotypic extreme': ti, ab, kw OR 'extreme phenotype': ti, ab, kw) AND ('rare variant': ti, ab, kw OR 'genetics': ti, ab, kw) AND [2009–2019]/py AND [english]/lim. Records published in the last 10 years, studies in English language, and only human studies were included in the literature search by configuring filters if available, e.g., on PubMed

4.1.2. Research Question and Selection Criteria

The objective of this systematic review was to assess the evidence supporting the design of genetic studies using extreme phenotype strategies to find rare or novel variants or genes involved in complex disorders. According to this hypothesis, the following research question was formulated as: "Are EP strategies useful to establish the genetic contribution in complex diseases?". To answer this question, we followed the "Population, Intervention, Comparison, Outcome, Study design" (PICOS) strategy:

1. **Population:** Patients with a complex disease or condition.

2. **Intervention:** Selection of individuals according to extreme phenotype criteria (i.e., early onset, fast progression of disease, very high or very low scores in psychometric tests, or extreme levels of a biomarker).
3. **Comparison:** Genetic association studies (genome-wide association studies (GWAS), WES, genotyping, Sanger sequencing, or targeted sequencing).
4. **Outcome:** genetic findings reported (rare variants, candidate genes, or pathways associated with the condition of interest).

4.1.3. Quality assessment of selected studies

The extracted records were screened to remove duplicate entries. The title and abstract of all articles were reviewed to exclude reviews, meta-analysis, and irrelevant records (non-genetic studies, pharmacogenomics or clinical studies). The search was conducted primarily for rare variants, but any type of variants were retained and included in this systematic review. After screening, the obtained records were considered for full-text assessment in the next step. To assess the quality of these articles, we formulated 8 questions for EP studies (Table 4). For each question, a positive answer was scored as 1 and a negative answer as 0. Each author classified and rated each record independently of each other. Differences in the scores were discussed to get a final consensus score. If a record achieved ≥60% of the total score, the response to Q8 was "yes", and the reported rare variants have a MAF < 0.05, then the record was selected for synthesis. So, only studies with significant results were included for synthesis. The outcome for each selected study was assessed according to Q8 and the following criteria: if a given study had found any rare or de novo variant, common variant, copy number variants, candidate genes, or pathways for EP subjects, then the major outcome was considered as positive.

Table 4. Criteria used to assess the quality of the selected genetic studies using an extreme phenotype approach.

No.	Question	Answer
Q1	Is there a thorough description of the study design?	Yes/No
Q2	Has the study described the method of sequencing/genotyping?	Yes/No
Q3	Has the study provided information about population ancestry?	Yes/No
Q4	Is there any information on the sex of the selected individuals?	Yes/No
Q5	Is there any information on the age of disease onset?	Yes/No

Q6	Has the study used extreme phenotype criteria for sample recruitment?	Yes/No
Q7	Has the study performed sex-specific analysis for genetic associations?	Yes/No
Q8	Has the study reported statistical analysis and significant genetic findings?	Yes/No

4.1.4. Data Extraction and Synthesis

The following information was extracted from each article selected for synthesis: first author's last name, publication year, disease/disorder name, population ancestry, study design, sequencing method, EP/disease phenotype criteria, sample size for cases, age of disease onset, sex of individuals, MAF, and main genetic findings. Moreover, the phenotype criteria and the main genetic findings for EP were of great interest for synthesis.

4.2. Subjects and definition of phenotype

For the aim 2, individuals were recruited through the Meniere disease Consortium (MeDiC), a non-profit multicentre crowdsourcing network of clinical reference hospitals that has built up a database with clinical information and biological samples to investigate the genetic basis of MD. The diagnosis of patients was performed according to the diagnostic criteria for MD stated by the International Classification Committee for Vestibular Disorders of the Barany society[36]. The Spanish version of the THI questionnaire[50] was used to assess the tinnitus severity and the functional impact of tinnitus on daily life[51]. A total of 59 Spanish patients with chronic and persistent tinnitus were selected from the MeDiC database according to THI score. Diagnosis and psychoacoustic characterization of chronic tinnitus in patients with MD was performed by experienced otoneurologists and has been previously reported [52]. Thirty individuals with THI score ≥76 were classified as Meniere disease-tinnitus extreme phenotype (MD-EP), 29 individuals with THI ≥ 56 and <76 were defined as Meniere disease-tinnitus almost extreme phenotype (MD-AEP). An in-house group of MD patients without persistent tinnitus (N=32) were used as internal controls for this study. The majority of the cases were sporadic (N=48), for the familial cases selected; only one case per family was included to prevent stratification. The clinical information of patients with MD and tinnitus phenotypes is detailed in Supplementary Tables S1 and S2.

To achieve the aim 3, a second independent tinnitus cohort from Sweden was selected as a replication cohort: the Swedish Tinnitus Outreach Project (STOP)[53]. STOP participants originate from the study[54], a population-based cohort. Thus, STOP participants are non-

clinical and representative of the general population. For this study, blood-derived DNA from 97 individuals with tinnitus "as a big problem" [TFI ≥ 48[55]] was used for genome sequencing. A subgroup of 34 individuals with severe tinnitus was also selected according to the THI ≥ 56 (Supplementary Table S3). We also retrieved rare variant summary statistics data from a third cohort of patients with epilepsy, the CoGIE cohort, that consisted of 701 individuals with the diagnosis of Generalised genetic Epilepsy, previously reported[56]. The CoGIE cohort was select as an external control to confirm that the genes associations reported in tinnitus were not observed in a non-related neurological disorder. All cases and controls were of European ancestry.

4.3. Whole-exome sequencing

Whole exome sequencing (WES) was performed on MD-EP, MD-AEP cases and in-house MD controls. DNA was extracted from blood or saliva samples using quality controls as previously described[57]. Exon capture was done with the SureSelectXT Human All Exon V6 (Mb) kit (Agilent), and the sequencing was done using HiSeq 4000 platform (Illumina) or NovaSeq 6000 platform (llumina). Paired-end reads were generated per sample to provide an on-target coverage of 100X minimum, with a total coverage of 10GB/sample in HiSeq4000 and 18GB/sample in Novaseq 6000. Read size was 100bp on HiSeq 4000 sequenced samples and 150 bp on Novase6000 sequenced samples.

4.4. Bioinformatics analysis

Raw reads were stored as FASTQ files for each individual. GATK best practices pipelines were utilized to generate Binary Alignment Map (BAM) and Variant Calling Format (VCF) files from raw unmapped reads[58]. Human reference genome GRCH37/hg19 was used to align the reads with the help of Burrows-Wheeler Aligner (BWA-MEM) algorithm. To filter out low quality single nucleotide variants (SNVs) the recommended hard filter was applied as "QD < 2.0 || MQ < 40.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -6.0". The called variants were further filtered out by an in-house MD control dataset composed of 32 individuals to exclude variants associated with MD. The final list of remaining variants was functionally annotated using KGGSeq suite [59] v1.0 and ANNOVAR tool[60] 2019Nov04. Variants were annotated for predicted effect on protein function; allele frequency in public databases (gnomAD, CSVS and dbSNP); and predicted pathogenicity

with in silico algorithms, including Combined Annotation Dependent Depletion scores (CADD).

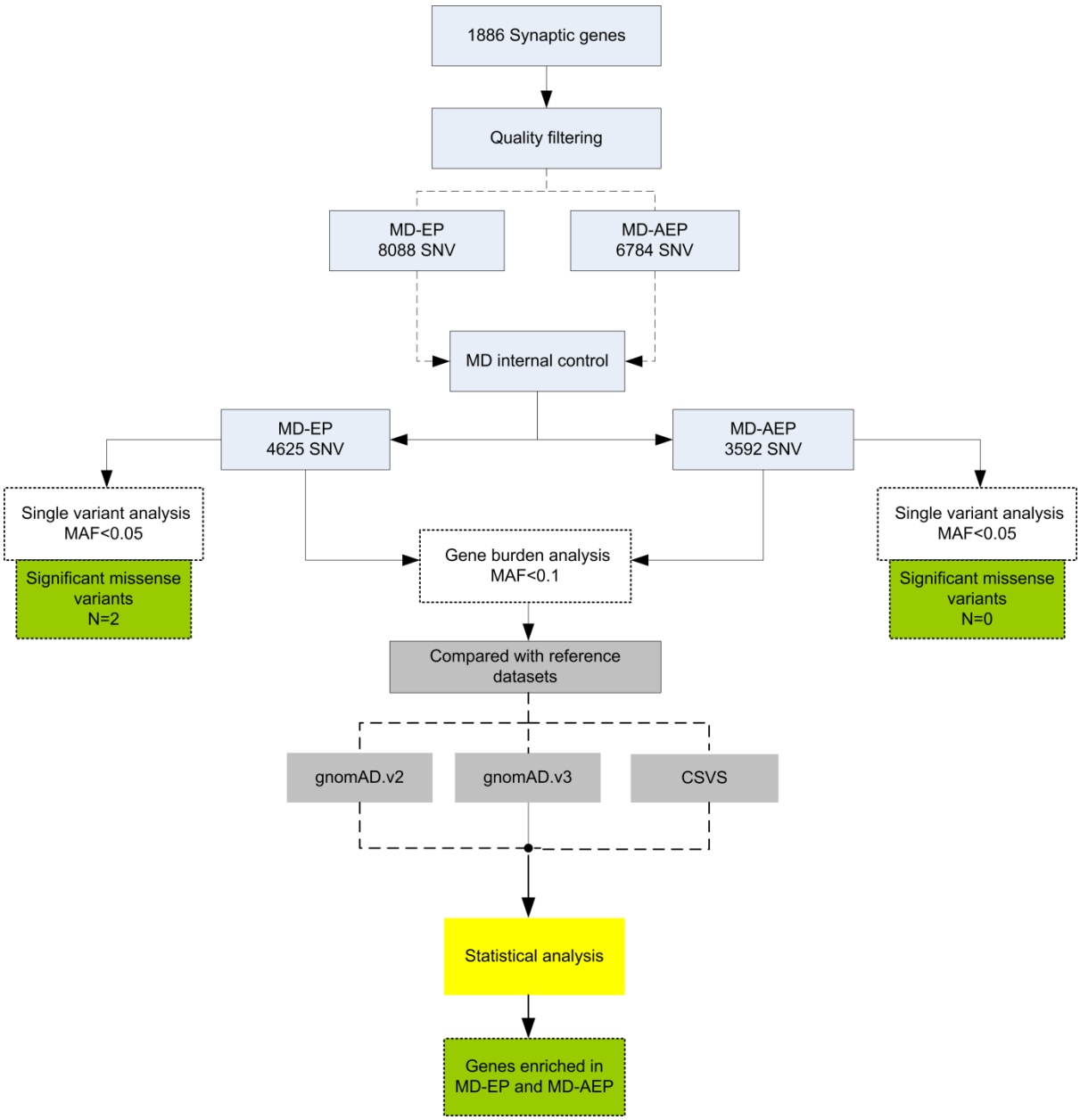
To search for target genes involved in tinnitus, we selected a total of 1886 synaptic genes (SG) from SynaptomeDB[61] (available at http://metamoodics.org/SynaptomeDB/citation.php), which is an ontology-based knowledge base for genes that are encoded in known proteins related with the synapse. The encoded components of synapse include scaffold proteins, membrane transporters, cytoskeletal/adhesion proteins, neurotransmitters and its receptors. Additionally, hearing loss genes (N=152) associated with syndromic and non-syndromic deafness from Deafness Variation Database (DVD) v.8.1(https://deafnessvariationdatabase.org/) were also analysed to separate the potential effect of rare variation in hearing loss genes on tinnitus[62] (Supplementary Table S4). In addition to these two gene sets, total number of all human genes (N=20,000) were also selected for final analysis.

In order to search for variants associated with tinnitus, two types of variant analysis were performed: single variant analysis (SVA) and gene burden analysis (GBA) for MD-EP and MD=AEP (Figure 8). We have used three independent datasets as reference population: Non-Finnish European (NFE) population dataset from gnomAD.v2, NFE from gnomAD.v3[63] and a Spanish dataset from Collaborative Spanish Variant Server (CSVS)[64].

We also called small insertions and deletions (indels) from MD-EP and MD-AEP patients and filtered out by in-house controls. The filtering criteria "QD < 2.0 || ReadPosRankSum < -20.0 || InbreedingCoeff < -0.8 || FS > 200.0 || SOR > 10.0" was applied according to GATK best-practice guidelines. The remaining list of indels was annotated for allele frequency in gnomAD and CSVS.

All variants were assessed and evaluated according to the guidelines provided by ACMG and AMP[37]. The final filtered list of candidate variants was checked through IGV v.2.8.9 using BAM files of variant carriers and further validated by Sanger sequencing. In addition, FLAGS[65] and pseudogenes were excluded to prioritize the genes. Missense variants and their associated amino acid change across protein sequence were represented using Illustrator for biological sequences (IBS)[66].

Figure 8: Flowchart for filtering and prioritization of variants associated with tinnitus in Spanish patients with Meniere disease-tinnitus extreme phenotype (MD-EP). Single variant and gene burden analyses were performed in a set of 1886 synaptic genes selected for EP and almost extreme phenotypes (AEP) for tinnitus. Individuals with MD and no persistent tinnitus were used as an internal control to filter variants associated with MD.



MD-EP= Meniere disease- tinnitus extreme phenotype-, MD-AEP= Meniere disease- tinnitus almost Extreme phenotype, CSVS=Collaborative Spanish Variant Server

4.5. Gene Ontology and gene-set enrichment analysis

For the objective 4, gene ontology (GO) analyses and gene enrichment analyses were performed using GSEA and MsigDB (https://www.gsea-msigdb.org/gsea/index.jsp) as previously described [67]. This tool uses gene sets or gene expression data to predict molecular pathways and biological processes. There are three key elements of the methods used by GSEA as enrichment score, estimation of p-value based on permutation and the calculation of false discovery rate. Two gene lists generated according to the GBA for SNVs and indels including 24 and 31 genes were used to retrieve signalling pathways and biological processes.

4.6. Gene expression analysis using the Allen Brain Atlas

In-situ hybridization (ISH) data in the mouse brain were obtained from the Allen Brain Atlas data set (http://www.brain-map.org); methods for data collection have been described previously[68].Antisense expression data were available in coronal and sagittal sectioned brains for both Ank2 and Tsc2 (4 mice: Ank2, sagittal section = Exp 68844707, coronal section = Exp 71924087 and Tsc2, sagittal section = Exp 70919985, coronal section = Exp 1431). Sagittal sections, for both brains, were visually examined and areas of high expression were noted. These regions were then confirmed via inspection of coronal sections and a good correspondence was found (every highlighted region was confirmed). To quantify these findings we obtained these data in a Matlab format[69].ISH data for 4,104 genes were downloaded as mouse brain-wide expression profiles partitioned into 49,742 cubic voxels of 200 micron size [68,70].In this format *expression energy* of a given gene, *g*, is a weighted sum of the greyscale-intensity of the pixels within a voxel:

$$E(v,g) = \frac{\sum_{p \epsilon v} M(p) I(p)}{\sum_{p \epsilon v} 1}$$

Where *p* denotes a given pixel, *v* a given voxel, *I(p)* the intensity within a given pixel and *M(p)* is a Boolean mask that equals 1 if the gene is expressed at pixel *p* or a 0 otherwise. Coronal data came pre-annotated to allow allocation of each voxel to a given brain region [68,70]. Mean expression energy for each brain region was simply the mean of the expression energy for all voxels annotated to fall within this region, likewise for standard deviations and counts used for calculation of the standard error of mean. The raw annotated voxel data were also transferred into SPSS to allow statistical testing of the variation in expression data. Voxels were treated as independent samples of expression within a given brain region and a

Kruskal-Wallis test performed to determine if expression within different regions differed statistically.

Expression of genes were also contrasted in order determine if visually observed strong correspondence of brain wide expression of Tsc2 and Ank2 was statistically significant. First, for each pair of genes within the set of 4,104 genes a co-expression value was calculated:

$$CoExpr(g, g') = \frac{\sum_{v=1}^{V} E(v, g)E(v, g')}{\sum_{u=1}^{V} E(u, g)^2 \sum_{w=1}^{V} E(w, g')^2}$$

Where *V* is the total number of voxels. The coexpression values from all gene pairs were used to estimate the probability distribution of a given coexpression value being obtained. This distribution was then used to determine if the coexpression of Ank2 and Tsc2 was statistically significant.

4.7. Statistical analysis

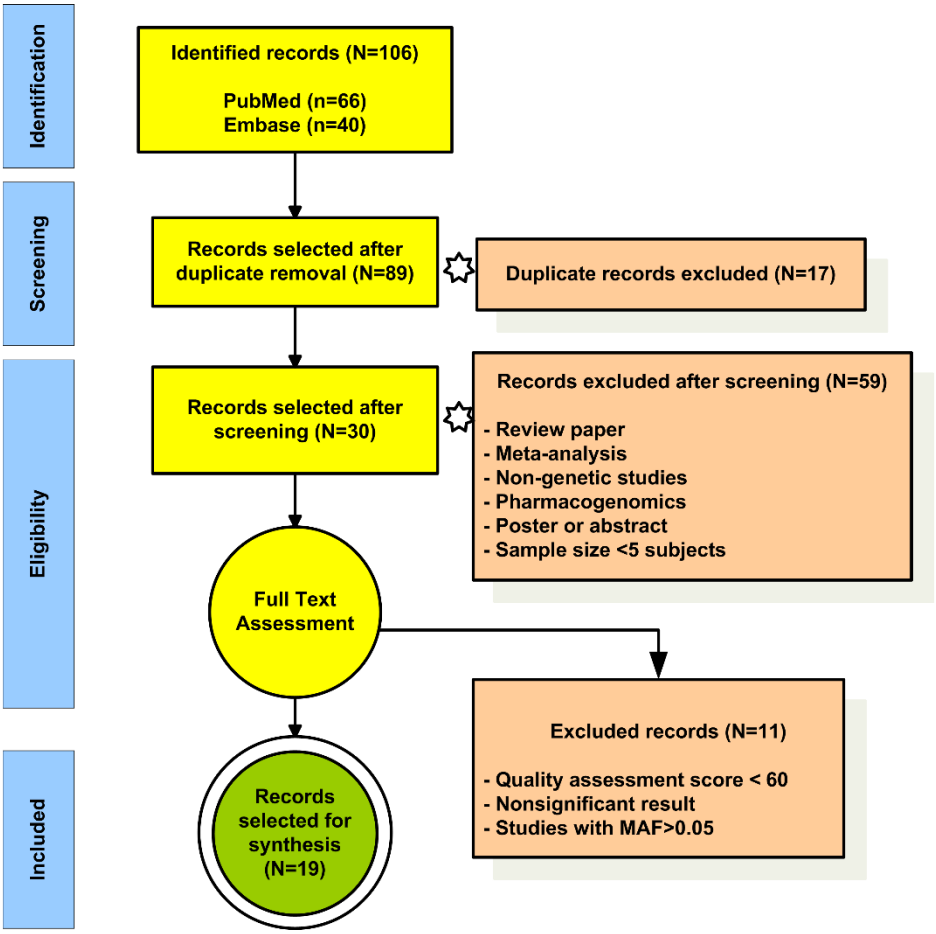
Non-Finnish European population datasets from gnomAD.v2 (Exomes=56,885; Genomes=7,718), and gnomAD.v3 (Genomes=32,399), a Spanish population dataset from CSVS (Exomes=1,942) and a Swedish population dataset from SweGen (Genomes=1000) were used as control groups (Ameur et al. 2017) to compare the minor allele frequency (MAF), to calculate the odds ratio (OR) for Spanish MD-EP, MD-AEP and Swedish tinnitus cohorts. For SVA the OR with 95% CI was calculated for each variant using three control datasets and p-values were corrected by the total number of variants being compared. For GBA, total alternate alleles per gene using 2x2 contingency matrixes were calculated for EP, AEP and control datasets. For each gene, the OR was calculated with 95% CI and two-tailed p-value was corrected for multiple testing by the total number of genes being compared following Bonferroni-correction. A corrected p-value <0.05 was considered significant. For each gene, the EF was also calculated as previously described[38].

5. Results

5.1. Selection and Characteristics of EP Studies

For the EP strategy, we retrieved 106 records in total, 66 records from PubMed and 40 from Embase, by using the search strings reported in the search strategy section. After duplicates' removal, we retained 89/106 records aggregated from the two databases. Next, after screening by title and abstract of the articles, we retrieved 30/89 records that were included for full-text assessment. The discarded records were reviews, meta-analyses, non-genetic studies, pharmacogenomics studies, posters, or abstracts presented at scientific meetings. All studies including variants with MAF > 0.05, single cases, or <5 patients with EP were excluded. We performed quality assessment for 30 articles, and 19/30 records surpassed the minimum quality assessment score and were considered for synthesis (Figure 9).

Figure 9. Flowchart to select extreme phenotype records for synthesis.



Among the 19 studies selected for synthesis, 16 records were related to physical conditions, 1 was on bipolar disorder, and 2 were related to neurological disorders including epilepsy and Alzheimer's disease. All of these studies reported rare variants, candidate genes, or potential pathways associated with a particular trait using an EP approach. These 19 EP studies covered 18 complex diseases. Information about population ancestry and sample size of cases was available for all 19 studies. Only 11/19 studies reported the age of disease onset, and 18/19 records reported the sex of the individuals. The most common criteria to define EP included early onset, late onset, family history, acute form, and/or fast progression of a disease. In addition, disease-specific features were also considered to define an EP, such as the worst score in biomarkers levels including Bone Mass Density (BMD) and spirometry-based severity according to Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade. The reported sample size was between 12 and 32,965 individuals. A summary of the characteristics of these 19 EP studies is shown in Supplementary Table S5.

5.2. Synthesized Findings of EP Studies

In the 19 EP studies, the combination of general and disease-specific EP criteria was used to select individuals. Information on the study design, sequencing technique, and ancestry population was available for all 19 studies. The reported sample size varied according to the design and sequencing method: 1711 ± 2513 (mean ± SD) for GWAS, 929 ± 2389 for genotyping, 1274 ± 9380 for WES, 29 ± 9 for targeted sequencing, and 949 ± 8742 for Sanger sequencing. All 19 examined studies using EP to select individuals reported significant findings including several rare variants, copy number variants, potential candidate genes or pathways associated with the condition of interest. WES was able to find rare variants in 13/19 studies (MAF = 0.00–0.05) in identified variants. It also helped in the identification of several novel candidate genes including *TACC2* [71], *PRKCD*, *C1QTNF4*, *DNMT3A* [72], *LOC728699*, and *FASTK* [32]. GWAS identified a rare variant in 1/19 study (MAF = 0.04). In addition, genotyping and targeted and Sanger sequencing contributed in the identification of many candidate genes and micro-deletions.

5.3. Rare missense variants in synaptic genes associated with tinnitus extreme phenotype in MD patients.

First, we performed a SVA in patients with MD-EP and MD-AEP. The total number of obtained variants with MAF <0.05 were 2287 for MD-EP and 1610 for MD-AEP, respectively. Two missense variants were found significantly associated in patients with MD-

EP after p-correction. The first was a heterozygous variant and it was found in 3 unrelated individuals located at exon 21 in *DAAM1* gene (chr14:59826182A>C; p.Asn875His; rs61740455) with MAF_{csvs}=0.002 and CADD=17.85). The associated second variant was located at exon 32 in *MYH10 gene* (chr17:8397065C>A; p.Ala1399Ser; rs149021341; MAF_{csvs} =0.001, CADD= 22), and it was found in 2 individuals and one of the carriers was homozygous. Both variants were classified as likely benign according to ACMG and AMP guidelines. These variants were not replicated in the Swedish or CoGIE cohorts and no further analyses were carried out.

Next, we carried out a GBA in the Spanish MD cohort with MD-EP and MD-AEP. For this, we selected variants with MAF<0.1 to analyse the combined effect of different common and rare variants in the same gene. The retained variants in patients were 4625 for MD-EP and 3592 for MD-AEP, respectively after filtering by MD in-house controls to rule out rare variants associated with hearing or vestibular phenotypes. The GBA of missense variants showed 24 significant genes in tinnitus MD-EP including *PRUNE2, AKAP9, SORBS1, ITGAX, ANK2, KIF20B, LRPPRS, SYNPO, TSC2* (Table 5), and 18 genes for MD-AEP (Table 6). Interestingly, none of these genes showed an enrichment of synonymous or 5`UTR variants in MD-EP (Supplementary Tables S6 and S7); additionally, the genes from synonymous analysis for MD-AEP are detailed (Supplementary Table S8).

Table 5: List of synaptic genes showing enrichment of missense variants in Spanish patients with Meniere disease (MD) and reference datasets (Non-Finnish European from gnomAD.v2 or gnomAD.v3, Spanish from CSVS) were used to compare allelic Listed genes were significant when they were compared against CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(C
PRUNE2	***9***	6.02(3.60-10.07)	0.83	1.44E-08	5.89(3.52-9.85)	0.83	2.75E-08	5.40(3
AKAP9	***6***	12.32(5.48-27.68)	0.92	2.2E-06	13.89(6.17-31.27)	0.93	4.04E-07	6.68(2
SORBS1	***6***	10.93(4.87-24.55)	0.91	1.31E-05	11.52(5.12-25.93)	0.91	6.57E-06	7.73(3
ITGAX	***5***	73.02(29.68-179.66)	0.99	<1.00E-15	61.68(24.88-152.94)	0.98	<1.00E-15	14.29(
ANK2	***4***	18.30(6.78-49.40)	0.95	1.80E-05	19.95(7.36-54.08)	0.95	7.55E-06	21.93(
KIF20B	***4***	7.76(3.45-17.49)	0.87	1.42E-03	8.43(3.74-19.01)	0.88	5.27E-04	16.57(
TSC2	***4***	63.73(23.35-173.96)	0.98	8.38E-13	53.56(19.47-147.30)	0.98	2.35E-11	21.93(
SPHK2	***4***	5.47(2.25-13.28)	0.82	*NS*	5.51(2.27-13.39)	0.82	*NS*	8.45(3
SYNPO	***4***	74.87(27.35-204.94)	0.99	<1.00E-15	78.43(28.21-218.03)	0.99	<1.00E-15	32.90(
LRPPRC	***4***	49.75(18.29-135.32)	0.98	3.73E-11	73.20(26.39-203.03)	0.99	4.19E-13	21.93(
XYLT1	***4***	2.00(0.74-5.38)	0.50	*NS*	2.09(0.78-5.61)	0.52	*NS*	10.95(
ALCAM	***3***	8.22(2.62-25.81)	0.88	*NS*	9.48(3.01-29.81)	0.89	*NS*	24.67(
CDH13	***3***	12.15(4.50-32.85)	0.92	1.60E-03	13.09(4.83-35.46)	0.92	8.05E-04	33.08(
DOCK7	***3***	52.57(16.54-167.10)	0.98	3.53E-08	70.09(21.61-227.27)	0.99	2.71E-09	24.67(
BIN1	***3***	56.70(17.82-180.42)	0.98	1.53E-08	73.20(22.54-237.74)	0.99	1.72E-09	49.36(
FLII	***3***	26.77(8.48-84.44)	0.96	3.87E-05	33.95(10.66-108.12)	0.97	4.64E-06	24.67(
HSPA4L	***3***	17.41(5.53-54.77)	0.94	1.95E-03	16.71(5.29-52.75)	0.94	2.98E-03	16.44(
IQSEC1	***3***	32.85(10.39-103.82)	0.97	5.12E-06	30.49(9.59-96.94)	0.97	1.32E-05	24.67(
IQSEC3	***3***	4.16(1.33-13.04)	0.76	*NS*	4.47(1.43-14.03)	0.78	*NS*	16.44(
LLGL1	***3***	27.29(10.06-74.03)	0.96	1.57E-07	25.08(9.21-68.31)	0.96	5.53E-07	13.92(
MADD	***3***	128.54(39.58-417.46)	0.99	1.26E-12	73.20(22.54-237.74)	0.99	1.72E-09	49.36(

MBP	***3***	170.13(51.78-559.02)	0.99	<1.00E-15	82.35(25.24-268.68)	0.99	4.99E-10	16.44(
MPRIP	***3***	82.62(25.77-264.88)	0.99	2.10E-10	78.43(24.09-255.39)	0.99	8.32E-10	49.36(
NRCAM	***3***	69.68(21.82-222.56)	0.99	1.49E-09	50.67(15.78-162.74)	0.98	8.07E-08	49.36(
TRAP1	***3***	13.72(4.37-43.13)	0.93	1.39E-02	11.27(3.58-35.47)	0.91	*NS*	24.67(
VCAN	**3**	90.37(28.13-290.35)	0.99	7.41E-11	76.61(23.55-249.22)	0.99	1.07E-09	24.67(
MYO18A	**3**	204.11(72.08-577.95)	0.99	<1.00E-15	169.90(58.69-491.88)	0.99	<1.00E-15	33.08(
MYO5A	**3**	13.72(4.37-43.13)	0.93	1.39E-02	16.97(5.37-53.37)	0.94	2.62E-03	16.44(
PPP1R9A	***2***	171.57(51.97-566.41)	0.99	<1.00E-15	127.78(38.15-427.99)	0.99	7.12E-12	24.87(
CCDC22	***2***	7.38(2.72-20.04)	0.86	*NS*	8.37(3.08-22.75)	0.88	*NS*	14.06(
EPX	***2***	8.17(3.01-22.19)	0.88	*NS*	8.69(3.20-23.64)	0.88	4.28E-02	11.61(

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

**Table 6: List of synaptic genes showing enrichment of missense variants in Spanish patients with Meniere disease (MD) and ti
Three reference datasets (Non-Finnish European from gnomAD.v2 or gnomAD.v3, Spanish from CSVS) were used to compar
cohort. Listed genes were significant when they were compared against CSVS reference dataset**

Gene	#Variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
ITGAX	5	65.65(28.87-149.28)	0.98	<1.00E-15	65.79(28.66-151.06)	0.98	<1.00E-15
KIAA1549	5	10.18(4.19-24.69)	0.90	5.46E-04	11.82(4.86-28.76)	0.92	9.62E-05
GOLGB1	4	41.14(15.16-111.66)	0.98	5.60E-10	41.32(15.11-113.01)	0.98	7.87E-10
GPR158	4	13.80(5.12-37.20)	0.93	4.06E-04	14.70(5.44-39.75)	0.93	2.24E-04
WFS1	4	10.77(4-29.02)	0.91	4.90E-03	12.51(4.63-33.80)	0.92	1.19E-03
WNK1	4	21.10(7.81-57.01)	0.95	3.40E-06	17.88(6.60-48.43)	0.94	2.63E-05
PPFIA1	4	92.82(33.78-255.07)	0.99	<1.00E-15	85.78(30.79-238.97)	0.99	<1.00E-15
RIN1	4	35(12.91-94.86)	0.97	5.22E-09	30.50(11.20-83.04)	0.97	4.25E-08
TAOK2	4	27.05(10-73.16)	0.96	1.57E-07	21.95(8.09-59.54)	0.95	2.46E-06
CAD	3	6.07(1.94-19.05)	0.84	*NS*	5.86(1.87-18.41)	0.83	*NS*
FASN	3	7.09(2.26-22.26)	0.86	*NS*	7.19(2.29-22.60)	0.86	*NS*
KIF5A	3	27.83(8.82-87.84)	0.96	2.65E-05	27.71(8.73-87.96)	0.96	3.27E-05
ANK1	3	30.85(9.77-97.44)	0.97	9.63E-06	25.06(7.90-79.44)	0.96	8.39E-05
LMO7	3	35.35(13.01-96.10)	0.97	5.24E-09	47.13(17.14-129.56)	0.98	1.54E-10
MYO1C	3	12.78(4.73-34.56)	0.92	9.65E-04	11.92(4.40-32.29)	0.92	2.06E-03
PLXNA2	3	99.12(35.92-273.54)	0.99	<1.00E-15	120.34(42.48-340.89)	0.99	<1.00E-15
PTPRS	3	16.95(5.39-53.32)	0.94	2.46E-03	17.75(5.62-56.06)	0.94	1.80E-03
RYR2	3	153.52(46.99-501.54)	0.99	<1.00E-15	106.56(32.32-351.29)	0.99	3.22E-11
SPHK2	3	6.09(2.50-14.84)	0.84	*NS*	6.35(2.60-15.50)	0.84	*NS*
TRAP1	3	42.75(13.49-135.49)	0.98	3.31E-07	32.77(10.30-104.30)	0.97	6.50E-06

UNC13A	3	18.75(5.96-59.03)	0.95	1.03E-03	11.16(3.55-35.15)	0.91	*NS*
ST14	2	60.38(18.87-193.23)	0.98	9.15E-09	61.41(18.94-199.08)	0.98	1.28E-08
CRMP1	2	3.34(1.75-6.39)	0.70	*NS*	3.56(1.86-6.81)	0.72	*NS*

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by n

The most significant finding for MD-EP was an enrichment of missense variants in the *ANK2* gene, against NFE population from gnomAD.v2 [OR=18.30(6.78-49.40), EF=0.95, corrected-p=1.80E-05], gnomAD.v3 [OR=19.95(7.36-54.08), EF=0.95, corrected-p=7.55E-06] and Spanish population from CSVS [OR=21.93(7.02-68.48), EF=0.95, corrected-p=2.02E-04]. In *ANK2*, four different missense rare variants were found in 3 different sporadic cases; three of the variants were novel and they have not been reported in gnomAD or CSVS databases. The variant 4:114294537G>A; exon 45 was found only in one case and two of the novel variants 4:114277102T>G; exon 38 and 4:114294509G>C; exon 45 were carried by the same patient. The third novel variant 4:114262911A>G was located at exon 33 (Table 7).

In the next step, we selected missense variants with CADD≥20 from synaptic genes in MD-EP (561 SNV) and MD-AEP (560 SNV) for the GBA. We obtained 7 genes including *ANK2, SPTB, BIN1, FLII, TSC2, CDH13 and MYO18A* with significant burden of rare pathogenic variants (Table 8), and 9 genes (*DMD, GOLGB1, MYO1C, OGDHL, PPFIA1, PTPRS, RYR2, ST14, TRAP1)* significant for AEP (Supplementary Table S9), when they were compared with reference datasets.

5.3.1. Indel analysis

We also performed a SVA and GBA of indels in SG from Spanish patients with MD-EP and MD-AEP. Indels were further filtered out by in-house controls. A total of 1565 indels (MAF<0.05) for SVA, and 2370 indels (MAF<0.1) for the GBA were retrieved for the MD-EP, and 1404 indels for SVA (MAF<0.05) and 1693 (MAF<0.05) for GBA in the MD-AEP group, respectively. We found an enrichment of indels in 31 genes in the MD-EP (Supplementary Table S10), including *TSC2, AKAP9* and several other genes and 48 genes in the MD-AEP (Supplementary Table S11), when data were compared with European reference datasets (gnomAD.v2 and gnomAD.v3). Unfortunately, we cannot compare the allelic frequencies in MD-EP or MD-AEP for indels CSVS, since the number of indels reported in CSVS dataset is low and it will overestimate the burden.

We also compared rare, Loss-of-Function (LoF) variants including nonsense, splice-site and frameshift small insertions and deletions in the SG set for MD-EP andMD-AEP. We found 61 LoF variants in the MD-EP (19 nonsense, 5 splice-site and 37 frameshift indels); and 25 LoF variants in the MD-EP (12 nonsense, 3 splice-site and 10 frameshift indels). However,

the number of nonsense or novel splice-site variants found was small, and no significant burden of LoF variants was found in MD-EP or in MD-AEP.

Table 7: Rare missense variants found in the gene burden analysis for *ANK2* gene in Spanish patients phenotype (MD-EP) and Swedish patients with severe tinnitus

Spanish MD-EP cohort (N=30)							
Pos	**Exon**	**rsID**	**MAF (MD-EP)**	**MAF NFE**		**MAF. CSVS**	**CADD**
				gnomAD.v2	**gnomAD.v3**		
4:114262911:A>G	33	-	0.0167	-	-	-	23.2
4:114277102:T>G	38	-	0.0167	-	-	-	24.1
4:114294509:G>C	45	-	0.0167	-	-	-	25.7
4:114294537:G>A	45	rs45454496	0.0167	0.0037	0.0034	0.003	25.4
Swedish tinnitus cohort (N=97)							
Pos	**Exon**	**rsID**	**MAF (Swedish cohort)**	**MAF NFE**		**MAF. SweGen**	**CADD**
				gnomAD.v2	**gnomAD.v3**		
4:114275980:G>A	38	rs149645600	0.0052	0.0013	0.00113	0.001	23.1
4:114276906:G>A	38	rs141191319	0.0206	0.0039	0.0034	0.0095	7.91
4:114277914:G>A	38	rs753223319	0.0052	2.65E-05	-	-	12.84
4:114278016:C>A	38	rs764914059	0.0052	8.95E-06	-	-	9.27
4:114278128:C>T	38	rs145895389	0.0052	0.003	0.0031	0.0035	3.78
4:114279628:T>C	38	rs36210417	0.0258	0.0107	0.0110	0.0065	25
4:114294462:C>T	45	rs121912706	0.0052	0.0017	0.0016	0.003	35

4:114294537:G>A	45	rs45454496	0.0052	0.0037	0.0034	0.004	25.4
Swedish severe tinnitus cohort (N=34)							
Pos	**Exon**	**rsID**	**MAF (Swedish cohort)**	**MAF NFE**		**MAF. SweGen**	**CADD**
				gnomAD.v2	**gnomAD.v3**		
4:114276906:G>A	38	rs141191319	0.0294	0.0039	0.0034	0.0095	7.91
4:114278016:C>A	38	rs764914059	0.0147	8.95E-06	-	-	9.27

VUS= Variant of uncertain significance

Table 8: Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with Meniere disea (MD-EP). Listed genes were significant when they were compared against CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(CI)
ANK2	4	18.30(6.78-49.40)	0.95	1.80E-05	19.95(7.36-54.08)	0.95	7.55E-06	21.93(7.02-6
SPTB	4	23.36(8.64-63.12)	0.96	9.86E-07	25.23(9.29-68.54)	0.96	4.60E-07	10.95(3.77-3
ARHGAP23	3	11.28(3.59-35.43)	0.91	*NS*	7.57(2.41-23.80)	0.87	*NS*	24.67(6.49-9
BIN1	3	56.70(17.83-180.42)	0.98	1.53E-08	73.20(22.54-237.74)	0.99	1.72E-09	49.36(10.97
FLII	3	26.77(8.48-84.44)	0.96	3.87E-05	33.95(10.66-108.12)	0.97	4.64E-06	24.67(6.49-9
TRAP1	3	13.72(4.37-43.13)	0.93	1.39E-02	11.27(3.58-35.47)	0.91	*NS*	24.67(6.49-9
TSC2	3	50.73(15.97-161.15)	0.98	5.22E-08	42.77(13.37-136.83)	0.98	4.60E-07	24.67(6.49-9
CCDC22	2	7.38(2.72-20.04)	0.86	*NS*	8.37(3.08-22.75)	0.88	*NS*	14.06(4.71-4
CDH13	2	19.69(6.22-62.27)	0.95	7.45E-04	21.28(6.69-67.65)	0.95	4.17E-04	49.77(11.02
FASN	2	8.94(2.83-28.18)	0.89	*NS*	7.20(2.28-22.73)	0.86	*NS*	16.57(4.62-5
MYO18A	2	153.51(46.73-504.26)	0.99	1.00E-15	127.78(38.15-427.99)	0.99	7.12E-12	24.87(6.52-9

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordere

5.4. Replication of *ANK2, AKAP9* and *TSC2* in a Swedish tinnitus cohort with severe tinnitus

First, we selected all significant genes from MD-EP analysis to investigate if the burden of missense variants found could be replicated in a Swedish tinnitus cohort including 97 individuals. We used three different population datasets as reference controls (gnomAD.v2, gnomAD.v3 and SweGen). The observed MAF for each gene was calculated and compared with controls, whilst p-values were corrected by the total number of variants per gene. Six genes *ANK2, MYO18A, MADD, KIF20B, MPRIP, MBP* and *NRCAM* showed an enrichment of missense variants. Subsequently, we selected a subset of 34 patients with severe tinnitus (THI score ≥56) and found a burden of missense variants in *ANK2, AKAP9* and *TSC*2 genes (Table 9). Missense variants identified in the GBA for *ANK2* gene are previously detailed in Table 7; most of these variants are clustered around exons 38 to 45 across the gene sequence (Figure 10). Tables 10 and 11 listed missense variants found in the GBA for *AKAP9* and *TSC2* genes in Spanish and Swedish patients with tinnitus. Rare variants found in *ANK2* and *TSC2* genes were also validated by Sanger sequencing (Supplementary Figures S1 and S2)

In addition, we used an independent cohort of generalised genetic epilepsy to determine if the association of *ANK2, TSC2 and AKAP9* genes with severe tinnitus was a non-specific finding, since some neurological disorders such as epilepsy could also share some common genetic background with tinnitus. For this, we performed a GBA using the same SG list in this epilepsy cohort, but none of the genes showed a significant enrichment of missense variants strongly suggesting the genes captured here are tinnitus-specific.

Lastly, we performed GBA of indels in the Swedish cohort using SG with MAF<0.1. We found 2 genes (*APC* and *CLASP2*) in the tinnitus cohort (N=97), and 6 genes (*AGL, APC, CLASP2, PC, ACACA* and *APPL2*) in subgroup with severe tinnitus (N=34), showing a significant burden of indels when compared with European (gnomAD) and Swedish reference data (SweGen) (Supplementary Table S12).

Table 9: Swedish Tinnitus replication cohort, Synaptic genes showing an enrichment of missense rare variants in Swedish patie

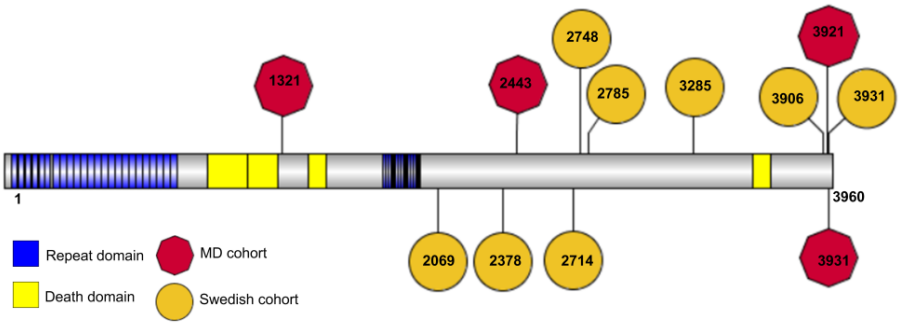
without diagnosis of MD

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[SweGen] OR(CI)
Non selected Tinnitus (N=97)								
ANK2	8	3.20(1.92-5.33)	0.69	6.25E-05	3.28(1.97-5.47)	0.70	4.08E-05	2.83(1.59-5.02)
MYO18A	5	5.99(2.84-12.64)	0.83	1.34E-05	5.94(2.81-12.57)	0.83	1.61E-05	6.60(2.55-17.07)
MADD	4	4.99(2.23-11.17)	0.80	3.78E-04	4.78(2.13-10.73)	0.79	6.04E-04	3.89(1.52-9.97)
KIF20B	4	4.99(2.06-12.06)	0.80	1.45E-03	4.89(2.02-11.87)	0.80	1.77E-03	3.70(1.33-10.30)
MPRIP	3	35.36(11-113.70)	0.97	6.53E-09	28.77(8.82-93.82)	0.97	7.60E-08	15.54(2.59-93.19)
MBP	2	12.95(3.18-52.76)	0.92	7.04E-04	35.34(8.20-152.23)	0.97	3.43E-06	10.36(1.45-73.74)
NRCAM	2	47.15(11.13-199.77)	0.98	3.37E-07	111.91(22.52-556.20)	0.99	1.62E-08	20.72(1.87-229.04)
Severe tinnitus (N=34)								
AKAP9	3	4.93(2.03-12)	0.80	1.29E-03	5.80(2.38-14.12)	0.83	3.20E-04	3.32(1.31-8.47)
TSC2	2	13.92(4.41-43.93)	0.93	1.4E-05	10.97(3.47-34.66)	0.91	9.02E-05	12.87(3.29-50.31)
ANK2	2	11.51(3.65-36.28)	0.91	6.06E-05	13.20(4.17-41.78)	0.92	2.26E-05	4.73(1.38-16.17)

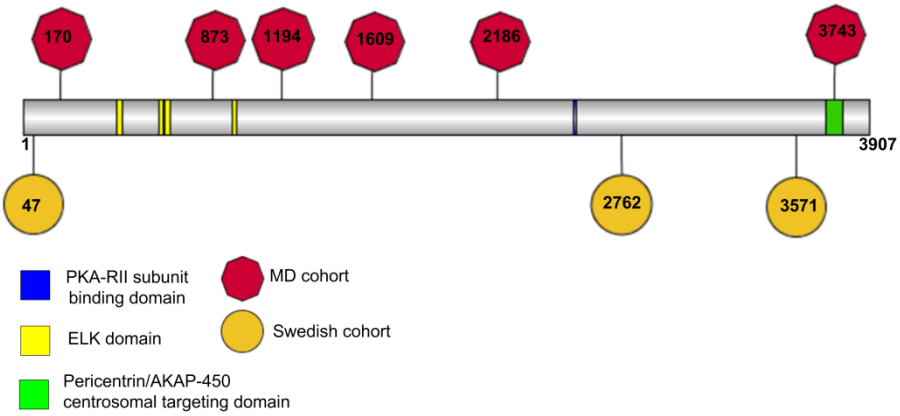
OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are

Figure 10: Distribution of rare variants across *ANK2*, *AKAP9* and *TSC2* genes found in the gene burden analysis in Spanish patients with Meniere disease-tinnitus extreme phenotype (MD-EP) and Swedish tinnitus cohorts. Each octagon/circle indicates the position of the involved amino acid in the protein sequence

ANK2



AKAP9



TSC2

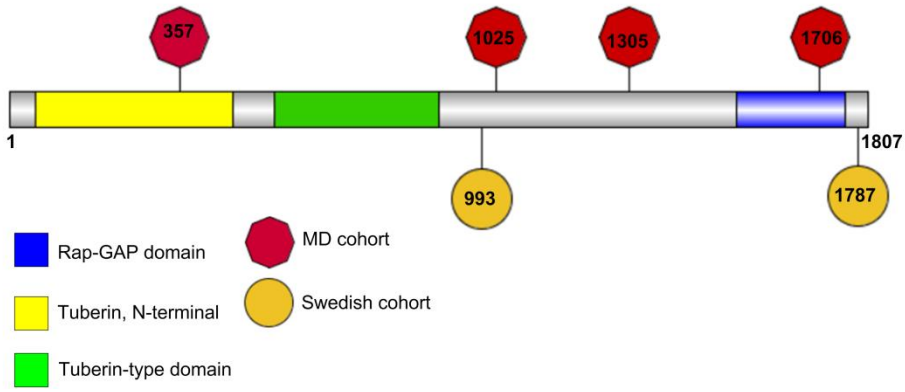


Table 10: Rare missense variants found in the gene burden analysis for *AKAP9* gene in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus

Spanish MD-EP tinnitus cohort (N=30)									
Pos	**Exon**	**rsID**	**MAF (MD-EP)**	**MAF NFE**		**MAF. CSVS**	**CADD**	**ACMG**	**Amino acid change**
				gnomAD.v2	**gnomAD.v3**				
7:91622303:G>C	5	rs144888041	0.0167	0.0026	0.0030	0.008	20.1	Benign (PS4,BS1,BS2,BP1,BP4)	E170D
7:91631849:A>G	8	rs746429266	0.0167	0	0	0.001	17.79	Benign (PS4,BS1,BS2,BP1,BP4)	K873R
7:91643610:G>A	10	rs139965373	0.0167	0.0004	0.0004	0.001	25	Benign (PS4,BS1,BS2,BP1,BP4)	A1194T
7:91670121:G>A	18	rs148146011	0.0167	0.0003	0.0001	0.002	22.8	Benign (PS4,BS1,BS2,BP1,BP4,BP6)	R1609K
7:91700267:T>C	28	rs76177450	0.0167	0.0049	0.0038	0.003	16.42	Benign (PS4,BS1,BS2,BP1,BP4,BP6)	S2186P
7:91732039:G>C	46	rs143306820	0.0167	4.48E-05	3.10E-05	-	24.4	Benign (PS4,PP3,BS1,BS2,BP1)	M3743I
7:91574215:CT/C-	-	rs1309343726	0.0179	0.0006	0.0006	-	-	Benign(PS4,BS1,BS2,BP4)	c.48+3755delT
7:91659313:AT/A-	-	rs779223487	0.0179	0.0002	0.0001	-	-	(Benign PS4,BS1,BS2,BP4)	c.4245+13delT
7:91706410:CT/C-	-	rs370936884	0.0179	0.0009	0.0003	-	-	(Benign PS4,BS1,BS2,BP4)	c.6765+106delT
Swedish tinnitus cohort (N=34)									
Pos	**Exon**	**rsID**	**MAF (Swedish cohort)**	**MAF NFE**		**MAF. SweGen**	**CADD**	**ACMG**	**Amino acid change**
				gnomAD.v2	**gnomAD.v3**				
7:91603115:C>T	2	rs35669569	0.0441	0.0133	0.011	0.0205	0.009	Benign (BS1,BS2,BP1,BP4,BP6)	H47Y
7:91712609:A>C	33	rs144875383	0.0147	0.0019	0.0018	0.002	0.211	Benign (BS1,BS2,BP1,BP4)	K2762N
7:91727526:G>A	43	-	0.0147	-	-	-	32	Pathogenic (PS4,PM2,PP3,BP1)	E3571K

ACMG=American college of medical genetics and genomics, VUS= Variant of uncertain significance

Table 11: Rare missense variants found in the gene burden analysis for *TSC2* gene in Spanish patients with Meniere dis

(MD-EP) and Swedish patients with severe tinnitus

Spanish MD-EP tinnitus cohort (N=30)								
Pos	**Exon**	**rsID**	**MAF (MD-EP)**	**MAF NFE**		**MAF. CSVS**	**CADD**	**ACMG**
				gnomAD.v2	**gnomAD.v3**			
16:2110765:C>T	11	rs150195368	0.0167	0.0006	0.0009	-	23.8	Likely benign (PS4,PP,BS2,BI
16:2129140:C>T	27	-	0.0167	-	-	-	21.7	Likely pathog (PS4,PM2,PP2,PP3)
16:2133726:C>T	33	rs45517320	0.0167	6.09E-05	7.74E-05	0.001	14.01	VUS (PS4,PM5,PP2,BS1,BS2,BP4,
16:2138096:C>T	40	rs45517391	0.0167	0.0004	0.0003	0.002	23.2	Likely pathog (PS4,PM1,PP2,PP3,BS2)
16:2114151:C/+TG	-	rs754285275	0.0179	0.0003	0.0002	-	-	Likely pathogenic (PS4,PM2,B
16:2123243: G/+T	-	rs141745833	0.0179	0.0003	0.0007	-	-	Benign (PS4,BA1,BP4)
16:2127041:C/+TA	-	rs200120767	0.0179	0.0018	0.0023	-	-	Likely pathogenic (PS4,M2,PP
16:2130492: G/+T	-	rs112025110	0.0179	0.0048	0.0051	-	-	Benign (BA1,BP4)
16:2136476:CA/C-	-	rs142421783	0.0179	0.0006	0.0006	-	-	Benign (PS4,BA1BP4,BP6)
Swedish tinnitus cohort(N=34)								
Pos	**Exon**	**rsID**	**MAF (Swedish cohort)**	**MAF NFE**		**MAF. SweGen**	**CADD**	**ACMG**
				gnomAD.v2	**gnomAD.v3**			
16:2129044:C>T	27	rs137854410	0.0147	3.60E-05	4.64E-05	-	33	VUS (PS4,PP2,PP3,BS2,BP6)
16:2138546:G>A	42	rs45517419	0.0294	0.0032	0.0040	0.0035	1.823	Benign (PP2,BS1,BS2,BP4,BP

VUS= Variant of uncertain significance

5.5. Rare missense variants in hearing loss genes associated with tinnitus extreme phenotype in Meniere disease

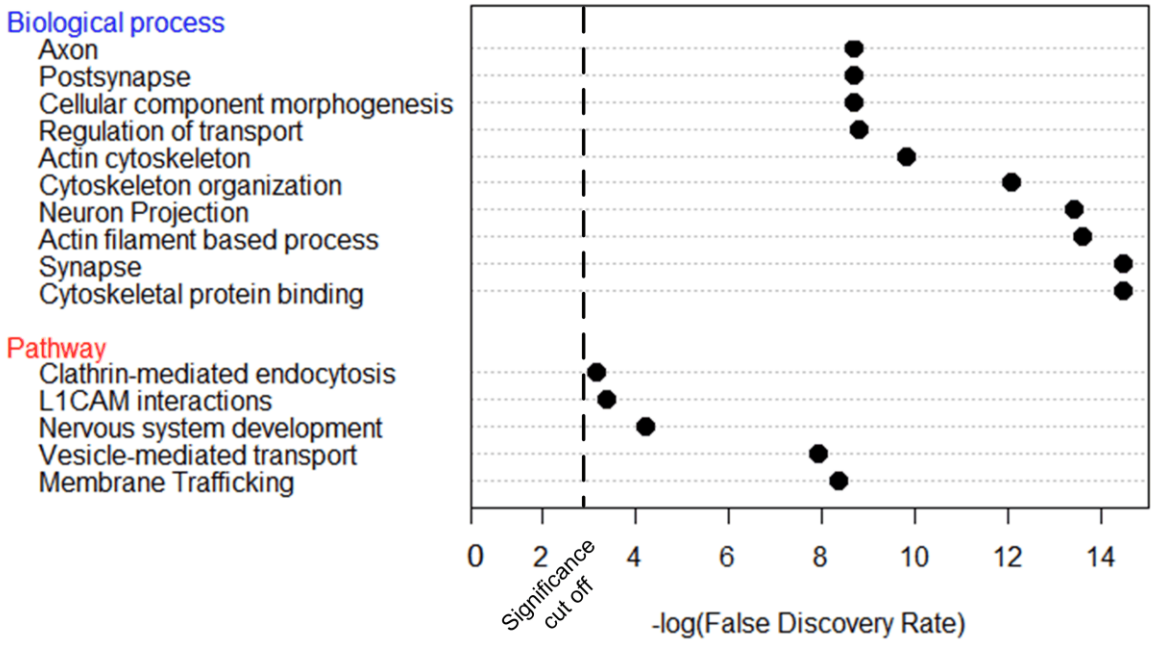
Next, we performed GBA of missense variants using hearing loss genes in patients with MD. We obtained 305 variants from MD-EP and 313 from MD-AEP with MAF<0.1, respectively. The 6 genes included *USH1G*, *ILDR1*, *OTOA*, *PCDH15*, *CACNA1D* and *NARS2* found significant in MD-EP (Supplementary Table S13) and 4 genes showed significant enrichment in MD-AEP (Supplementary Table S14)

To replicate the burden of rare variants found in hearing loss genes in MD-EP patients, we selected a subset of 62 patients with self-reported hearing problems from the Swedish cohort. Then, we performed a GBA in the 6 hearing loss genes *USH1G*, *ILDR1*, *OTOA*, *PCDH15*, *CACNA1D* and *NARS2*, however, none of these genes showed an enrichment of missense variants in this cohort.

5.6. Gene ontology and Gene-set enrichment analysis in patients with tinnitus

We selected a total of all 55 significant genes from MD-EP to perform GO and gene-set enrichment analysis including 24 genes with enrichment of missense variants and 31 genes with enrichment of indels analysis. Two genes (*TCS2* and *AKAP9)* had a burden of missense and structural variants and were in both gene lists. Using this selection, we found a total of 5 significant molecular pathways enriched in Reactome and 10 significant biological processes enriched in GO database. The significant pathways are membrane trafficking, vesicle-mediated transport, nervous system development, L1CAM interactions and Clathrin-mediated endocytosis (Figure 11, Table 12). GO biological processes analysis revealed a significant enrichment of genes involved in the regulation of cytoskeletal protein binding (GO:0008092*;* p=4.49E-19), synapse (GO:0045202;p=6.48E-19), actin filament-based process (GO:0030029;p=7.71E-18*)*, neuron projection (GO:0043005;p=1.48E-17*)*, cytoskeleton organization (GO:0007010;p=4.11E-16*)*, actin cytoskeleton (GO:0015629;p=8.77E-14*)*, regulation of transport (GO:0051049;p=1.07E-12*)*, cellular component morphogenesis (GO:0032989;p=1.67E-12*)*, postsynapse (GO:0098794;p=1.83E-12*)* and axon formation (GO:0030424;p=1.94E-12)(Supplementary Table S15).

Figure 11: Gene ontology (GO) and gene-set enrichment analysis in synaptic genes



These analyses were performed using 55 genes obtained in the gene burden analysis of MD-EP. GSEA tool was used to obtained molecular pathways and biological processes.

Table 12: The detailed information of GO results obtained for biological processes and pathways in MD-EP and MD-AEP in synaptic genes

Biological process	**# Genes in Gene Set (K)**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
Axon	643	13	0.0202	1.94E-12	1.99E-09
Postsynapse	640	13	0.0203	1.83E-12	1.99E-09
Cellular component morphogenesis	800	14	0.0175	1.67E-12	1.99E-09
Regulation of transport	1856	19	0.0102	1.07E-12	1.57E-09
Actin cytoskeleton	503	13	0.0258	8.77E-14	1.50E-10
Cytoskeleton organization	1396	20	0.0143	4.11E-16	8.44E-13
Neuron projection	1366	21	0.0154	1.48E-17	3.80E-14
Actin filament based process	804	18	0.0224	7.71E-18	2.64E-14
Synapse	1357	22	0.0162	6.48E-19	3.33E-15
Cytoskeletal protein binding	979	20	0.0204	4.49E-19	3.33E-15
Pathway	**# Genes in Gene Set (K)**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
Clathrin-mediated endocytosis	145	5	0.0345	1.44E-06	6.90E-04
L1CAM interactions	121	5	0.0413	5.90E-07	4.23E-04
Nervous system development	580	9	0.0155	6.59E-08	6.30E-05
Vesicle-mediated transport	724	13	0.018	8.52E-12	1.22E-08
Membrane Trafficking	629	13	0.0207	1.47E-12	4.22E-09

5.7. *ANK2* gene expression profile in the mouse brain

In-situ hybridization (ISH) data in the mouse (Ank2: n=2, Tsc2: n=2, one coronally and one sagittally sectioned each) obtained from the Allen brain atlas demonstrated strong *Ank2* and *Tsc2* expression in a number of brain regions (no such data were available for *Akap9*, Figures 10&11). Visual inspection revealed strong expression of both genes in the cortex, hippocampus (pyramidal layer of CA1, CA2 and CA3 and the granule cell layer of the dentate gyrus), olfactory bulb (the granule and mitral layers) and cerebellum. In addition, to subregions of other brain regions, notably: Tenia tecta, the epithalamus (especially the medial habenula), piriform area (layer 2) and the magnocellular mucleus (Supplementary Figure 3A-D). These ISH data are available pre-rendered into a three-dimensional annotated reference volume[68] of 200 micron voxels containing the maximal-intensity value within each[69].

This allowed quantitative analysis of gene expression across 209 registered brain regions (e.g. see Figure 4C-D). *Ank2* ISH data revealed significant variations in expression across brain regions (voxels were grouped by brain region and compared, Kruskal-Wallis, $p<0.001$). Mean expression was ranked across brain regions confirming the strongest expression in a number of regions (Figure 12E), including: olfactory areas (piriform area, tenia tecta and accessory olfactory bulb), the pallidum (ventral regions, particularly: the magnocellular nucleus and caudal regions, particularly: the bed nucleus of the anterior commissure and the bed nuclei of the stria terminalis), the epithalamus (particularly: medial habenula), the hippocampus (dentate gyrus and the pyramidal layers of Ammon's horn, i.e. CA) the hypothalamus (periventricular regions, particularly: anteroventral periventricular nucleus, parastrial nucleus, anteroventral preoptic nucleus and medial preoptic nucleus, and the hypothalamic medial zone, particularly: the ventral premammillary nucleus), the striatum (ventral regions, particularly: the olfactory tubercle and the lateral septal complex, particularly the lateral septal nucleus) and the cortex.

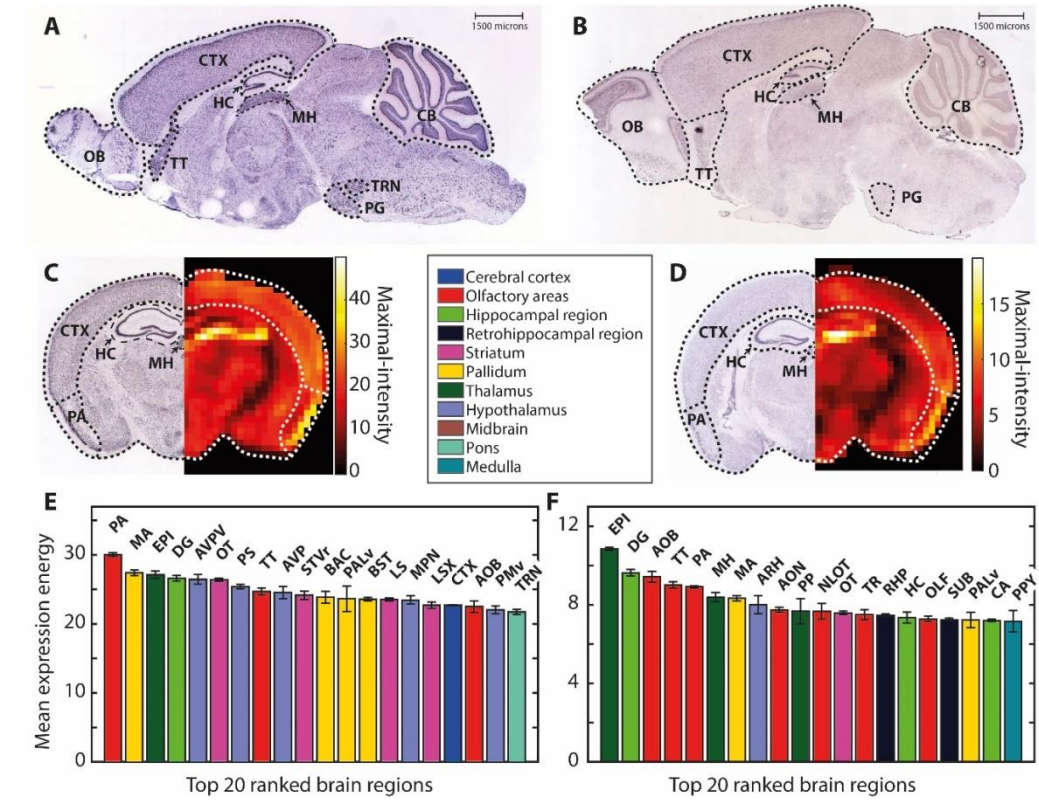
In addition, strong expression was observed in pontine gray and tegmental reticular nucleus. Similarly, *Tsc2* showed significant variations in expression (Kruskal-Wallis, $p<0.001$) with the strongest expression in: olfactory areas (accessory olfactory bulb, tenia tecta, piriform area, the nucleus of the lateral olfactory tract, anterior olfactory nucleus and the postpiriform transition area), the pallidum (ventral regions, particularly: the magnocellular nucleus), the thalamus (epithalamus, particularly: medial habenula, and the peripeduncular nucleus), the

hippocampus (dentate gyrus and the pyramidal layers of Ammon's horn, i.e. CA) the hypothalamus (arcuate hypothalamic nucleus), the striatum (the olfactory tubercle) and the medulla (parapyramidal nucleus).

There was a marked similarity in the brain wide expression of Ank2 and Tsc2, this could potentially suggest a common mechanism or brain regions of interest. To confirm this coexpression of 4,104 genes in the mouse brain were compared. These data were used to build a probability distribution for deriving a given amount of coexpression, based on this coexpression of Ank2 and Tsc2 was found to be highly significant (coexpression = 0.9031, p = 0.0091). In order to identify the brain regions with high coexpression brain wide expression profiles were normalised and multiplied (see methods) allowing a brain wide coexpression map (see Figure 13A).

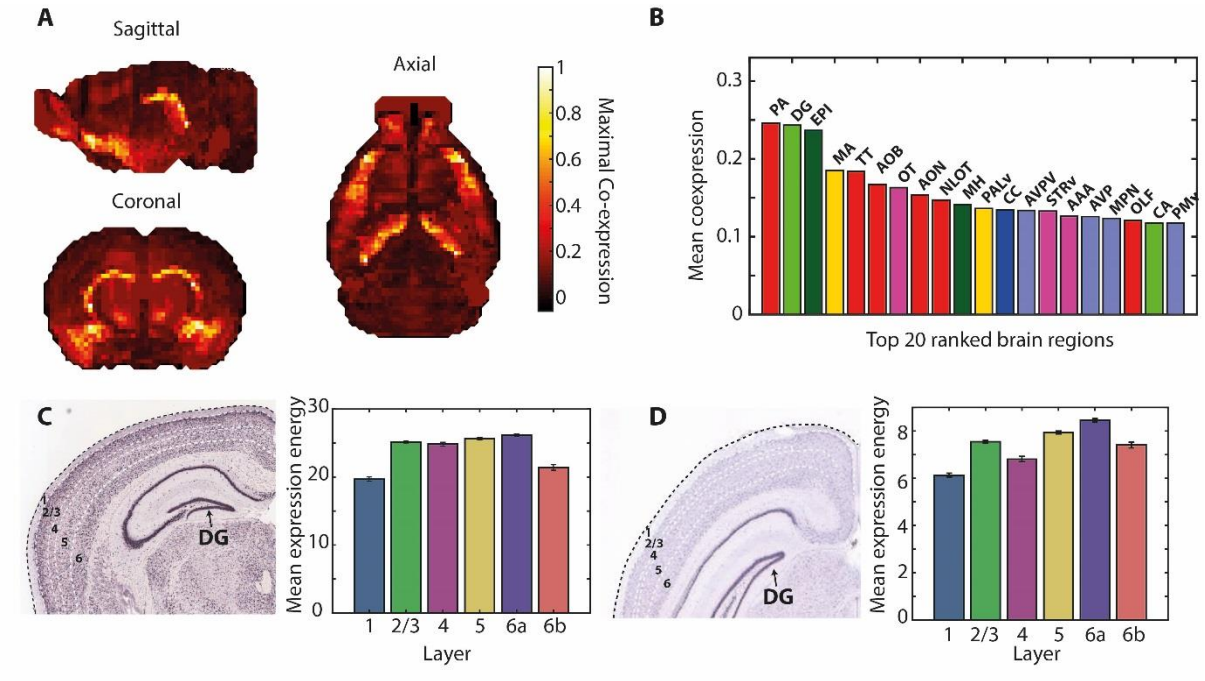
This revealed the strongest coexpression in: olfactory areas (piriform area, tenia tecta, accessory olfactory bulb, anterior olfactory nucleus and the nucleus of the lateral olfactory tract), the hippocampus (dentate gyrus and the pyramidal layers of ammon's horn), the epithalamus (particularly: medial habenula), the pallidum (ventral regions, particularly: the magnocellular nucleus), the striatum (ventral regions, particularly: the olfactory tubercle and anterior amygdalar area), the cerebral cortex (analysis of cortical layers revealed strong expression for both in layers 2/3, 5 and 6a but not in layers 1 and 4) and the hypothalamus (periventricular regions, particularly: anteroventral periventricular nucleus, anteroventral preoptic nucleus and medial preoptic nucleus, and the hypothalamic medial zone, particularly: the ventral premammillary nucleus).

Figure 12: Brain wide expression profiles of Ank2 and Tsc2 in the mouse brain taken from in-situ hybridization data from the Allen Brain Atlas data set (http://www.brain-map.org).



Sagittal sections of expression in the adult mouse (P56) brain for both A) Ank2 and B) Tsc2. Strong expression for both genes is found in a number of brain regions, including: CTX = Cortex, HC = Hippocampus, CB = Cerebellum, MH = Medial habenula, TT = Taenia tecta. Coronal sections (C and D, left panels) were taken from Allen Brain Atlas in a pre-rendered to fit an annotated format (C and D, right panels) allowing easy identification expression in different brain regions. This was used to identify brain regions demonstrating strongest expression (E and F, see text for details). PA = piriform area, MA = magnocellular nucleus, EPI = epithalamus, DG = dentate gyrus, AVPV = anteroventral periventricular nucleus, OT = olfactory tubercle, PS = parastrial nucleus, TT = tenia tecta, AVP = anteroventral preoptic nucleus, STVr = Striatum, ventral region, BAC = bed nucleus of the anterior commissure, PALv = pallidum, ventral region, BST = bed nuclei of the stria terminalis, LS = lateral septal nucleus, MPN = medial preoptic nucleus, LSX = lateral septal complex, CTX = cortex, AOB = accessory olfactory bulb, PMv = ventral premammillary nucleus, TRN = tegmental reticular nucleus, PP = peripeduncular nucleus, NLOT = nucleus of the lateral olfactory tract, ARH = arcuate hypothalamic nucleus, PP = peripenducular nucleus, TR = postpiriform transition area, RHP = retrohippocampal region, PPY = Parapyramidal nucleus.

Figure 13: Significant brain wide co-expression of Ank2 and Tsc2 was found in the mouse brain (see text).



A) Mean across-section (sagittal, coronal and axial planes) co-expression of Ank2 and Tsc2. Colorbar indicates strong coexpression in yellow. B) Mean coexpression in 209 brain regions was calculated and ranked revealing the top 20 brain regions where Ank2 and Tsc2 were co-expressed. Coronal sections revealing layer specific Ank2 expression in the cortex (C, left panel), mean expression for each cortical layer (across the whole brain) was calculated (C, right panel) revealing strongest expression in layers 2/3, 5 and 6a. Coronal sections revealing layer specific Tsc2 expression in the cortex (D, left panel), mean expression for each cortical layer (across the whole brain) was calculated (D, right panel) revealing strongest expression in layers 2/3, 5 and 6a.

5.8. Hypothesis free data-driven approach with tinnitus extreme phenotype in Meniere disease

In hypothesis free data-driven approach the phenotype, disease or pathways associated with genes are completely unknown. For hypothesis free analysis, we have performed GBA with missense and synonymous variants for MD-EP and MD-AEP against all genes in the human genome (N=20,000)[73]. The bioinformatics analyses, MAF and statistical approach were same as of synaptic gene approach. The p-values were corrected by the total number of genes. The most significant gene from GBA for MD-EP is *ADGRV1* which showed enrichment of missense variants, but not for synonymous variants (Supplementary Table S16 and S17). The genes in Supplementary Table S16 were further filtered out to exclude genes previously associated with hearing loss in familial and sporadic MD cases. The genes with burden of rare variants in familial MD include *ACAN*, *SPTA1*, *ALDH16A1*, *ZNF142*, *CFAP65*,

ARHGAP8, CACNA1S, MYO7A, TICRR, CCDC40, LAMC3, KIF17, OPRM1, KIF26A, KIF14, OTOG, NOS1, KIF26B, SPTB, UNC5B, GPR179, LY75, MYBPC2, MYH7B, LRRN4, ANKAR, ATR, CFH, DNAH14 and *GEMIN4*. The genes with burden of rare variants in sporadic MD include *ADGRV1, SEC16A, STARD9, PIEZO1, SPTA1, TRIOBP, ZNF469, COL18A1, DLEC1, MYH7B, NBEAL1, PNPLA7, PRUNE2, COL20A1, SCNN1D, ABCC12, ADAMTSL4, AFF1, ALKBH8, BDP1, CAPN15, CCDC171, CPAMD8, MYCBPAP, NID1, OTOGL, PTPN21, TNRC18*, TRPV1 and *ABCA5*.

The genes excluded from Supplementary Table S16 were *ACAN, MYH7B, SPTB, CACNA1S, NOS1, ADGRV1, PRUNE2, ZNF469, SEC16A* and *TRPV1*. The list of genes obtained is detailed in Supplementary Table S18 and it includes only genes with burden of rare and common variants for severe tinnitus, but not for hearing loss.

The GBA results for MD-AEP with missense and synonymous variants are detailed in Supplementary Tables S19 and S20, respectively. The Supplementary Table S19 was also filtered out to exclude genes previously associated with hearing loss in familial and sporadic MD cases. The excluded genes from Supplementary Table S19 include *ZNF469, STARD9, CFH, NBEAL1* and *CCDC40*. The list of genes obtained is detailed in Supplementary Table S21 and it includes only genes with burden of rare and common variants for severe tinnitus, but not for hearing loss.

5.8.1. Main genes found in the hypothesis-free driven analysis

The main gene identified in the hypothesis-free driven approach was *ADGRV1*. We have identified 9 missense variants carried by 7 unrelated individuals with MD-EP. This gene encodes adhesion G protein-coupled receptor v1 and highly expressed in central nervous system, but also in the ankle links among stereocilia in the cochlear and vestibular hair cells[74]. The G protein contains 7-transmembrane receptor domain and is responsible for calcium binding[75]. *ADGRV1* has unestablished function but in ectodomain the multiple calcium exchangers b-repeat may mediate the calcium through protein-protein interaction. *ADGRV1* has been previously reported for Usher syndrome type[76] 2, nonsyndromic deafness in 12 Chines Han families[77] , epilepsy/familial febrile Seizure[75].

5.8.2. Gene ontology and Gene-set enrichment analysis in patients with tinnitus

The gene ontology and gene set enrichment was performed using GSEA as detailed previously in section 4.5. All significant genes that showed an enrichment of missense variants from MD-EP and MD-AEP were combined together (N=201). This final list of genes

was used for gene ontology and gene-set enrichment analysis. The top pathways include cytoskeletal protein binding, and most significant biological processes were post translational protein binding and membrane trafficking, Figure 14. These results show an overlap and consistency with the previous GO analysis results obtained with the synaptic gene set, since the overlap pathways are (membrane trafficking and vesicle-mediated transport) and biological process are cytoskeletal protein binding, cytoskeleton organization and neuron projections).

Figure 14: Gene ontology (GO) and gene-set enrichment analysis in severe tinnitus by using all genes with enrichment of rare variants in the human genome

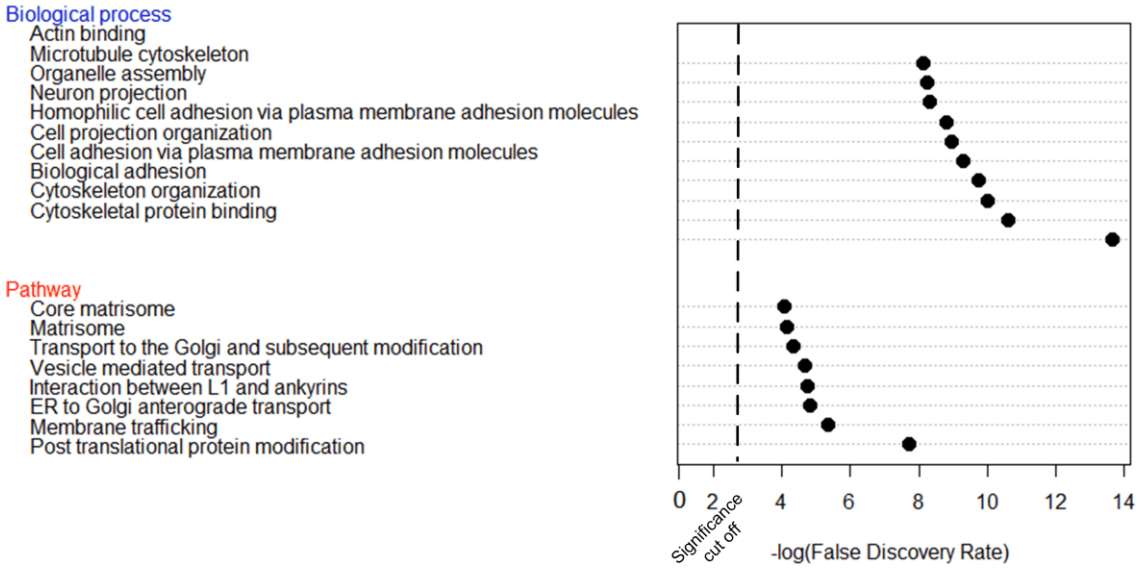


Table 12: The detailed information of results obtained for biological processes and pathways in MD-EP and MD-AEP by using all genes with enrichment of rare variants in the human genome

Biological process	**# Genes in Gene Set (K)**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
Cytoskeletal protein binding	979	29	0.0296	2.10E-14	2.16E-10
Cytoskeleton organization	1396	30	0.0215	2.39E-11	1.23E-07
Biological adhesion	1481	30	0.0203	9.96E-11	3.41E-07
Cell adhesion via plasma membrane adhesion molecules	277	14	0.0505	1.75E-10	4.50E-07
Cell projection organization	1588	30	0.0189	5.21E-10	1.07E-06
Homophilic cell adhesion via plasma membrane adhesion molecules	168	11	0.0655	1.10E-09	1.88E-06
Neuron projection	1366	27	0.0198	1.60E-09	2.35E-06
Organelle assembly	878	21	0.0239	4.67E-09	6.00E-06
Microtubule cytoskeleton	1256	25	0.0199	5.92E-09	6.76E-06
Actin binding	437	15	0.0343	7.73E-09	7.94E-06
Pathway	**# Genes in Gene Set (K)**	**# Genes in Overlap (k)**	**k/K**	**p-value**	**FDR q-value**
Post translational protein modification	1432	26	0.0182	1.85E-08	5.31E-05
Membrane trafficking	629	14	0.0223	4.37E-06	6.27E-03
ER to Golgi anterograde transport	155	7	0.0452	1.48E-05	1.03E-02
Interaction between L1 and Ankyrins	31	4	0.129	1.77E-05	1.03E-02
Vesicle mediated transport	724	14	0.0193	2.10E-05	1.03E-02
Transport to the Golgi and subsequent modification	186	7	0.0376	4.75E-05	1.70E-02
Matrisome	1026	16	0.0156	7.20E-05	2.29E-02
Core matrisome	275	8	0.0291	8.44E-05	2.42E-02

6. Discussion

6.1. Extreme phenotype strategies to uncover rare variants

6.1.1. Summary of the main findings from systematic review

The systematic review shows that individuals with an EP may reveal rare variants that can influence genetic susceptibility in most complex disorders. Complex disorders have a heterogeneous spectrum of symptoms, with variable expressivity observed in each patient. By cluster analysis, it is possible to identify subgroups of patients, and by selecting patients with EP (high expressivity), we would expect to find an enrichment of rare variations associated with the EP[78]. However, we cannot recommend a particular EP strategy to select patients, although the selection of individuals with an early-onset disease and/or a severe phenotype (genetic anticipation) will probably help in the search of rare variations. In contrast, elderly patients can show mutations associated with exposure to environmental factors along life (ultraviolet radiation, chemical agents, pollutants)[79]. In general, the criteria to define EP combine common and disease-specific features such as the chronic state of a disease, very high or low biomarker levels such as BMD, spirometry-based severity level according to GOLD, family history, and early/late age of disease onset.

Of note, a large sample size was not required in WES studies for the discovery cohort, and 10/19 records had a number of cases <100. Therefore, a moderate sample size of individuals with EP was sufficient to identify candidate rare variants or genes. These individuals with EP were carriers of rare variants with a high effect size to target new candidate genes. The EP approach was reproducible across different populations, since the selected studies recruited cases with different ethnic backgrounds including Asian, African, and European ancestry and with monogenic diseases such as cystic fibrosis[29] with an EP (persistent tracheobronchial infection with early onset)[80]. Therefore, the information about age of disease onset and sex of the selected individuals may result essential to define an EP[81].

6.1.2. How to select EP in quantitative traits

Individuals with EP are characterized by extreme clinically relevant attributes, toxic effects, or extreme responses to a treatment[16]. From a theoretical perspective, a very EP is more informative than an almost EP, but in practice there are several limitations associated with the very EP individuals, such as vulnerability to phenotype heterogeneity and measurement

errors. If a significant proportion from both sides of an extreme is discarded, the almost EP can still be more powerful than random sampling of the same size. The benefits of EP sampling were demonstrated by proposing power calculation methods with the help of the maximum likelihood approach[27,82]. It was also indicated that EP sampling to detect rare variants is more cost efficient as compared to traditional study designs[83]. Replication in a second independent EP cohort to enhance the power of a study is highly recommended, but it is unlikely to obtain a large sample size of EP subjects from a single region and multicenter studies are needed[84].However, the EP approach is considered more efficient than random sampling for the detection of rare variants associated with the trait over a random sampling[27].

6.1.3. Familial disorder as EP

Some common disorders show rare familial phenotypes with Mendelian inheritance associated with rare variants with large effect size. There are many studies using the EP strategy for familial cases of complex disorders, such as Alzheimer's disease(AD)[85], polyautoimmunity disorder[86], and congenital hypothyroidism[87]. For example, a recent study using linkage analysis demonstrated that by selecting individuals with familial autoimmunity and polyautoimmunity as EP, it was possible to identify the *SRA1* gene (LOD score = 5.48)[86]. Furthermore, a WES study on AD analyzed non-Hispanic White patients and Caribbean Hispanic families to find genes associated with early-onset AD. Heterozygous non-synonymous variants with global MAF < 0.001 were selected for variant prioritization and showed autosomal-dominant segregation in these families. Several genes such as *RUFY*, *TCIRG1*, *PSD2*, and *RIN3* were identified that could be involved in endolysosomal transport in both early- and late-onset AD[85].

In some complex diseases such as MD, there is also a strong evidence of genetic predisposition, as suggested by multiple reports describing affected families, showing an autosomal-dominant inheritance with 60% penetrance. By using WES in 46 families with MD, a burden of multiplex rare missense variants in the *OTOG* gene was found in 15 families (30%) of familial cases[88], which illustrates the success of considering familial cases as EP. Furthermore, a study on genetic epilepsy with hay febrile seizures plus (Dravet syndrome) has reported a *SCN1A* missense variant in a large Jewish family (14/17 cases) with epilepsy syndrome at both extremes (low and high)[31], and a study on thyroid dysgenesis with congenital hypothyroidism found a familial variant in *PAX8* gene associated with EP[87].

6.1.4. An EP Strategy to Investigate the Genetic Contribution to Tinnitus

Tinnitus is the perception of noise in the absence of an external acoustic stimulation, affecting more than 15% of the population and causing a decrease in health-related quality of life [48]. Several specific instruments have been defined to characterize chronic or severe tinnitus, and these instruments have been proposed to measure tinnitus annoyance to define EP for genetic studies[89]. Epidemiological evidence to support a genetic contribution to tinnitus is still weak because of the heterogeneous nature of this condition. In fact, tinnitus can occur together with multiple comorbidities including hearing loss, migraine, sleep disorders, anxiety, several psychological conditions, and some rare monogenic disorders[1]. The careful selection of phenotypes for genetic studies is crucial. The inclusion criteria should consider young individuals with severe forms of bilateral tinnitus to investigate the genetic contribution of rare variations to tinnitus. These individuals may carry a greater susceptibility and lower environmental load; however, severe forms of tinnitus in young individuals are rare[90] and multicenter studies are needed to reach a minimum sample size[91].

6.2. Synaptic genes involved in tinnitus

The present study reports a burden of rare missense and structural variants in several SG in patients with severe tinnitus. These genes are involved in cytoskeleton organization and cytoskeleton protein binding in neurons suggesting a novel mechanisms involved in tinnitus severity. In particular, a burden of missense rare and novel variants in *ANK2*, *AKAP9* and *TSC2* genes was found in Spanish MD patients with severe tinnitus (MD-EP), and this burden was replicated in a Swedish cohort of individuals with severe tinnitus. In addition, when we included all human genes in the burden analysis (hypothesis-free analysis) we also confirmed the association of *ANK2* and *TSC2* gene with MD-EP and MD-AEP. Using a large genetic generalized epilepsy cohort, we could confirm the specificity of these new genes to tinnitus.

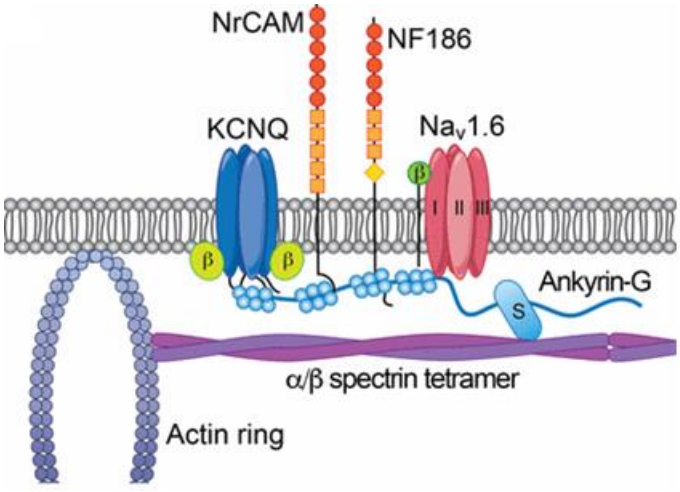
The synapse between sensory inner hair cells, primary auditory neurons and these neurons itself are potential candidates for tinnitus, but its perception and long term maintenance involves complex networks in the central nervous system, both in auditory and in non-auditory structures[92]. GO analyses suggest that membrane trafficking and cytoskeletal protein binding in neurons are involved at the molecular level. Future ongoing studies in a larger cohort of patients with severe tinnitus will confirm these predictions.

Tinnitus is associated with hearing loss in 90% of cases, according to standard pure tone audiograms. The most accepted causative model of tinnitus is based on the reduction in the auditory input associated with hearing loss, which leads to increased gain in the auditory pathway; that is, an amplification of spontaneous activity in the auditory neurons will lead to the perception of tinnitus[93]. This change in the intrinsic neuronal excitability after sensory deprivation occurs at the axon initial segments (AISs), the site of initiation of the action potential, which increase in length, and expression of voltage-dependent Na+ channels and Ankyrin-G, a membrane scaffolding protein encoded by the *ANK2* gene in the AISs[94].

6.2.1. Association of *ANK2* and axon initial segment with tinnitus severity

The *ANK2* (ENSG00000145362) gene, which is located at chromosome 4q25-q26, encodes Ankyrin-2, a large structural protein that carries death and ankyrin repeat containing domains. The Ankyrin gene has 46 exons in total and exon 37/38 is brain specific[95,96]. It belongs to the ankyrin family that links the integral proteins to the fundamental spectrin-actin cytoskeleton and plays an important role in different activities including micrometer scale organization of vertebrate plasma membranes in a broad spectrum of physiological contexts. *ANK2* encodes two different polypeptide including Ankyrin-2 (expressed in different tissues) and giant Ankyrin-2, a neuro-specific isoform variant expressed broadly in the central nervous system, with 2133 residues encoded by exon 37 between death and spectrin-binding domains[96].

Figure 15: Organization of the AIS and of Ankyrin-G[97].



Schematic of key components of the AIS. These include cell adhesion molecules (NrCAM and NF186) and ion channels (KCNQ and NaV) all bound to ankyrin repeats in the amino terminus. AnkG is, in turn, linked to the spectrin tetramer which is shown associated with an actin ring. Tetramers and the associated actin rings are spaced ~190 nm apart[97].

Giant Ankyrin-2 is a key protein to keep connectivity and neural activity in the central nervous system. It contributes to the development, maintenance and the refinement of neural circuits in different brain areas. The neural signals that arise at the AIS site regulate the neural activity. However, the lack of auditory input can cause an increase in the length of AISs ultimately leading to increase connectivity the auditory neurons in avian brainstem [94]. In addition, this is accompanied with an increase in whole-cell Na^+ current, membrane excitability and spontaneous firing. After auditory deprivation, the preservation of auditory function indicates that the change may have occurred at synaptic functionality level rather than at the structural level. However, the homeostatic changes occurring at AIS might play an important role to maintain the integrity of the remaining neurons in auditory circuits[94], something that may also occur in severe tinnitus.

Rare variations in ankB isoform may produce an increase of axonal branching[96]. In humans, rare variants in *ANK2* gene have previously been reported in individuals with autism spectrum disorder (ASD) [96] and long QT syndrome[98]. Although hearing phenotypes are poorly investigated in ASD, reports point towards increased loudness sensitivity, or hyperacusis in ASD[99]. Since severe tinnitus is highly associated with hyperacusis[53], it is thus plausible that *ANK2* may be involved in sensory gating processes, which are thought to underlie severe tinnitus with increased loudness[100]. A variant in *ANK2* (4: 114294537 G>A) has been previously reported for autosomal cardiac arrhythmia in 2 heterozygous Ashkenazi Jewish centenarians[101]. This same variant was also found in one individual with MD-EP in our cohort.

In another study of long QT, a novel variant c.1937 C>T, exon 18 in the membrane-binding domain was found in 2 Canadian Gitxsan multigenerational families. This variant showed a loss of function activity in myocytes affected expression and localization of downstream binding partners[98]. In addition, three mutations in human *ANK2*, exon 37 have previously been reported for ASD, affecting only the giant ankB isoform. Further, mouse models with these human *ANK2* mutations have evidenced that the giant ankB may contribute to maintain the normal structure connectivity in central nervous system. The increase of axon branching might be a potential cellular mechanism to explain ASD[96], and thus plausibly severe tinnitus.

Epidemiological and genetic studies consistently support that severe tinnitus has a genetic contribution, and common and rare variants with epistatic effects shape the phenotype[39,4]. A recent GWAS using a broad definition of tinnitus found a small number of loci and common variants with small effect sizes[41]. Moreover, the MVP cohort used for replication does not represent the general US population, since these individuals report a history of occupational chronic noise exposure and have an increased prevalence of traumatic brain injury, both known risk factors for tinnitus. Although we did not find any of these genes in our burden analysis, this GWAS is a major contribution to define the genetic architecture of noise-induced hearing loss and tinnitus, but their findings cannot be extended to any tinnitus phenotype. These findings are probable related to noise-induced hearing loss and tinnitus, but not necessarily to severe tinnitus.

Our study emphasizes the need of larger genetic studies using severe tinnitus as inclusion criteria. However, the proportion of individuals in existing biobanks with such definition of tinnitus is rare, and thus new efforts to biobank tinnitus are needed[102]. The present study also calls for the need of new definitions of tinnitus as a disorder that would distinguish the rare pathogenic form of tinnitus (present in 1% of the population) from the more symptomatic and common form reported by a large proportion of the population.

Tinnitus as a neurological disorder may not only result from sensory deprivation as it probably occurs in high-frequency hearing loss or MD, or after synaptic reorganization that lead to changes on the neuronal excitability at different brain areas, but also from enhanced connectivity with non-auditory brain regions as it is often observed in tinnitus patients or individuals with severe tinnitus[103].We observed expression of *ANK2* in a number of distinct auditory and non-auditory brain regions within the mouse brain. Interestingly marked expression was found in the auditory pathway, including the auditory cortex, the dorsal cortex of the inferior colliculus, the medial nucleus of the trapezoid body and the dorsal cochlear nucleus. These finding support that multiple nuclei in the auditory pathway could be involved in the development of severe tinnitus.

However, we found stronger expression of *ANK2* in a number of non-auditory brain regions that have been associated with tinnitus including the cortex, hippocampus and the cerebellum [104,105]. As previously mentioned, in neurons, the expression of *ANK2* is found in the AIS [106] and is believed to play a role in the development and function of axonal branching[96] and synaptic connectivity[107]. This role in the regulation of neuronal connectivity could potentially

explain why it has been associated with the neurological disorders such as epilepsy[108]and ASD[109]**.** Interestingly, the cortal expression for *TSC2* was found to be strongest in projection layers (2/3, 5 and 6) and weakest in input layers (1 and 4), potentially indicating a role in cortical connectivity[110]. This, combined with the fact tinnitus has been associated with abnormal functional connectivity between brain regions[111,103], may be a tantalising hint for a link between neural connectivity and tinnitus. Rare missense variants in *ANK2* gene may thus reflect subtle structural differences in the axonal cytoskeleton protein organization involving axonal branching and neuronal connectivity, which would be exacerbated in individuals with severe tinnitus. Mice lacking ank2 constitutively die at birth from developmental defects, and thus conditional models with depletion in adulthood would be required to back-translate these findings in animal models. Further work is thus required to understand the role of *ANK2* in the presentation of tinnitus symptoms.

In addition, we have found *PRAMEF1* which is a PRAME Family Member 1 gene and responsible for reproductive tissues during development and many types of cancer. *PRAMEF1* was previously reported for Lynch syndrome endometrial cancer in Chinese population[112]. However, this gene is not expressed in the brain according to GTEx portal (https://www.gtexportal.org/home/).

6.2.2. Contribution of *AKAP9* and *TSC2* in tinnitus

Interestingly, we have found a burden of rare variation in *AKAP9*, another gene previously associated with long-QT syndrome[113]. *AKAP9* encodes A-kinase anchor protein 9, a member of the A-kinase anchor protein family member, whose known function is binding to the protein kinase A (PKA) regulatory subunit with the objective of enclose it to different parts of the cell where phosphorylation is needed[114]. Moreover, *AKAP9* is known to form a multiprotein complex with *KCNQ1*, PKARII and PP1 to translate PKA to the ion channel and phosphorylate a serine residue for its activation[114,115]. Known mutations in *AKAP9* gene disrupt this behaviour and difficult the interaction between a-kinase anchor protein 9, Potassium voltage-gated channel subfamily Q member 1, leading to a dysfunction in the regulation of K^+ entry through this channel and, therefore, pathological consequences such as long-QT syndrome.

Our study also reveals a significant enrichment of rare variants in *TSC2* gene in patients with MD-EP and severe tinnitus. Tuberous Sclerosis Complex 2 (also known as TSC2 or Tuberin)

is a known tumour suppressor protein part of the tuberous sclerosis complex (TSC) along with TSC1. This complex is involved in the negative regulation of mTORC1 activity. Loss of tuberin function causes constitutive activation of the mTORC1 signalling pathway leading to tuberous sclerosis tumours[116]. *TSC2* loss of function mutations cause misregulation of the endocytosis processes and disruption on autophagy-lysosome pathway [116].

Of note, the TSC complex has also been related with several neurological abnormalities such epilepsy or autism[117]. In fact, some studies show that overexpression of the TSC complex supresses axon formation while the lack of either TSC1 or TSC2 induces the appearance of ectopic axons[118]. Similar to *ANK2,* ISH data revealed, in the mouse brain, that *TSC2* expression was greatest in projection layers (2/3, 5 and 6) and weakest in the input layers (1 and 4). Given these are the neurons demonstrating long-range projecting axons this expression profile seemingly fits with this believed function.

The regulation of the Mtorc1 pathway via the TSC complex has been found to be a key part in some age-associated diseases, including age-related hearing loss. Some studies describe how mtorc1 constitutive activity in the aging cochlear neurosensory epithelium leads to hearing loss in mice, while inactivation of this pathway reduces age-related HL in those mice[119]. Regarding axonal formation, tsc2 heterozygous mice also have a specific deficit in protein synthesis in neurons of the hippocampus causing synaptic depression because of the lack of glutamate receptors. Those mice showed abnormal synaptic plasticity and cognition due to the suboptimal metabotropic glutamate receptor-mediated protein synthesis leading to behavioural impairments[120]. Finally, we found highly significant co-expression of ank2 and tsc2 across the mouse brain, potentially suggesting they are expressed in similar neuronal subtypes.

Mouse brain co-expression of ank2 and tsc2 was particularly strong in limbic brain regions (i.e. the hypothalamus, epithalamus, striatum, pallidum and hippocampus), that form a complex circuit distributed across the brain. In addition, strong expression was found across cortex particularly in cortical layers generally associated with cortical projections (i.e. layers 2/3, 5 and 6). Limbic (particularly ventral striatum and the hippocampus) and cortical contributions have been previously implicated in the manifestation of tinnitus[121]. The strong expression in cortical projection layers fits with the existing literature demonstrating the involvement of ank2 and tsc2 in axonal function. While more work is required these data

together hint at a potential network based dysfunction of tinnitus, involving projections to and between auditory (e.g. cortical) and non-auditory (e.g. limbic) brain regions.

6.2.3. Severe tinnitus and sex differences

Few studies have been conducted to analyse the effect of sex in genetic contribution to tinnitus. A study on Swedish twins found that the heritability in women with bilateral tinnitus increased from 0.41 to 0.62 when they selected only young women, suggesting that early onset is probably associated with a higher heritability [90]. Another study on Swedish families revealed a higher recurrence risk ratio in women indicating that this group is more susceptible to severe tinnitus as compared to men. These studies highlight the genetic contribution to severe tinnitus is different in men and woman in sex dependent manner[40].

In addition to these studies, the genetic findings with MD-EP cohort also provide the evidence that women are at high risk to develop severe tinnitus. This cohort has 23/30 women cases and the carriers of rare and novel variants in *ANK2* are all unrelated women, suggesting sexual differences in the heritability of tinnitus. However, the carriers of *ANK2* variants in Swedish tinnitus cohort are both women and men. Further studies are needed in a larger sample size to establish if the burden of rare variants in *ANK2* is specific for women.

6.2.4. Tinnitus, hearing loss and cognitive decline

It is believed that individuals with tinnitus may face difficulties with memory and several studies have reported the poor cognitive performance in individuals with tinnitus[122]. In addition to tinnitus, hearing loss is accountable for 9% of dementia cases worldwide[123]. Hearing loss is reported as an independent risk factor for dementia but the association between tinnitus an cognitive decline is still unclear [122].

The possible mechanism of dementia associated with tinnitus and hearing loss are briefly explained as follows.

- As previously explained in section 6.2, the auditory input deprivation associated with hearing loss is one of the possible mechanisms involved in tinnitus generation. This

auditory deprivation causes a gain increase in the auditory pathway amplifying the spontaneous activity of neurons which is perceived as tinnitus.

- Cochlear damage is the most common cause of hearing loss whereas dementia is caused by cortical degeneration with loss of neurons, particularly of the multimodal cortex. However, what is the link between these conditions? To answer this crucial question several mechanisms have been proposed in different studies.
    1. The first mechanism could be the common pathology that affects cochlea, nuclei in the ascending auditory pathway and cortex.
    2. The second possible mechanism could be the decreased stimulation of the cognitive processing caused by hearing loss. The auditory deprivation can create an impoverished environment particularly involving language and speech, affecting negatively several brain structures and its functions.
    3. The third mechanism focuses on the unavailability of cognitive resources being utilized due to hearing loss. Individuals with hearing loss require high utilization of cognitive resources for listening, making unavailability of resources for other tasks i.e. attention, working memory and language processing.

However, the main difference between second and third mechanism is that; the second mechanism cause changes in brain structure and neuronal mechanism before onset of dementia; and the third mechanism cause changes brain function during dementia explaining cognitive decline[123,124].

In addition to these studies, we have found a burden of rare variants in 4 genes that have been previously associated with dementia including *BIN1*[125], *NOTCH3*[126] and *SRRM4*[127] in MD-EP cohort and *ANK1*[128] in MD-AEP cohort,. However, future replication studies are required to validate the association of tinnitus as a risk factor in the development of neurodegenerative disorders.

6.3. Limitations

Some weaknesses were found in the design of EP strategies; therefore, further research is required. The replication of the genetic studies across different populations with different ethnic backgrounds has enough potential to validate genetic associations; however, the frequency of allelic variants is different across different populations, and specific reference data for allelic frequencies are needed for each population. The rare variants reported in simplex families with EPs should be validated in more patients with a severe phenotype. Most of the studies used WES rather whole-genome sequencing (WGS) and this can cause the loss of useful genetic information and erroneous results in calculating the effect size of rare variants at the individual level across a particular phenotype.

As a first limitation, our design is an exome-based study and we might be missing some other variants in non-coding regions that may have a significant contribution to severe tinnitus. Future studies will require whole genome sequencing in order to fine-map regulatory regions and their involvement in this neurological disorder.

Second, the EP strategy is based on the selection of individuals at the extremes of the phenotype distribution, who are expected to carry an enrichment of rare pathogenic variants, and they are not representative of the phenotypic variance observed in MD.

Third, the Swedish cohort used for replication does not represent a clinical sample, and therefore the assessment did not involve a physician.

Fourth, most of our MD patients were females and future studies should consider gender differences in the genetic contribution to tinnitus.

Finally, indel analysis is highly dependent on the callers and annotation tools used for each frequency database compared, and reference databases need to be extended. Most of the indels in our dataset were not found in either the gnomAD population database or the CSVS database and were ruled out to avoid an inflation factor in the GBA.

Conclusions

1. A systematic review of genetic studies using extreme phenotype approach has shown evidence that the extreme phenotype strategy is a useful approach to establish the genetic contribution of rare variations to complex diseases.

2. Several synaptic genes have a burden of rare missense variants in coding regions in patients with severe tinnitus and Meniere disease. The main genes showing this enrichment include *ANK2*, *AKAP9* and *TSC2*.

3. The burden of rare variants in *ANK2* was replicated in an independent non-Meniere disease severe tinnitus cohort from Sweden, but not in a generalised Epilepsy cohort, demonstrating the specificity of our findings for tinnitus, regardless of the underlying disease.

4. The potential pathways and biological processes predicted include membrane trafficking and cytoskeletal protein binding.

Conclusiones

1. Una revisión sistemática de estudios genéticos que han utilizado el fenotipo extremo ha confirmado la efectividad de la estrategia de seleccionar individuos con fenotipo extremo para establecer la contribución genética de las variantes raras en las enfermedades complejas.

2. Varios genes sinápticos tienen una carga de variantes raras sin sentido en las regiones de codificación en pacientes con acúfeno severo y enfermedad de Meniere. Entre los principales genes que muestran este enriquecimiento se incluyen *ANK2*, *AKAP9* y *TSC2*.

3. La carga de variantes raras en *ANK2* se replicó en una cohorte independiente de acúfeno severo sin enfermedad de Meniere de Suecia, pero no en la cohorte de epilepsia generalizada genética, lo cual demuestra la especificidad de nuestros hallazgos para el acúfeno, independientemente de la enfermedad subyacente.

4. Las posibles vías y procesos biológicos predichos incluyen el tranporte a traves de membranas y la unión a proteínas citoesqueléticas.

References

1 Vona B, Nanda I, Shehata-Dieler W, Haaf T. Genetics of tinnitus: Still in its infancy. *Front. Neurosci.* 2017. doi:10.3389/fnins.2017.00236.

2 Møller AR. Pathophysiology of tinnitus. *Otolaryngol Clin North Am.* 2003; **36**: 249–266.

3 Amanat S, Gallego-Martinez A, Lopez-Escamez JA. Genetic Inheritance and Its Contribution to Tinnitus. *Curr Top Behav Neurosci*.2020. doi:10.1007/7854_2020_155.

4 Lopez-Escamez JA, Amanat S. Heritability and Genetics contribution to tinnitus. *Otolaryngol Clin NA.* 2020; **53**: 501–513.

5 Ridder D De, Schlee W, Vanneste S, Londero A, Bel A, Li LP *et al.* Tinnitus and tinnitus disorder : Theoretical and operational definitions (an international multidisciplinary proposal). *Elsevier*.2021. doi:10.1016/bs.pbr.2020.12.002.

6 Pickles JO. Auditory Pathways : Anatomy and Physiology. *Handbook Clin Neurol*. 2015; **129**; p. 3-25.

7 Peterson DC, Reddy V, Hamel RN. 2020. Neuroanatomy, auditory pathway. *StatPearls [Internet]* . 2020.

8 Møller AR, Langguth B, DeRidder D, KleinjungT. eds. Textbook of tinnitus .*Springer Science & Business Media*. 2010.Chapter 8: Anatomy and Physiology of the Auditory System; p.51-68.

9 Møller AR, Langguth B, DeRidder D, KleinjungT. eds.Textbook of tinnitus .*Springer Science & Business Media*.2010.Chapter 10: Pathology of the Auditory System that Can Cause Tinnitus; p.77-93

10 Møller AR, Langguth B, DeRidder D, KleinjungT. eds. Textbook of tinnitus .*Springer Science & Business Media*. 2010.Chapter 11: The Role of Auditory Deprivation; p.95-98.

11 Møller AR, Langguth B, DeRidder D, KleinjungT. eds. Textbook of tinnitus .*Springer Science & Business Media*. 2010.Chapter 12: The Role of Neural Plasticity in Tinnitus; p.99-102.

12 Møller AR, Langguth B, DeRidder D, KleinjungT. eds. Textbook of tinnitus .*Springer Science & Business Media*. 2010.Chapter 13:Neural Synchrony and Neural Plasticity in Tinnitus; p.103-112.

13 Ueberfuhr MA, Wiegrebe L, Krause E, Gürkov R, Drexl M. Tinnitus in normal-hearing participants after exposure to intense low-frequency sound and in Ménière's disease patients. *Front Neurol* 2017; **7**: 1–11.

14 Herraiz C, Tapia MC, Plaza G. Tinnitus and Ménière's disease: Characteristics and prognosis in a tinnitus clinic sample. *Eur Arch Oto-Rhino-Laryngology.* 2006; **263**: 504–509.

15 Møller AR, Langguth B, DeRidder D, KleinjungT. eds. Textbook of tinnitus.*Springer Science & Business Media.* 2010.Chapter 38:Tinnitus and Meniere disease; p.311-316.

16 Pérez-Gracia JL, Gúrpide A, Ruiz-Ilundain MG, Alegría CA, Colomer R, García-Foncillas J *et al.* Selection of extreme phenotypes: The role of clinical observation in translational research. *Clin Transl Oncol* 2010; **12**: 174–180.

17 Zimmermann E, Gamborg M, Sørensen TIA, Baker JL. Sex differences in the association between birth weight and adult type 2 diabetes. *Diabetes.* 2015; **64**: 4220–4225.

18 Amanat S, Requena T and Lopez-Escamez J A.A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders. *Genes.* 2020, **11**: 987.

19 Gibson G. Rare and common variants: Twenty arguments. *Nat Rev Genet.* 2012; **13**: 135–145.

20 Lohmann K, Klein C. Next Generation Sequencing and the Future of Genetic Diagnosis. *Neurotherapeutics.* 2014; **11**: 699–707.

21 Srivastava S, Cohen JS, Vernon H, Barañano K, McClellan R, Jamal L *et al.* Clinical whole exome sequencing in child neurology practice. *Ann Neurol.* 2014; **76**: 473–483.

22 Jansen IE, Ye H, Heetveld S, Lechler MC, Michels H, Seinstra RI *et al.* Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing. *Genome Biol.* 2017;**18**:1–26.

23 Johnson MB, Patel KA, De Franco E, Houghton JAL, McDonald TJ, Ellard S *et al.* A type 1 diabetes genetic risk score can discriminate monogenic autoimmunity with diabetes from early-onset clustering of polygenic autoimmunity with diabetes. *Diabetologia.* 2018; **61**: 862–869.

24 Verma SS, Ritchie MD. Another round of "clue" to uncover the mystery of complex traits. *Genes (Basel).* 2018; **9**. doi:10.3390/genes9020061.

25 Craddock N, Kendler K, Neale M, Nurnberger J, Purcell S, Rietschel M *et al.* Dissecting the phenotype in genome-wide association studies of psychiatric illness. *Br J Psychiatry*.2009; **195**: 97–99.

26 Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA *et al.* Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet.* 2011; **12**: 745–755.

27 Bjørnland T, Bye A, Ryeng E, Wisløff U, Langaas M. Improving power of genetic association studies by extreme phenotype sampling: a review and some new results. *arXiv preprint.*2017. arXiv:1701.01286. 2017

28 Barnett IJ, Lee S, Lin X. Detecting Rare Variant Effects Using Extreme Phenotype Sampling in Sequencing Association Studies. *Genet Epidemiol.* 2013; **37**: 142–151.

29 Emond MJ, Louie T, Emerson J, Zhao W, Mathias RA, Knowles MR *et al.* Exome sequencing of extreme phenotypes identifies DCTN4 as a modifier of chronic Pseudomonas aeruginosa infection in cystic fibrosis. *Nat Genet*.2012; **44**: 886–889.

30 Bjørnland T, Langaas M, Grill V, Mostad IL. Assessing gene-environment interaction effects of FTO, MC4R and lifestyle factors on obesity using an extreme phenotype sampling design: Results from the HUNT study. *PLoS One.* 2017; **12**: 1–16.

31 Goldberg-Stern H, Aharoni S, Afawi Z, Bennett O, Appenzeller S, Pendziwiat M *et al.* Broad phenotypic heterogeneity due to a novel SCN1A mutation in a family with genetic epilepsy with febrile seizures plus. *J Child Neurol.* 2014; **29**: 221–226.

32 Shtir C, Aldahmesh MA, Al S, Emad D, Alkuraya H, Abouammoh MA *et al.* Exome - based case – control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum Genet.* 2016; **135**: 193–200.

33 Husson T, Duboc JB, Quenez O, Charbonnier C, Rotharmel M, Cuenca M *et al.* Identification of potential genetic risk factors for bipolar disorder by whole-exome sequencing. *Transl Psychiatry.* 2018; **8**. doi:10.1038/s41398-018-0291-7.

34 Blue E, Louie TL, Chong JX, Hebbring SJ, Barnes KC, Rafaels NM *et al.* Variation in cilia protein genes and progression of lung disease in cystic fibrosis. *Ann Am Thorac Soc.* 2018; **15**: 440–448.

35 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* Finding the missing heritability of complex diseases. *Nature.* 2009; **461**: 747–753.

36 Lopez-Escamez JA, Carey J, Chung WH, Goebel JA, Magnusson M, Mandalà M *et al.* Diagnostic criteria for Menière's disease. *J Vestib Res Equilib Orientat.* 2015; **25**: 1–7.

37 Richards S, Aziz N, Bale S, Bick D, Das S. ACMG Standards and Guidelines Standards and guidelines for the interpretation of sequence variants : a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.*Genet Med*.2015;**17**:405-423.

38 Walsh R, Mazzarotto F, Whiffin N, Buchan R, Midwinter W, Wilk A *et al.* Quantitative approaches to variant classification increase the yield and precision of genetic testing in Mendelian diseases : the case of hypertrophic cardiomyopathy. *Genome Med* .2019;**11**: 1–18.

39 Cederroth CR, Pirouzifard M, Trpchevska N, Idrizbegovic E, Canlon B, Sundquist J *et al.* Association of Genetic vs Environmental Factors in Swedish Adoptees with Clinically Significant Tinnitus. *JAMA Otolaryngol - Head Neck Surg.* 2019; **145**: 222–229.

40 Trpchevska N, Bulla J, Hellberg MP, Edvall NK, Lazar A, Mehraei G *et al.* Sex-Dependent Aggregation of Tinnitus in Swedish Families. *J Clin Med.* 2020;**9:**3812

41 Clifford RE, Maihofer AX, Stein MB, Ryan AF, Nievergelt CM. Novel Risk Loci in Tinnitus and Causal Inference With Neuropsychiatric Disorders Among Adults of European Ancestry. *JAMA Otolaryngol - Head Neck Surg.* 2020; **146**:1015-1025. doi:10.1001/jamaoto.2020.2920.

42 Sand PG, Luettich A, Kleinjung T, Hajak G, Langguth B. An examination of KCNE1 mutations and common variants in chronic tinnitus. *Genes (Basel.)* 2010; **1**: 23–37.

43 Deniz M, Bayazit YA, Celenk F, Karabulut H, Yilmaz A, Gunduz B *et al.* Significance of serotonin transporter gene polymorphism in tinnitus. *Otol Neurotol* 2010; **31**: 19–24.

44 Sand PG, Langguth B, Kleinjung T. Deep resequencing of the voltage-gated potassium channel subunit KCNE3 gene in chronic tinnitus. *Behav Brain Funct.* 2011; **7**: 1-5.

45 Sand PG, Langguth B, Schecklmann M, Kleinjung T. GDNF and BDNF gene interplay in chronic tinnitus. *Int J Mol Epidemiol Genet.* 2012; **3**: 245–251.

46 Sand PG, Langguth B, Itzhacki J, Bauer A, Geis S, Cárdenas-Conejo ZE *et al.* Resequencing of the auxiliary GABAB receptor subunit gene KCTD12 in chronic tinnitus. *Front Syst Neurosci.* 2012; **6**: 1–5.

47 Orenay-Boyacioglu S, Coskunoglu A, Caki Z, Cam FS. Relationship Between Chronic Tinnitus and Glial Cell Line-Derived Neurotrophic Factor Gene rs3812047, rs1110149, and rs884344 Polymorphisms in a Turkish Population. *Biochem Genet.* 2016; **54**: 552–563.

48 Gilles A, Camp G, Van de Heyning P, Fransen E. A pilot genome-wide association study identifies potential metabolic pathways involved in tinnitus. *Front Neurosc.i* 2017; **11**: 1–10.

49 Moher D. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement David. *Syst Rev.* 2015; **207**: 1–9.

50 Herráiz C, Plaza G, Tapia MC.Evaluación de la incapacidad en pacientes con acúfenos. *Acta Otorrinolaringoló Esp.* 2001; **52**: 534-538.

51 Newman CW, Sandridge SA, Jacobson GP. Psychometric adequacy of the Tinnitus Handicap Inventory (THI) for evaluating treatment outcome. *J Am Acad Audiol.* 1998; **9**: 153–160.

52 Perez-Carpena P, Martinez-Martinez M, Martínez Carranza R A, Batuecas-Caletrio A, Lopez-Escamez JA. A tinnitus symphony in 100 patients with Meniere's disease. *Clin Otolaryngol.* 2019; **44;**1176-1180.

53 Cederroth CR, Lugo A, Edvall NK, Lazar A, Bulla J, Uhlen I *et al.* Association between Hyperacusis and Tinnitus. *J Clin Med.*2020;**9** : 2412

54 Almqvist C, Paul HA, Leif WF, Erik G. LifeGene -a large prospective population-based study of global relevance. *Eur J Epidemiol.* 2011; **26**: 67–77.

55 Fackrell K, Hall DA, Barry JG, Hoare DJ. Performance of the Tinnitus Functional Index as a diagnostic instrument in a UK clinical population. *Hear Res.* 2018;**358** : 74-85

56 May P, Girard S, Harrer M, Bobbili DR, Schubert J, Wolking S *et al.* Rare coding variants in genes encoding GABA A receptors in genetic generalised epilepsies: an exome-based case-control study. *Lancet Neurol.* 2018; **17**: 699–708.

57 Szczepek AJ, Frejo L, Vona B, Trpchevska N, Cederroth CR, Caria H *et al.* Recommendations on Collecting and Storing Samples for Genetic Studies in Hearing and Tinnitus Research. *Ear Hear.* 2019;**40** : 219–226.

58 Depristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C *et al.* technical reports A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat genet.* 2011; **43**. doi:10.1038/ng.806.

59 Li MX, Gui HS, Kwan JSH, Bao SY, Sham PC. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. *Nucleic Acids Res.* 2012; **40**. doi:10.1093/nar/gkr1257.

60 Wang K, Li M, Hakonarson H. ANNOVAR : functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010; **38**.

61 Pirooznia M, Wang T, Avramopoulos D, Valle D, Thomas G, Huganir RL *et al.* SynaptomeDB : an ontology-based knowledgebase for synaptic genes. *Bioinformatics.* 2012; **28**: 897–899.

62 Azaiez H, Booth KT, Ephraim SS, Crone B, Black-ziegelbein EA, Marini RJ *et al.* Genomic Landscape and Mutational Signatures of Deafness-Associated Genes. *Am J Hum Genet.* 2018;

103: 484–497.

63 Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans.*Nature.* 2020; **581**. doi:10.1038/s41586-020-2308-7.

64 Peña-Chilet M, Roldán G, Perez-Florido J, Ortuño FM, Carmona R, AquinoV *et al*. CSVS, a crowdsourcing database of the Spanish population genetic variability. *Nucleic Acids Res.* 2020;**49**.

65 Shyr C, Tarailo-graovac M, Gottlieb M, Lee JJY, Karnebeek C Van. FLAGS , frequently mutated genes in public exomes. *BMC Med Genet.* 2014;**7** : 1–14.

66 Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z *et al.* Sequence analysis IBS : an illustrator for the presentation and visualization of biological sequences.*Bioinformatics.* 2015; **31**: 3359–3361.

67 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005; **102**: 15545–15550.

68 Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A *et al.* Genome-wide atlas of gene expression in the adult mouse brain.*Nature.* 2007; **445**. doi:10.1038/nature05453.

69 Grange P, Bohland JW, Hawrylycz M, Mitra PP. Brain Gene Expression Analysis: a MATLAB toolbox for the analysis of brain-wide gene-expression data. *arXiv preprint* .2012.http://arxiv.org/abs/1211.6177.

70 Goldowitz D. Allen Reference Atlas. A Digital Color Brain Atlas of the C57BL/6J Male Mouse - by H. W. Dong. *Genes, Brain Behav.* 2010; **9**: 128–128.

71 Bruse S, Moreau M, Bromberg Y, Jang JH, Wang N, Ha H *et al.* Whole exome sequencing identifies novel candidate genes that modify chronic obstructive pulmonary disease susceptibility. *Hum Genomics.* 2016; **10**: 1–12.

72 Pullabhatla V, Roberts AL, Lewis MJ, Mauro D, Morris DL, Odhams CA *et al.* De novo mutations implicate novel genes in systemic lupus erythematosus.*Hum Mol Genet.* 2018; **27**: 421–429.

73 International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature.* 2004;**431**: 931–945.

74 Zou J, Mathur PD, Zheng T, Wang Y, Almishaal A, Park AH *et al.* Individual USH2 proteins make distinct contributions to the ankle link complex during development of the mouse cochlear stereociliary bundle. *Hum Mol Genet.*2015; **24**: 6944–6957.

75 Han JY, Lee HJ, Lee Y, Park J. Identification of Missense ADGRV1 Mutation as a Candidate Genetic Cause of Familial Febrile Seizure 4.*Children.* 2020;**7**:144

76 Koenekoop R, Arriaga M, Trzupek KM, et al. Usher Syndrome Type II. 1999 Dec 10 [Updated 2020 Oct 22]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. *GeneReviews® [Internet]*. Seattle (WA): University of Washington, Seattle; 1993-2021.

77 Wu D, Huang W, Xu Z, Chen X, Tang Y, Qiu J *et al.* Clinical and genetic study of 12 Chinese Han families with nonsyndromic deafness. *Mol Genet Genomic Med.* 2020; **8**.

78 Allen AS, Bellows ST, Berkovic SF, Bridgers J, Burgess R, Cavalleri G *et al.* Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol.* 2017; **16**: 135–143.

79 Forsberg LA, Rasi C, Razzaghian HR, Pakalapati G, Waite L, Thilbeault KS *et al.* Age-related somatic structural changes in the nuclear genome of human blood cells. *Am J Hum Genet* .2012; **90**: 217–228.

80 Castellani C, Assael BM. Cystic fibrosis : a clinical view. *Cell Mol Life Sci.* 2016. doi:10.1007/s00018-016-2393-9.

81 Fuseini H, Newcomb DC. Mechanisms Driving Gender Differences in Asthma. *Curr Allergy Asthma Rep.* 2017; **17**. doi:10.1007/s11882-017-0686-1.

82 Kang G, Lin D, Hakonarson H, Chen J. Two-stage extreme phenotype sequencing design for discovering and testing common and rare genetic variants: Efficiency and power. *Hum Hered.* 2012; **73**: 139–147.

83 Li D, Lewinger JP, Gauderman WJ, Murcray CE, Conti D. Using extreme phenotype sampling to identify the rare causal variants of quantitative traits in association studies. *Genet Epidemiol.* 2011; **35**: 790–799.

84 Guey LT, Kravic J, Melander O, Burtt NP, Laramie JM, Lyssenko V *et al.* Power in the phenotypic extremes: A simulation study of power in discovery and replication of rare variants. *Genet Epidemiol.* 2011; **35**: 236–246.

85 Kunkle BW, Vardarajan BN, Naj AC, Whitehead PL, Rolati S, Slifer S *et al.* Early-onset Alzheimer disease and candidate risk genes involved in endolysosomal transport. *JAMA Neurol.* 2017; **74**: 1113–1122.

86 Johar A, Sarmiento-Monroy JC, Rojas-Villarraga A, Silva-Lara MF, Patel HR, Mantilla RD *et al.* Definition of mutations in polyautoimmunity. *J Autoimmun.* 2016; **72**: 65–72.

87 De Sanctis L, Corrias A, Romagnolo D, Di Palma T, Biava A, Borgarello G *et al.* Familial PAX8 small deletion (c.989_992delACCC) associated with extreme phenotype variability. *J Clin Endocrinol Metab.* 2004; **89**: 5669–5674.

88 Roman-naranjo, P. Gallego-Martinez, A.Soto-Varela, A.; Aran, I. del Carmen Moleon, M. Espinosa-Sanchez, J.M. Amor-Dorado, J.C. Batuecas-Caletrio, A. Perez-Vazquez, P. Lopez-Escamez, J.A. Rare variants in the OTOG gene are a frequent cause of familial Meniere ' s disease. *Ear Hear.* 2020.

89 Lopez-Escamez JA, Bibas T, Cima RFF, Van de Heyning P, Knipper M, Mazurek B *et al.* Genetics of tinnitus: An emerging area for molecular diagnosis and drug development. *Front Neurosci.* 2016; **10**: 1–13.

90 Maas IL, Brüggemann P, Requena T, Bulla J, Edvall NK, Hjelmborg JVB *et al.* Genetic susceptibility to bilateral tinnitus in a Swedish twin cohort. *Genet Med.* 2017; **19**: 1007–1012.

91 Schlee W, Hall DA, Canlon B, Cima RFF, Kleine E De, Hauck F *et al.* Innovations in Doctoral Training and Research on Tinnitus : The European School on Interdisciplinary Tinnitus Research (ESIT) Perspective. *Front Aging Neurosci.* 2018; **9**: 447

92 Guitton MJ. Tinnitus : pathology of synaptic plasticity at the cellular and system

levels.*Neuroscience.* 2012; **6**: 12.

93 Schaette R, Kempter R. Development of tinnitus-related neuronal hyperactivity through homeostatic plasticity after hearing loss : a computational model. *Eur J Neurosci.*2006; **23**: 3124–3138.

94 Kuba H, Oichi Y, Ohmori H. Presynaptic activity regulates Na^+ channel distribution at the axon initial segment. *Nature.* 2010; **465**: 1075–1078.

95 Mohler PJ, Gramolini AO, Bennett V. The ankyrin-B C-terminal domain determines activity of ankyrin-B/G chimeras in rescue of abnormal inositol 1, 4, 5-trisphosphate and ryanodine receptor distribution in ankyrin-B (−/−) neonatal cardiomyocytes.*J Biol Chem.* 2002; **277**: 10599–10607.

96 Yang R, Walder-christensen KK, Kim N, Wu D, Lorenzo DN. *ANK2* autism mutation targeting giant ankyrin-B promotes axon branching and ectopic connectivity. *Proc Natl Acad Sci*. 2019; **116**. doi:10.1073/pnas.1904348116.

97 Salzer JL. An unfolding role for ankyrin-G at the axon initial segment. *Proc Natl Acad Sci.* 2019; **116**: 19228–19230.

98 Swayne LA, Murphy NP, Asuri S, Chen L, Xu X, Mcintosh S *et al.* Novel Variant in the ANK2 Membrane-Binding Domain Is Associated With Ankyrin-B Syndrome and Structural Heart Disease in a First Nations Population With a High Rate of Long QT Syndrome. *Circ Cardiovasc Genet.* 2017;**10**.

99 Bruneau N, Roge B, Veuillet E, Adrien J, Barthe C. Increased perception of loudness in autism. *Hear Res.*2004; **198**: 87–92.

100 Vanneste S, Song JJ, Ridder D De. Thalamocortical dysrhythmia detected by machine learning. *Nat Commun.*2018;**9**:1-13doi:10.1038/s41467-018-02820-0.

101 Freudenberg-hua Y, Freudenberg J, Vacic V, Abhyankar A, Emde A, Ben-avraham D *et al.* Disease variants in genomes of 44 centenarians. *Mol Genet Genomic Med* .2014;**2**:438-450. doi:10.1002/mgg3.86.

102 Cederroth CR, Kähler AK, Sullivan PF, Lopez-Escamez JA. Genetics of tinnitus: Time to biobank phantom sounds. *Front Genet.* 2017; **8**: 1–3.

103 Schlee W, Weisz N, Bertrand O, Hartmann T, Elbert T. Using Auditory Steady State Responses to Outline the Functional Connectivity in the Tinnitus Brain. *PLoS One.*2008; **3**. doi:10.1371/journal.pone.0003720.

104 Mennink LM, Dijk JMC Van, Dijk P Van. The cerebellar (para) flocculus : A review on its auditory function and a possible role in tinnitus. *Hear Res.* 2020; **398**: 108081.

105 Landgrebe M, Langguth B, Rosengarth K, Braun S, Koch A, Kleinjung T *et al*. Structural brain changes in tinnitus: grey matter decrease in auditory and non-auditory brain areas. *Neuroimage*. 2009;**46**:213-8. doi: 10.1016/j.neuroimage.2009.01.069.

106 Galiano MR, Jha S, Ho TS, Zhang C, Ogawa Y, Chang K *et al.* A Distal Axonal Cytoskeleton Forms an Intra-Axonal Boundary that Controls Axon Initial Segment Assembly. *Cell.* 2012; **149**: 1125–1139.

107 Pielage J, Cheng L, Fetter RD, Carlton PM, Sedat JW, Davis GW. Article A Presynaptic Giant Ankyrin Stabilizes the NMJ through Regulation of Presynaptic Microtubules and Transsynaptic Cell Adhesion.*Neuron.*2008;**58**: 195–209.

108 Guo W, Shang D, Cao J, Feng K, He Y, Jiang Y *et al.* Identifying and Analyzing Novel Epilepsy-Related Genes Using Random Walk with Restart Algorithm. *Biomed Res Int.* 2017

109 Iossifov I, Roak BJO, Sanders SJ, Ronemus M, Krumm N, Levy D *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature.* 2014; **515**: 216–221.

110 Harris KD, Shepherd GMG. The neocortical circuit: Themes and variations. *Nat Neurosci.* 2015; **18**: 170–181.

111 Vanneste S, Alsalman O, Ridder D De. COMT and the neurogenetic architecture of hearing loss induced tinnitus. *Hear Res.* 2018; **365**: 1–15.

112 Gao Y, Zhang X, Wang T, Zhang Y, Wang Q, Hu Y. Common Potential Biomarkers for Sporadic and Suspected Lynch Syndrome Endometrial Cancer. *Cancer Manag Res.* 2020; **12**: 11231–11241.

113 Hedley PL, Jørgensen P, Schlamowitz S, Wangari R, Moolman-Smook J, Brink PA *et al.* The genetic basis of long QT and short QT syndromes: A mutation update. *Hum Mutat.* 2009; **30**: 1486–1511.

114 Chen L, Kass RS. A-Kinase anchoring protein 9 and IKs channel regulation. *J Cardiovasc Pharmacol.* 2011; **58**: 459–461.

115 Chen L, Marquardt ML, Tester DJ, Sampson KJ, Ackerman MJ, Kass RS. Mutation of an A-kinase-anchoring protein causes long-QT syndrome. *Proc Natl Acad Sci.* 2007; **104**: 20990–20995.

116 Ozcan U, Ozcan L, Yilmaz E, Düvel K, Sahin M, Manning BD *et al.* Loss of the Tuberous Sclerosis Complex Tumor Suppressors Triggers the Unfolded Protein Response to Regulate Insulin Signaling and Apoptosis. *Mol Cell.* 2008; **29**: 541–551.

117 Specchio N, Pietrafusa N, Trivisano M, Moavero R, De Palma L, Ferretti A *et al.* Autism and Epilepsy in Patients With Tuberous Sclerosis Complex. *Front Neurol.* 2020; **11**. doi:10.3389/fneur.2020.00639.

118 Choi YJ, Di Nardo A, Kramvis I, Meikle L, Kwiatkowski DJ, Sahin M *et al.* Tuberous sclerosis complex proteins control axon formation. *Genes Dev.* 2008; **22**: 2485–2495.

119 Fu X, Sun X, Zhang L, Jin Y, Chai R, Yang L *et al.* Tuberous sclerosis complex-mediated mTORC1 overactivation promotes age-related hearing loss. *J Clin Invest.* 2018; **128**: 4938–4955.

120 Auerbach BD, Osterweil EK, Bear MF. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature.* 2011; **480**: 63–68.

121 Rauschecker JP, May ES, Maudoux A, Ploner M. Frontostriatal Gating of Tinnitus and Chronic Pain. *Trends Cogn Sci.* 2015; **19**: 567–578.

122 Lee HY. Beyond Hearing Loss : Does Tinnitus Cause Cognitive Impairment ? *Clin Exp Otorhinolaryngol*. 2020; **13**: 2–3.

123 Griffiths TD, Lad M, Kumar S, Holmes E, Mcmurray B, Maguire EA *et al.* How Can Hearing Loss Cause Dementia ? *Neuron.* 2020; **108**: 401–412.

124 Gao J, Armstrong NM, Deal JA, Lin FR, He P. Hearing loss and cognitive function among Chinese older adults : the role of participation in leisure activities. *BMC Geriatr.* 2020;**20** : 1–10.

125 Karch CM, Cruchaga C, Goate A. Alzheimer's Disease Genetics : From the bench to the clinic. *Neuron.*2014; **83**: 11–26.

126 Cochran JN, Mckinley EC, Cochran M, Amaral MD, Moyers BA, Lasseigne BN *et al.* Genome sequencing for early-onset or atypical dementia : high diagnostic yield and frequent observation of multiple contributory alleles. *Mol Case Stud.*2019; **5**: 1–18.

127 Chung J, Wang X, Maruyama T, Zhang X, Sherva R, Takeyama H *et al.* Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. *Alzheimers Dement*.2018; **14**: 623–633.

128 Pimenova AA, Raj T, Goate AM, Sciences G. Untangling Genetic Risk for Alzheimer's Disease. *Biol Psychiatry* .2018; **83**: 300–310.

129 Liu Y, Kheradmand F, Davis CF, Scheurer ME, Wheeler D, Tsavachidis S *et al.* Focused analysis of exome sequencing data for rare germline mutations in familial and sporadic lung cancer. *J Thorac Oncol.* 2016; **11**: 52–61.

130 Johar AS, Mastronardi C, Rojas-Villarraga A, Patel HR, Chuah A, Peng K *et al.* Novel and rare functional genomic variants in multiple autoimmune syndrome and Sjögren's syndrome. *J Transl Med.* 2015; **13**: 1–11.

131 Hiekkala ME, Vuola P, Artto V, Häppölä P, Häppölä E, Vepsäläinen S *et al.* The contribution of CACNA1A, ATP1A2 and SCN1A mutations in hemiplegic migraine: A clinical and genetic study in Finnish migraine families. *Cephalalgia.* 2018; **38**: 1849–1863.

132 Qiao D, Ameli A, Prokopenko D, Chen H, Kho AT, Parker MM *et al.* Whole exome sequencing analysis in severe chronic obstructive pulmonary disease. *Hum Mol Genet.* 2018; **27**: 3801–3812.

133 Nuytemans K, Ortel TL, Gomez L, Hofmann N, Alves N, Dueker N *et al.* Variants in chondroitin sulfate metabolism genes in thrombotic storm. *Thromb Res.* 2018; **161**: 43–51.

134 Aubart M, Gazal S, Arnaud P, Benarroch L, Gross MS, Buratti J *et al.* Association of modifiers and other genetic factors explain Marfan syndrome clinical variability. *Eur J Hum Genet.* 2018; **26**: 1759–1772.

135 Gregson CL, Newell F, Leo PJ, Clark GR, Paternoster L, Marshall M *et al.* Genome-wide association study of extreme high bone mass: Contribution of common genetic variation to extreme BMD phenotypes and potential novel BMD-associated genes. *Bone.* 2018; **114**: 62–71.

136 Lee HS, Yang SK, Hong M, Jung S, Kim BM, Moon JW *et al.* An intergenic variant rs9268877 between HLA-DRA and HLA-DRB contributes to the clinical course and long-term outcome of ulcerative colitis. *J Crohn's Colitis.* 2018; **12**: 1113–1121.

137 Tomaiuolo R, Bellia C, Caruso A, Di Fiore R, Quaranta S, Noto D *et al.* Prothrombotic gene

variants as risk factors of acute myocardial infarction in young women. *J Transl Med.* 2012; **10**: 1–5.

138 Shen Y, Xu J, Yang X, Liu Y, Ma Y, Yang D *et al.* Evidence for the involvement of the proximal copy of the MAGEA9 gene in Xq28-linked CNV67 specific to spermatogenic failure. *Biol Reprod.* 2017; **96**: 610–616.

139 Uzun A, Schuster J, McGonnigal B, Schorl C, Dewan A, Padbury J. Targeted sequencing and meta-analysis of preterm birth. *PLoS One.* 2016; **11**: 1–17.

Supplementary Material

Table S1: Clinical profile of 30 Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

No.	Sex	Age at onset	THI score	Familial MD	Type of migraine	High blood pressure	Type 2 diabetes	Diagnosis of autoimmune disease	HADS score	Others
1	X	34	76	Yes	MA	-	-	-	31	-
2	X	33	98	-	MA	-	-	-	34	Arthrosis
3	X	40	76	-	MA	Yes	-	-	21	-
4	X	31	82	-	MA	-	-	-	28	-
5	X	41	78	-	-	-	-	-	27	Arthrosis
6	X	22	76	-	MO	Yes	-	-	23	-
7	X	56	82	-	-	-	-	-	20	-
8	X	14	80	-	-	-	-	-	34	Asthma
9	X	25	76	-	MA	Yes	Yes	-	NA	Anxiety Arthrosis
10	X	20	86	Yes	-	-	Yes	-	27	ArthrosisAsthma
11	X	55	86	-	-	Yes	-	-	NA	Arthrosis
12	X	35	82	-	-	NA	NA	-	NA	-
13	X	38	76	-	-	NA	NA	NA	NA	-
14	X	50	88	-	-	-	-	-	17	-
15	X	37	96	-	-	-	-	-	31	-
16	X	40	82	Yes	-	-	Yes	-	NA	-
17	X	31	80	-	-	-	-	-	NA	-
18	X	33	90	Yes	MA	-	-	Hypothyroidism	NA	Asthma, seasonal allergy
19	X	22	95	-	MO	-	-	-	24	Asthma, seasonal allergy
20	Y	29	78	-	-	Yes	-	-	30	-
21	Y	35	88	-	-	-	-	Psoriasis	16	-
22	Y	25	84	-	-	-	-	-	NA	Asthma, seasonal allergy
23	Y	57	90	-	-	-	-	-	21	-
24	Y	36	82	-	-	-	-	-	23	-
25	Y	31	94	Yes	-	Yes	-	-	10	_
26	Y	49	82	-	-	-	-	-	26	-
27	X	20	96	-	MO	-	-	-	15	-
28	X	46	88	-	-	-	-	-	18	-
29	X	54	76	-	-	Yes	-	-	28	-
30	X	40	90	Yes	-	Yes	-	-	32	-

THI= Tinnitus handicap inventory, MA=migraine with aura, MO= migraine without aura, HADS=Hospital anxiety and depression scale, NA= data not available

Table S2: Clinical profile of 29 Spanish patients with Meniere disease and tinnitus almost extreme phenotype (MD-AEP)

No.	Sex	Age at onset	THI score	Familial MD	Type of migraine	High blood pressure	Type 2 diabetes	Diagnosis of autoimmune disease	HADS score	Others
1	X	40	74	-	MO	Yes	-	Psoriasis	14	-
2	Y	50	68	-	-	-	-	Spondylitis, ulcerative colitis	20	Anxiety, depression
3	X	31	64	-	-	-	-	Vitiligo	NA	Artrosis
4	Y	53	72	-	-	Yes	-	-	NA	-
5	X	45	60	-	-	Yes	-	-	4	Headache
6	X	22	70	-	-	-	-	-	NA	-
7	X	31	70	-	-	Yes	Yes	Hypothryrodism Rheumatoid arthritis	NA	-
8	X	33	70	-	MA	Yes	-	Antiphospholipid syndrome	18	-
9	X	39	60	-	-	-	-	-	17	-
10	Y	58	60	-	MO	-	-	-	18	-
11	X	29	66	-	MA	Yes	-	-	NA	Arthrosis
12	Y	39	72	-	-	-	-	-	NA	-
13	X	57	70	-	NA	NA	NA	NA	NA	-
14	X	33	62	-	-	-	-	Celico	13	-
15	X	42	68	Yes	-	Yes	-	-	12	-
16	X	42	72	Yes	-	-	-	-	8	-
17	X	41	66	Yes	-	Yes	Yes	-	25	-
18	X	15	62	Yes	-	-	-	-	11	-
19	X	48	64	-	-	-	-	-	NA	Arthrosis
20	X	51	72	-	-	Yes	-	-	NA	-
21	Y	24	58	-	-	-	-	-	NA	-
22	X	39	72	-	MA	-	-	Spondylitis	NA	-
23	Y	33	56	-	-	-	-	Psoriasis	2	-
24	X	45	56	-	-	-	-	Spondylitis	NA	Arthrosis
25	Y	47	74	-	MO	Yes	Yes	-	27	-
26	X	56	74	-	-	-	-	-	22	-
27	Y	30	74	-	-	-	-	-	26	-
28	Y	21	74	-	-	-	-	-	NA	-
29	Y	48	74	Yes	-	-	-	-	10	-

THI= Tinnitus handicap inventory, MA=migraine with aura, MO= migraine without aura, HADS=Hospital anxiety and depression scale, NA= data not available

Table S3: Clinical profile of 97 Swedish individuals with tinnitus from STOP cohort

No.	Age	Sex	THI score	Hearing disorder	Headaches	Vertigo	TMJ	Neck pain	Other pain	Under treatment for psychiatric disorder
1	53	Y	14	Yes	-	-	-	-	Yes	-
2	54	Y	18	Yes	-	-	-	-	-	-
3	69	X	24	NA	Yes	NA	-	Yes	Yes	-
4	51	Y	24	Yes	-	-	-	-	-	-
5	68	X	28	Yes	-	-	-	-	-	-
6	50	X	28	NA	-	NA	NA	NA	-	-
7	43	X	28	NA	Yes	Yes	NA	Yes	-	Yes
8	63	X	30	Yes	Yes	-	-	-	Yes	-
9	35	Y	30	Yes	-	Yes	Yes	Yes	-	-
10	37	Y	32	Yes	-	Yes	-	-	-	-
11	52	Y	32	Yes	-	-	-	-	-	-
12	39	X	32	Yes	-	-	-	Yes	Yes	-
13	60	X	32	NA	Yes	Yes	NA	Yes	-	-
14	77	X	34	Yes	-	-	-	-	-	-
15	53	Y	34	-	Yes	-	-	-	-	-
16	41	Y	34	Yes	-	-	-	-	-	Yes
17	47	Y	34	Yes	-	-	-	-	-	-
18	47	X	36	Yes	-	Yes	-	-	-	-
19	39	X	36	Yes	Yes	Yes	Yes	Yes	Yes	-
20	41	Y	36	-	-	-	-	-	-	Yes
21	37	Y	36	Yes	-	-	-	NA	-	-
22	45	X	38	Yes	Yes	-	-	Yes	Yes	-
23	48	X	38	NA	Yes	Yes	-	Yes	Yes	Yes
24	34	X	38	-	Yes	Yes	NA	Yes	Yes	Yes
25	66	X	40	Yes	-	-	-	-	-	-
26	54	Y	40	Yes	-	-	-	-	Yes	-
27	70	Y	40	Yes	-	-	-	-	-	-
28	77	Y	40	Yes	-	-	-	-	-	-
29	48	X	42	-	Yes	Yes	Yes	Yes	Yes	-
30	48	X	42	Yes	-	-	-	-	-	-
31	39	X	42	Yes	-	-	-	-	-	Yes
32	33	Y	42	-	Yes	Yes	Yes	Yes	Yes	-
33	49	Y	42	Yes	-	-	-	Yes	Yes	-
34	32	X	42	Yes	Yes	-	-	Yes	Yes	-
35	51	Y	42	NA	-	-	Yes	-	-	Yes
36	59	X	44	Yes	-	-	-	-	-	-
37	49	X	44	NA	Yes	Yes	Yes	Yes	Yes	-
38	72	X	46	Yes	-	Yes	-	-	-	-
39	52	X	48	Yes	-	Yes	Yes	Yes	Yes	-
40	35	X	48	NA	Yes	-	-	Yes	-	-
41	33	Y	48	NA	-	-	Yes	-	Yes	-
42	52	X	48	-	-	-	-	Yes	-	-
43	58	Y	48	Yes	-	-	-	-	Yes	-
44	30	X	50	Yes	Yes	NA	Yes	Yes	Yes	-
45	48	X	50	NA	Yes	Yes	Yes	Yes	NA	-
46	32	Y	50	-	-	-	Yes	-	-	-
47	33	X	50	Yes	Yes	Yes	-	-	-	-
48	54	X	52	NA	-	Yes	NA	Yes	Yes	Yes
49	77	X	52	NA	-	Yes	-	-	-	-
50	63	X	52	Yes	-	NA	Yes	-	Yes	Yes
51	37	Y	52	Yes	-	-	-	Yes	-	-
52	52	Y	52	Yes	Yes	Yes	-	-	Yes	-
53	77	Y	52	Yes	-	-	-	NA	-	-
54	79	X	52	Yes	-	-	-	-	-	-
55	56	Y	54	Yes	-	Yes	-	-	Yes	-
56	58	X	54	NA	-	Yes	-	-	-	-
57	60	X	54	NA	Yes	Yes	Yes	Yes	Yes	-
58	50	Y	54	NA	-	-	-	-	Yes	-
59	72	X	54	Yes	-	-	-	Yes	Yes	-
60	43	X	56	-	Yes	Yes	-	-	-	-
61	57	X	56	Yes	-	NA	NA	Yes	-	-
62	30	Y	56	Yes	-	-	-	-	-	-
63	37	Y	56	Yes	-	Yes	-	Yes	-	-
64	53	X	58	Yes	Yes	Yes	-	-	NA	Yes

65	35	Y	58	Yes	-	-	-	-	-	Yes
66	52	X	58	NA	NA	NA	-	-	-	-
67	29	Y	58	Yes	Yes	Yes	NA	-	Yes	-
68	52	Y	60	Yes	Yes	-	Yes	-	-	-
69	78	X	60	Yes	Yes	Yes	-	-	Yes	-
70	48	X	60	Yes	-	-	-	-	-	-
71	52	Y	60	Yes	Yes	Yes	Yes	Yes	Yes	-
72	52	Y	62	Yes	-	-	-	Yes	Yes	-
73	40	Y	64	-	-	-	Yes	Yes	-	-
74	39	X	66	Yes	Yes	Yes	Yes	Yes	-	-
75	37	X	68	Yes	-	Yes	Yes	-	-	-
76	32	Y	68	NA	-	-	-	-	-	-
77	66	X	68	NA	-	Yes	-	Yes	Yes	-
78	35	Y	68	Yes	-	-	-	-	-	-
79	50	Y	70	Yes	Yes	-	Yes	Yes	Yes	-
80	37	X	70	-	NA	-	Yes	Yes	-	Yes
81	67	Y	70	Yes	-	-	-	-	-	-
82	31	X	72	NA	Yes	Yes	-	Yes	-	-
83	58	X	72	-	Yes	-	Yes	Yes	Yes	Yes
84	53	X	72	NA	Yes	Yes	Yes	Yes	Yes	-
85	56	X	72	Yes	-	Yes	-	-	Yes	-
86	50	Y	76	Yes	Yes	-	-	Yes	Yes	-
87	40	X	78	Yes	Yes	Yes	Yes	-	Yes	Yes
88	32	X	78	NA	Yes	Yes	-	Yes	Yes	-
89	61	Y	82	-	-	-	Yes	Yes	Yes	Yes
90	52	X	82	Yes	-	NA	Yes	Yes	Yes	-
91	30	X	82	Yes	-	-	-	-	-	-
92	37	Y	86	-	-	Yes	-	-	-	-
93	35	Y	88	-	Yes	-	-	-	-	-
94	33	X	88	Yes	-	NA	-	-	-	-
95	29	Y	90	Yes	Yes	Yes	Yes	Yes	Yes	Yes
96	27	X	92	Yes	-	Yes	-	Yes	Yes	-
97	55	X	96	Yes	-	-	-	-	-	-

THI= Tinnitus handicap inventory, TMJ=Temporo-mandibular joint dysfunction, NA= data not available

Table S4: Hearing loss genes obtained from associated with syndromic and non-syndromic deafness from Deafness Variation Database (DVD)

Hearing loss genes (N=152)			
ESPN	*ESRRB*	*CLRN1*	*ILDR1*
GJB3	*STRC*	*CCDC50*	*MCM2*
COL9A2	*CIB2*	*OPA1*	*SLITRK6*
KCNQ4	*HOMER2*	*WFS1*	*COCH*
BSND	*TBC1D24*	*GRXCR1*	*SIX1*
ROR1	*CRYM*	*DSPP*	*TNC*
CDC14A	*OTOA*	*CISD2*	*MT-TS1*
COL11A1	*KARS*	*MARVELD2*	*SMPX*
GPSM2	*MYO15A*	*BDP1*	*POU3F4*
KCNJ10	*TMEM132E*	*ADGRV1 (GPR98)*	*TIMM8A*
USH2A	*USH1G*	*HSD17B4*	*PRPS1*
NLRP3	*ACTG1*	*SLC22A4*	*COL4A6*
GATA3	*LOXHD1*	*HARS2*	*COL4A5*
MYO3A	*GIPC3*	*DIAPH1*	*AIFM1*
PCDH15	*CLPP*	*GRXCR2*	*MT-TS1*
CDH23	*S1PR2*	*POU4F3*	*SMPX*
C10orf2	*SYNE4*	*TCOF1*	*TPRN*
PDZD7	*CEACAM16*	*FOXI1*	*MT-RNR1*
TECTB	*SIX5*	*SERPINB6*	*MT-TL1*
FGFR2	*MYH14*	*DCDC2*	
EPS8L2	*OTOF*	*FAM65B*	
KCNQ1	*PNPT1*	*COL11A2*	
USH1C	*ATP6V1B1*	*LHFPL5*	
OTOG	*ALMS1*	*PEX6*	
CABP2	*LOXL3*	*POLR1C*	
FGF3	*ELMOD3*	*CLIC5*	
LRTOMT	*DFNB59 (PJVK)*	*COL9A1*	
MYO7A	*PAX3*	*MYO6*	
NARS2	*COL4A4*	*CD164*	
RDX	*COL4A3*	*EYA4*	
TECTA	*EDN3*	*DFNA5*	
EPS8	*OSBPL2*	*ADCY1*	
COL2A1	*KCNE1*	*HGF*	
MSRB3	*CLDN14*	*PEX1*	
OTOGL	*TMPRSS3*	*SLC26A5*	
PTPRQ	*TSPEAR*	*SLC26A4*	
KITLG	*TBX1*	*MET*	
SLC17A8	*MYH9*	*MIR96*	
DIABLO	*TRIOBP*	*FGFR1*	
P2RX2	*SOX10*	*SNAI2*	
GJB2	*ATP2B2*	*EYA1*	
GJB6	*LARS2*	*GRHL2*	
POLR1D	*TMIE*	*TJP2*	
DIAPH3	*CACNA1D*	*TMC1*	
EDNRB	*MITF*	*DFNB31 (WHRN)*	

Table S5: Summary of the 19 genetic studies using an extreme phenotype approach selected for synthesis.

Reference	Disease	EP Criteria	Study Design	Sequencing Method	Ancestry	Number of Patients	Onset	Sex
Pullabhatla et al. (2017) [72]	Systemic lupus erythematosus	Proband with early onset and clinical features with poor outcome	Family trios, Replication cohort	WES	EU	30 trios, 10995	<25 y	Not repo d
Johar et al.(2016) [86]	Polyautoimmunity	Polyautoimmunity and familial autoimmunity\	Case–control, Cross-sectional	WES	Colombian	47	Not reported	M,F
Kunkle et al. (2017) [85]	Alzheimer's disease	Early-onset Alzheimer's disease, familial or sporadic	Case–control, Replication cohort	WES	NHW and Caribbean Hispanic	93, 8570	<65 y	M,F
Emond et al.(2012) [29]	Cystic fibrosis (CF)	CF with early onset of persistent pseudomonas aeruginosa infection	Case–control, Replication cohort	WES	EU America, African American, White Hispanic, NHW, Asian, Aleut	43, 696	≤2.5 y	M,F
Shtir et al. (2016) [32]	Diabetes	Diabetes for at least 10 years without diabetic retinopathy	Case–control, Cross-sectional	WES	Saudi	43	Not reported	M,F
Liu et al. (2016) [129]	Lung cancer	Familial or sporadic lung cancer cases, ever smokers or severe chronic obstructive pulmonary disease (COPD)	Case–control, Cross-sectional	WES	NHW	48 sporadic 54 familial	56 y familial 61 y sporadic	M,F
Husson et al. (2018) [33]	Bipolar I disorder	Family history of mood disorder and early onset	Case–control, Cross-sectional	WES	EU	92	mean: 24 y	M,F
Johar et al. (2015) [130]	Multiple autoimmune syndrome	Multiple autoimmune syndrome with Sjögren's syndrome	Case–control, Cross-sectional	WES	Colombian	12	28–67 y	F
Hiekkala et al. (2018) [131]	Hemiplegic migraine	≥2 migraine attacks, completely reversible motor weakness	Case report, Cross-sectional	WES	Finnish	293	median: 12 y	M,F
Qiao et al. (2018) [132]	COPD	COPD cases with GOLD grade 3 or 4	Case–control, Cross-sectional	WES	EU, NHW, African American	≈1769	>45 y, ≤65 y	M/F
Bruse et al.	COPD	COPD cases with GOLD grade	Case–control,	WES	NHW	62	Not	M/F

(2016) [71]		3 or 4	Cross-sectional				reported	
Nuytemans et al. (2018) [133]	Thrombotic storm (TS)	Severe onset of ≥2 arterial, unusual clot location, refractory, reoccurrence	Case report, Cross-sectional	WES, Targeted sequencing	White and Indian	26 (13 trios)	Not reported	M,F
Aubart et al. (2018) [134]	Marfan syndrome	Severe aortic features (dissection or preventive thoracic aortic aneurysm rupture surgery at a young age) or sib pairs	Case–control, Cross-sectional	WES	EU	51 EP and 8 sib-pairs	≈10–30 y	M,F
Gregson et al. (2018) [135]	Bone mass density	Extremely high or moderately high bone mass density	Case–control, Replication cohort	GWAS	EU	1258, 32965	Not reported	M,F
Lee et al. (2018) [136]	Ulcerative colitis	Ulcerative colitis patients with good or poor prognosis	Case–control, Replication cohort	Genotyping	Korean	881, 274	35.6 ± 13.9 y	M,F
Tomaiuolo et al. (2012) [137]	Acute myocardial infarction (AMI)	AMI patients with first episode before or after 45 years of age	Case–control, Replication cohort	Genotyping	EU	1653, 909	Not reported	M,F
Goldberg-Stern et al. (2013) [31]	Epilepsy with febrile seizures plus	Generalized epilepsy with febrile seizures plus, a proband with Dravet syndrome	Case-control, Cross-sectional	Sanger sequencing	Ashkenazi Jewish	14 familial cases	infancy to 7 y	M,F
Shen et al. (2017) [138]	Spermatogenic failure	Spermatogenic failure with azoospermia, mild oligozoospermia or severe oligozoospermia	Case–control, Cross sectional	Sanger sequencing	Chinese Han	884	Not reported	M
Uzun et al. (2016) [139]	Preterm birth	Patients delivering <34 weeks	Case report, Cross-sectional	Targeted Sequencing of 329 genes	African-American; Asian; Hispanic; White; Native American	32	Not reported	F

Legend: Non-Hispanic White, NHW, European, EU, Whole-Exome Sequencing, WES, GWAS, genome-wide association studies, EP, extreme phenoty
frequency

Table S6: Synaptic genes showing enrichment of synonymous variants in Spanish patients with Meniere disease and genes were significant when they were compared against CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
SYNJ2	5	3.44(1.70-6.95)	0.71	*NS*	3.39(1.68-6.85)	0.70	*NS*
BRSK2	5	6.02(2.48-14.58)	0.83	1.34E-01	6.60(2.72-16.01)	0.85	*NS*
TLN1	5	4.42(1.82-10.71)	0.77	*NS*	4.44(1.83-10.76)	0.77	*NS*
PC	4	233.72(82.15-664.89)	0.99	<1.00E-15	141.70(49.63-404.55)	0.99	<1.00E-15
CRIP2	4	46.17(16.99-125.48)	0.98	1.09E-10	42.63(15.58-116.70)	0.98	5.26E-10
RIMBP2	4	26.73(10.96-65.20)	0.96	9.71E-10	25.51(10.42-62.48)	0.96	2.58E-09
FGD4	4	32.25(13.21-78.78)	0.97	4.65E-11	26.36(10.76-64.60)	0.96	1.56E-09
HGS	4	7.60(2.82-20.45)	0.87	*NS*	7.87(2.92-21.22)	0.87	*NS*
IARS	4	35.69(13.17-96.75)	0.97	4.00E-09	35.70(13.08-97.42)	0.97	5.58E-09
IQSEC2	3	22.66(8.36-61.39)	0.96	1.59E-06	25.37(9.31-69.11)	0.96	4.81E-07
AP3D1	3	525.88(145.47-1901.09)	0.9	<1.00E-15	366.07(98.29-1363.47)	0.99	<1.00E-15
MYH14	3	10.58(3.37-33.22)	0.91	*NS*	9.40(2.99-29.56)	0.89	*NS*
HTT	3	74.15(23.19-237.13)	0.99	7.31E-10	8.45(2.69-26.58)	0.88	*NS*
SBF1	3	11.28(3.59-35.43)	0.91	*NS*	11(3.50-34.64)	0.91	*NS*
DPP3	3	12.53(3.99-39.37)	0.92	2.83E-02	11.11(3.53-34.99)	0.91	*NS*
SYNM	3	152.22(46.56-497.67)	0.99	<1.00E-15	106.27(32.19-350.79)	0.99	3.56E-11
UNC13A	3	15.37(4.89-48.32)	0.93	5.56E-03	13.32(4.23-41.99)	0.92	1.85E-02
KIAA1217	2	9.96(3.67-27.05)	0.90	1.23E-02	10.12(3.72-27.55)	0.90	1.10E-02
LLGL1	2	97.21(30.06-314.36)	0.99	3.98E-11	237.33(67.31-836.75)	0.99	<1.00E-15
SLC25A3	2	25.56(8.07-81.01)	0.96	6.84E-05	23.21(7.29-73.87)	0.96	1.91E-04

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Table S7: Synaptic genes showing enrichment of 5´UTR variants in Spanish patients with Meniere disease and Listed genes were significant when they were compared against CSVS reference dataset

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[CSVS] OR(CI)
TUBB3	5	62.28(23.33-166.26)	0.98	4.19E-13	59.67(24.09-147.84)	0.98	<1.00E-15	20.56(7.48-5
ACTG1	4	261.61(65.04-1052.23)	0.99	8.38E-12	399.35(126.26-1263.12)	0.99	<1.00E-15	32.90(9.84-1
DPYSL2	4	7.18(3.99-12.93)	0.86	9.20E-08	6.85(3.83-12.26)	0.85	1.74E-07	6.34(3.46-11
RTN4	3	49.04(14.16-169.79)	0.98	1.53E-06	39.21(12.28-125.20)	0.97	1.11E-06	24.67(6.49-9
TRIO	3	31.38(9.39-104.87)	0.97	4.10E-05	23.69(7.48-75.05)	0.96	1.41E-04	49.36(10.97-
MYO1D	2	43.95(12.77-151.23)	0.98	3.71E-06	34.23(10.70-109.55)	0.97	4.95E-06	24.87(6.52-9
PDE4D	2	15.75(6.27-39.53)	0.94	8.18E-06	11.41(4.64-28.05)	0.91	2.15E-04	10.85(4.15-2
RPLP1	2	27.27(8.19-90.77)	0.96	1.34E-04	41(12.77-131.64)	0.98	8.28E-07	24.87(6.52-9
SEPT2	2	16.13(4.96-52.48)	0.94	7.25E-03	12.80(4.04-40.54)	0.92	2.73E-02	16.57(4.62-5

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, ge

NS=Nonsignificant

**Table S8: Synaptic genes showing enrichment of synonymous variants in Spanish patients with Meniere disease and ti
Listed genes were significant when they were compared against CSVS reference dataset**

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CS OR
PTPRS	6	5.19(2.31-11.65)	0.81	*NS*	5.11(2.27-11.47)	0.80	*NS*	7.0
KALRN	5	38.66(15.84-94.39)	0.97	2.09E-12	44.39(18.03-109.29)	0.98	4.19E-13	7.9
AGAP1	4	14.96(6.15-36.41)	0.93	4.69E-06	14.43(5.92-35.20)	0.93	8.32E-06	9.7
PIP5K1C	4	130.87(47.19-362.89)	0.99	<1.0E-15	137.78(48.42-392.07)	0.99	<1.00E-15	17
TSC2	4	34.69(12.80-94.03)	0.97	5.88E-09	49.41(18.01-135.58)	0.98	6.91E-11	13
PKP4	4	5.41(2.40-12.17)	0.82	*NS*	5.24(2.33-11.81)	0.81	*NS*	6.9
ANK2	4	133.05(47.96-369.12)	0.99	<1.00E-15	108.25(38.50-304.37)	0.99	<1.00E-15	17
REV3L	4	70.64(25.84-193.09)	0.99	0<1.00E-15	101.03(36.04-283.24)	0.99	<1.00E-15	17
SHANK3	4	8.79(3.62-21.37)	0.89	3.02E-03	8.52(3.50-20.75)	0.88	4.39E-03	14
FARP1	3	105.03(32.58-338.63)	0.99	1.26E-11	106.56(32.32-351.29)	0.99	3.22E-11	25
AKAP9	3	99.78(30.99-321.24)	0.99	2.26E-11	136.40(40.80-456.01)	0.99	2.51E-12	25
PLXNA3	3	126.19(50.49-315.40)	0.99	<1.00E-15	112.74(44.45-285.97)	0.99	<1.00E-15	86
PRUNE2	3	19.24(6.11-60.56)	0.95	8.20E-04	19.14(6.06-60.51)	0.95	9.39E-04	51
RIMBP2	3	16.85(5.36-53.01)	0.94	2.58E-03	15.92(5.04-50.24)	0.94	4.46E-03	17
ROCK1	3	171.07(52.11-561.52)	0.99	<1.00E-15	227.34(65.22-792.44)	0.99	<1.00E-15	25
VCPIP1	3	88.04(27.43-282.52)	0.99	9.80E-11	126.29(37.96-420.24)	0.99	5.86E-12	51
SCN1A	2	14.85(5.46-40.40)	0.93	2.39E-04	12.51(4.59-34.09)	0.92	1.48E-03	10
SLC2A1	2	54.87(19.98-150.70)	0.98	1.47E-11	42.04(15.24-115.97)	0.98	9.70E-10	9.8
KEL	2	33.97(12.43-92.83)	0.97	1.19E-08	34(12.36-93.47)	0.97	1.56E-08	10
CASK	2	183.03(55.34-605.39)	0.99	<1.00E-15	181.06(52.84-620.41)	0.99	4.19E-13	25
CAD	2	54.87(19.98-150.70)	0.98	1.47E-11	79.76(28.48-223.43)	0.99	<1.00E-15	34

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordere

Table S9: Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with Meniere di (MD-AEP). Listed genes were significant when they were compared against CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(CI)
DMD	3	8.04(3.30-19.61)	0.88	8.55E-03	7.49(3.07-18.28)	0.87	1.84E-02	28.70(10-82
GOLGB1	3	30.85(9.77-97.44)	0.97	9.63E-06	30.99(9.75-98.52)	0.97	1.12E-05	25.54(6.72-9
MYO1C	3	12.78(4.73-34.56)	0.92	9.65E-04	11.92(4.40-32.29)	0.92	2.06E-03	13.68(4.63-4
PPFI A1	3	74.83(23.40-239.27)	0.99	6.46E-10	66.85(20.67-216.27)	0.99	4.31E-09	51.09(11.35
PTPRS	3	16.95(5.39-53.32)	0.94	2.46E-03	17.75(5.62-56.06)	0.94	1.80E-03	51.09(11.35
RYR2	3	153.52(46.99-501.54)	0.99	1.00E-15	106.56(32.32-351.29)	0.99	3.22E-11	17.02(4.76-6
TRAP1	3	42.75(13.49-135.49)	0.98	3.31E-07	32.77(10.30-104.30)	0.97	6.50E-06	17.02(4.76-6
CRMP1	2	3.34(1.75-6.39)	0.70	*NS*	3.56(1.86-6.81)	0.72	*NS*	5.29(2.71-10
OGDHL	2	105.95(32.70-343.29)	0.99	1.42E-11	64.89(19.98-210.70)	0.98	7.17E-09	25.75(6.74-9
ST14	2	60.38(18.87-193.23)	0.98	9.15E-09	61.41(18.94-199.08)	0.98	1.28E-08	17.16(4.78-6

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Table S10: Synaptic genes showing enrichment of indels in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
GPSM1	6	8.90(4.37-18.15)	0.89	3.33E-06	6.08(3.01-12.29)	0.84	9.14E-04
SGTA	6	4.91(2.52-9.58)	0.80	5.65E-03	5.11(2.63-9.92)	0.80	2.77E-03
CACNA2D1	5	13.88(5.61-34.33)	0.93	2.37E-05	24.32(9.95-59.43)	0.96	4.77E-09
TSC2	5	11.30(4.59-27.84)	0.91	2.58E-04	10.21(4.20-27.82)	0.90	5.57E-04
AAK1	4	9.50(3.48-25.95)	0.89	2.15E-02	14.01(5.18-37.87)	0.93	3.73E-04
MRAS	4	17.25(6.23-47.77)	0.94	7.95E-05	12.07(4.47-32.60)	0.92	1.71E-03
SIPA1L1	4	11.94(5.22-27.32)	0.92	8.08E-06	11.07(4.90-25)	0.91	1.39E-05
WASL	4	86.34(27.93-266.86)	0.99	1.84E-11	56.76(20.63-156.15)	0.98	9.63E-12
PIP4K2A	3	26.29(7.97-86.71)	0.96	1.49E-04	20.18(6.38-63.82)	0.95	5.93E-04
PLXNA2	3	16.59(5.98-45.99)	0.94	1.28E-04	14.34(5.29-38.90)	0.93	3.17E-04
GSK3B	3	26.29(7.97-86.71)	0.96	1.49E-04	22.78(7.20-72.15)	0.96	2.00E-04
FARSA	3	23.37(7.13-76.64)	0.96	3.74E-04	18.20(5.76-57.51)	0.95	1.45E-03
AKAP9	3	33.66(10.06-112.57)	0.97	2.15E-05	55.21(17.17-177.49)	0.98	3.16E-08
ANXA11	3	33.66(10.06-112.57)	0.97	2.15E-05	30.45(9.58-96.78)	0.97	1.32E-05
SH3PXD2A	3	17.90(5.51-58.07)	0.94	2.95E-03	13.22(4.19-41.67)	0.92	1.97E-02
STK32C	3	22.14(6.77-72.43)	0.95	5.71E-04	18.58(5.88-58.73)	0.95	1.21E-03
AP1G1	2	62.35(13.32-291.88)	0.98	2.91E-04	61.99(14.77-260.12)	0.98	3.21E-05
AP2A2	2	40.08(9-178.42)	0.98	2.40E-03	22.64(5.52-92.88)	0.96	2.79E-02
ATF7IP	2	13.25(4.10-42.82)	0.92	2.98E-02	13.64(4.30-43.22)	0.93	1.70E-02
BAZ1B	2	56.11(12.15-259.06)	0.98	4.65E-04	52.34(12.54-218.44)	0.98	1.06E-04
CORO1C	2	52.90(21.71-128.88)	0.98	<1.00E-15	103.26(43.93-242.75)	0.99	1.00E-15
DNM3	2	46.76(10.34-211.37)	0.98	1.11E-03	33.17(8.04-136.91)	0.97	2.44E-03
HSPA12A	2	112.24(21.55-584.75)	0.99	3.91E-05	196.34(43.43-887.54)	0.99	1.30E-08
ICA1	2	35.06(7.97-154.32)	0.97	4.79E-03	22.64(5.52-92.88)	0.96	2.79E-02
NDRG2	2	70.15(14.73-334.07)	0.99	1.77E-04	61.99(14.77-260.12)	0.98	3.21E-05
RGS8	2	10.33(4.15-25.72)	0.90	9.85E-04	8.86(3.60-21.79)	0.89	3.82E-03
SNAP47	2	47.18(13.70-162.48)	0.98	1.90E-06	26.20(8.22-83.52)	0.96	6.35E-05
SNX5	2	28.05(6.48-121.45)	0.96	1.56E-02	43.62(10.50-181.12)	0.98	3.81E-04
TAOK2	2	62.35(13.32-291.88)	0.98	2.91E-04	94.23(22.05-402.69)	0.99	1.61E-06
TRAP1	2	280.64(39.18-2010.18)	0.99	3.78E-05	235.61(51.03-1087.74)	0.99	4.87E-09
TUBB2A	2	190.53(53.02-684.73)	0.99	1.68E-12	92.27(32.80-259.57)	0.99	<1.00E-15

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by number of indels.

Table S11: Synaptic genes showing enrichment of indels in Spanish patients with Meniere disease and tinnitus almost extreme phenotype (MD-AEP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
RYR2	6	6.30(2.95-13.42)	0.84	3.62E-03	5.34(2.52-11.33)	0.81	2.39E-02
TSC2	4	39.86(14.04-113.21)	0.97	8.49E-09	36.07(13.232-98.33)	0.97	4.58E-09
ANK2	4	9.81(3.59-26.79)	0.90	1.58E-02	11.18(4.13-30.24)	0.91	3.73E-03
SFXN5	3	71.77(20.53-250.91)	0.99	4.14E-08	88.62(27.31-287.58)	0.99	1.55E-10
BRSK1	3	119.64(32.02-446.97)	0.99	2.11E-09	54.44(16.99-174.50)	0.98	3.28E-08
ENAH	3	153.82(39.34-601.44)	0.99	8.54E-10	71.74(22.24-231.37)	0.99	1.61E-09
EPB41L2	3	28.32(8.63-92.90)	0.96	6.55E-05	21.82(6.89-69.09)	0.95	3.01E-04
GRIP1	3	89.72(25.02-321.70)	0.99	9.66E-09	113.00(34.52-369.91)	0.99	1.05E-11
HSPA9	3	215.36(50.93-910.60)	0.99	5.29E-10	188.34(56.02-633.25)	0.99	<1.00E-15
JUP	3	8.79(3.21-24.07)	0.89	4.38E-02	10.23(3.77-27.77)	0.90	9.43E-03
VPS8	3	23.91(7.34-77.92)	0.96	2.63E-04	29.14(9.18-92.54)	0.97	2.01E-05
AARS	2	29.89(6.96-128.45)	0.97	9.31E-03	65.50(15.65-274.08)	0.98	1.94E-05
ACTR2	2	65.25(14.25-298.75)	0.98	1.39E-04	66.95(15.99-280.35)	0.99	1.65E-05
ANXA6	2	358.95(49.99-2577.53)	0.99	9.32E-06	376.71(78.85-1799.65)	0.99	1.99E-10
ATP2A1	2	47.84(10.78-212.39)	0.98	6.89E-04	34.62(8.39-142.88)	0.97	1.80E-03
ATP6V1C2	2	119.64(23.81-601.07)	0.99	1.18E-05	502.29(99.98-2523.52)	0.99	8.12E-11
BASP1	2	71.77(15.50-332.42)	0.99	8.77E-05	47.07(11.34-195.35)	0.98	2.13E-04
BIN1	2	39.86(9.11-174.44)	0.97	1.87E-03	42.43(10.24-175.72)	0.98	4.44E-04
CSNK2A1	2	27.59(6.45-118.06)	0.96	1.46E-02	35.86(8.68-148.07)	0.97	1.42E-03
DAAM1	2	24.73(5.81-105.28)	0.96	2.67E-02	25.74(6.26-105.80)	0.96	1.26E-02
DNM1L	2	23.91(5.63-101.61)	0.96	3.22E-02	23.16(5.64-95.10)	0.96	2.45E-02
DOCK9	2	79.75(16.98-374.52)	0.99	5.42E-05	65.50(15.65-274.08)	0.98	1.94E-05
DPYSL2	2	57.31(16.65-197.29)	0.98	2.58E-07	54.42(16.87-175.52)	0.98	4.24E-08
EIF3C	2	119.64(23.81-601.07)	0.99	1.18E-05	158.60(36.38-691.39)	0.99	2.89E-08
NRXN3	3	9.81(3.58-26.90)	0.90	1.70E-02	8.89(3.28-24.12)	0.89	3.36E-02
PDE10A	3	12.95(4.04-41.51)	0.92	3.08E-02	11.81(3.74-37.27)	0.92	4.78E-02
EPS15L1	2	44.85(10.16-198.04)	0.98	9.78E-04	57.94(13.89-241.63)	0.98	4.77E-05
GAPVD1	2	12.31(4.44-34.10)	0.92	2.61E-03	18.26(6.66-50.08)	0.95	3.13E-05
HIBCH	2	37.76(8.66-164.63)	0.97	2.52E-03	42.43(10.24-175.72)	0.98	4.44E-04
HPCAL1	2	239.29(39.49-1450.11)	0.99	4.79E-06	94.16(22.22-399.07)	0.99	1.30E-06
ITSN2	2	35.11(10.53-117.05)	0.97	1.31E-05	21.95(6.89-69.99)	0.95	3.34E-04
MCCC1	2	51.26(11.18-228.96)	0.98	4.77E-04	35.86(8.68-148.07)	0.97	1.42E-03
MYO5A	2	239.29(39.49-1450.11)	0.99	4.79E-06	111.60(26.13-476.64)	0.99	3.68E-07
NOMO1	2	47.84(10.78-212.39)	0.98	6.89E-04	27.38(6.65-112.62)	0.96	8.50E-03
PARP1	2	65.25(14.25-298.75)	0.98	1.39E-04	38.13(9.22-157.61)	0.97	9.33E-04
PDHB	2	358.95(49.99-2577.53)	0.99	9.32E-06	1506.91(209.86-10820.36)	0.99	6.50E-10
PFN2	2	143.57(27.48-750.14)	0.99	7.39E-06	215.25(48.19-961.45)	0.99	3.76E-09
POR	2	20.52(6.29-66.96)	0.95	1.04E-03	18.56(5.83-59.09)	0.95	1.45E-03
PYGB	2	358.95(49.99-2577.53)	0.99	9.32E-06	251.13(55.38-1138.89)	0.99	1.48E-09
RAB5B	2	33.36(11.73-94.94)	0.97	9.24E-08	35.42(12.85-97.64)	0.97	1.01E-08
REV3L	2	119.64(23.81-601.07)	0.99	1.18E-05	115.89(27.08-495.90)	0.99	2.78E-07
RPL30	2	102.54(21-500.67)	0.99	1.97E-05	79.29(18.83-333.84)	0.99	4.69E-06

SAE1	2	179.47(34.44-992.81)	0.99	5.16E-06	273.96(59.83-1254.41)	0.99	9.04E-10
SH3GL3	2	90.76(25.16-327.42)	0.99	1.07E-08	101.61(30.98-333.30)	0.99	4.61E-11
SORBS2	2	37.76(8.66-164.63)	0.97	2.52E-03	42.43(10.24-175.72)	0.98	4.44E-04
SYNPO	2	43.55(12.90-146.97)	0.98	2.25E-06	33.85(10.57-108.37)	0.97	5.65E-06
TCP11L1	2	15.10(4.67-48.86)	0.93	1.11E-02	21.34(6.69-68.01)	0.95	4.30E-04
VPS41	2	143.57(27.48-750.14)	0.99	7.39E-06	111.60(26.13-476.64)	0.99	3.68E-07

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by number of indels.

Table S12: List of synaptic genes showing enrichment of indels in Swedish tinnitus cohort compared with SweGen and datasets

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	SweGe OR(CI
Non selected Tinnitus (N=97)								
APC	4	2.58(2.08-3.20)	0.61	<1.00E-15	1.93(1.56-2.39)	0.48	2.58E-06	1.73(1.
CLASP2	4	3.13(1.92-5.10)	0.68	9.48E-03	2.82(1.74-4.56)	0.64	4.92E-02	11.18(4
Severe tinnitus(N=34)								
AGL	4	6.93(4.01-11.95)	0.86	6.6E-09	7.91(4.61-13.58)	0.87	1.11E-10	5.37(2.
APC	4	3.48(2.52-4.81)	0.71	9.67E-11	2.61(1.89-3.61)	0.62	1.08E-05	2.33(1.
CLASP2	4	5.51(2.99-10.13)	0.82	8.00E-05	5.06(2.77-9.27)	0.80	2.75E-04	16.82(7
PC	4	4.22(3.07-5.81)	0.76	<1.00E-15	2.99(2.17-4.10)	0.67	2.67E-08	4.37(3.
ACACA	3	172.27(92.86-319.59)	0.99	<1.00E-15	4.13(2.54-6.70)	0.76	1.88E-05	4.02(2.
APPL2	2	5.16(3.41-7.80)	0.81	1.42E-11	5.28(3.50-7.96)	0.81	4.19E-12	5.65(3.

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ord

Table S13: List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients with Me (MD)-EP when they were compared against CSVS reference dataset

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[CSVS] OR(CI)
ADGRV1	16	1.74(1.24-2.44)	0.42	*NS*	1.70(1.21-2.38)	0.41	*NS*	2.18(1.54-3
USH1G	6	19.56(8.69-44.01)	0.95	1.01E-10	22.63(10.02-51.12)	0.96	9.65E-12	7.16(3.06-1
ILDR1	3	24.09(7.64-75.93)	0.96	8.50E-06	28.14(8.86-89.37)	0.96	2.29E-06	16.44(4.60
MYO3A	3	3.40(1.40-8.28)	0.71	*NS*	3.48(1.43-8.49)	0.71	*NS*	5.71(2.26-1
OTOA	3	96.40(29.95-310.25)	0.99	2.80E-12	102.95(31.24-339.27)	0.99	3.98E-12	16.44(4.60
PCDH15	3	13.99(4.45-43.97)	0.93	9.61E-04	14.37(4.56-45.32)	0.93	8.22E-04	12.33(3.56
NARS2	2	116.85(27.72-492.54)	0.99	1.34E-08	122.01(28-531.73)	0.99	2.41E-08	32.90(5.97
CACNA1D	2	37.43(9.13-153.43)	0.97	7.37E-05	26.13(6.36-107.46)	0.96	9.24E-04	32.90(5.97
CDC14A	2	5.69(1.40-23.06)	0.82	*NS*	6.48(1.60-26.33)	0.85	*NS*	32.90(5.97

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered

Table S14: List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients with M phenotype (MD-AEP) when they were compared against CSVS reference dataset

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[CSVS] OR(CI)
TRIOBP	13	2.09(1.35-3.23)	0.52	*NS*	2.19(1.42-3.39)	0.54	*NS*	2.42(1.56-
DSPP	12	2.26(1.43-3.57)	0.56	*NS*	1.52(0.96-2.40)	0.34	*NS*	2.92(1.83-
DFNB31	3	35.21(11.13-111.34)	0.97	2.04E-07	22.87(7.22-72.43)	0.96	1.56E-05	25.54(6.72
DIAPH1	2	58.69(14.21-242.34)	0.98	2.77E-06	43.71(10.52-181.58)	0.98	3.05E-05	34.05(6.17

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered

Table S15: Gene Ontology analyses showing the list of genes found in Reactome pathways and GO biological proc

Pathway	**Total genes (N)**	**Candidate genes (N)**	**Gene(s) name**	**P-value**
Membrane trafficking	629	13	*AP2A2,TUBB2A,DNM3,ANK2,WASL,TSC2,AP1G1,MYO5A, BIN1,AAK1,KIF20B,SNX5,MADD*	1.47E-12
Vesicle mediated transport	724	13	*AP2A2,TUBB2A,DNM3,ANK2,WASL,TSC2,AP1G1,MYO5A, BIN1,AAK1,KIF20B,SNX5,MADD*	8.52E-12
Nervous system development	580	9	*AP2A2,TUBB2A,DNM3,ANK2,WASL,,NRCAM,GSK3B,PLX NA2,MBP*	6.59E-08
L1cam interactions	121	5	*AP2A2,TUBB2A,DNM3,ANK2,WASL,NRCAM*	5.90E-07
Clathrin mediated endocytosis	145	5	*AP2A2,DNM3,WASL,BIN1,AAK1*	1.44E-06
Biological process	**Total genes (N)**	**Candidate genes (N)**	**Gene(s) name**	**P-value**
Cytoskeletal protein binding	979	20	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C, GSK3B,DNM3,NRCAM,LLGL1,TAOK2,WASL,SORBS1,FLII, MPRIP,KIF20B,AP1G1,SNX5,LRPPRC*	4.49E-19
Synapse	1357	22	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C, GSK3B,DNM3,NRCAM, AKAP9,SIPA1L1,IQSEC1,AAK1,CDH13,SNAP47,MBP,RGS 8,TSC2,ICA1,MADD,SGTA*	6.48E-19
Actin filament based process	804	18	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C,L LGL1,TAOK2,WASL,SORBS1,FLII,MPRIP,AKAP9,SIPA1L1, IQSEC1,MRAS, CACNA2D1*	7.71E-18
Neuron projection	1366	21	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2, GSK3B,DNM3,NRCAM,LLGL1,TAOK2,KIF20B, AKAP9, SIPA1L1, AAK1,CDH13,SNAP47,MBP,RGS8,DOCK7,NDRG2,VCAN*	1.48E-17
Cytoskeleton organization	1396	20	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C, GSK3B,LLGL1,TAOK2,WASL,SORBS1,FLII,MPRIP, AKAP9,SIPA1L1,IQSEC,MARS,DOCK7,TUBB2A*	4.11E-16
Actin cytoskeleton	503	13	*PPP1R9A,MYO5A,BIN1,SYNPO, MYO18A,CORO1C,LLGL1,TAOK2,WASL,SORBS1,FLII,MP RIP, SH3PXD2A*	8.77E-14

Regulation of transport	1856	19	*PPP1R9A,MYO5A,BIN1, ANK2,MYO18A, GSK3B,DNM3, LLGL1, WASL,SORBS1, KIF20B,AP1G1,SNX5,AKAP9,AAK1,CDH13,TSC2,ICA1, CACNA2D1*	1.07E-12
Cellular component morphogenesis	800	14	*PPP1R9A,ANK2, CORO1C,GSK3B,DNM3,NRCAM,LLGL1,TAOK2,WASL, FLII,KIF20B, SIPA1L1,DOCK7, PLXNA2*	1.67E-12
Postsynapse	640	13	*PPP1R9A,MYO5A,SYNPO,ANK2,MYO18A, GSK3B,DNM3,NRCAM, AKAP9,SIPA1L1,IQSEC1, SNAP47,TSC2*	1.83E-12
Axon	643	13	*PPP1R9A,MYO5A,BIN1, GSK3B,DNM3,NRCAM,LLGL1,TAOK2,KIF20B,AAK1,MBP, DOCK7, NDRG2*	1.94E-12

FDR=False discovery rate

Table S16: Hypothesis-free data driven approach. Genes showing enrichment of missense variants in Spanish pat extreme phenotype (MD-EP). Listed genes were significant when they were compared against gnomAD Non-Finnish Eu

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CS OI
ACAN	13	5.51(3.45-8.8)	0.82	1.77E-08	3.88(2.43-6.2)	0.74	2.78E-04	4.(
PRAMEF1	12	7.49(4.94-11.36)	0.87	<1.00E-15	9.89(6.52-15.01)	0.90	<1.00E-15	13
MYH7B	10	6.45(3.64-11.43)	0.84	3.55E-06	5.46(3.08-9.69)	0.82	1.25E-04	7.2
ADGRV1	9	21.5(12.1-38.23)	0.95	<1.00E-15	23.19(13.01-41.34)	0.96	<1.00E-15	18
DCHS1	9	7.11(4.01-12.61)	0.86	4.02E-07	7.32(4.13-13)	0.86	2.11E-07	7.(
PRUNE2	9	6.02(3.6-10.07)	0.83	1.53E-07	5.89(3.52-9.85)	0.83	2.92E-07	5.4
ZNF469	9	5.35(2.77-10.35)	0.81	1.25E-02	5.24(2.7-10.13)	0.81	1.79E-02	7.2
CEP290	8	6.99(3.73-13.1)	0.86	2.48E-05	6.67(3.56-12.5)	0.85	6.40E-05	5.4
CEP295	8	117.74(57.34-241.77)	0.99	<1.00E-15	95.48(46.1-197.76)	0.99	<1.00E-15	16
C5orf42	7	57.67(27.02-123.06)	0.98	0.00E+00	58.67(27.26-126.26)	0.98	<1.00E-15	28
CASZ1	7	5.92(2.94-11.92)	0.83	1.32E-02	5.88(2.92-11.87)	0.83	1.47E-02	9.(
ENOSF1	7	38.11(17.93-81.03)	0.97	<1.00E-15	19.06(8.97-40.49)	0.95	3.51E-10	16
MCM8	7	7(3.31-14.8)	0.86	6.90E-03	7.18(3.39-15.2)	0.86	5.05E-03	7.3
SEC16A	7	12.82(6.36-25.88)	0.92	2.10E-08	13.35(6.6-26.98)	0.93	1.07E-08	9.(
WDR62	7	20.82(11.08-39.11)	0.95	<1.00E-15	21.67(11.5-40.82)	0.95	<1.00E-15	11
ABCA1	6	17.51(7.79-39.39)	0.94	8.83E-08	16.42(7.28-37.01)	0.94	3.01E-07	16
ADAMTS10	6	31.68(14.05-71.46)	0.97	<1.00E-15	33.78(14.89-76.61)	0.97	<1.00E-15	14
ARHGAP39	6	32.16(15.87-65.19)	0.97	<1.00E-15	28.85(14.19-58.67)	0.97	<1.00E-15	22
BAIAP3	6	6.32(3.13-12.75)	0.84	5.26E-03	5.91(2.93-11.93)	0.83	1.45E-02	7.(
CEP250	6	9.16(4.72-17.77)	0.89	1.19E-06	8.58(4.41-16.66)	0.88	4.55E-06	5.2
FLNB	6	11.67(5.19-26.21)	0.91	5.35E-05	12.83(5.7-28.88)	0.92	1.43E-05	10
PRRC2C	6	29.27(12.98-66)	0.97	8.88E-12	32.93(14.52-74.68)	0.97	<1.00E-15	10
SHROOM2	6	131.4(60.68-284.55)	0.99	<1.00E-15	148.24(66.88-328.59)	0.99	<1.00E-15	57
AKAP9	**6**	**12.32(5.48-27.68)**	**0.92**	**2.40E-05**	**13.89(6.17-31.27)**	**0.93**	**4.28E-06**	**6.(
CHRNG	5	9.19(3.79-22.29)	0.89	1.85E-02	9.84(4.05-23.91)	0.90	8.93E-03	17
CIC	5	97.54(42.61-223.24)	0.99	<1.00E-15	83.68(36.2-193.41)	0.99	<1.00E-15	33
EML6	5	166.22(66.2-417.35)	0.99	<1.00E-15	203.37(77.79-531.69)	1.00	<1.00E-15	27
FAM71E2	5	16.44(6.77-39.93)	0.94	1.27E-05	14.86(6.11-36.19)	0.93	5.53E-05	27
ITGAX	5	73.02(29.68-179.66)	0.99	<1.00E-15	61.68(24.88-152.94)	0.98	<1.00E-15	14
MPDZ	5	36.37(14.9-88.74)	0.97	5.77E-11	33.88(13.81-83.1)	0.97	2.84E-10	11
PELP1	5	8.42(3.74-18.92)	0.88	5.09E-03	7.65(3.4-17.21)	0.87	1.76E-02	11
RTTN	5	16.27(6.7-39.52)	0.94	1.46E-05	15.41(6.33-37.52)	0.94	3.43E-05	20
SPATA31D1	5	120.5(48.47-299.57)	0.99	<1.00E-15	21.19(8.68-51.71)	0.95	3.98E-07	27
SPTB	5	15.97(6.58-38.8)	0.94	1.88E-05	18.47(7.58-45.04)	0.95	2.84E-06	13
TRPV1	5	25.82(12.73-52.35)	0.96	<1.00E-15	31.12(15.27-63.42)	0.97	<1.00E-15	13
USP17L10	5	14.73(7.82-27.72)	0.93	<1.00E-15	17.93(9.5-33.84)	0.94	<1.00E-15	21
ANK2	4	18.3(6.78-49.4)	0.95	1.91E-04	19.95(7.36-54.08)	0.95	8.00E-05	21
ARHGAP9	4	12.28(4.56-33.11)	0.92	1.42E-02	12.75(4.72-34.46)	0.92	1.03E-02	16
BCAR3	4	22.24(9.12-54.19)	0.96	1.76E-07	19.82(8.11-48.43)	0.95	1.15E-06	14
C7orf33	4	14.11(5.23-38.04)	0.93	3.38E-03	11.86(4.39-32.02)	0.92	2.14E-02	21
CACNA1S	4	14.65(5.43-39.5)	0.93	2.27E-03	16.69(6.17-45.16)	0.94	6.02E-04	65
CCDC178	4	21.06(7.8-56.87)	0.95	3.68E-05	16.13(5.96-43.66)	0.94	8.71E-04	65
CHAD	4	254.78(99.41-652.99)	1.00	<1.00E-15	367.63(132.54-1019.7)	1.00	<1.00E-15	41

DSCAML1	4	105.64(38.3-291.38)	0.99	<1.00E-15	122.01(43.09-345.5)	0.99	<1.00E-15	32
FRMPD1	4	33.37(13.66-81.52)	0.97	2.80E-10	40.83(16.57-100.6)	0.98	1.33E-11	11
GBP5	4	11.81(4.38-31.83)	0.92	2.10E-02	13.54(5.01-36.6)	0.93	5.59E-03	32
KANK1	4	48.19(17.72-131.04)	0.98	6.26E-10	32.29(11.85-87.98)	0.97	2.19E-07	21
KIF20B	4	7.76(3.45-17.49)	0.87	1.50E-02	8.43(3.74-19.01)	0.88	5.59E-03	16
KLHDC4	4	14.01(5.2-37.76)	0.93	3.65E-03	14.72(5.45-39.82)	0.93	2.33E-03	65
LRPPRC	4	49.75(18.29-135.32)	0.98	3.95E-10	73.2(26.39-203.03)	0.99	4.44E-12	21
LSG1	4	7.34(3.89-13.84)	0.86	1.46E-05	7.47(3.96-14.11)	0.87	1.08E-05	5.2
MYO3B	4	85.15(37.22-194.78)	0.99	<1.00E-15	79.09(34.21-182.86)	0.99	<1.00E-15	12
MYO7B	4	12.28(4.56-33.11)	0.92	1.42E-02	14.72(5.45-39.82)	0.93	2.33E-03	13
NOTCH3	4	74.87(27.35-204.94)	0.99	<1.00E-15	82.87(29.76-230.81)	0.99	<1.00E-15	21
POLQ	4	29.2(10.79-79.03)	0.97	6.17E-07	20.61(7.6-55.87)	0.95	5.50E-05	21
POM121L12	4	65.91(24.13-180.01)	0.98	4.44E-12	48.26(17.59-132.42)	0.98	1.03E-09	32
PRDM2	4	27.14(10.03-73.42)	0.96	1.58E-06	25.09(9.24-68.15)	0.96	5.23E-06	16
RECQL4	4	31.42(12.86-76.72)	0.97	7.59E-10	31.49(12.83-77.31)	0.97	1.03E-09	16
SEMA5B	4	37.43(13.8-101.5)	0.97	2.21E-08	38.18(13.98-104.32)	0.97	2.44E-08	16
SLFN12	4	122.42(44.2-339.03)	0.99	<1.00E-15	129.19(45.49-366.93)	0.99	0.00E+00	65
SPTBN4	4	65.91(24.13-180.01)	0.98	4.44E-12	72(25.97-199.6)	0.99	4.44E-12	65
SSFA2	4	86.65(31.56-237.87)	0.99	0.00E+00	104.58(37.2-293.97)	0.99	0.00E+00	16
SYDE2	4	265.96(92.77-762.42)	1.00	0.00E+00	99.83(35.58-280.04)	0.99	0.00E+00	21
SYNPO	4	74.87(27.35-204.94)	0.99	0.00E+00	78.43(28.21-218.03)	0.99	0.00E+00	32
TAS2R30	4	90.75(51.28-160.6)	0.99	0.00E+00	11.63(6.63-20.4)	0.91	0.00E+00	13
TBC1D8	4	265.96(92.77-762.42)	1.00	0.00E+00	199.67(68.28-583.85)	0.99	0.00E+00	21
TNKS1BP1	4	18.57(6.88-50.12)	0.95	1.62E-04	17.14(6.33-46.41)	0.94	4.47E-04	16
TSC2	4	63.73(23.35-173.96)	0.98	8.88E-12	53.56(19.47-147.3)	0.98	2.49E-10	21
USP25	4	49.75(18.29-135.32)	0.98	3.95E-10	45.27(16.52-124.07)	0.98	2.47E-09	21
B4GALNT3	3	7.62(3.74-15.5)	0.87	4.19E-04	6.86(3.37-13.96)	0.85	2.23E-03	6.3
BIN1	3	56.7(17.82-180.42)	0.98	1.62E-07	73.2(22.54-237.74)	0.99	1.82E-08	49
C11orf80	3	304.45(89.3-1038.02)	1.00	<1.00E-15	173.39(50.86-591.18)	0.99	4.44E-12	49
CCDC88B	3	304.45(89.3-1038.02)	1.00	<1.00E-15	235.33(67.04-826.03)	1.00	<1.00E-15	49
CDH13	3	12.15(4.5-32.85)	0.92	1.69E-02	13.09(4.83-35.46)	0.92	8.54E-03	33
COG5	3	101.47(31.48-327.05)	0.99	2.04E-10	126.71(38-422.44)	0.99	6.66E-11	49
DHX34	3	525.88(145.47-1901.09)	1.00	<1.00E-15	253.43(71.6-897.05)	1.00	<1.00E-15	49
FAAP100	3	13.9(5.14-37.59)	0.93	4.27E-03	13.57(5.01-36.78)	0.93	5.89E-03	33
IQCC	3	12.14(4.49-32.79)	0.92	1.72E-02	11.73(4.33-31.76)	0.91	2.55E-02	16
LAMA4	3	310.26(106.87-900.72)	1.00	<1.00E-15	736.32(206-2631.85)	1.00	<1.00E-15	13
LLGL1	3	27.29(10.06-74.03)	0.96	1.67E-06	25.08(9.21-68.31)	0.96	5.86E-06	13
MADD	3	128.54(39.58-417.46)	0.99	1.33E-11	73.2(22.54-237.74)	0.99	1.82E-08	49
MBD6	3	28.2(8.94-89.01)	0.96	2.47E-04	26.77(8.44-84.95)	0.96	4.82E-04	49
MPRIP	3	82.62(25.77-264.88)	0.99	2.22E-09	78.43(24.09-255.39)	0.99	8.82E-09	49
MTMR8	3	14.29(7.54-27.1)	0.93	8.88E-12	13.89(7.31-26.37)	0.93	1.78E-11	9.2
MYO18A	3	204.11(72.08-577.95)	1.00	<1.00E-15	169.9(58.69-491.88)	0.99	<1.00E-15	33
NLRP6	3	517.11(169.94-1573.52)	1.00	<1.00E-15	883.59(235.31-3317.87)	1.00	<1.00E-15	66
NOS1	3	83.82(26.14-268.81)	0.99	1.88E-09	109.81(33.21-363.11)	0.99	2.71E-10	49
NR1D1	3	73.21(22.9-234.07)	0.99	8.97E-09	78.43(24.09-255.39)	0.99	8.82E-09	49
NRCAM	3	69.68(21.82-222.56)	0.99	1.58E-08	50.67(15.78-162.74)	0.98	8.56E-07	49
NXPE1	3	19.53(6.2-61.47)	0.95	7.59E-03	16.97(5.37-53.57)	0.94	2.78E-02	49
PCDHA7	3	22.76(7.22-71.71)	0.96	1.90E-03	23.69(7.48-75.05)	0.96	1.50E-03	49
PLXNB2	3	79.23(24.74-253.74)	0.99	3.61E-09	76.61(23.55-249.22)	0.99	1.13E-08	49

SHANK1	3	20.02(7.39-54.21)	0.95	7.41E-05	19.97(7.35-54.27)	0.95	8.74E-05	33
SMTN	3	72.47(26.42-198.79)	0.99	<1.00E-15	57.35(20.77-158.41)	0.98	1.11E-10	22
SPOCK1	3	63.55(19.93-202.63)	0.98	4.49E-08	82.35(25.24-268.68)	0.99	5.29E-09	49
SRRM4	3	29.76(13.11-67.56)	0.97	8.88E-12	25.95(11.39-59.09)	0.96	1.78E-10	25
TOGARAM1	3	72.29(22.62-231.08)	0.99	1.04E-08	73.2(22.54-237.74)	0.99	1.82E-08	49
VWA5B1	3	198.88(70.32-562.43)	0.99	<1.00E-15	220.88(74.74-652.81)	1.00	<1.00E-15	33
WNK2	3	482.06(134.87-1722.95)	1.00	<1.00E-15	366.07(98.29-1363.47)	1.00	<1.00E-15	49

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ord

Table S17: Hypothesis-free data driven approach. Genes showing enrichment of synonymous variants in Spanish pa
extreme phenotype (MD-EP). Listed genes were significant when they were compared against gnomAD Non-Finnish Eu

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
DNHD1	8	5.4(3.1-9.3)	0.81	4.67E-05	5.6(3.2-9.7)	0.82	2.32E-05
SPEN	8	24.9(12.3-50.2)	0.96	<1.00E-15	25.5(12.6-51.7)	0.96	<1.00E-15
CFAP46	7	5.4(2.8-10.4)	0.81	1.20E-02	5.7(2.9-11)	0.82	5.65E-03
INTS1	7	8.8(4.4-17.8)	0.89	2.39E-05	9.4(4.7-19)	0.89	7.71E-06
TUBGCP6	7	21.5(10.6-43.5)	0.95	<1.00E-15	18.9(9.3-38.2)	0.95	4.44E-12
CCDC40	6	41.1(19.3-87.6)	0.98	<1.00E-15	47.6(22.2-102.1)	0.98	<1.00E-15
EPG5	6	6.5(3.5-12.2)	0.85	1.09E-04	6.8(3.6-12.8)	0.85	4.80E-05
FN1	6	19.1(9-40.5)	0.95	2.93E-10	21.6(10.1-45.9)	0.95	3.11E-11
LTK	6	7.5(3.3-16.7)	0.87	2.20E-02	11.6(5.2-26.2)	0.91	5.98E-05
NCOR2	6	35.6(15.8-80.3)	0.97	<1.00E-15	32(14.1-72.5)	0.97	<1.00E-15
TECPR1	6	5.2(3-9.1)	0.81	9.15E-05	5.6(3.2-9.7)	0.82	2.68E-05
AP5B1	5	11.9(4.9-28.9)	0.92	8.76E-04	17.3(7.1-42)	0.94	7.39E-06
CCDC168	5	6.5(3-13.7)	0.85	2.24E-02	6.1(2.9-13)	0.84	4.41E-02
CSPG4	5	55.7(22.7-136.5)	0.98	<1.00E-15	28.7(11.7-70.3)	0.97	3.91E-09
KCNT1	5	15.1(6.2-36.7)	0.93	3.97E-05	14.1(5.8-34.4)	0.93	1.07E-04
MIB2	5	47(20.7-106.4)	0.98	<1.00E-15	42.4(18.6-96.5)	0.98	<1.00E-15
MTCL1	5	13.7(5.7-33.4)	0.93	1.41E-04	17(7-41.5)	0.94	8.78E-06
NFATC1	5	51.1(22.5-115.9)	0.98	<1.00E-15	57(24.9-130.5)	0.98	<1.00E-15
OR51E2	5	6.7(4.6-9.9)	0.85	<1.00E-15	6.2(4.2-9.1)	0.84	<1.00E-15
SPTBN2	5	44.1(19.5-99.9)	0.98	<1.00E-15	45.6(20-104)	0.98	<1.00E-15
CRIP2	4	46.2(17-125.5)	0.98	1.16E-09	42.6(15.6-116.7)	0.98	5.57E-09
IARS	4	35.7(13.2-96.8)	0.97	4.24E-08	35.7(13.1-97.4)	0.97	5.92E-08
JPH3	4	15.3(6.8-34.4)	0.93	1.02E-06	16.3(7.2-37)	0.94	3.91E-07
PC	4	233.7(82.2-664.9)	1.00	<1.00E-15	141.7(49.6-404.6)	0.99	<1.00E-15
PHLPP1	4	47.2(19.3-115.7)	0.98	<1.00E-15	51.5(20.8-127.5)	0.98	<1.00E-15
RIMBP2	4	26.7(11-65.2)	0.96	1.03E-08	25.5(10.4-62.5)	0.96	2.74E-08
TRRAP	4	45.4(16.7-123.2)	0.98	1.50E-09	34.6(12.7-94.3)	0.97	9.00E-08
ZFPM1	4	16.7(6.8-40.6)	0.94	1.13E-05	23.7(9.7-58.1)	0.96	8.08E-08
AP3D1	3	525.9(145.5-1901.1)	1.00	<1.00E-15	366.1(98.3-1363.5)	1.00	<1.00E-15
ARHGAP31	3	40.6(16.5-99.7)	0.98	1.33E-11	34.3(13.9-84.4)	0.97	3.24E-10
CCDC154	3	231.4(69.2-773.3)	1.00	<1.00E-15	156.9(46.4-530.7)	0.99	8.88E-12
DEAF1	3	28.3(9-89.4)	0.96	2.36E-04	22.1(7-70)	0.95	2.81E-03
FAM90A1	3	31.3(13.8-71)	0.97	4.44E-12	29.6(13-67.6)	0.97	1.33E-11
FGD4	3	26.8(9.9-72.7)	0.96	2.08E-06	21.7(8-59.1)	0.95	3.24E-05
HIST1H1C	3	76(27.7-208.7)	0.99	<1.00E-15	133.9(46.9-381.8)	0.99	<1.00E-15
IQSEC2	3	22.7(8.4-61.4)	0.96	1.69E-05	25.4(9.3-69.1)	0.96	5.10E-06
MAGEC1	3	105.2(32.6-339.3)	0.99	1.33E-10	67.2(20.8-217.7)	0.99	4.46E-08
MASP1	3	37.8(13.9-102.8)	0.97	2.21E-08	35.9(13.1-98.2)	0.97	6.28E-08
NEK10	3	33.9(12.5-92)	0.97	9.94E-08	33.7(12.3-92.1)	0.97	1.43E-07
SCN4A	3	6.8(4.2-11.1)	0.85	2.58E-10	6.8(4.2-11)	0.85	3.33E-10
SETD2	3	170.1(51.8-559)	0.99	<1.00E-15	173.4(50.9-591.2)	0.99	4.44E-12
SETMAR	3	16.4(7.2-37.1)	0.94	3.98E-07	21.5(9.4-48.8)	0.95	5.13E-09
SIN3A	3	8.9(4.4-18.2)	0.89	2.97E-05	9.2(4.5-18.7)	0.89	2.01E-05

YTHDF1	3	53.8(21.9-132.6)	0.98	<1.00E-15	50.5(20.3-125.2)	0.98	<1.00E-15	
ZEB1	3	41(12.9-129.9)	0.98	5.47E-06	37.9(11.9-120.8)	0.97	1.67E-05	

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, gen

Table S18: Hypothesis-free data driven approach. The filter was applied to exclude hearing loss genes previously assoc
Genes showing enrichment of missense variants in Spanish patients with Meniere disease and tinnitus extreme phenoty
when they were compared against gnomAD Non-Finnish European and CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CS OR
PRAMEF1	12	7.49(4.94-11.36)	0.87	<1.00E-15	9.89(6.52-15.01)	0.90	<1.00E-15	13
DCHS1	9	7.11(4.01-12.61)	0.86	4.02E-07	7.32(4.13-13)	0.86	2.11E-07	7.0
CEP290	8	6.99(3.73-13.1)	0.86	2.48E-05	6.67(3.56-12.5)	0.85	6.40E-05	5.4
CEP295	8	117.74(57.34-241.77)	0.99	<1.00E-15	95.48(46.1-197.76)	0.99	<1.00E-15	16
C5orf42	7	57.67(27.02-123.06)	0.98	0.00E+00	58.67(27.26-126.26)	0.98	<1.00E-15	28
CASZ1	7	5.92(2.94-11.92)	0.83	1.32E-02	5.88(2.92-11.87)	0.83	1.47E-02	9.0
ENOSF1	7	38.11(17.93-81.03)	0.97	<1.00E-15	19.06(8.97-40.49)	0.95	3.51E-10	16
MCM8	7	7(3.31-14.8)	0.86	6.90E-03	7.18(3.39-15.2)	0.86	5.05E-03	7.3
WDR62	7	20.82(11.08-39.11)	0.95	<1.00E-15	21.67(11.5-40.82)	0.95	<1.00E-15	11
ABCA1	6	17.51(7.79-39.39)	0.94	8.83E-08	16.42(7.28-37.01)	0.94	3.01E-07	16
ADAMTS10	6	31.68(14.05-71.46)	0.97	<1.00E-15	33.78(14.89-76.61)	0.97	<1.00E-15	14
ARHGAP39	6	32.16(15.87-65.19)	0.97	<1.00E-15	28.85(14.19-58.67)	0.97	<1.00E-15	22
BAIAP3	6	6.32(3.13-12.75)	0.84	5.26E-03	5.91(2.93-11.93)	0.83	1.45E-02	7.0
CEP250	6	9.16(4.72-17.77)	0.89	1.19E-06	8.58(4.41-16.66)	0.88	4.55E-06	5.2
FLNB	6	11.67(5.19-26.21)	0.91	5.35E-05	12.83(5.7-28.88)	0.92	1.43E-05	10
PRRC2C	6	29.27(12.98-66)	0.97	8.88E-12	32.93(14.52-74.68)	0.97	<1.00E-15	10
SHROOM2	6	131.4(60.68-284.55)	0.99	<1.00E-15	148.24(66.88-328.59)	0.99	<1.00E-15	57
AKAP9	6	12.32(5.48-27.68)	0.92	2.40E-05	13.89(6.17-31.27)	0.93	4.28E-06	6.0
CHRNG	5	9.19(3.79-22.29)	0.89	1.85E-02	9.84(4.05-23.91)	0.90	8.93E-03	17
CIC	5	97.54(42.61-223.24)	0.99	<1.00E-15	83.68(36.2-193.41)	0.99	<1.00E-15	33
EML6	5	166.22(66.2-417.35)	0.99	<1.00E-15	203.37(77.79-531.69)	1.00	<1.00E-15	27
FAM71E2	5	16.44(6.77-39.93)	0.94	1.27E-05	14.86(6.11-36.19)	0.93	5.53E-05	27
ITGAX	5	73.02(29.68-179.66)	0.99	<1.00E-15	61.68(24.88-152.94)	0.98	<1.00E-15	14
MPDZ	5	36.37(14.9-88.74)	0.97	5.77E-11	33.88(13.81-83.1)	0.97	2.84E-10	11
PELP1	5	8.42(3.74-18.92)	0.88	5.09E-03	7.65(3.4-17.21)	0.87	1.76E-02	11
RTTN	5	16.27(6.7-39.52)	0.94	1.46E-05	15.41(6.33-37.52)	0.94	3.43E-05	20
SPATA31D1	5	120.5(48.47-299.57)	0.99	<1.00E-15	21.19(8.68-51.71)	0.95	3.98E-07	27
USP17L10	5	14.73(7.82-27.72)	0.93	<1.00E-15	17.93(9.5-33.84)	0.94	<1.00E-15	21
ANK2	4	18.3(6.78-49.4)	0.95	1.91E-04	19.95(7.36-54.08)	0.95	8.00E-05	21
ARHGAP9	4	12.28(4.56-33.11)	0.92	1.42E-02	12.75(4.72-34.46)	0.92	1.03E-02	16
BCAR3	4	22.24(9.12-54.19)	0.96	1.76E-07	19.82(8.11-48.43)	0.95	1.15E-06	14
C7orf33	4	14.11(5.23-38.04)	0.93	3.38E-03	11.86(4.39-32.02)	0.92	2.14E-02	21
CCDC178	4	21.06(7.8-56.87)	0.95	3.68E-05	16.13(5.96-43.66)	0.94	8.71E-04	65
CHAD	4	254.78(99.41-652.99)	1.00	<1.00E-15	367.63(132.54-1019.7)	1.00	<1.00E-15	41
DSCAML1	4	105.64(38.3-291.38)	0.99	<1.00E-15	122.01(43.09-345.5)	0.99	<1.00E-15	32
FRMPD1	4	33.37(13.66-81.52)	0.97	2.80E-10	40.83(16.57-100.6)	0.98	1.33E-11	11
GBP5	4	11.81(4.38-31.83)	0.92	2.10E-02	13.54(5.01-36.6)	0.93	5.59E-03	32
KANK1	4	48.19(17.72-131.04)	0.98	6.26E-10	32.29(11.85-87.98)	0.97	2.19E-07	21
KIF20B	4	7.76(3.45-17.49)	0.87	1.50E-02	8.43(3.74-19.01)	0.88	5.59E-03	16
KLHDC4	4	14.01(5.2-37.76)	0.93	3.65E-03	14.72(5.45-39.82)	0.93	2.33E-03	65
LRPPRC	4	49.75(18.29-135.32)	0.98	3.95E-10	73.2(26.39-203.03)	0.99	4.44E-12	21
LSG1	4	7.34(3.89-13.84)	0.86	1.46E-05	7.47(3.96-14.11)	0.87	1.08E-05	5.2

MYO3B	4	85.15(37.22-194.78)	0.99	<1.00E-15	79.09(34.21-182.86)	0.99	<1.00E-15	12
MYO7B	4	12.28(4.56-33.11)	0.92	1.42E-02	14.72(5.45-39.82)	0.93	2.33E-03	13
NOTCH3	4	74.87(27.35-204.94)	0.99	<1.00E-15	82.87(29.76-230.81)	0.99	<1.00E-15	21
POLQ	4	29.2(10.79-79.03)	0.97	6.17E-07	20.61(7.6-55.87)	0.95	5.50E-05	21
POM121L12	4	65.91(24.13-180.01)	0.98	4.44E-12	48.26(17.59-132.42)	0.98	1.03E-09	32
PRDM2	4	27.14(10.03-73.42)	0.96	1.58E-06	25.09(9.24-68.15)	0.96	5.23E-06	16
RECQL4	4	31.42(12.86-76.72)	0.97	7.59E-10	31.49(12.83-77.31)	0.97	1.03E-09	16
SEMA5B	4	37.43(13.8-101.5)	0.97	2.21E-08	38.18(13.98-104.32)	0.97	2.44E-08	16
SLFN12	4	122.42(44.2-339.03)	0.99	<1.00E-15	129.19(45.49-366.93)	0.99	0.00E+00	65
SPTBN4	4	65.91(24.13-180.01)	0.98	4.44E-12	72(25.97-199.6)	0.99	4.44E-12	65
SSFA2	4	86.65(31.56-237.87)	0.99	0.00E+00	104.58(37.2-293.97)	0.99	0.00E+00	16
SYDE2	4	265.96(92.77-762.42)	1.00	0.00E+00	99.83(35.58-280.04)	0.99	0.00E+00	21
SYNPO	4	74.87(27.35-204.94)	0.99	0.00E+00	78.43(28.21-218.03)	0.99	0.00E+00	32
TAS2R30	4	90.75(51.28-160.6)	0.99	0.00E+00	11.63(6.63-20.4)	0.91	0.00E+00	13
TBC1D8	4	265.96(92.77-762.42)	1.00	0.00E+00	199.67(68.28-583.85)	0.99	0.00E+00	21
TNKS1BP1	4	18.57(6.88-50.12)	0.95	1.62E-04	17.14(6.33-46.41)	0.94	4.47E-04	16
TSC2	4	63.73(23.35-173.96)	0.98	8.88E-12	53.56(19.47-147.3)	0.98	2.49E-10	21
USP25	4	49.75(18.29-135.32)	0.98	3.95E-10	45.27(16.52-124.07)	0.98	2.47E-09	21
B4GALNT3	3	7.62(3.74-15.5)	0.87	4.19E-04	6.86(3.37-13.96)	0.85	2.23E-03	6.3
BIN1	3	56.7(17.82-180.42)	0.98	1.62E-07	73.2(22.54-237.74)	0.99	1.82E-08	49
C11orf80	3	304.45(89.3-1038.02)	1.00	<1.00E-15	173.39(50.86-591.18)	0.99	4.44E-12	49
CCDC88B	3	304.45(89.3-1038.02)	1.00	<1.00E-15	235.33(67.04-826.03)	1.00	<1.00E-15	49
CDH13	3	12.15(4.5-32.85)	0.92	1.69E-02	13.09(4.83-35.46)	0.92	8.54E-03	33
COG5	3	101.47(31.48-327.05)	0.99	2.04E-10	126.71(38-422.44)	0.99	6.66E-11	49
DHX34	3	525.88(145.47-1901.09)	1.00	<1.00E-15	253.43(71.6-897.05)	1.00	<1.00E-15	49
FAAP100	3	13.9(5.14-37.59)	0.93	4.27E-03	13.57(5.01-36.78)	0.93	5.89E-03	33
IQCC	3	12.14(4.49-32.79)	0.92	1.72E-02	11.73(4.33-31.76)	0.91	2.55E-02	16
LAMA4	3	310.26(106.87-900.72)	1.00	<1.00E-15	736.32(206-2631.85)	1.00	<1.00E-15	13
LLGL1	3	27.29(10.06-74.03)	0.96	1.67E-06	25.08(9.21-68.31)	0.96	5.86E-06	13
MADD	3	128.54(39.58-417.46)	0.99	1.33E-11	73.2(22.54-237.74)	0.99	1.82E-08	49
MBD6	3	28.2(8.94-89.01)	0.96	2.47E-04	26.77(8.44-84.95)	0.96	4.82E-04	49
MPRIP	3	82.62(25.77-264.88)	0.99	2.22E-09	78.43(24.09-255.39)	0.99	8.82E-09	49
MTMR8	3	14.29(7.54-27.1)	0.93	8.88E-12	13.89(7.31-26.37)	0.93	1.78E-11	9.2
MYO18A	3	204.11(72.08-577.95)	1.00	<1.00E-15	169.9(58.69-491.88)	0.99	<1.00E-15	33
NLRP6	3	517.11(169.94-1573.52)	1.00	<1.00E-15	883.59(235.31-3317.87)	1.00	<1.00E-15	66
NR1D1	3	73.21(22.9-234.07)	0.99	8.97E-09	78.43(24.09-255.39)	0.99	8.82E-09	49
NRCAM	3	69.68(21.82-222.56)	0.99	1.58E-08	50.67(15.78-162.74)	0.98	8.56E-07	49
NXPE1	3	19.53(6.2-61.47)	0.95	7.59E-03	16.97(5.37-53.57)	0.94	2.78E-02	49
PCDHA7	3	22.76(7.22-71.71)	0.96	1.90E-03	23.69(7.48-75.05)	0.96	1.50E-03	49
PLXNB2	3	79.23(24.74-253.74)	0.99	3.61E-09	76.61(23.55-249.22)	0.99	1.13E-08	49
SHANK1	3	20.02(7.39-54.21)	0.95	7.41E-05	19.97(7.35-54.27)	0.95	8.74E-05	33
SMTN	3	72.47(26.42-198.79)	0.99	<1.00E-15	57.35(20.77-158.41)	0.98	1.11E-10	22
SPOCK1	3	63.55(19.93-202.63)	0.98	4.49E-08	82.35(25.24-268.68)	0.99	5.29E-09	49
SRRM4	3	29.76(13.11-67.56)	0.97	8.88E-12	25.95(11.39-59.09)	0.96	1.78E-10	25
TOGARAM1	3	72.29(22.62-231.08)	0.99	1.04E-08	73.2(22.54-237.74)	0.99	1.82E-08	49
VWA5B1	3	198.88(70.32-562.43)	0.99	<1.00E-15	220.88(74.74-652.81)	1.00	<1.00E-15	33
WNK2	3	482.06(134.87-1722.95)	1.00	<1.00E-15	366.07(98.29-1363.47)	1.00	<1.00E-15	49

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Table S19: Hypothesis-free data driven approach. Genes showing enrichment of missense variants in Spanish patients extreme phenotype (MD-AEP). Listed genes were significant when they were compared against gnomAD Non-Finnish E

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(C
ZNF729	13	7.26(4.83-10.92)	0.86	<1.00E-15	3.3(2.2-4.96)	0.70	1.84E-04	8.61(5
ATM	8	9.66(5.3-17.59)	0.90	2.43E-09	10.51(5.76-19.16)	0.90	3.33E-10	5.07(2
CCDC168	8	12.24(6.07-24.67)	0.92	5.07E-08	12.53(6.2-25.3)	0.92	3.58E-08	11.58(
NWD1	8	8.06(4.16-15.61)	0.88	1.23E-05	7.84(4.04-15.19)	0.87	2.18E-05	9.89(4
TOPAZ1	8	5.73(2.96-11.1)	0.83	4.42E-03	5.19(2.68-10.05)	0.81	2.11E-02	5.23(2
ATG2B	7	9.73(4.6-20.58)	0.90	5.24E-05	15.4(7.26-32.68)	0.94	2.08E-08	29.79(
CNTRL	7	12.54(5.92-26.53)	0.92	7.60E-07	11.9(5.61-25.21)	0.92	2.08E-06	13.23(
EYS	7	11.86(5.88-23.94)	0.92	9.92E-08	36.16(17.77-73.6)	0.97	<1.00E-15	22.75(
IGFN1	7	7.71(4.23-14.05)	0.87	5.17E-07	8.64(4.73-15.76)	0.88	4.22E-08	4.62(2
MUC6	7	355.09(206.91-609.41)	0.99	<1.00E-15	7.63(4.62-12.59)	0.87	4.00E-11	4.92(2
ARHGAP23	6	12.95(6.11-27.44)	0.92	4.55E-07	6.26(2.96-13.27)	0.84	3.30E-02	10.16(
DSCAML1	6	11.78(5.24-26.47)	0.92	4.69E-05	11.7(5.2-26.34)	0.91	5.60E-05	8.87(3
FBN1	6	73.01(32.1-166.01)	0.99	<1.00E-15	70.3(30.61-161.44)	0.99	<1.00E-15	51.09(
MGAM	6	9.01(4.01-20.24)	0.89	1.99E-03	31.13(13.74-70.53)	0.97	4.44E-12	8(3.41
NCKAP5	6	30.77(13.64-69.39)	0.97	4.44E-12	41.57(18.28-94.54)	0.98	<1.00E-15	34.05(
NCOR2	6	6.9(3.42-13.92)	0.86	1.41E-03	7.6(3.76-15.36)	0.87	3.18E-04	10.13(
POM121L2	6	3.32(2.19-5.02)	0.70	2.99E-04	3.52(2.33-5.34)	0.72	5.59E-05	3.31(2
POTEE	6	33.53(14.86-75.65)	0.97	<1.00E-15	19.87(8.8-44.84)	0.95	1.23E-08	17.02(
SDK1	6	9.27(4.59-18.71)	0.89	1.06E-05	11.08(5.48-22.41)	0.91	4.40E-07	9.43(4
TAS2R30	6	5946.11(2497.31-14157.73)	0.99	<1.00E-15	19.93(12.64-31.41)	0.95	<1.00E-15	36.37(
TLR5	6	6.46(3.63-11.49)	0.85	4.66E-06	8.89(4.99-15.85)	0.89	2.51E-09	7.6(4.1
ZNF469	6	12.68(5.99-26.87)	0.92	6.60E-07	10.88(5.13-23.08)	0.91	9.77E-06	12.25(
ATP5J2-PTCD1; PTCD1	5	32.92(13.5-80.27)	0.97	3.11E-10	23.28(9.53-56.86)	0.96	9.84E-08	12.15(
AXIN1	5	8.3(3.69-18.67)	0.88	6.12E-03	37.8(16.62-85.95)	0.97	<1.00E-15	20.49(
CEP350	5	32.49(13.33-79.22)	0.97	3.86E-10	48.15(19.53-118.72)	0.98	<1.00E-15	17.02(
CFTR	5	36.41(14.92-88.84)	0.97	5.77E-11	111.43(44.15-281.25)	0.99	<1.00E-15	14.8(5
CROCC2	5	9.61(4.27-21.61)	0.90	8.91E-04	8.78(3.9-19.76)	0.89	3.12E-03	8.02(3
ITGAX	5	65.65(28.87-149.28)	0.98	<1.00E-15	21.5(9.51-48.63)	0.95	3.42E-09	10.5(4
MCF2L	5	19.67(8.09-47.82)	0.95	9.96E-07	28.98(11.84-70.95)	0.97	3.38E-09	42.57(
MPHOSPH9	5	10.75(4.78-24.18)	0.91	1.88E-04	19.48(8.62-44.02)	0.95	1.89E-08	8.71(3
PCDHA12	5	12.88(5.3-31.26)	0.92	3.27E-04	10.06(4.14-24.45)	0.90	6.96E-03	21.28(
PCDHAC1	5	9.41(3.88-22.82)	0.89	1.43E-02	13.29(5.46-32.35)	0.92	2.36E-04	28.37(
PCDHB4	5	12.5(5.15-30.36)	0.92	4.75E-04	14.37(5.9-34.99)	0.93	8.62E-05	28.37(
SEC16B	5	399.18(151.75-1050.06)	0.99	<1.00E-15	811.99(256.2-2573.46)	1.00	<1.00E-15	85.16(
SELP	5	33.25(13.63-81.08)	0.97	2.62E-10	9.38(3.86-22.79)	0.89	1.55E-02	10.63(
STARD9	5	8.17(3.63-18.36)	0.88	7.55E-03	13.37(5.93-30.16)	0.93	8.19E-06	20.49(
WDR49	5	18.03(7.42-43.82)	0.94	3.49E-06	15.26(6.27-37.17)	0.93	3.90E-05	10.63(
ZXDA	5	21.72(10.72-44.02)	0.95	<1.00E-15	23.18(11.4-47.13)	0.96	0.00E+00	27.52(
ALPK2	4	24.25(8.97-65.55)	0.96	6.59E-06	18.24(6.74-49.41)	0.95	2.23E-04	22.7(7
APOBR	4	46.4(17.07-126.1)	0.98	1.07E-09	36.07(13.22-98.43)	0.97	5.10E-08	13.61(
ATG2A	4	32.98(14.57-74.68)	0.97	<1.00E-15	95.55(41.12-221.98)	0.99	<1.00E-15	103.09
CACNA2D4	4	25.01(9.25-67.62)	0.96	4.48E-06	116.58(41.32-328.9)	0.99	<1.00E-15	34.05(

CC2D2A	4	76.75(28.04-210.11)	0.99	<1.00E-15	62.27(22.57-171.81)	0.98	3.11E-11	68.12(
CFH	4	23.07(10.21-52.14)	0.96	8.97E-10	25.84(11.39-58.65)	0.96	1.47E-10	15.25(
COL6A2	4	42.91(15.8-116.51)	0.98	3.28E-09	59.81(21.7-164.85)	0.98	5.33E-11	68.12(
CTDP1	4	85.83(31.28-235.48)	0.99	<1.00E-15	89.14(31.95-248.69)	0.99	<1.00E-15	34.05(
DIDO1	4	56.61(20.78-154.22)	0.98	5.77E-11	39.52(14.46-108.01)	0.97	1.52E-08	34.05(
DMXL1	4	32.05(11.83-86.79)	0.97	1.81E-07	26.89(9.89-73.09)	0.96	2.21E-06	22.7(7
FBXL18	4	14.2(5.83-34.54)	0.93	9.97E-05	13.97(5.73-34.07)	0.93	1.34E-04	9.76(3
GOLGB1	4	41.14(15.16-111.66)	0.98	5.94E-09	46.86(17.1-128.46)	0.98	1.51E-09	34.05(
GSE1	4	26.95(9.96-72.91)	0.96	1.74E-06	30.3(11.13-82.47)	0.97	4.91E-07	68.12(
HABP2	4	27.7(10.24-74.95)	0.96	1.22E-06	75.77(27.31-210.22)	0.99	<1.00E-15	34.05(
JMJD1C	4	43.85(16.14-119.1)	0.98	2.41E-09	28.05(10.31-76.29)	0.96	1.30E-06	22.7(7
LEXM	4	40.51(14.93-109.94)	0.98	7.36E-09	113.66(40.34-320.3)	0.99	0.00E+00	68.12(
MGA	4	20.95(8.6-51.05)	0.95	4.33E-07	49.62(20.08-122.64)	0.98	0.00E+00	42.75(
MYLK	4	18.29(6.78-49.38)	0.95	1.93E-04	17.14(6.33-46.4)	0.94	4.48E-04	17.02(
MYO3B	4	13.33(5.48-32.42)	0.92	2.27E-04	96.74(38.47-243.28)	0.99	0.00E+00	12.65(
NBEAL1	4	12.56(4.66-33.84)	0.92	1.14E-02	13.72(5.08-37.09)	0.93	4.89E-03	22.7(7
NEK1	4	25.99(9.61-70.28)	0.96	2.76E-06	20.56(7.58-55.73)	0.95	5.65E-05	22.7(7
NOD2	4	9.47(4.2-21.35)	0.89	1.17E-03	18.93(8.36-42.86)	0.95	3.50E-08	34.35(
NUP214	4	72.09(29.26-177.6)	0.99	0.00E+00	83.93(33.53-210.09)	0.99	0.00E+00	14.86(
PCDHA5	4	35.72(15.77-80.91)	0.97	0.00E+00	32.59(14.32-74.13)	0.97	0.00E+00	34.35(
PIK3C2B	4	12.39(5.83-26.34)	0.92	1.25E-06	13.94(6.54-29.72)	0.93	1.75E-07	10.25(
PLXND1	4	8.19(3.64-18.46)	0.88	7.77E-03	73.96(32.07-170.61)	0.99	0.00E+00	15.25(
POM121C	4	12.74(4.72-34.33)	0.92	9.86E-03	13.56(5.02-36.64)	0.93	5.54E-03	13.61(
PXDN	4	145.54(62.91-336.71)	0.99	0.00E+00	16.16(7.15-36.57)	0.94	4.72E-07	11.43(
RHD	4	21.49(8.82-52.37)	0.95	2.97E-07	35(14.24-86.04)	0.97	1.87E-10	42.75(
RIN1	4	35(12.91-94.86)	0.97	5.53E-08	32.01(11.75-87.19)	0.97	2.43E-07	13.61(
SBF1	4	295.68(102.64-851.81)	0.99	<1.00E-15	23.42(8.63-63.57)	0.96	1.20E-05	17.02(
SFI1	4	532.24(175.3-1615.93)	0.99	<1.00E-15	57.54(20.9-158.44)	0.98	8.88E-11	17.02(
SRCAP	4	266.11(93-761.47)	0.99	<1.00E-15	85.78(30.79-238.97)	0.99	<1.00E-15	68.12(
TAOK2	4	27.05(10-73.16)	0.96	1.66E-06	11.7(4.33-31.61)	0.91	2.44E-02	17.02(
TGM4	4	48.17(19.65-118.06)	0.98	<1.00E-15	172.98(66.93-447.08)	0.99	<1.00E-15	9.76(3
TMEM131	4	614.13(198.75-1897.64)	0.99	<1.00E-15	47.35(17.27-129.82)	0.98	1.31E-09	34.05(
TPO	4	52.5(23.1-119.32)	0.98	<1.00E-15	48.43(21.18-110.77)	0.98	<1.00E-15	11.43(
WNK1	4	21.1(7.81-57.01)	0.95	3.60E-05	68.88(24.9-190.56)	0.99	8.88E-12	34.05(
ZNF850	4	26.95(9.96-72.91)	0.96	1.74E-06	156.78(54.68-449.58)	0.99	<1.00E-15	17.02(
AADACL4	3	460.59(130.08-1630.81)	0.99	<1.00E-15	682.07(161.73-2876.54)	0.99	<1.00E-15	51.09(
ALDH3B2	3	64.37(20.19-205.22)	0.98	3.84E-08	50.88(15.85-163.35)	0.98	8.03E-07	25.54(
ANK1	3	30.85(9.77-97.44)	0.97	1.02E-04	27.05(8.52-85.84)	0.96	4.36E-04	25.54(
ARID1A	3	19.11(6.07-60.17)	0.95	9.20E-03	19.93(6.3-63.01)	0.95	7.02E-03	25.54(
ATP8B4	3	53.45(16.81-169.88)	0.98	3.09E-07	42.61(13.33-136.25)	0.98	5.00E-06	25.54(
BPIFC	3	49.06(15.45-155.76)	0.98	7.93E-07	55.89(17.37-179.86)	0.98	3.01E-07	51.09(
BSCL2	3	997.96(247.57-4022.8)	0.99	<1.00E-15	16.95(5.37-53.52)	0.94	2.81E-02	25.54(
CCDC40	3	30.39(12.4-74.43)	0.97	1.62E-09	24.76(10.08-60.82)	0.96	5.15E-08	12.74(
CP	3	25.04(7.94-78.95)	0.96	7.78E-04	179.48(52.62-612.12)	0.99	4.44E-12	25.54(
DVL1	3	127.38(39.27-413.21)	0.99	1.33E-11	113.66(34.36-375.97)	0.99	1.78E-10	51.09(
EHHADH	3	171.07(52.11-561.52)	0.99	<1.00E-15	227.34(65.22-792.44)	0.99	<1.00E-15	25.54(
F3	3	22.47(8.29-60.89)	0.96	1.88E-05	34.11(12.48-93.26)	0.97	1.21E-07	14.41(
FHOD1	3	26.25(8.32-82.79)	0.96	4.97E-04	179.48(52.62-612.12)	0.99	4.44E-12	51.09(
FILIP1L	3	36.49(11.54-115.45)	0.97	1.85E-05	31.86(10.02-101.32)	0.97	9.10E-05	51.09(

GANAB	3	32.88(10.4-103.92)	0.97	5.38E-05	29.13(9.17-92.54)	0.97	2.16E-04	51.09(
GIGYF2	3	93.36(33.88-257.29)	0.99	<1.00E-15	71.44(25.73-198.39)	0.99	4.44E-12	68.52(
GRB7	3	26.93(9.93-73.04)	0.96	1.99E-06	24.44(8.97-66.54)	0.96	8.04E-06	22.82(
GRIK4	3	108.85(33.73-351.32)	0.99	8.44E-11	426.29(112.14-1620.48)	0.99	<1.00E-15	51.09(
HLA-DQB1	3	5.72(3.31-9.88)	0.83	8.23E-06	7.11(4.11-12.29)	0.86	4.52E-08	5.14(2
KIAA0319L	3	86.76(27.05-278.33)	0.99	1.23E-09	64.33(19.91-207.86)	0.98	6.86E-08	51.09(
MKL1	3	44.01(13.88-139.52)	0.98	2.57E-06	131.15(39.33-437.4)	0.99	4.00E-11	25.54(
NPIPB8	3	105.03(32.58-338.63)	0.99	1.33E-10	87.43(26.76-285.63)	0.99	2.69E-09	51.09(
OR5P3	3	16.04(5.93-43.39)	0.94	9.29E-04	175.9(60.74-509.43)	0.99	<1.00E-15	22.82(
PIGN	3	146.03(44.79-476.1)	0.99	4.44E-12	103.33(31.39-340.12)	0.99	4.71E-10	51.09(
PRRC2C	3	176.1(53.57-578.82)	0.99	<1.00E-15	262.32(74.09-928.81)	0.99	<1.00E-15	51.09(
PTPRS	3	16.95(5.39-53.32)	0.94	2.61E-02	18.22(5.77-57.57)	0.95	1.52E-02	51.09(
RAB44	3	29.19(9.25-92.16)	0.97	1.76E-04	25.82(8.14-81.88)	0.96	6.74E-04	25.54(
RGPD4	3	15.94(7.04-36.1)	0.94	6.31E-07	16.61(7.32-37.72)	0.94	3.70E-07	51.98(
RNF207	3	74.83(23.4-239.27)	0.99	6.85E-09	58.78(18.24-189.43)	0.98	1.77E-07	25.54(
SEC16A	3	21.84(6.93-68.8)	0.95	2.78E-03	23.5(7.42-74.45)	0.96	1.60E-03	25.54(
SLC2A7	3	213.84(64.4-710.02)	0.99	<1.00E-15	136.4(40.8-456.01)	0.99	2.66E-11	25.54(
SNAI3	3	18.13(5.76-57.06)	0.94	1.46E-02	16.95(5.37-53.52)	0.94	2.81E-02	25.54(
TAX1BP1	3	157.56(48.18-515.3)	0.99	<1.00E-15	33.75(10.6-107.45)	0.97	5.17E-05	25.54(
TMEM102	3	54.24(19.86-148.1)	0.98	1.33E-10	58.62(21.22-161.91)	0.98	7.99E-11	34.25(
TOP3B	3	176.1(53.57-578.82)	0.99	<1.00E-15	106.56(32.32-351.29)	0.99	3.42E-10	25.54(
TPSAB1	3	176.1(53.57-578.82)	0.99	<1.00E-15	113.66(34.36-375.97)	0.99	1.78E-10	25.54(
VN1R2	3	11.37(4.21-30.72)	0.91	3.30E-02	28.04(10.28-76.47)	0.96	1.48E-06	17.11(
XKR9	3	73.01(22.84-233.31)	0.99	9.10E-09	37.05(11.62-118.15)	0.97	2.05E-05	25.54(

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are

Table S20: Hypothesis-free data driven approach. Genes showing enrichment of synonymous variants in Spanish pa almost extreme phenotype (MD-AEP). Listed genes were significant when they were compared against gnomAD Non dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(CI)
PIEZO1	14	7.11(4.56-11.09)	0.86	<1.00E-15	6.97(4.47-10.88)	0.86	<1.00E-15	6.61(4.15
ABCC12	11	53.87(31.5-92.13)	0.98	<1.00E-15	43.43(25.32-74.49)	0.98	<1.00E-15	11.53(6.5
ABCA2	9	4.55(2.57-8.08)	0.78	4.33E-03	4.43(2.5-7.86)	0.77	7.28E-03	4.35(2.41
PCDHA9	8	13.69(7.7-24.34)	0.93	<1.00E-15	4.2(2.36-7.45)	0.76	1.97E-02	6.68(3.67
ARHGEF10L	7	22.09(10.93-44.65)	0.95	<1.00E-15	20.47(10.1-41.47)	0.95	<1.00E-15	7.36(3.53
BRD9	6	22(9.77-49.52)	0.95	1.67E-09	22.06(9.76-49.82)	0.95	1.99E-09	10.2(4.3-
GREB1	6	9.03(4.27-19.12)	0.89	1.77E-04	8.77(4.14-18.59)	0.89	2.94E-04	13.65(6.0
TRRAP	6	11.76(5.23-26.42)	0.91	4.82E-05	12.32(5.47-27.73)	0.92	2.64E-05	34.05(12
BCL9L	5	10.95(4.51-26.57)	0.91	2.43E-03	9.97(4.1-24.23)	0.90	7.71E-03	12.15(4.6
KMT2B	5	34.71(15.36-78.46)	0.97	<1.00E-15	33.86(14.91-76.91)	0.97	<1.00E-15	20.49(8.1
PCDHA10	5	8.59(3.54-20.83)	0.88	3.95E-02	8.82(3.63-21.43)	0.89	3.05E-02	17.02(6.3
TPR	5	44.33(19.58-100.37)	0.98	<1.00E-15	44.14(19.37-100.6)	0.98	<1.00E-15	25.62(9.9
ANK2	4	133.05(47.96-369.12)	0.99	<1.00E-15	108.25(38.5-304.37)	0.99	<1.00E-15	17.02(5.6
ARHGAP12	4	14.34(6.36-32.35)	0.93	2.80E-06	14.71(6.51-33.25)	0.93	2.11E-06	8.06(3.42
CACNA1G	4	26.07(9.64-70.51)	0.96	2.65E-06	25.53(9.4-69.35)	0.96	4.20E-06	13.61(4.6
CELSR2	4	285.12(99.21-819.41)	1.00	<1.00E-15	267.47(89.31-801.05)	1.00	<1.00E-15	34.05(10
FRY	4	11.67(4.33-31.46)	0.91	2.37E-02	11.29(4.18-30.5)	0.91	3.45E-02	13.61(4.6
HTR6	4	6.31(3.11-12.78)	0.84	6.40E-03	5.72(2.82-11.6)	0.83	2.65E-02	7.08(3.38
LTK	4	14.32(5.31-38.61)	0.93	2.91E-03	349.77(113.19-1080.79)	1.00	<1.00E-15	34.05(10
MED16	4	28.54(11.69-69.65)	0.96	3.65E-09	27.96(11.41-68.55)	0.96	6.65E-09	21.37(7.7
MYO9B	4	147.83(53.1-411.57)	0.99	<1.00E-15	168.4(58.45-485.13)	0.99	<1.00E-15	68.12(16
PACS2	4	63.35(23.21-172.86)	0.98	8.88E-12	44.13(16.12-120.83)	0.98	3.43E-09	22.7(7.26
PCDHB7	4	319.34(110.26-924.91)	0.99	<1.00E-15	239.31(80.78-708.95)	0.99	<1.00E-15	22.7(7.26
PDZD2	4	9.05(3.72-21.99)	0.89	2.33E-02	9.05(3.72-22.04)	0.89	2.42E-02	11.02(4.2
PIP5K1C	4	130.87(47.19-362.89)	0.99	<1.00E-15	137.78(48.42-392.07)	0.99	<1.00E-15	17.02(5.6
POLR3A	4	86.76(31.62-238.09)	0.99	<1.00E-15	64.94(23.51-179.37)	0.98	1.78E-11	17.02(5.6
PPP6R1	4	48.37(17.79-131.52)	0.98	5.91E-10	66.85(24.19-184.79)	0.99	8.88E-12	68.12(16
PREPL	4	48.3(21.27-109.68)	0.98	<1.00E-15	53.32(23.27-122.14)	0.98	<1.00E-15	10.55(4.4
REV3L	4	70.64(25.84-193.09)	0.99	<1.00E-15	101.03(36.04-283.24)	0.99	<1.00E-15	17.02(5.6
RGS3	4	30.34(11.21-82.14)	0.97	3.74E-07	24.17(8.9-65.62)	0.96	8.21E-06	13.61(4.6
SHANK3	4	8.79(3.62-21.37)	0.89	3.20E-02	8.52(3.5-20.75)	0.88	4.66E-02	14.24(5.3
TRPM5	4	25.9(9.58-70.05)	0.96	2.88E-06	22.61(8.33-61.34)	0.96	1.84E-05	13.61(4.6
TSC2	4	34.69(12.8-94.03)	0.97	6.23E-08	49.41(18.01-135.58)	0.98	7.33E-10	13.61(4.6
WNK2	4	109.35(39.64-301.69)	0.99	<1.00E-15	68.88(24.9-190.56)	0.99	8.88E-12	13.61(4.6
ZFYVE19	4	12.75(5.24-31.01)	0.92	3.99E-04	12.92(5.3-31.5)	0.92	3.62E-04	14.24(5.3
AKAP9	3	99.78(30.99-321.24)	0.99	2.40E-10	136.4(40.8-456.01)	0.99	2.66E-11	25.54(6.7
CAPN15	3	65.07(20.41-207.49)	0.98	3.40E-08	64.33(19.91-207.86)	0.98	6.86E-08	25.54(6.7
CDC23	3	193.14(58.49-637.76)	0.99	<1.00E-15	682.07(161.73-2876.54)	<1.00E-15	<1.00E-15	25.54(6.7
COL28A1	3	18.52(5.88-58.3)	0.95	1.21E-02	21.3(6.73-67.39)	0.95	3.90E-03	51.09(11
CPSF1	3	125.46(45.18-348.37)	0.99	<1.00E-15	106.35(37.76-299.54)	0.99	<1.00E-15	22.82(7.2
CTU2	3	80.11(32.36-198.31)	0.99	<1.00E-15	80.97(32.29-203.05)	0.99	<1.00E-15	21.52(7.7

DGCR2	3	22.41(7.11-70.62)	0.96	2.19E-03	21.03(6.65-66.55)	0.95	4.36E-03	51.09(11
EHD2	3	213.84(64.4-710.02)	0.99	<1.00E-15	179.48(52.62-612.12)	0.99	4.44E-12	25.54(6.7
EPN1	3	19.88(6.31-62.59)	0.95	6.51E-03	18.03(5.71-56.95)	0.94	1.67E-02	25.54(6.7
FARP1	3	105.03(32.58-338.63)	0.99	1.33E-10	106.56(32.32-351.29)	0.99	3.42E-10	25.54(6.7
FBRSL1	3	16.81(6.21-45.5)	0.94	5.51E-04	18.57(6.84-50.44)	0.95	2.01E-04	22.82(7.2
FGFR4	3	161.82(49.42-529.84)	0.99	<1.00E-15	126.29(37.96-420.24)	0.99	6.22E-11	25.54(6.7
FIGNL2	3	25.34(10.36-62.01)	0.96	2.88E-08	25.31(10.3-62.17)	0.96	3.70E-08	28.7(10-8
GLIS3	3	37.64(11.89-119.12)	0.97	1.34E-05	35.14(11.03-111.95)	0.97	3.47E-05	51.09(11
GRM6	3	27.7(8.78-87.43)	0.96	2.94E-04	21.99(6.95-69.59)	0.95	2.94E-03	51.09(11
JPH3	3	75.03(27.34-205.88)	0.99	<1.00E-15	61.79(22.34-170.89)	0.98	4.00E-11	68.52(16
LRRC3	3	166.31(50.73-545.22)	0.99	<1.00E-15	65.57(20.28-211.98)	0.98	5.61E-08	25.54(6.7
MIDN	3	8.89(3.93-20.12)	0.89	3.06E-03	9.45(4.17-21.4)	0.89	1.45E-03	8.82(3.72
MYH1	3	127.38(39.27-413.21)	0.99	1.33E-11	162.38(47.99-549.47)	0.99	4.44E-12	51.09(11
NADSYN1	3	74.83(23.4-239.27)	0.99	6.85E-09	43.71(13.66-139.82)	0.98	3.86E-06	51.09(11
NTRK3	3	65.07(20.41-207.49)	0.98	3.40E-08	113.66(34.36-375.97)	0.99	1.78E-10	51.09(11
PARP4	3	130.15(40.09-422.52)	0.99	8.88E-12	106.56(32.32-351.29)	0.99	3.42E-10	51.09(11
PLXNA3	3	126.19(50.49-315.4)	0.99	<1.00E-15	112.74(44.45-285.97)	0.99	<1.00E-15	86.15(22
PRUNE2	3	19.24(6.11-60.56)	0.95	8.70E-03	19.14(6.06-60.51)	0.95	9.96E-03	51.09(11
PTPRQ	3	11.2(4.59-27.33)	0.91	2.19E-03	11.2(4.58-27.38)	0.91	2.32E-03	12.28(4.6
ROCK1	3	171.07(52.11-561.52)	0.99	<1.00E-15	227.34(65.22-792.44)	0.99	<1.00E-15	25.54(6.7
SLC2A9	3	12.51(4.63-33.81)	0.92	1.28E-02	13.51(4.98-36.62)	0.93	6.21E-03	17.11(5.6
SNRNP200	3	130.15(40.09-422.52)	0.99	8.88E-12	106.56(32.32-351.29)	0.99	3.42E-10	25.54(6.7
SNTG2	3	115.13(35.61-372.23)	0.99	4.44E-11	126.29(37.96-420.24)	0.99	6.22E-11	25.54(6.7
TCFL5	3	598.77(163.36-2194.68)	0.99	<1.00E-15	341.02(93.04-1249.96)	0.99	<1.00E-15	51.09(11
TET3	3	47.51(14.97-150.75)	0.98	1.13E-06	56.82(17.65-182.94)	0.98	2.53E-07	51.09(11
TEX15	3	47.89(15.09-151.97)	0.98	1.03E-06	37.46(11.74-119.47)	0.97	1.84E-05	25.54(6.7
TMEM270	3	113.09(40.84-313.11)	0.99	<1.00E-15	207.88(70.88-609.67)	0.99	<1.00E-15	34.25(10
UNG	3	54.24(19.86-148.1)	0.98	1.33E-10	47.13(17.14-129.56)	0.98	1.63E-09	22.82(7.2
VCPIP1	3	88.04(27.43-282.52)	0.99	1.04E-09	126.29(37.96-420.24)	0.99	6.22E-11	51.09(11
ZCCHC11	3	997.96(247.57-4022.8)	0.99	<1.00E-15	682.07(161.73-2876.54)	0.99	<1.00E-15	51.09(11
ZRANB1	3	239.5(71.64-800.69)	0.99	<1.00E-15	243.58(69.37-855.27)	0.99	<1.00E-15	51.09(11

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are

Table S21: Hypothesis-free data driven approach. The filter was applied to exclude hearing loss genes previously assoc .Genes showing enrichment of missense variants in Spanish patients with Meniere disease and tinnitus almost extreme significant when they were compared against gnomAD Non-Finnish European and CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	CSVS OR(C
ZNF729	13	7.26(4.83-10.92)	0.86	<1.00E-15	3.3(2.2-4.96)	0.70	1.84E-04	8.61(5
ATM	8	9.66(5.3-17.59)	0.90	2.43E-09	10.51(5.76-19.16)	0.90	3.33E-10	5.07(2
CCDC168	8	12.24(6.07-24.67)	0.92	5.07E-08	12.53(6.2-25.3)	0.92	3.58E-08	11.58(
NWD1	8	8.06(4.16-15.61)	0.88	1.23E-05	7.84(4.04-15.19)	0.87	2.18E-05	9.89(4
TOPAZ1	8	5.73(2.96-11.1)	0.83	4.42E-03	5.19(2.68-10.05)	0.81	2.11E-02	5.23(2
ATG2B	7	9.73(4.6-20.58)	0.90	5.24E-05	15.4(7.26-32.68)	0.94	2.08E-08	29.79(
CNTRL	7	12.54(5.92-26.53)	0.92	7.60E-07	11.9(5.61-25.21)	0.92	2.08E-06	13.23(
EYS	7	11.86(5.88-23.94)	0.92	9.92E-08	36.16(17.77-73.6)	0.97	<1.00E-15	22.75(
IGFN1	7	7.71(4.23-14.05)	0.87	5.17E-07	8.64(4.73-15.76)	0.88	4.22E-08	4.62(2
MUC6	7	355.09(206.91-609.41)	0.99	<1.00E-15	7.63(4.62-12.59)	0.87	4.00E-11	4.92(2
ARHGAP23	6	12.95(6.11-27.44)	0.92	4.55E-07	6.26(2.96-13.27)	0.84	3.30E-02	10.16(
DSCAML1	6	11.78(5.24-26.47)	0.92	4.69E-05	11.7(5.2-26.34)	0.91	5.60E-05	8.87(3
FBN1	6	73.01(32.1-166.01)	0.99	<1.00E-15	70.3(30.61-161.44)	0.99	<1.00E-15	51.09(
MGAM	6	9.01(4.01-20.24)	0.89	1.99E-03	31.13(13.74-70.53)	0.97	4.44E-12	8(3.41
NCKAP5	6	30.77(13.64-69.39)	0.97	4.44E-12	41.57(18.28-94.54)	0.98	<1.00E-15	34.05(
NCOR2	6	6.9(3.42-13.92)	0.86	1.41E-03	7.6(3.76-15.36)	0.87	3.18E-04	10.13(
POM121L2	6	3.32(2.19-5.02)	0.70	2.99E-04	3.52(2.33-5.34)	0.72	5.59E-05	3.31(2
POTEE	6	33.53(14.86-75.65)	0.97	<1.00E-15	19.87(8.8-44.84)	0.95	1.23E-08	17.02(
SDK1	6	9.27(4.59-18.71)	0.89	1.06E-05	11.08(5.48-22.41)	0.91	4.40E-07	9.43(4
TAS2R30	6	5946.11(2497.31-14157.73)	0.99	<1.00E-15	19.93(12.64-31.41)	0.95	<1.00E-15	36.37(
TLR5	6	6.46(3.63-11.49)	0.85	4.66E-06	8.89(4.99-15.85)	0.89	2.51E-09	7.6(4.1
ATP5J2-PTCD1; PTCD1	5	32.92(13.5-80.27)	0.97	3.11E-10	23.28(9.53-56.86)	0.96	9.84E-08	12.15(
AXIN1	5	8.3(3.69-18.67)	0.88	6.12E-03	37.8(16.62-85.95)	0.97	<1.00E-15	20.49(
CEP350	5	32.49(13.33-79.22)	0.97	3.86E-10	48.15(19.53-118.72)	0.98	<1.00E-15	17.02(
CFTR	5	36.41(14.92-88.84)	0.97	5.77E-11	111.43(44.15-281.25)	0.99	<1.00E-15	14.8(5
CROCC2	5	9.61(4.27-21.61)	0.90	8.91E-04	8.78(3.9-19.76)	0.89	3.12E-03	8.02(3
ITGAX	5	65.65(28.87-149.28)	0.98	<1.00E-15	21.5(9.51-48.63)	0.95	3.42E-09	10.5(4
MCF2L	5	19.67(8.09-47.82)	0.95	9.96E-07	28.98(11.84-70.95)	0.97	3.38E-09	42.57(
MPHOSPH9	5	10.75(4.78-24.18)	0.91	1.88E-04	19.48(8.62-44.02)	0.95	1.89E-08	8.71(3
PCDHA12	5	12.88(5.3-31.26)	0.92	3.27E-04	10.06(4.14-24.45)	0.90	6.96E-03	21.28(
PCDHAC1	5	9.41(3.88-22.82)	0.89	1.43E-02	13.29(5.46-32.35)	0.92	2.36E-04	28.37(
PCDHB4	5	12.5(5.15-30.36)	0.92	4.75E-04	14.37(5.9-34.99)	0.93	8.62E-05	28.37(
SEC16B	5	399.18(151.75-1050.06)	0.99	<1.00E-15	811.99(256.2-2573.46)	1.00	<1.00E-15	85.16(
SELP	5	33.25(13.63-81.08)	0.97	2.62E-10	9.38(3.86-22.79)	0.89	1.55E-02	10.63(
WDR49	5	18.03(7.42-43.82)	0.94	3.49E-06	15.26(6.27-37.17)	0.93	3.90E-05	10.63(
ZXDA	5	21.72(10.72-44.02)	0.95	<1.00E-15	23.18(11.4-47.13)	0.96	0.00E+00	27.52(
ALPK2	4	24.25(8.97-65.55)	0.96	6.59E-06	18.24(6.74-49.41)	0.95	2.23E-04	22.7(7
APOBR	4	46.4(17.07-126.1)	0.98	1.07E-09	36.07(13.22-98.43)	0.97	5.10E-08	13.61(
ATG2A	4	32.98(14.57-74.68)	0.97	<1.00E-15	95.55(41.12-221.98)	0.99	<1.00E-15	103.09
CACNA2D4	4	25.01(9.25-67.62)	0.96	4.48E-06	116.58(41.32-328.9)	0.99	<1.00E-15	34.05(
CC2D2A	4	76.75(28.04-210.11)	0.99	<1.00E-15	62.27(22.57-171.81)	0.98	3.11E-11	68.12(

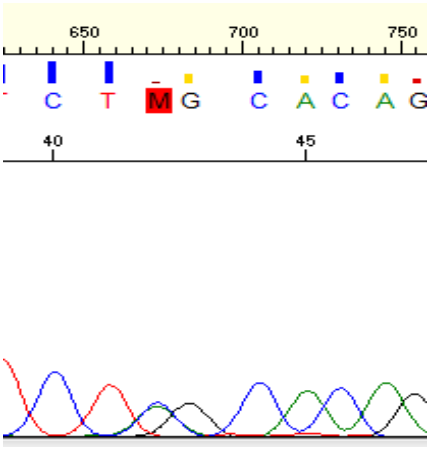
COL6A2	4	42.91(15.8-116.51)	0.98	3.28E-09	59.81(21.7-164.85)	0.98	5.33E-11	68.12(
CTDP1	4	85.83(31.28-235.48)	0.99	<1.00E-15	89.14(31.95-248.69)	0.99	<1.00E-15	34.05(
DIDO1	4	56.61(20.78-154.22)	0.98	5.77E-11	39.52(14.46-108.01)	0.97	1.52E-08	34.05(
DMXL1	4	32.05(11.83-86.79)	0.97	1.81E-07	26.89(9.89-73.09)	0.96	2.21E-06	22.7(7
FBXL18	4	14.2(5.83-34.54)	0.93	9.97E-05	13.97(5.73-34.07)	0.93	1.34E-04	9.76(3
GOLGB1	4	41.14(15.16-111.66)	0.98	5.94E-09	46.86(17.1-128.46)	0.98	1.51E-09	34.05(
GSE1	4	26.95(9.96-72.91)	0.96	1.74E-06	30.3(11.13-82.47)	0.97	4.91E-07	68.12(
HABP2	4	27.7(10.24-74.95)	0.96	1.22E-06	75.77(27.31-210.22)	0.99	<1.00E-15	34.05(
JMJD1C	4	43.85(16.14-119.1)	0.98	2.41E-09	28.05(10.31-76.29)	0.96	1.30E-06	22.7(7
LEXM	4	40.51(14.93-109.94)	0.98	7.36E-09	113.66(40.34-320.3)	0.99	0.00E+00	68.12(
MGA	4	20.95(8.6-51.05)	0.95	4.33E-07	49.62(20.08-122.64)	0.98	0.00E+00	42.75(
MYLK	4	18.29(6.78-49.38)	0.95	1.93E-04	17.14(6.33-46.4)	0.94	4.48E-04	17.02(
MYO3B	4	13.33(5.48-32.42)	0.92	2.27E-04	96.74(38.47-243.28)	0.99	0.00E+00	12.65(
NEK1	4	25.99(9.61-70.28)	0.96	2.76E-06	20.56(7.58-55.73)	0.95	5.65E-05	22.7(7
NOD2	4	9.47(4.2-21.35)	0.89	1.17E-03	18.93(8.36-42.86)	0.95	3.50E-08	34.35(
NUP214	4	72.09(29.26-177.6)	0.99	0.00E+00	83.93(33.53-210.09)	0.99	0.00E+00	14.86(
PCDHA5	4	35.72(15.77-80.91)	0.97	0.00E+00	32.59(14.32-74.13)	0.97	0.00E+00	34.35(
PIK3C2B	4	12.39(5.83-26.34)	0.92	1.25E-06	13.94(6.54-29.72)	0.93	1.75E-07	10.25(
PLXND1	4	8.19(3.64-18.46)	0.88	7.77E-03	73.96(32.07-170.61)	0.99	0.00E+00	15.25(
POM121C	4	12.74(4.72-34.33)	0.92	9.86E-03	13.56(5.02-36.64)	0.93	5.54E-03	13.61(
PXDN	4	145.54(62.91-336.71)	0.99	0.00E+00	16.16(7.15-36.57)	0.94	4.72E-07	11.43(
RHD	4	21.49(8.82-52.37)	0.95	2.97E-07	35(14.24-86.04)	0.97	1.87E-10	42.75(
RIN1	4	35(12.91-94.86)	0.97	5.53E-08	32.01(11.75-87.19)	0.97	2.43E-07	13.61(
SBF1	4	295.68(102.64-851.81)	0.99	<1.00E-15	23.42(8.63-63.57)	0.96	1.20E-05	17.02(
SFI1	4	532.24(175.3-1615.93)	0.99	<1.00E-15	57.54(20.9-158.44)	0.98	8.88E-11	17.02(
SRCAP	4	266.11(93-761.47)	0.99	<1.00E-15	85.78(30.79-238.97)	0.99	<1.00E-15	68.12(
TAOK2	4	27.05(10-73.16)	0.96	1.66E-06	11.7(4.33-31.61)	0.91	2.44E-02	17.02(
TGM4	4	48.17(19.65-118.06)	0.98	<1.00E-15	172.98(66.93-447.08)	0.99	<1.00E-15	9.76(3
TMEM131	4	614.13(198.75-1897.64)	0.99	<1.00E-15	47.35(17.27-129.82)	0.98	1.31E-09	34.05(
TPO	4	52.5(23.1-119.32)	0.98	<1.00E-15	48.43(21.18-110.77)	0.98	<1.00E-15	11.43(
WNK1	4	21.1(7.81-57.01)	0.95	3.60E-05	68.88(24.9-190.56)	0.99	8.88E-12	34.05(
ZNF850	4	26.95(9.96-72.91)	0.96	1.74E-06	156.78(54.68-449.58)	0.99	<1.00E-15	17.02(
AADACL4	3	460.59(130.08-1630.81)	0.99	<1.00E-15	682.07(161.73-2876.54)	0.99	<1.00E-15	51.09(
ALDH3B2	3	64.37(20.19-205.22)	0.98	3.84E-08	50.88(15.85-163.35)	0.98	8.03E-07	25.54(
ANK1	3	30.85(9.77-97.44)	0.97	1.02E-04	27.05(8.52-85.84)	0.96	4.36E-04	25.54(
ARID1A	3	19.11(6.07-60.17)	0.95	9.20E-03	19.93(6.3-63.01)	0.95	7.02E-03	25.54(
ATP8B4	3	53.45(16.81-169.88)	0.98	3.09E-07	42.61(13.33-136.25)	0.98	5.00E-06	25.54(
BPIFC	3	49.06(15.45-155.76)	0.98	7.93E-07	55.89(17.37-179.86)	0.98	3.01E-07	51.09(
BSCL2	3	997.96(247.57-4022.8)	0.99	<1.00E-15	16.95(5.37-53.52)	0.94	2.81E-02	25.54(
CP	3	25.04(7.94-78.95)	0.96	7.78E-04	179.48(52.62-612.12)	0.99	4.44E-12	25.54(
DVL1	3	127.38(39.27-413.21)	0.99	1.33E-11	113.66(34.36-375.97)	0.99	1.78E-10	51.09(
EHHADH	3	171.07(52.11-561.52)	0.99	<1.00E-15	227.34(65.22-792.44)	0.99	<1.00E-15	25.54(
F3	3	22.47(8.29-60.89)	0.96	1.88E-05	34.11(12.48-93.26)	0.97	1.21E-07	14.41(
FHOD1	3	26.25(8.32-82.79)	0.96	4.97E-04	179.48(52.62-612.12)	0.99	4.44E-12	51.09(
FILIP1L	3	36.49(11.54-115.45)	0.97	1.85E-05	31.86(10.02-101.32)	0.97	9.10E-05	51.09(
GANAB	3	32.88(10.4-103.92)	0.97	5.38E-05	29.13(9.17-92.54)	0.97	2.16E-04	51.09(
GIGYF2	3	93.36(33.88-257.29)	0.99	<1.00E-15	71.44(25.73-198.39)	0.99	4.44E-12	68.52(
GRB7	3	26.93(9.93-73.04)	0.96	1.99E-06	24.44(8.97-66.54)	0.96	8.04E-06	22.82(
GRIK4	3	108.85(33.73-351.32)	0.99	8.44E-11	426.29(112.14-1620.48)	0.99	<1.00E-15	51.09(

HLA-DQB1	3	5.72(3.31-9.88)	0.83	8.23E-06	7.11(4.11-12.29)	0.86	4.52E-08	5.14(2
KIAA0319L	3	86.76(27.05-278.33)	0.99	1.23E-09	64.33(19.91-207.86)	0.98	6.86E-08	51.09(
MKL1	3	44.01(13.88-139.52)	0.98	2.57E-06	131.15(39.33-437.4)	0.99	4.00E-11	25.54(
NPIPB8	3	105.03(32.58-338.63)	0.99	1.33E-10	87.43(26.76-285.63)	0.99	2.69E-09	51.09(
OR5P3	3	16.04(5.93-43.39)	0.94	9.29E-04	175.9(60.74-509.43)	0.99	<1.00E-15	22.82(
PIGN	3	146.03(44.79-476.1)	0.99	4.44E-12	103.33(31.39-340.12)	0.99	4.71E-10	51.09(
PRRC2C	3	176.1(53.57-578.82)	0.99	<1.00E-15	262.32(74.09-928.81)	0.99	<1.00E-15	51.09(
PTPRS	3	16.95(5.39-53.32)	0.94	2.61E-02	18.22(5.77-57.57)	0.95	1.52E-02	51.09(
RAB44	3	29.19(9.25-92.16)	0.97	1.76E-04	25.82(8.14-81.88)	0.96	6.74E-04	25.54(
RGPD4	3	15.94(7.04-36.1)	0.94	6.31E-07	16.61(7.32-37.72)	0.94	3.70E-07	51.98(
RNF207	3	74.83(23.4-239.27)	0.99	6.85E-09	58.78(18.24-189.43)	0.98	1.77E-07	25.54(
SEC16A	3	21.84(6.93-68.8)	0.95	2.78E-03	23.5(7.42-74.45)	0.96	1.60E-03	25.54(
SLC2A7	3	213.84(64.4-710.02)	0.99	<1.00E-15	136.4(40.8-456.01)	0.99	2.66E-11	25.54(
SNAI3	3	18.13(5.76-57.06)	0.94	1.46E-02	16.95(5.37-53.52)	0.94	2.81E-02	25.54(
TAX1BP1	3	157.56(48.18-515.3)	0.99	<1.00E-15	33.75(10.6-107.45)	0.97	5.17E-05	25.54(
TMEM102	3	54.24(19.86-148.1)	0.98	1.33E-10	58.62(21.22-161.91)	0.98	7.99E-11	34.25(
TOP3B	3	176.1(53.57-578.82)	0.99	<1.00E-15	106.56(32.32-351.29)	0.99	3.42E-10	25.54(
TPSAB1	3	176.1(53.57-578.82)	0.99	<1.00E-15	113.66(34.36-375.97)	0.99	1.78E-10	25.54(
VN1R2	3	11.37(4.21-30.72)	0.91	3.30E-02	28.04(10.28-76.47)	0.96	1.48E-06	17.11(
XKR9	3	73.01(22.84-233.31)	0.99	9.10E-09	37.05(11.62-118.15)	0.97	2.05E-05	25.54(

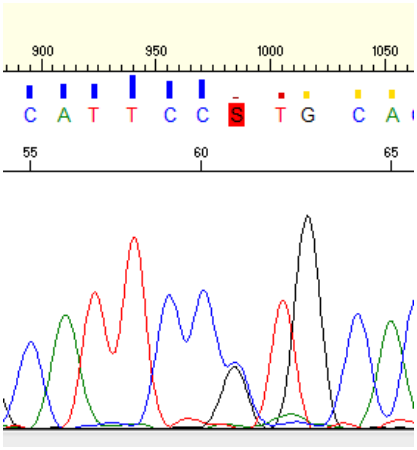
OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are

Figure S1: Sanger sequencing of rare variants in *ANK2* gene in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

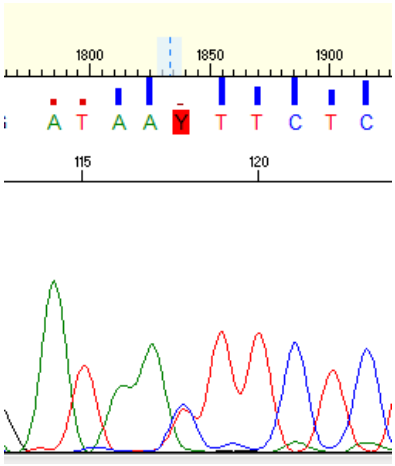
Case#8 (4:114277102 T>G)



Case#8 (4:114294509 G>C)



Case#13(4:114262911 A>G)



Case#6(4:1 114294537 G>A)

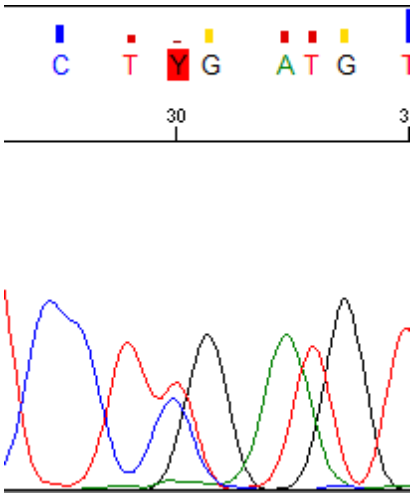
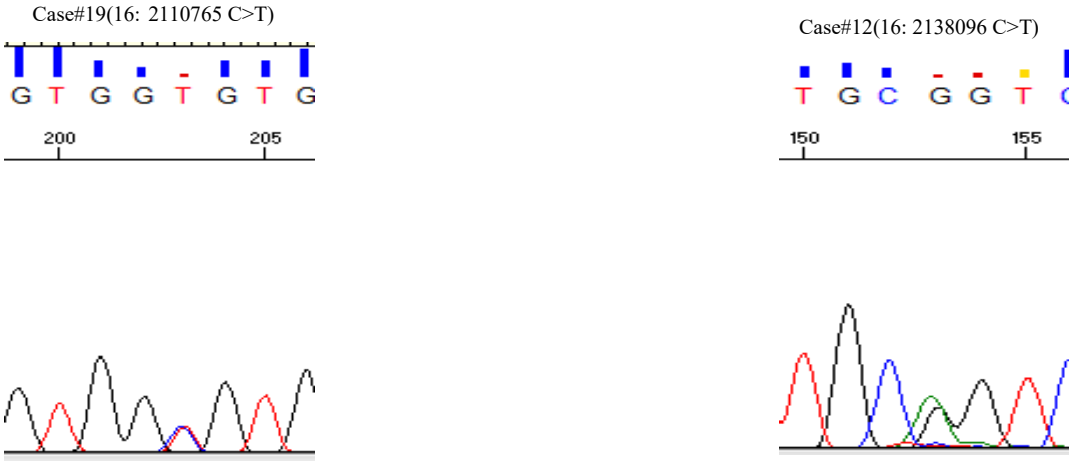


Figure S2: Sanger sequencing of rare variants in *TSC2* gene in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)



Original Research Articles

Open access scientific Publications

1. Amanat, S., Requena, T., & Lopez-Escamez, J. A. (2020). **A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders**. *Genes*, *11*(9), 987. doi.org/10.3390/genes11090987

   IF 3.759 JCR 53/177 (T1/Q2) Category Genetics & Heredity

2. Amanat, S., Gallego-Martinez, A., Sollini, J., Perez-Carpena, P., Espinosa-Sanchez, J. M., Aran, I., ... & Lopez Escamez, J. A. A.(2021). **Burden of Rare Variants in Synaptic Genes in Patients with Severe Tinnitus: An Exome Based Extreme Phenotype Study**. (2021). *EBioMedicine, 66(103309),* doi.org/10.1016/j.ebiom.2021.103309

   IF 5.736 JCR 17/139 (T1/Q1) Category Medicine, Research & Experimental.

A Systematic Review of Extreme Phenotype Strategies to Search for Rare Variants in Genetic Studies of Complex Disorders

Sana Amanat [1], Teresa Requena [2] and Jose Antonio Lopez-Escamez [1,3,4,*]

[1] Otology & Neurotology Group CTS495, Department of Genomic Medicine, GENYO—Centre for Genomics and Oncological Research—Pfizer/University of Granada/Junta de Andalucía, PTS, Granada 18016, Spain; sana.amanat@genyo.es

[2] Centre for Discovery Brain Sciences, Edinburgh Medical School: Biomedical Sciences, University of Edinburgh, Edinburgh EH8 9JZ, UK; mrequena@ed.ac.uk

[3] Department of Otolaryngology, Instituto de Investigación Biosanitaria ibs.GRANADA, Hospital Universitario Virgen de las Nieves, Universidad de Granada, Granada 18016, Spain

[4] Department of Surgery, Division of Otolaryngology, Universidad de Granada, Granada 18016, Spain

* Correspondence: antonio.lopezescamez@genyo.es; Tel.: +34-958-715-500-160

Received: 17 July 2020; Accepted: 19 August 2020; Published: date

Abstract: Exome sequencing has been commonly used to characterize rare diseases by selecting multiplex families or singletons with an extreme phenotype (EP) and searching for rare variants in coding regions. The EP strategy covers both extreme ends of a disease spectrum and it has been also used to investigate the contribution of rare variants to the heritability of complex clinical traits. We conducted a systematic review to find evidence supporting the use of EP strategies in the search for rare variants in genetic studies of complex diseases and highlight the contribution of rare variations to the genetic structure of polygenic conditions. After assessing the quality of the retrieved records, we selected 19 genetic studies considering EPs to demonstrate genetic association. All studies successfully identified several rare or de novo variants, and many novel candidate genes were also identified by selecting an EP. There is enough evidence to support that the EP approach for patients with an early onset of a disease can contribute to the identification of rare variants in candidate genes or pathways involved in complex diseases. EP patients may contribute to a better understanding of the underlying genetic architecture of common heterogeneous disorders such as tinnitus or age-related hearing loss.

Keywords: Genetic epidemiology, genetic association studies, extreme phenotype, exome sequencing, tinnitus

1. Introduction

A clinical phenotype is a set of observable signs, symptoms, and behavioral features associated with a human disorder. The phenotype includes multiple features or traits and it may be categorical (male or female sex) or quantitative (glucose levels or hearing thresholds). These observable variations in the phenotype of a disorder is known in Mendelian genetics as expressivity and it may range from mild to severe [1,2] Phenotypic variation in quantitative traits can be represented by a bell-shaped graph where mild and severe phenotypes are located at the tails of the distribution. However, the majority of the subjects show an intermediate phenotype (Figure 1).

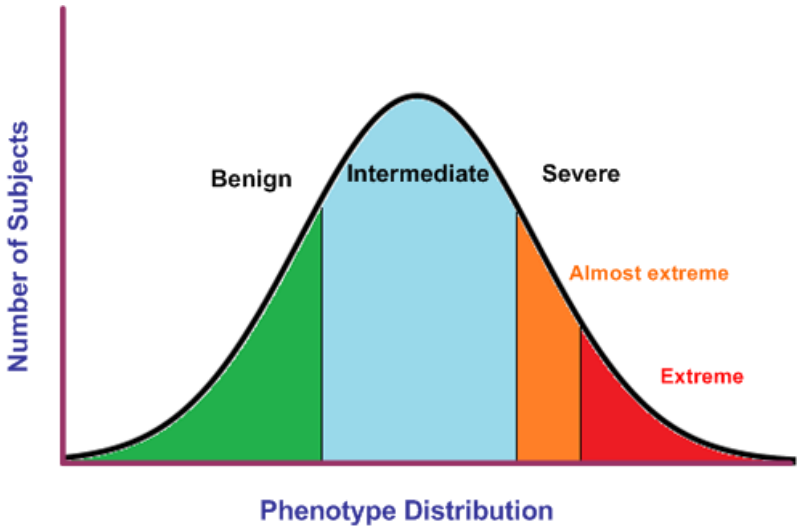


Figure 1. Phenotypic variation in quantitative traits. Individuals' phenotypes can be classified as benign, intermediate, or severe according to general and disease-specific criteria. Extreme phenotypes are identified at the ends of the normal distribution (green, orange, and red areas).

The genetic architecture of human diseases allows a better understanding of the genetic variants that can influence the phenotype in complex diseases [3]. Next-Generation Sequencing (NGS) technology has been used to uncover missing heritability and elucidate the genetic contribution to common and rare diseases with underlying heterogeneity. In particular, Whole-Exome Sequencing (WES) provides an opportunity to capture rare and ultra-rare alleles of protein-coding genes, which highly influence disease risk. In the last few years, several novel genes have been identified by utilizing WES for various neurological diseases, such as epileptic encephalopathies (*KCNQ2*, *STXBP1*, and *KCNB1)* and Parkinson's disease (*VPS13C*, *ARSB*, *PTPRH*, *GPATCH2L*, and *UHRF1BP1L*) [4–6].

A significant increase in the prevalence of complex diseases such as bipolar disorder, coronary artery disease [7], type 2 diabetes, hypertension, obesity, and cancer has been reported the last decades [8]. This increase could be related to environmental factors such as diet or lifestyle changes. However, the genetic contribution to complex conditions is still largely unknown, since the contribution of rare variations to heritability is still undetermined. There are several factors that limit the power of gene discovery approaches, such as phenotypic variance [9], the overlap of clinical features observed for similar conditions, minor allelic frequency (MAF), the heterogeneous nature of loci, and the low effect size of potential risk alleles [10].

There is a well-established inverse relationship between the allelic frequency of a given variant and its effect size on the phenotype (Figure 2). The underlying hypothesis is that extreme phenotypes (EP) will occur in extreme cases with an excess of rare variants, with a moderate effect size on the phenotype in addition to the effect of the common variants for the trait of interest. The EP strategy aims to identify rare genetic variants causing a large effect on disease risk [11,12]. The EP study design includes the selection of individuals whose phenotypes are at the extreme ends of a disease phenotype distribution. These extreme subjects may be characterized by early or late age of onset, benign or severe forms of disease, family history, fast progression of symptoms, very high or very low scores in psychometric tests, or extreme levels of a biomarker [13–15]. This strategy may identify rare genetic variants by sequencing a relatively small sample size and it can target novel candidate genes, since rare variants that contribute to a particular trait are enriched at the two extremes of a disease distribution [10]. The combination of EP with WES has successfully identified several rare variants and candidate genes for diabetic retinopathy [16], bipolar disorder [17], and cystic fibrosis [18] across diverse ethnic groups.

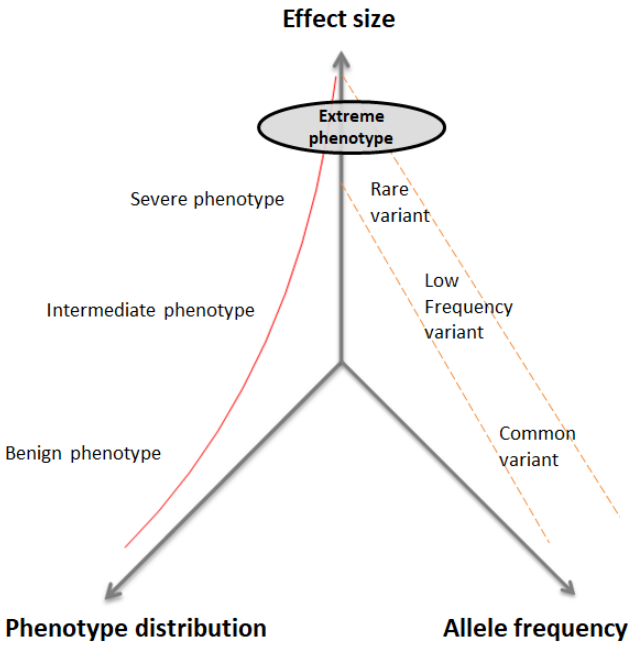


Figure 2. Distribution of genetic variants according to allelic frequency and effect size on the phenotype in quantitative traits. Individuals with extreme phenotypes will show a burden of rare variations with a moderate to large effect size (modified from Manolio et al., 2008 [19]).

The aim of this systematic review is to critically analyze the contribution of strategies based on EPs to uncover rare or novel variants or candidate genes in genetic studies of complex disorders.

2. Materials and Methods

2.1. Study Design

This is a systematic review of genetic studies in complex diseases and it follows Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines (Table S1) [20] and recommendations from the Human Genome Epidemiology Network (HuGENet) review handbook (https://www.cdc.gov/genomics/hugenet/).

2.2. Search Strategies

Literature search for EP strategies was performed on 12 December 2019 using two bibliographic databases (PubMed and Embase). For EP strategies the keywords "phenotypic extreme", "extreme phenotype", "rare variant" and "genetics" were used to formulate the search string. The selected keywords could appear in the title, abstract, text word, author keywords, or MeSH Terms of the articles. The keyword string used for the literature search in PubMed was: ((((“phenotypic extreme”[Title/Abstract] OR “extreme phenotype”)[Title/Abstract] AND (“rare variant”[Title/Abstract] OR “genetics”)[Title/Abstract])) OR ((“phenotypic extreme”[Text Word] OR “extreme phenotype”)[Text Word] AND (“rare variant”[Text Word] OR “genetics”)[Text Word])) OR ((“phenotypic extreme” OR “extreme phenotype”) AND (“rare variant” OR “genetics”) [MeSH Terms]); that for Embase was: (‘phenotypic extreme’: ti, ab, kw OR ‘extreme phenotype’: ti, ab, kw) AND (‘rare variant’: ti, ab, kw OR ‘genetics’: ti, ab, kw) AND [2009–2019]/py AND [english]/lim. Records published in the last 10 years, studies in English language, and only human studies were included in the literature search by configuring filters if available, e.g., on PubMed.

2.3. Research Question and Selection Criteria

The objective of this systematic review is to assess the evidence supporting the design of genetic studies using extreme phenotype strategies to find rare or novel variants or genes involved in complex disorders. According to this hypothesis, we formulated the following research question: "Are EP strategies useful to establish the genetic contribution in complex diseases?". To answer this question, we followed the "Population, Intervention, Comparison, Outcome, Study design" (PICOS) strategy:

1. Population: Patients with a complex disease or condition.
2. Intervention: Selection of individuals according to extreme phenotype criteria (i.e., early onset, fast progression of disease, very high or very low scores in psychometric tests, or extreme levels of a biomarker).
3. Comparison: Genetic association studies (genome-wide association studies (GWAS), WES, genotyping, Sanger sequencing, or targeted sequencing).
4. Outcome: genetic findings reported (rare variants, candidate genes, or pathways associated with the condition of interest).
5. Study design: case–control, case report, case–cohort, or trios.

2.4. Exclusion Criteria

- Studies in non-human models.
- Studies not published in English.
- Studies with a publication date ≥10 years.

2.5. Quality Assessment of Selected Studies

The extracted records were screened to remove duplicate entries. The title and abstract of all articles were reviewed to exclude reviews, meta-analysis, and irrelevant records (non-genetic studies, pharmacogenomics or clinical studies). The search was conducted primarily for rare variants, but any type of variants were retained and included in this systematic review. After screening, the obtained records were considered for full-text assessment in the next step. To assess the quality of these articles, we formulated 8 questions for EP studies (Table 1). For each question, a positive answer was scored as 1 and a negative answer as 0. Each author classified and rated each record independently of each other. Differences in the scores were discussed to get a final consensus score. If a record achieved ≥60% of the total score, the response to Q8 was "yes", and the reported rare variants have a MAF < 0.05, then the record was selected for synthesis. So, only studies with significant results were included. Two of the authors carried out the synthesis (SA, JALE). The outcome for each selected study was assessed according to Q8 and the following criteria: if a given study had found any rare or de novo variant, common variant, copy number variants, candidate genes, or pathways for EP subjects, then the major outcome was considered as positive.

Table 1. Criteria used to assess the quality of the selected genetic studies using an extreme phenotype approach.

No.	Question	Answer
Q1	Is there a thorough description of the study design?	Yes/No
Q2	Has the study described the method of sequencing/genotyping?	Yes/No
Q3	Has the study provided information about population ancestry?	Yes/No
Q4	Is there any information on the sex of the selected individuals?	Yes/No
Q5	Is there any information on the age of disease onset?	Yes/No
Q6	Has the study used extreme phenotype criteria for sample recruitment?	Yes/No
Q7	Has the study performed sex-specific analysis for genetic associations?	Yes/No

Q8	Has the study reported significant genetic findings?	Yes/No

2.6. Data Extraction and Synthesis

The following information was extracted from each article selected for synthesis: first author's last name, publication year, disease/disorder name, population ancestry, study design, sequencing method, EP/disease phenotype criteria, sample size for cases, age of disease onset, sex of individuals, MAF, and main genetic findings. Moreover, the phenotype criteria and the main genetic findings for EP were of great interest for synthesis.

2.7. Risk of Bias

The Cochrane collaboration tool [21] was used to assess the risk of bias for each selected study (Table S2).

3. Results

3.1. Selection and Characteristics of EP Studies

For the EP strategy, we retrieved 106 records in total, 66 records from PubMed and 40 from Embase, by using the search strings reported in the search strategy section. After duplicates' removal, we retained 89/106 records aggregated from the two databases. Next, after screening by title and abstract of the articles, we retrieved 30/89 records that were included for full-text assessment. The discarded records were reviews, meta-analyses, non-genetic studies, pharmacogenomics studies, posters, or abstracts presented at scientific meetings. All studies including variants with MAF > 0.05, single cases, or <5 patients with EP were excluded. We performed quality assessment for 30 articles, and 19/30 records surpassed the minimum quality assessment score and were considered for synthesis. (Figure 2, Table S3).

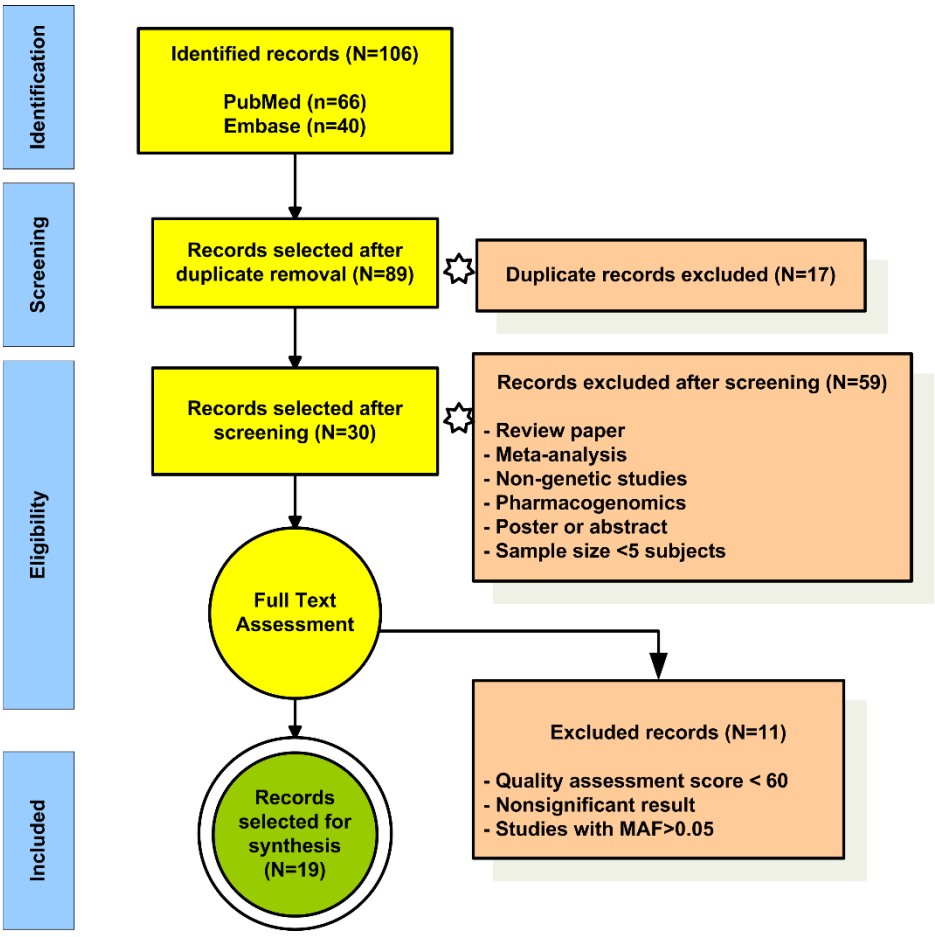


Figure 3. Flowchart to select extreme phenotype records for synthesis.

Among the 19 studies selected for synthesis, 16 records were related to physical conditions, 1 was on bipolar disorder, and 2 were related to neurological disorders including epilepsy and Alzheimer's disease. All of these studies reported rare variants, candidate genes, or potential pathways associated with a particular trait using an EP approach. These 19 EP studies covered 18 complex diseases.

Information about population ancestry and sample size of cases was available for all 19 studies. Only 11/19 studies reported the age of disease onset, and 18/19 records reported the sex of the individuals. The most common criteria to define EP included early onset, late onset, family history, acute form, and/or fast progression of a disease. In addition, disease-specific features were also considered to define an EP, such as the worst score in biomarkers levels including Bone Mass Density (BMD) and spirometry-based severity according to Global Initiative for Chronic Obstructive Lung Disease (GOLD) grade. The reported sample size was between 12 and 32,965 individuals. A summary of the characteristics of these 19 EP studies is shown in Table 2.

3.2. Synthesized Findings of EP Studies

In the 19 EP studies, the combination of general and disease-specific EP criteria was used to select individuals. Information on the study design, sequencing technique, and ancestry population was available for all 19 studies. The reported sample size varied according to the design and sequencing method: 1711 ± 2513 (mean ± SD) for GWAS, 929 ± 2389 for genotyping, 1274 ± 9380 for WES, 29 ± 9 for targeted sequencing, and 949 ± 8742 for Sanger sequencing. All 19 examined studies using EP to select individuals reported significant findings including several rare variants, copy number variants, potential candidate genes or pathways associated with the condition of interest. WES was able to find rare variants in 13/19 studies (MAF = 0.00–0.05) in identified variants. It also helped in the identification of several novel candidate genes including *TACC2* [22], *PRKCD, C1QTNF4, DNMT3A* [23], *LOC728699*, and *FASTK* [16]. GWAS identified a rare variant in 1/19 study (MAF = 0.04). In addition, genotyping and targeted and Sanger sequencing contributed in the identification of many candidate genes and micro-deletions.

Table 2. Summary of the 19 genetic studies using an extreme phenotype approach selected for synthe

Reference	Disease	EP Criteria	Study Design	Sequencing Method	Ancestry	Number of Patients	Onset	Se
Pullabhatla et al. (2017) [23]	Systemic lupus erythematosus	Proband with early onset and clinical features with poor outcome	Family trios, Replication cohort	WES	EU	30 trios, *10995*	<25 y	No rep te
Johar et al.(2016) [24]	Polyautoimmunity	Polyautoimmunity and familial autoimmunity\	Case–control, Cross-sectional	WES	Colombian	47	Not reported	M
Kunkle et al. (2017) [25]	Alzheimer's disease	Early-onset Alzheimer's disease, familial or sporadic	Case–control, Replication cohort	WES	NHW and Caribbean Hispanic	93, *8570*	<65 y	M

Emond et al.(2012) [13]	Cystic fibrosis (CF)	CF with early onset of persistent *pseudomonas aeruginosa* infection	Case–control, Replication cohort	WES	EU America, African American, White Hispanic, NHW, Asian, Aleut	43, *696*	≤2.5 y	M
Shtir et al. (2016) [16]	Diabetes	Diabetes for at least 10 years without diabetic retinopathy	Case–control, Cross-sectional	WES	Saudi	43	Not reported	M
Liu et al. (2016) [26]	Lung cancer	Familial or sporadic lung cancer cases, ever smokers or severe chronic obstructive pulmonary disease (COPD)	Case–control, Cross-sectional	WES	NHW	48 sporadic 54 familial	56 y familial 61 y sporadic	M
Husson et al. (2018) [17]	Bipolar I disorder	Family history of mood disorder and early onset	Case–control, Cross-sectional	WES	EU	92	mean: 24 y	M
Johar et al. (2015) [27]	Multiple autoimmune syndrome	Multiple autoimmune syndrome with Sjögren's syndrome	Case–control, Cross-sectional	WES	Colombian	12	28–67 y	F
Hiekkala et	Hemiplegic	≥2 migraine attacks,	Case report,	WES	Finnish	293	median:	M

al. (2018) [28]	migraine	completely reversible motor weakness	Cross-sectional				12 y	
Qiao et al. (2018) [29]	COPD	COPD cases with GOLD grade 3 or 4	Case–control, Cross-sectional	WES	EU, NHW, African American	≈1769	>45 y, ≤65 y	M
Bruse et al. (2016) [22]	COPD	COPD cases with GOLD grade 3 or 4	Case–control, Cross-sectional	WES	NHW	62	Not reported	M
Nuytemans et al. (2018) [30]	Thrombotic storm (TS)	Severe onset of ≥2 arterial, unusual clot location, refractory, reoccurrence	Case report, Cross-sectional	WES, Targeted sequencing	White and Indian	26 (13 trios)	Not reported	M

Aubart et al. (2018) [31]	Marfan syndrome	Severe aortic features (dissection or preventive thoracic aortic aneurysm rupture surgery at a young age) or sib pairs	Case–control, Cross-sectional	WES	EU	51 EP and 8 sib-pairs	≈10–30 y	M
Gregson et al. (2018) [32]	Bone mass density	Extremely high or moderately high bone mass density	Case–control, Replication cohort	GWAS	EU	1258, *32965*	Not reported	M
Lee et al. (2018) [33]	Ulcerative colitis	Ulcerative colitis patients with good or poor prognosis	Case–control, Replication cohort	Genotyping	Korean	881, *274*	35.6 ± 13.9 y	M
Tomaiuolo	Acute	AMI patients with first	Case–control,	Genotyping	EU	1653,	Not	M

et al. (2012) [34]	myocardial infarction (AMI)	episode before or after 45 years of age	Replication cohort			*909*	reported	
Goldberg-Stern et al. (2013) [15]	Epilepsy with febrile seizures plus	Generalized epilepsy with febrile seizures plus, a proband with Dravet syndrome	Case-control, Cross-sectional	Sanger sequencing	Ashkenazi Jewish	14 familial cases	infancy to 7 y	M
Shen et al. (2017) [35]	Spermatogenic failure	Spermatogenic failure with azoospermia, mild oligozoospermia or severe oligozoospermia	Case–control, Cross sectional	Sanger sequencing	Chinese Han	884	Not reported	M
Uzun et al. (2016) [36]	Preterm birth	Patients delivering <34 weeks	Case report, Cross-sectional	Targeted Sequencing of 329 genes	African-American; Asian; Hispanic; White; Native American	32	Not reported	F

Legend: Non-Hispanic White, NHW, European, EU, Whole-Exome Sequencing, WES, GWAS, genome-wide association studies, EP, ex
Polymorphism AF, allelic frequency.

4. Discussion

4.1. Summary of the Main Findings

Our systematic review shows that individuals with an EP may reveal rare variants that can influence genetic susceptibility in most complex disorders. Complex disorders have a heterogeneous spectrum of symptoms, with variable expressivity observed in each patient. By cluster analysis, it is possible to identify subgroups of patients, and by selecting patients with EP (high expressivity), we would expect to find an enrichment of rare variations associated with the EP [37]. However, we cannot recommend a particular EP strategy to select patients, although the selection of individuals with an early-onset disease and/or a severe phenotype (genetic anticipation) will probably help in the search of rare variations. In contrast, elderly patients can show mutations associated with exposure to environmental factors along life (ultraviolet radiation, chemical agents, pollutants) [38]. In general, the criteria to define EP combine common and disease-specific features such as the chronic state of a disease, very high or low biomarker levels such as BMD, spirometry-based severity level according to GOLD, family history, and early/late age of disease onset.

Of note, a large sample size was not required in WES studies for the discovery cohort, and 10/19 records had a number of cases <100. Therefore, a moderate sample size of individuals with EP was sufficient to identify candidate rare variants or genes. These individuals with EP were carriers of rare variants with a high effect size to target new candidate genes. The EP approach was reproducible across different populations, since the selected studies recruited cases with different ethnic backgrounds including Asian, African, and European ancestry and with monogenic diseases such as cystic fibrosis [13] with an extreme phenotype (persistent tracheobronchial infection with early onset) [39]. Therefore, the information about age of disease onset and sex of the selected individuals is essential to define an EP [40].

4.2. Selection of EP in Quantitative Traits

Individuals with EP are characterized by extreme clinically relevant attributes, toxic effects, or extreme responses to a treatment [1]. From a theoretical perspective, a very EP is more informative than an almost EP, but in practice there are several limitations associated with the very EP, such as vulnerability to phenotype heterogeneity and measurement errors. If a significant proportion from both sides of an extreme is discarded, the almost EP can still be more powerful than random sampling of the same size. The benefits of EP sampling were demonstrated by proposing power calculation methods with the help of the maximum likelihood approach [11,41]. It was also indicated that EP sampling to detect rare variants is more cost-efficient as compared to traditional study designs with a large cohort [42]. Replication in a second independent EP cohort to enhance the power of a study is highly recommended, but it is unlikely to obtain a large sample size of EP subjects from a single region [43]. However, the EP approach is considered more efficient than random sampling for the detection of rare variants associated with a trait [11]

4.3. Familial Disorders and EP Strategy

Some common disorders show rare familial phenotypes with Mendelian inheritance associated with rare variants with large effect size. There are many studies using the EP strategy for familial cases of complex disorders, such as Alzheimer's disease (AD) [25], polyautoimmunity disorder [24], and congenital hypothyroidism [44]. For example, a recent study using linkage analysis demonstrated that by selecting individuals with familial autoimmunity and polyautoimmunity as EP, it was possible to identify the *SRA1* gene (LOD score = 5.48) [24]. Furthermore, a WES study on AD analyzed non-Hispanic White patients and Caribbean Hispanic families to find genes associated with early-onset AD. Heterozygous non-synonymous variants with global MAF < 0.001 were selected for variant prioritization and showed autosomal-dominant segregation in these families. Several genes such as *RUFY*, *TCIRG1*, *PSD2*, and *RIN3* were identified that could be involved in endolysosomal transport in both early- and late-onset AD [25]. In some complex diseases such as Meniere disease (MD), a syndrome characterized by hearing loss, episodic vertigo, and tinnitus, there is also a strong evidence of genetic predisposition in most affected families, showing an autosomal-dominant inheritance with almost 60% penetrance. By using WES in familial MD analysis, a burden of multiplex rare missense variants in the *OTOG* gene was reported in 30% of familial cases [45], which illustrates the success of considering familial cases as EP. Furthermore, a study on genetic epilepsy with hay febrile seizures plus (Dravet syndrome) has reported a *SCN1A* missense variant in a large Jewish family (14/17 cases) with epilepsy syndrome at both extremes (low and high) [15], and a study on thyroid dysgenesis with congenital hypothyroidism found a familial *PAX8* variant associated with EP [44].

4.4. An EP Strategy to Investigate the Genetic Contribution to Tinnitus

Tinnitus is the perception of noise in the absence of an external acoustic stimulation, affecting more than 15% of the population and causing a decrease in health-related quality of life [46]. Several specific instruments have been defined to characterize chronic or severe tinnitus, and these instruments have been proposed to measure tinnitus annoyance to define EP for genetic studies [47]. Epidemiological evidence to support a genetic contribution to tinnitus is still weak because of the heterogeneous nature of this condition. In fact, tinnitus can occur together with multiple comorbidities including hearing loss, migraine, sleep disorders, anxiety, other psychological conditions, and some rare monogenic disorders [48]. The careful selection of phenotypes for genetic studies is crucial. The inclusion criteria should consider young individuals with severe forms of bilateral tinnitus to investigate the genetic contribution of rare variations to tinnitus. These individuals may carry a greater susceptibility and lower environmental load; however, severe forms of tinnitus in young individuals are rare [49] and multicenter studies are needed to reach a minimum sample size [50]

4.5. Limitations

Some weaknesses were found in the design of EP strategies; therefore, further research is required. The replication of the genetic studies across different populations with different ethnic backgrounds has enough potential to validate genetic associations [13,36]; however, the frequency of allelic variants is different across different populations, and specific reference data for allelic frequencies are needed for each population. The rare variants reported in simplex families with EPs should be validated in more patients with a severe phenotype [24]. Most of the studies used WES rather whole-genome sequencing (WGS) and this can cause the loss of useful genetic information and erroneous results in calculating the effect size of rare variants at the individual level across a particular phenotype [17].

5. Conclusions

Genetic studies have confirmed the effectiveness of the EP strategies to establish the genetic contribution of rare variations to complex diseases.

Supplementary Materials: The following is available online at www.mdpi.com/xxx/s1, Table S1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), Table S2: Risk of bias of 19 EP studies; Table S3: Quality assessment of EP studies.

Author Contributions: J.A.L.E. conceived the study design and develop the scientific arguments. J.A.L.E. and S.A. performed literature search, quality assessment of the studies, interpretation of the data, drafting of the manuscript, and revision of the final version. T.R. also helped in the interpretation of the data, developing the scientific arguments, and revision of the final draft. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by H2020 MSCA-ITN-2016–722046 Grant (JALE). The project leading to these results has received funding from “la Caixa” Foundation (ID 100010434), under agreement LCF/PR/DE18/52010002 (JALE). This project is part of the European School of Interdisciplinary Tinnitus (ESIT) research; Sana Amanat is a PhD student in the Biomedicine Program at the University of Granada.

Acknowledgments: We would like to thank Patricia Perez-Carpena, MD, PhD for the critical review and the useful comments to improve the manuscript and Marisa Flook for English language editing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pérez-Gracia, J.L.; Gúrpide, A.; Ruiz-Ilundain, M.G.; Alegría, C.A.; Colomer, R.; García-Foncillas, J.; Bermejo, I.M. Selection of extreme phenotypes: The role of clinical observation in translational research. *Clin. Transl. Oncol.* **2010**, *12*, 174–180.
2. Zimmermann, E.; Gamborg, M.; Sørensen, T.I.A.; Baker, J.L. Sex differences in the association between birth weight and adult type 2 diabetes. *Diabetes* **2015**, *64*, 4220–4225.
3. Gibson, G. Rare and common variants: Twenty arguments. *Nat. Rev. Genet.* **2012**, *13*, 135–145.
4. Lohmann, K.; Klein, C. Next generation sequencing and the future of genetic diagnosis. *Neurotherapeutics* **2014**, *11*, 699–707.
5. Srivastava, S.; Cohen, J.S.; Vernon, H.; Barañano, K.; McClellan, R.; Jamal, L.; Naidu, S.; Fatemi, A. Clinical whole exome sequencing in child neurology practice. *Ann. Neurol.* **2014**, *76*, 473–483.
6. Jansen, I.E.; Ye, H.; Heetveld, S.; Lechler, M.C.; Michels, H.; Seinstra, R.I.; Lubbe, S.J.; Drouet, V.; Lesage, S.; Majounie, E.; et al. Discovery and functional prioritization of Parkinson ’ s disease candidate genes from large-scale whole exome sequencing. *Genome Biol.* **2017**, *18*, 1–26, doi:10.1186/s13059-017-1147-9.

7. Johnson, M.B.; Patel, K.A.; De Franco, E.; Houghton, J.A.; McDonald, T.J.; Ellard, S.; Flanagan, S.E.; Hattersley, A.T. A type 1 diabetes genetic risk score can discriminate monogenic autoimmunity with diabetes from early-onset clustering of polygenic autoimmunity with diabetes. *Diabetologia* **2018**, *61*, 862–869.
8. Verma, S.S.; Ritchie, M.D. Another round of "clue" to uncover the mystery of complex traits. *Genes* **2018**, *9*, 61.
9. Craddock, N.; Kendler, K.; Neale, M.; Nurnberger, J.; Purcell, S.; Rietschel, M.; Perlis, R.; Santangelo, S.L.; Schulze, T.; Smoller, J.W.; et al. Dissecting the phenotype in genome-wide association studies of psychiatric illness. *Br. J. Psychiatry* **2009**, *195*, 97–99.
10. Bamshad, M.J.; Ng, S.B.; Bigham, A.W.; Tabor, H.K.; Emond, M.J.; Nickerson, D.A.; Shendure, J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* **2011**, *12*, 745–755.
11. Bjørnland, T.; Bye, A.; Ryeng, E.; Wisløff, U.; Langaas, M. Improving power of genetic association studies by extreme phenotype sampling: A review and some new results. *arXiv* **2017**, arXiv:1701.01286.
12. Barnett, I.J.; Lee, S.; Lin, X. Detecting rare variant effects using extreme phenotype sampling in sequencing association studies. *Genet. Epidemiol.* **2013**, *37*, 142–151.
13. Emond, M.J.; Louie, T.; Emerson, J.; Zhao, W.; Mathias, R.A.; Knowles, M.R.; Wright, F.A.; Rieder, M.J.; Tabor, H.K.; Nickerson, D.A.; et al. Exome sequencing of extreme phenotypes identifies *DCTN4* as a modifier of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis. *Nat. Genet.* **2012**, *44*, 886–889.
14. Bjørnland, T.; Langaas, M.; Grill, V.; Mostad, I.L. Assessing gene-environment interaction effects of FTO, MC4R and lifestyle factors on obesity using an extreme phenotype sampling design: Results from the HUNT study. *PLoS ONE* **2017**, *12*, e0175071.
15. Goldberg-Stern, H.; Aharoni, S.; Afawi, Z.; Bennett, O.; Appenzeller, S.; Pendziwiat, M.; Kuhlenbäumer, G.; Basel-Vanagaite, L.; Shuper, A.; Korczyn, A.D.; et al. Broad phenotypic heterogeneity due to a novel SCN1A mutation in a family with genetic epilepsy with febrile seizures plus. *J. Child Neurol.* **2014**, *29*, 221–226.
16. Shtir, C.; Aldahmesh, M.A.; Al-Dahmash, S.; Abboud, E.; Alkuraya, H.; Abouammoh, M.A.; Nowailaty, S.R.; Al-Thubaiti, G.; Naim, E.A.; ALYounes, B.; et al. Exome-based case—Control association study using extreme phenotype design reveals novel candidates with protective effect in diabetic retinopathy. *Hum. Genet.* **2016**, *135*, 193–200.
17. Husson, T.; Duboc, J.B.; Quenez, O.; Charbonnier, C.; Rotharmel, M.; Cuenca, M.; Jegouzo, X.; Richard, A.C.; Frebourg, T.; Deleuze, J.F.; et al. Identification of potential genetic risk factors for bipolar disorder by whole-exome sequencing. *Transl. Psychiatry* **2018**, *8*, 1–7.
18. Blue, E.; Louie, T.L.; Chong, J.X.; Hebbring, S.J.; Barnes, K.C.; Rafaels, N.M.; Knowles, M.R.; Gibson, R.L.; Bamshad, M.J.; Emond, M.J. Variation in cilia protein genes and progression of lung disease in cystic fibrosis. *Ann. Am. Thorac. Soc.* **2018**, *15*, 440–448,.
19. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753.
20. Moher, D. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement David. *Syst. Rev.* **2015**, *4*, 1.
21. Higgins, J.P.; Altman, D.G.; Gøtzsche, P.C.; Jüni, P.; Moher, D.; Oxman, A.D.; Savović, J.; Schulz, K.F.; Weeks, L.; Sterne, J.A. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. **2011**, *343*, d5928, doi:10.1136/bmj.d5928.
22. Bruse SMoreau, M.; Bromberg, Y.; Jang, J.H.; Wang, N.; Ha, H.; Picchi, M.; Lin, Y.; Langley, R.J.; Qualls, C.; et al. Whole exome sequencing identifies novel candidate genes that modify chronic obstructive pulmonary disease susceptibility. *Hum. Genom.* **2016**, *10*, 1.
23. Pullabhatla, V.; Roberts, A.L.; Lewis, M.J.; Mauro, D.; Morris, D.L.; Odhams, C.A.; Tombleson, P.; Liljedahl, U.; Vyse, S.; Simpson, M.A.; et al. De novo mutations implicate novel genes in systemic lupus erythematosus. *Hum. Mol. Genet.* **2018**, *27*, 421–429.
24. Johar, A.; Sarmiento-Monroy, J.C.; Rojas-Villarraga, A.; Silva-Lara, M.F.; Patel, H.R.; Mantilla, R.D.; Velez, J.I.; Schulte, K.M.; Mastronardi, C.; Arcos-Burgos, M.; et al. Definition of mutations in polyautoimmunity. *J. Autoimmun.* **2016**, *72*, 65–72.
25. Kunkle, B.W.; Vardarajan, B.N.; Naj, A.C.; Whitehead, P.L.; Rolati, S.; Slifer, S.; Carney, R.M.; Cuccaro, M.L.; Vance, J.M.; Gilbert, J.R.; et al. Early-onset Alzheimer disease and candidate risk genes involved in endolysosomal transport. *JAMA Neurol.* **2017**, *74*, 1113–1122.
26. Liu, Y.; Kheradmand, F.; Davis, C.F.; Scheurer, M.E.; Wheeler, D.; Tsavachidis, S.; Armstrong, G.; Simpson, C.; Mandal, D.; Kupert, E.; et al. Focused analysis of exome sequencing data for rare germline mutations in familial and sporadic lung cancer. *J. Thorac. Oncol.* **2016**, *11*, 52–61.
27. Johar, A.S.; Mastronardi, C.; Rojas-Villarraga, A.; Patel, H.R.; Chuah, A.; Peng, K.; Higgins, A.; Milburn, P.; Palmer, S.; Silva-Lara, M.F.; et al. Novel and rare functional genomic variants in multiple autoimmune syndrome and Sjögren's syndrome. *J. Transl. Med.* **2015**, *13*, 173.

28. Hiekkala, M.E.; Vuola, P.; Artto, V.; Häppölä, P.; Häppölä, E.; Vepsäläinen, S.; Cuenca-Leon, E.; Lal, D.; Gormley, P.; Hämäläinen, E.; et al. The contribution of CACNA1A, ATP1A2 and SCN1A mutations in hemiplegic migraine: A clinical and genetic study in Finnish migraine families. *Cephalalgia* **2018**, *38*, 1849–1863.
29. Qiao, D.; Ameli, A.; Prokopenko, D.; Chen, H.; Kho, A.T.; Parker, M.M.; Morrow, J.; Hobbs, B.D.; Liu, Y.; Beaty, T.H.; et al. Whole exome sequencing analysis in severe chronic obstructive pulmonary disease. *Hum. Mol. Genet.* **2018**, *27*, 3801–3812.
30. Nuytemans, K.; Ortel, T.L.; Gomez, L.; Hofmann, N.; Alves, N.; Dueker, N.; Beecham, A.; Whitehead, P.; Estabrooks, S.H.; Kitchens, C.S.; et al. Variants in chondroitin sulfate metabolism genes in thrombotic storm. *Thromb. Res.* **2018**, *161*, 43–51.
31. Aubart, M.; Gazal, S.; Arnaud, P.; Benarroch, L.; Gross, M.S.; Buratti, J.; Boland, A.; Meyer, V.; Zouali, H.; Hanna, N.; et al. Association of modifiers and other genetic factors explain Marfan syndrome clinical variability. *Eur. J. Hum. Genet.* **2018**, *26*, 1759–1772.
32. Gregson, C.L.; Newell, F.; Leo, P.J.; Clark, G.R.; Paternoster, L.; Marshall, M.; Forgetta, V.; Morris, J.A.; Ge, B.; Bao, X.; et al. Genome-wide association study of extreme high bone mass: Contribution of common genetic variation to extreme BMD phenotypes and potential novel BMD-associated genes. *Bone* **2018**, *114*, 62–71.
33. Lee, H.S.;Yang, S.K.; Hong, M.; Jung, S.; Kim, B.M.; Moon, J.W.; Park, S.H.; Ye, B.D.; Oh, S.H.; Kim, K.M.; et al. An intergenic variant rs9268877 between HLA-DRA and HLA-DRB contributes to the clinical course and long-term outcome of ulcerative colitis. *J. Crohn's Colitis* **2018**, *12*, 1113–1121.
34. Tomaiuolo, R.; Bellia, C.; Caruso, A.; Di Fiore, R.; Quaranta, S.; Noto, D.; Cefalù, A.B.; Di Micco, P.; Zarrilli, F.; Castaldo, G.; et al. Prothrombotic gene variants as risk factors of acute myocardial infarction in young women. *J. Transl. Med.* **2012**, *10*, 1–5.
35. Shen, Y.; Xu, J.; Yang, X.; Liu, Y.; Ma, Y.; Yang, D.; Dong, Q.; Yang, Y. Evidence for the involvement of the proximal copy of the *MAGEA9* gene in Xq28-linked CNV67 specific to spermatogenic failure. *Biol. Reprod.* **2017**, *96*, 610–616.
36. Uzun, A.; Schuster, J.; McGonnigal, B.; Schorl, C.; Dewan, A.; Padbury, J. Targeted sequencing and meta-analysis of preterm birth. *PLoS ONE* **2016**, *11*, e0155021.
37. Allen, A.S.; Bellows, S.T.; Berkovic, S.F.; Bridgers, J.; Burgess, R.; Cavalleri, G.; Chung, S.K.; Cossette, P.; Delanty, N.; Dlugos, D.; et al. Ultra-rare genetic variation in common epilepsies: A case-control sequencing study. *Lancet Neurol.* **2017**, *16*, 135–143.
38. Forsberg, L.A.; Rasi, C.; Razzaghian, H.R.; Pakalapati, G.; Waite, L.; Thilbeault, K.S.; Ronowicz, A.; Wineinger, N.E.; Tiwari, H.K.; Boomsma, D.; et al. Age-related somatic structural changes in the nuclear genome of human blood cells. *Am. J. Hum. Genet.* **2012**, *90*, 217–228.
39. Castellani, C.; Assael, B.M. Cystic fibrosis : A clinical view. *Cell. Mol. Life Sci.* **2016**, *74*, 129–140, doi:10.1007/s00018-016-2393-9.
40. Fuseini, H.; Newcomb, D.C. Mechanisms driving gender differences in asthma. *Curr. Allergy Asthma Rep.* **2017**, *17*, 19.
41. Kang, G.; Lin, D.; Hakonarson, H.; Chen, J. Two-stage extreme phenotype sequencing design for discovering and testing common and rare genetic variants: Efficiency and power. *Hum. Hered.* **2012**, *73*, 139–147.
42. Li, D.; Lewinger, J.P.; Gauderman, W.J.; Murcray, C.E.; Conti, D. Using extreme phenotype sampling to identify the rare causal variants of quantitative traits in association studies. *Genet. Epidemiol.* **2011**, *35*, 790–799.
43. Guey, L.T.; Kravic, J.; Melander, O.; Burtt, N.P.; Laramie, J.M.; Lyssenko, V.; Jonsson, A.; Lindholm, E.; Tuomi, T.; Isomaa, B.; et al. Power in the phenotypic extremes: A simulation study of power in discovery and replication of rare variants. *Genet. Epidemiol.* **2011**, *35*, 236–246.
44. De Sanctis, L.; Corrias, A.; Romagnolo, D.; Di Palma, T.; Biava, A.; Borgarello, G.; Gianino, P.; Silvestro, L.; Zannini, M.; Dianzani, I. Familial PAX8 small deletion (c.989_992delACCC) associated with extreme phenotype variability. *J. Clin. Endocrinol. Metab.* **2004**, *89*, 5669–5674.
45. Roman-naranjo, P.; Gallego-Martinez, A.; Soto-Varela, A.; Aran, I.; del Carmen Moleon, M.; Espinosa-Sanchez, J.M.; Amor-Dorado, J.C.; Batuecas-Caletrio, A.; Perez-Vazquez, P.; Lopez-Escamez, J.A. Rare variants in the *OTOG* gene are a frequent cause of familial Meniere ' s disease. *Ear Hear* **2020**, published ahead of print.
46. Gilles, A.; Camp, G.; Van de Heyning, P.; Fransen, E. A pilot genome-wide association study identifies potential metabolic pathways involved in tinnitus. *Front. Neurosci.* **2017**, *11*, 71.
47. Lopez-Escamez, J.A.; Bibas, T.; Cima, R.F.; Van de Heyning, P.; Knipper, M.; Mazurek, B.; Szczepek, A.J.; Cederroth, C.R. Genetics of tinnitus: An emerging area for molecular diagnosis and drug development. *Front. Neurosci.* **2016**, *10*, 377.
48. Vona, B.; Nanda, I.; Shehata-Dieler, W.; Haaf, T. Genetics of tinnitus: Still in its infancy. *Front. Neurosci.* **2017**, *11*, 236, doi:10.3389/fnins.2017.00236.
49. Maas, I.L.; Brüggemann, P.; Requena, T.; Bulla, J.; Edvall, N.K.; vBHjelmborg, J.; Szczepek, A.J.; Canlon, B.; Mazurek, B.; Lopez-Escamez, J.A.; et al. Genetic susceptibility to bilateral tinnitus in a Swedish twin cohort. *Genet. Med.* **2017**, *19*, 1007–1012.

50. Schlee, W.; Hall, D. A.; Canlon, B.; Cima, R. F. F.; Kleine, E. De; Hauck, F.; Huber, A.; Gallus, S.; Kleinjung, T.; Kypraios, T.; Langguth, B.; Lopez-escamez, J. A.; Lugo, A. Innovations in Doctoral Training and Research on Tinnitus : The European School on Interdisciplinary Tinnitus Research (ESIT). *Perspective*. **2018**, *9*, 1–7.

Supplementary Material

Table S1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Section/topic	#	Checklist item
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligi participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions a findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, c and study design (PICOS).
METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years cor publication status) used as criteria for eligibility, giving rationale.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to id studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it cou
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if ap the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and an obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumpti made.

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether study or outcome level) and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of c each meta-analysis.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, sel studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done pre-specified.
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for ex ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study: (a) simple summary data for eac effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see I
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their re (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias) and at review level (e.g., incomplete re research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence and implications for future
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of fun review.

Table S2: Risk of bias of 19 EP studies

Reference	Study design	Sample size	Bias				
			Selection	Performance	Detection	Attrition	Reporting
Pullabhatla et al. (2017)	Family trios, Replication cohort	30 trios, *10995*	High	Low	High	Low	High
Johar et al.(2016)	Case-control, Cross-Sectional	47	Low	Low	Low	High	Low
Kunkle et al. (2017)	Case-control, Replication cohort	93, *8570*	High	Low	High	Low	High
Emond et al.(2012)	Case-control, Replication cohort	43, *696*	Low	Low	Low	Low	Low
Shtir et al. (2016)	Case-control, Cross-Sectional	43	Low	Low	Low	High	Low
Liu et al. (2016)	Case-control, Cross-Sectional	48 sporadic and 54 familial	Low	Low	Low	Low	Low
Husson et al.(2018)	Case-control, Cross-Sectional	92	Low	Low	Low	High	Low
Johar et al.(2015)	Case-control, Cross-Sectional	12	Low	Low	Low	Low	Low
Hiekkala et al.(2018)	Case report, Cross sectional	293	Low	Low	Low	High	Low
Qiao et al.(2018)	Case-control, Cross-Sectional	≈1769	High	Low	High	High	High
Bruse et al.(2016)	Case-control, Cross-Sectional	62	Low	Low	Low	Low	Low
Nuytemans et al.(2018)	Case report, Cross sectional	26(13 trios)	High	Low	High	High	High
Aubart et al.(2018)	Case-control, Cross sectional	51 EP and 8 sib-pairs	Low	Low	Low	High	Low
Gregson et al. (2018)	Case-control, Replication cohort	1258, *32965*	Low	Low	Low	Low	Low
Lee et al. (2018)	Case-control, Replication cohort	881, *274*	Low	Low	Low	Low	Low
Tomaiuolo et al. (2012)	Case-control, Replication cohort	1653, *909*	Low	Low	Low	High	Low
Goldberg-Stern et al. (2013)	Case-control, Cross sectional	14 familial cases	Low	Low	Low	High	Low
Shen et al. (2017)	Case-control, Cross sectional	884	Low	Low	Low	High	Low
Uzun et al. (2016)	Case report, Cross sectional	32	High	Low	High	High	High

Table S3: Quality assessment of EP studies

Reference	Inclusion & exclusion criteria								Percentage	Qualified
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8		
Kolek et al.(2014)	1	1	0	1	0	1	0	0	50	No
Amin et al .(2012)	1	1	1	1	0	1	0	1	75	No
Coassin et al. (2017)	1	1	1	1	0	1	0	1	75	No
Renaud et al.(2016)	1	1	0	1	1	1	0	0	62.5	No
Lee et al. (2018)	1	1	1	1	1	1	0	1	87.5	Yes
Charles et al. (2018)	1	1	1	1	0	1	0	1	75	No
Bjørnland et al. (2017)	1	1	1	1	0	1	1	1	87.5	No
Aubart et al.(2018)	1	1	1	1	1	1	0	1	87.5	Yes
Goldberg-Stern et al. (2013)	1	1	1	1	1	1	0	1	87.5	Yes
Pullabhatla et al. (2017)	1	1	1	0	1	1	0	1	75	Yes
Johar et al.(2016)	1	1	1	1	0	1	0	1	75	Yes
Kunkle et al. (2017)	1	1	1	1	1	1	0	1	87.5	Yes
Shen et al. (2017)	1	1	1	1	0	1	0	1	75	Yes
Emond et al.(2012)	1	1	1	0	1	1	0	1	75	Yes
Emond et al.(2015)	1	1	1	0	1	1	0	1	75	No
Shtir et al. (2016)	1	1	1	0	0	1	0	1	62.5	Yes
Liu et al. (2016)	1	1	1	1	1	1	0	1	87.5	Yes
Eerde et al. (2012)	1	1	1	1	0	1	0	1	75	No
Gregson et al. (2018)	1	1	1	1	0	1	0	1	75	Yes
Paternoster et al.(2011)	1	1	1	1	0	1	0	1	75	No
Husson et al.(2018)	1	1	1	1	1	1	0	1	87.5	Yes
Limou et al.(2010)	1	1	1	1	0	1	0	1	75	No
Johar et al.(2015)	1	1	1	1	1	1	0	1	87.5	Yes
Peloso et al. (2016)	1	1	1	1	0	1	0	1	75	No
Tomaiuolo et al. (2012)	1	1	1	1	0	1	1	1	87.5	Yes
Uzun et al. (2016)	1	1	1	1	0	1	0	1	75	Yes
Hiekkala et al.(2018)	1	1	1	1	1	1	0	1	87.5	Yes
Nuytemans et al.(2018)	1	1	1	1	0	1	0	1	75	Yes
Qiao et al.(2018)	1	1	1	1	1	1	0	1	87.5	Yes
Bruse et al.(2016)	1	1	1	1	0	1	0	1	75	Yes

Legend: 1= Yes, 0= No

Burden of rare variants in synaptic genes in patients with severe tinnitus: an exome based extreme phenotype study

Short title: Synaptic genes in tinnitus patients

Sana Amanat[1], Alvaro Gallego-Martinez[1], Joseph Sollini[2], Patricia Perez-Carpena[1,3], Juan M. Espinosa-Sanchez[1,3], Ismael Aran[4], Andres Soto-Varela[5], Angel Batuecas-Caletrio[6], Barbara Canlon[7], Patrick May[8], Christopher R. Cederroth[2,7,9], Jose A. Lopez-Escamez[1,3,10]

[1]Otology & Neurotology Group CTS495, Department of Genomic Medicine, GENYO-Centre for Genomics and Oncological Research–Pfizer/University of Granada/ Junta de Andalucía, PTS, Granada, Spain.
[2]Hearing Sciences, Division of Clinical Neuroscience, School of Medicine, University of Nottingham,Nottingham, UK
[3]Department of Otolaryngology, Instituto de Investigación Biosanitaria, ibs.Granada, Hospital Universitario Virgen de las Nieves, Granada, Spain
[4]Department of Otolaryngology, Complexo Hospitalario de Pontevedra, Pontevedra, Spain.
[5]Division of Otoneurology, Department of Otorhinolaryngology, Complexo Hospitalario Universitario, Santiago de Compostela, Spain
[6]Department of Otolaryngology, Hospital Universitario de Salamanca, IBSAL Salamanca Spain, Granada, Spain
[7]Laboratory of Experimental Audiology, Department of Physiology and Pharmacology, Karolinska Institute, Stockholm, Sweden
[8]Bioinformatics Core, Luxembourg Centre for System Biomedicine, University of Luxemburg, Esch-sur-Alzette, Luxembourg
[9]National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust, Ropewalk House, Nottingham, UK
[10]Department of Surgery, Division of Otolaryngology, University of Granada, Granada, Spain

Corresponding author:
Dr. Jose A. Lopez-Escamez
Otology & Neurotology Group CTS495,
GENYO. Centre for Genomics and Oncological Research: Pfizer / University of Granada / Andalusian Regional Government
PTS Granada
Avenida de la Ilustración, 114
18016, Granada, Spain.
Phone. +34 958 715 500-160
E-mail: antonio.lopezescamez@genyo.es

Abstract

Background Tinnitus is a heterogeneous condition associated with audiological and/or mental disorders. Chronic, severe tinnitus is reported in 1% of the population and it shows a relevant heritability, according to twins, adoptees and familial aggregation studies. The genetic contribution to severe tinnitus is unknown since large genomic studies include individuals with self-reported tinnitus and large heterogeneity in the phenotype. The aim of this study was to identify genes for severe tinnitus in patients with extreme phenotype.

Methods For this extreme phenotype study, we used three different cohorts with European ancestry (Spanish with Meniere disease (MD), Swedish tinnitus and European genetic generalised epilepsy). In addition, four independent control datasets were also used for comparisons. Whole-exome sequencing was performed for the MD and epilepsy cohorts and whole-genome sequencing was carried out in Swedish with tinnitus.

Findings We found an enrichment of rare missense variants in 24 synaptic genes in a Spanish cohort, the most significant being *PRUNE2, AKAP9, SORBS1, ITGAX, ANK2, KIF20B* and *TSC2* ($p < 2E^{-04}$), when they were compared with reference datasets. This burden was replicated for *ANK2* gene in a Swedish cohort with 97 tinnitus individuals, and in a subset of 34 Swedish patients with severe tinnitus for *ANK2*, *AKAP9* and *TSC2* genes ($p < 2E^{-02}$). However, these associations were not significant in a third cohort of 701 genetic generalized epilepsy individuals without tinnitus. Gene ontology (GO) and gene-set enrichment analyses revealed several pathways and biological processes involved in severe tinnitus, including membrane trafficking and cytoskeletal protein binding in neurons.

Interpretation A burden of rare variants in *ANK2, AKAP9* and *TSC2* is associated with severe tinnitus. *ANK2,* encodes a cytoskeleton scaffolding protein that coordinates the assembly of several proteins, drives axonal branching and influences connectivity in neurons.

Funding

This study has been funded by H2020 MSCA-ITN-2016–722046, the H2020-SC1-2019-848261, and the GNP-182 GENDER-Net Co-Plus Fund (JALE and CRC). The project leading to these results has received funding from “la Caixa” Foundation (ID 100010434), under agreement LCF/PR/DE18/52010002 (JALE). This project is a part of European School of Interdisciplinary Tinnitus (ESIT) research and Sana Amanat is a student in Biomedicine Program at the University of Granada. CRC received additional funding from Svenska Läkaresällskapet (SLS-779681), Hörselforskningsfonden (503). The data handling for STOP and SweGen cohorts were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The time provided by JS was funded by the University of Nottingham, Nottingham Research Fellowship. PM was supported by the BMBF Treat-ION grant (01GM1907) and the DFG Research Unit FOR2715 (FNR INTER/DFG/17/11583046).

Keywords: Tinnitus, extreme phenotype, axon initial segment, exome sequencing

Research in context

Evidence before this study
Tinnitus is the perception of noise or 'ringing in ear' in the absence of an external acoustic stimulation affecting more than 15% of population. Severe tinnitus disturbing quality of life is experienced by 1% of the population and it has a significant heritability according to twins, adoptees and familial aggregation studies. A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders has found evidence to support a high effectiveness to reveal rare pathogenic variants and target novel candidate genes; particularly in neurological disorders such epilepsy or Parkinson disease. By selecting individuals with tinnitus extreme phenotype, we should expect a burden of rare variation in certain genes associated with severe tinnitus.

Added value of this study
We have found a significant enrichment of missense rare variants in synaptic genes including *ANK2, TSC2* and *AKAP9* in patients with tinnitus extreme phenotype in MD patients. We also replicated these findings in an independent cohort of tinnitus patients from Sweden. Gene ontology (GO) and gene-set enrichment analyses revealed several pathways and biological processes involved in severe tinnitus, the top GO terms being membrane trafficking and cytoskeletal protein binding in neurons.

Implications of all the available evidence
This is the first study reporting the association of rare variation in *ANK2, TSC2* and, *AKAP9* genes with severe tinnitus and supports the involvement of membrane trafficking and cytoskeletal protein binding in the pathophysiology of severe tinnitus.

Introduction

Tinnitus is the perception of noise in the absence of an external acoustic stimulation. The symptom is reported by more than 15% of the world population; however, tinnitus is considered a disorder when it is associated with emotional distress, cognitive dysfunction, and/or autonomic arousal, leading to behavioural changes and functional disability[1,2]. The main risk factor of tinnitus is hearing loss, but it is often associated with other conditions including hyperacusis, anxiety, depression, hypertension, insomnia or migraine[3]. Meniere disease (MD) is a rare inner ear disorder with a significant genetic contribution [4], characterised by episodes of vertigo, tinnitus and sensorineural hearing loss[5]. Although vertigo attacks are considered as the main symptom in the first years of the disease, persistent tinnitus is described as a most troublesome symptom by many MD patients [6,7].

Evidence for a genetic contribution to severe tinnitus is unknown since large genomic studies include individuals with self-reported tinnitus and large heterogeneity in the phenotype [8,9].The extreme phenotype (EP) strategy has been used in exome sequencing studies to investigate the genetic contribution of rare variants in rare and complex disorders[10,11]. Individuals with EP are characterized by extreme clinically relevant attributes, toxic effects, or extreme responses to a treatment. EP covers both extreme ends of a phenotype distribution in quantitative traits and a burden of rare variation is expected in certain genes in individuals with a severe tinnitus [12].

The aim of this study was to identify rare variants in synaptic genes by exome sequencing in patients with severe tinnitus. For this, we performed a gene burden analysis (GBA) in Spanish patients with MD and tinnitus EP. Candidate genes *ANK2, TSC2* and *AKAP9* found in the MD-EP cohort were replicated in a Swedish tinnitus cohort, but not in a third generalised genetic epilepsy cohort, overall identifying the first putative genes involved in severe tinnitus.

Materials

Subjects and definition of phenotype

Individuals were recruited through the Meniere disease Consortium (MeDiC), and the diagnosis of patients was performed according to the diagnostic criteria for MD stated by the Barany society[13]. The Spanish version of the Tinnitus Handicap Inventory (THI) questionnaire [14] was used to assess the tinnitus severity and the functional impact of tinnitus on daily life [15]. A total of 59 Spanish patients with chronic and persistent tinnitus were selected among 1890 individuals from the MeDiC cohort, according to percentile 90 in the THI score (extreme cases). Diagnosis and psychoacoustic characterization of chronic tinnitus in patients with MD was performed as previously reported [16].Tinnitus EP was defined in MD patients (MD-EP) with an early onset and severe persistent tinnitus according to Tinnitus Handicap Inventory (THI) score[16]. Thirty individuals with THI score ≥76 were classified as extreme phenotype (EP), 29 individuals with THI ≥ 56 and <76 were defined as almost extreme phenotype (AEP). An in-house group of patients with MD without persistent tinnitus (N=32) were used as internal controls for this study.The clinical information of patients with MD and

tinnitus phenotypes is detailed in (Supplementary Tables 1 and 2). A second independent tinnitus cohort of 97 individuals from Sweden was selected as a replication cohort: the Tinnitus Swedish Tinnitus Outreach Project (STOP)[17] which originates from the LifeGene study [18]. A subgroup of 34 individuals with severe tinnitus was also selected according to the THI ≥ 56 (Supplementary Table 3). We also retrieved rare variant summary statistics data from a third cohort of patients with epilepsy, the CoGIE cohort, that consisted of 701 individuals (152 Generalised genetic epilepsy cases and 549 controls), previously reported[19]. The CoGIE cohort was select as an external control to confirm that the genes associations reported in tinnitus were not observed in a non-related neurological disorder. All cases and controls were of European ancestry.

Procedures

Whole exome sequencing (WES) was performed on MD-EP, MD-AEP cases and in-house MD controls. DNA was extracted from blood or saliva samples using quality controls as previously described[20]. Exon capture was done with the SureSelectXT Human All Exon V6 (Mb) kit (Agilent), and the sequencing was done using HiSeq 4000 platform (Illumina) or NovaSeq 6000 platform (llumina). Paired-end reads were generated per sample to provide an on-target coverage of 100X minimum, with a total coverage of 10GB/sample in HiSeq4000 and 18GB/sample in Novaseq 6000. Read size was 100bp on HiSeq 4000 sequenced samples and 150 bp on Novase6000 sequenced samples.

Raw reads were stored as FASTQ files for each individual. GATK best practices pipelines were utilized to generate Binary Alignment Map (BAM) and Variant Calling Format (VCF) files from raw unmapped reads[21]. Human reference genome GRCH37/hg19 was used to align the reads with the help of Burrows-Wheeler Aligner (BWA-MEM) algorithm. To filter out low quality single nucleotide variants (SNVs) the recommended hard filter was applied as " quality by depth (QD), < 2.0 || fisher strand (FS) > 60.0 || root mean square quality mapping (MQ) < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -6.0". QD is used to normalize the quality of variants to evade the inflation in the existence of deep coverage. FS determines the probability which is based on Phred-scale of the site in case there is strand biasness. This score describes if the alternate alleles are more or less on the reverse or forward strand as compared to reference allele. MQ parameter explains the mapping quality of a site.

The called variants were further filtered out by an in-house MD control dataset composed of 32 individuals to exclude variants associated with MD. The final list of remaining variants was functionally annotated using KGGSeq suite [22] v1.0 and ANNOVAR tool[23] 2019Nov04.

To search for target genes involved in tinnitus, we have used the list of genes included in SynaptomeDB (N=1886). Genes encoding synaptic components included scaffold proteins, membrane transporters, cytoskeletal/adhesion proteins, neurotransmitters and its receptors (hereafter referred as synaptic genes, SG). [24] Additionally, hearing loss genes (N=152) from Deafness Variation Database (DVD) v.8.1(http://deafnessvariationdatabase.org) were also analysed to separate the potential effect of rare variation in hearing loss genes on tinnitus [25].

In order to search for variants associated with tinnitus, two types of variant analysis were performed: single variant analysis (SVA) and gene burden analysis (GBA) for EP and AEP (Supplementary Fig.1). A flowchart for variant analysis according to the type of variant, location in coding or non-coding regions and effect on the protein is described(Supplementary Fig.2).We have used three independent datasets as reference population: Non-Finnish European (NFE) population dataset from gnomAD.v2, NFE from gnomAD.v3 [26] and a Spanish dataset from Collaborative Spanish Variant Server (CSVS)[27].We also called small insertions and deletions (indels) from MD-EP and MD-AEP patients and filtered out by in-house controls and the filtering criteria was applied according to GATK best-practice guidelines. All variants were assessed and evaluated according to the guidelines provided by American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) [28]. The final filtered list of candidate variants was checked through IGVv.2.8.9,further validated by Sanger sequencing and represented using Illustrator for biological sequences (IBS)[29].

Gene ontology (GO) analyses and gene enrichment analyses were performed using GSEA and MsigDB (https://www.gsea-msigdb.org/gsea/index.jsp) as previously described [30]. Two gene lists generated according to the GBA for rare SNVs and indels including 24 and 31 genes were used to retrieve signalling pathways and biological processes. For Gene expression analysis, In-situ hybridization (ISH) data in the mouse brain were obtained from the Allen Brain Atlas data set (http://www.brain-map.org), methods for data collection have been described previously [31](Supplementary Note 1).

Statistical analysis

NFE population datasets from gnomAD.v2 (Exomes=56,885; Genomes=7,718), and gnomAD.v3 (Genomes=32,399), a Spanish population dataset from CSVS (Exomes=1,942) and a Swedish population dataset from SweGen

(Genomes=1000) were used as control groups[32] to compare the minor allele frequency (MAF) and to calculate the odds ratio (OR) for Spanish MD-EP and Swedish tinnitus cohorts. For SVA the OR with 95% CI was calculated for each variant using the three control datasets and p-values were corrected by the total number of variants being compared. For GBA, total alternate alleles per gene using 2x2 contingency matrixes were calculated for EP, AEP and control datasets. For each gene, the OR was calculated with 95% CI and two-tailed p-value was corrected for multiple testing by the total number of genes being compared following Bonferroni-correction. A corrected p-value <0.05 was considered significant. For each gene, the Etiological Fraction (EF) was also calculated as it was previously described [33].

Role of the funding source
The funders of this study had no role in study design, patient recruitment, data analysis, its interpretation or writing the manuscript. The authors had full access to data used in this study with the responsibility to submit it for publication.

Ethical approval

This study has been approved by the Andalucian Ethical Review Board and written consent was obtained from all subjects to conduct genetic studies (Protocol number 722046).

Results

Synaptic genes in Spanish patients with tinnitus extreme phenotype
First, we performed a SVA in patients with tinnitus EP in MD. The total number of obtained variants with MAF <0.05 were 2287 for EP and 1610 for AEP, respectively. Two missense variants were found significantly associated in patients with MD and EP after p-correction. The first was a heterozygous variant and it was found in 3 unrelated individuals located at exon 21 in *DAAM1* gene (chr14:59826182A>C; p.Asn875His; rs61740455) with MAF_{csvs}=0.002 and CADD=17.85). The associated second variant was located at exon 32 in *MYH10 gene* (chr17:8397065C>A; p.Ala1399Ser; rs149021341; MAF_{csvs} =0.001, CADD= 22), and it was found in 2 individuals and one of the carriers was homozygous. Next, we carried out a GBA in the Spanish MD cohort with EP and AEP. For this, we selected variants with MAF<0.1 to analyse the combined effect of different common and rare variants in the same gene. The retained variants in patients with MD were 4625 for EP and 3592 for AEP, respectively after filtering by MD in-house controls to rule out rare variants associated with hearing or vestibular phenotypes. The GBA of missense variants showed 24 significant genes in tinnitus MD-EP including *PRUNE2, AKAP9, SORBS1, ITGAX, ANK2, KIF20B, LRPPRS, SYNPO, TSC2* (Table 1), and 18 genes for MD-AEP (Supplementary Table 4). Interestingly, none of these genes showed an enrichment of synonymous or 5`UTR variants in EP (Supplementary Tables 5 and 6); additionally, the genes from synonymous analysis for AEP are detailed (Supplementary Table 7) The most significant finding for EP was an enrichment of missense variants in the *ANK2* gene, against NFE population from gnomAD.v2 [OR=18.30(6.78-49.40), EF=0.95, corrected-p=1.80E-05], gnomAD.v3 [OR=19.95(7.36-54.08), EF=0.95, corrected-p=7.55E-06] and Spanish population from CSVS [OR=21.93(7.02-68.48), EF=0.95, corrected-p=2.02E-04]. In *ANK2*, four different missense rare variants were found in 3 different sporadic cases; three of the variants were novel and they have not been reported in gnomAD or CSVS databases. The variant 4:114294537G>A; exon 45 was found only in one case and two of the novel variants 4:114277102T>G; exon 38 and 4:114294509G>C; exon 45 were carried by the same patient. The third novel variant 4:114262911A>G was located at exon 33 (Supplementary Table 8 and Supplementary Fig.3).

In the next step, we selected missense variants with CADD≥20 from SG in MD-EP (561 SNV) and MD-AEP (560 SNV) for the GBA. CADD score describes the deleteriousness of SNVs and can be used to prioritize the disease causal variants explaining the underlying genetic architecture and effect size. We obtained 7 genes with significant burden of rare pathogenic variants (Supplementary Table 9), and 9 genes significant for AEP (Supplementary Table 10), when they were compared with reference datasets.

Finally, we performed a SVA and GBA of indels in SG from Spanish patients with MD-EP and MD-AEP. Indels were further filtered out by in-house controls. A total of 1565 indels (MAF<0.05) for SVA, and 2370 indels (MAF<0.1) for the GBA were retrieved for the MD-EP, and 1404 indels for SVA (MAF<0.05) and 1693 (MAF<0.05) for GBA in the MD-AEP group, respectively. We found an enrichment of indels in 31 genes in the MD-EP (Supplementary Table 11), including *TSC2*, *AKAP9* and several other genes and 48 genes in the MD-AEP (Supplementary Table 12), when data were compared with European reference datasets (gnomAD.v2 and gnomAD.v3). Unfortunately, we cannot compare the allelic frequencies in MD-EP or MD-AEP for indels using the Spanish reference data (CSVS), since the number of indels reported in CSVS dataset is low and it will overestimate the burden. We also compared rare with Loss-of-Function (LoF) variants including nonsense, splice-site and frameshift small insertions and deletions in the SG set for MD-EP and MD-AEP. We found 61 LoF variants in the MD-EP and 25 LoF variants in the MD-AEP. However, the number of nonsense or novel splice-site variants found was small, and no significant burden of LoF variants was found in MD-EP and in MD-AEP.

Replication in Swedish patients with severe tinnitus

To replicate the findings in a Swedish cohort with severe tinnitus, we selected all the significant genes from MD-EP analysis. We used three different population datasets as reference controls (gnomAD.v2, gnomAD.v3 and SweGen). The observed MAF for each gene was calculated and compared with controls, whilst p-values were corrected by the total number of variants per gene. Six genes showed an enrichment of missense variants. Subsequently, we selected a subset of 34 patients with severe tinnitus (THI score ≥56) and found a burden of missense variants in *ANK2, AKAP9* and *TSC*2 genes (Table 2). Missense variants identified in the GBA for *ANK2* gene are detailed in (Supplementary Table 8); most of these variants are clustered around exons 38 to 45 across the gene sequence (Supplementary Fig.2).Supplementary Tables 13 and 14, list missense variants found in the GBA for *AKAP9* and *TSC2* genes in Spanish and Swedish patients with tinnitus. Rare variants found in *ANK2* and *TSC2* genes were also validated by Sanger sequencing (Supplementary Fig. 4 and 5).

In addition, we used an independent cohort of generalised genetic epilepsy to determine if the association of *ANK2, TSC2 and AKAP9* genes with severe tinnitus was a non-specific finding, since some neurological disorders such as epilepsy could also share some common genetic background with tinnitus. For this, we performed a GBA using the same SG list in this epilepsy cohort, but none of the genes showed a significant enrichment of missense variants strongly suggesting the genes captured here are tinnitus-specific.

Lastly, we performed GBA of indels in the Swedish cohort using synaptic genes with MAF<0.1. We found 2 genes in the tinnitus cohort (N=97), and 6 genes in subgroup with severe tinnitus (N=34), showing a significant burden of indels (Supplementary Table 15).

To investigate the association of rare missense variants in hearing loss genes with MD-EP, we performed a GBA of missense variants using hearing loss genes in patients with MD. We obtained 305 variants from EP and 313 from AEP with MAF<0.1, respectively. The 6 genes included *USH1G, ILDR1, OTOA, PCDH15, CACNA1D* and *NARS2* were found significant in EP (Supplementary Table 16), and 4 genes showed significant enrichment in AEP (Supplementary Table 17). To replicate the burden of rare variants found in hearing loss genes in MD-EP patients, we selected a subset of 62 patients with self-reported hearing problems from the Swedish cohort. Then, we performed a GBA in the 6 significant hearing loss genes of MD-EP, however, none of these genes showed an enrichment of missense variants in this cohort.

Gene ontology and Gene-set enrichment analysis in patients with tinnitus

We selected 55 significant genes from MD-EP to perform GO and gene-set enrichment analysis including 24 genes with enrichment of missense variants and 31 genes with enrichment of indels analysis. The most significant pathway and GO biological processes involved were the membrane trafficking and cytoskeletal protein binding (Fig. 1,Supplementary Table 18).

***ANK2* and *TSC2* gene expression profile in the mouse brain**

In-situ hybridization (ISH) data in the mouse (Ank2: n=2, Tsc2: n=2, one coronally and one sagittally sections each) obtained from the Allen Mouse Brain Atlas[31] demonstrated strong *Ank2* and *Tsc2* expression in a number of brain regions (Fig. 2 and 3). Visual inspection revealed strong expression of both genes in the cortex, hippocampus (pyramidal layer of CA1, CA2 and CA3 and the granule cell layer of the dentate gyrus), olfactory bulb (the granule and mitral layers), hypothalamus and cerebellum. In addition, to subregions of other brain regions, notably: Tenia tecta, the epithalamus (especially the medial habenula), piriform area (layer 2) and the magnocellular mucleus (Fig. 3). There was a marked similarity in the brain wide expression of Ank2 and Tsc2, this could potentially suggest a common mechanism or brain regions of interest. To confirm this co-expression, 4,104 genes in the mouse brain were compared (n = 4104). These data were used to build a probability distribution for deriving a given amount of coexpression, based on this coexpression of Ank2 and Tsc2 was found to be highly significant (coexpression = 0.9031, p = 0.0091), Fig. 3, (Supplementary Note 2). Somewhat comparable human data (though in a much lower quantity, i.e. Ank2: n = 3, Tsc2 = 2 brains) were also found via the BioGPS gene portal system (http://biogps.org). These data demonstrated strong expression of both Ank2 and TSC2 in cortex (occipital and parietal lobes, and prefrontal cortex), hypothalamus, cerebellum (peduncles) and the amygdala. Interestingly, the profiles of normalized brain expression of Ank2 and Tsc2 were also significantly correlated suggesting similar expression in the human brain also (Pearsons, r = 0.507, p = 0.0031). Single cell RNA-seq data from the Allen Cell Types Database also revealed similarities between human and mouse expression[34].For both humans and mice Ank2 and Tsc2 expression was significantly differentially distributed in cortical neurons (KS test, humans: Ank2, p $=1.1\text{x}10^{-6}$, Tsc2, p = $1.4\text{x}10^{-12}$, mouse: Ank2, p = $1.8\text{x}10^{-20}$, TSC2, p = $1.2\text{x}10^{-5}$), where stronger expression was commonly observed in pyramidal neurons when compared to inhibitory interneurons (Supplementary Fig. 6).

Discussion

The present study reports for the first time a burden of rare missense and structural variants in several SG in patients with severe tinnitus. These genes are involved in cytoskeleton organization and cytoskeleton protein binding in neurons suggesting novel mechanisms involved in tinnitus severity. In particular, a burden of missense rare and novel variants in *ANK2*, *AKAP9* and *TSC2* genes in Spanish MD patients with severe tinnitus (MD-EP), which was replicated in a Swedish cohort of individuals with severe tinnitus. Using a large genetic generalized epilepsy cohort, we could confirm the specificity of these new genes to tinnitus.

The synapse between sensory inner hair cells, primary auditory neurons and these neurons itself are potential candidates for tinnitus, but its perception and long term maintenance involves complex networks in the central nervous system, both in auditory and in non-auditory structures[35]. GO analyses suggest that membrane trafficking and cytoskeletal protein binding in neurons are involved at the molecular level. Future studies in a larger cohort of tinnitus patients will confirm these predictions.

Tinnitus is associated with hearing loss in 90% of cases, according to standard pure tone audiograms. The most accepted causative model of tinnitus is based on the reduction in the auditory input associated with hearing loss, which leads to increased gain in the auditory pathway; that is, an amplification of spontaneous activity in the auditory neurons will lead to the perception of tinnitus[36]. This change in the intrinsic neuronal excitability after sensory deprivation occurs at the axon initial segments (AISs), the site of initiation of the action potential, which increase in length, and expression of voltage-dependent Na+ channels and Ankyrin-G. a membrane scaffolding protein encoded by the *ANK2* gene in the AISs.

The *ANK2* (ENSG00000145362) gene, which is located at chromosome 4q25-q26, encodes Ankyrin-2, is a large structural protein that carries death and ankyrin repeat containing domains. The Ankyrin gene has 46 exons in total and exon 37/38 is brain specific[37,38] . It belongs to the ankyrin family that links the integral proteins to the fundamental spectrin-actin cytoskeleton and plays an important role in different activities including micrometer scale organization of plasma membranes in a broad spectrum of physiological context. *ANK2* encodes two different polypeptide including Ankyrin-2 (expressed in different tissues) and giant Ankyrin-2, a neuro-specific isoform variant expressed broadly in the central nervous system, with 2133 residues encoded by exon 37 between death and spectrin-binding domains [38]. Giant Ankyrin-2 is a key protein to keep connectivity and neural activity in the central nervous system. It contributes to the development, maintenance and the refinement of neural circuits in different brain areas. The neural signals that arise at AISs site regulate the neural activity. However, the lack of auditory input can cause an increase in the length of AISs ultimately exciting the auditory neurons in avian brainstem [39]. In addition, this is accompanied with an increase in whole-cell Na^+ current, membrane excitability and spontaneous firing. After auditory deprivation, the preservation of auditory function indicates that the change may have occurred at synaptic functionality level rather than at the structural level. However, the homeostatic changes occurring at AISs might play an important role to maintain the integrity of the remaining neurons in auditory circuits [39], something that may also occur in severe tinnitus.

Rare variations in ankB isoform may produce an increase of axonal branching [38]. In humans, rare variants in *ANK2* gene have previously been reported in individuals with autism spectrum disorder [38] and long QT syndrome[40].

Epidemiological and genetic studies consistently support that severe tinnitus has a genetic contribution and common and rare variants with epistatic effects shape the phenotype[3,41]. A recent GWAS using a broad definition for tinnitus found a small number of loci and common variants with small effect sizes[42].
Tinnitus as a neurological disorder may not only result from sensory deprivation as it probably occurs in high-frequency hearing loss or MD, or after synaptic reorganization that lead to changes on the neuronal excitability at different brain areas, but also from enhanced connectivity with non-auditory brain regions as it is often observed in tinnitus patients or individuals with severe tinnitus [43].

Ank2 is expressed in a number of distinct auditory and non-auditory brain regions within the mouse brain. We investigated ank2 and tsc2 co-expression profile by selecting RNAseq data, confirming that both genes have a significant co-expression in the mouse. Human data (on a much grosser) scale appears to confirm this with significant correlation of these genes across the brain regions for which data was available. In addition, both genes demonstrated significantly greater expression in excitatory neurons than inhibitory neurons, potentially suggesting their importance in the function of this class of neurons.

Using https://shield.hms.harvard.edu/index.html, we have checked the expression of ank2, tsc2 and akap9 in Spiral ganglion neurons and they are expressed. It is difficult to assess quantitatively about expression throughout the auditory

system as the Allen Brain ISH data are not annotated to include subregions, and not all auditory sub-nuclei are categorized. However, visual inspection shows noticeably stronger expression of ank2 in the dorsal Inferior Colliculus. Expression of tsc2 does not stand out relative to other non-auditory nuclei.

Interestingly, we have found a burden of rare variation in *AKAP9*, another gene previously associated with long-QT syndrome[44]. *AKAP9* encodes A-kinase anchor protein 9, a member of the A-kinase anchor protein family, whose known function is binding to the protein kinase A (PKA) regulatory subunit with the objective of enclose it to different parts of the cell where phosphorylation is needed[45].

Our study also reveals a significant enrichment of rare variants in *TSC2* gene in patients with MD-EP and severe tinnitus. Tuberous Sclerosis Complex 2 (also known as TSC2 or Tuberin) is a known tumour suppressor protein part of the tuberous sclerosis complex (TSC) along with TSC1. This complex is involved in the negative regulation of mTORC1 activity. Loss of tuberin function causes constitutive activation of the mTORC1 signalling pathway leading to tuberous sclerosis tumours [46]. The regulation of the Mtorc1 pathway via the TSC complex has been found to be a key part in some age-associated diseases, including age-related hearing loss. Finally, we found highly significant co-expression of ank2 and tsc2 across the mouse brain, potentially suggesting they are expressed in similar neuronal subtypes. Mouse brain co-expression of ank2 and tsc2 was particularly strong in limbic brain regions (i.e. the hypothalamus, epithalamus, striatum, pallidum and hippocampus), that form a complex circuit distributed across the brain. In addition, strong expression was found across cortex particularly in cortical layers generally associated with cortical projections (i.e. layers 2/3, 5 and 6).

Limitations

Our study has several limitations. The EP strategy is not a representative of complete phenotypic variance observed in MD. Secondly, most of our MD patients were females and future genetic studies should consider gender differences in tinnitus. The third limitation of the study is that other neural pathways not related to synapses may be implicated in tinnitus. Our extreme phenotype approach will be extended to investigate these pathways in future studies. Finally, indel analysis is highly dependent on the callers and annotation tools and most of the indels in our dataset were not found in either the gnomAD population database or the CSVS database.

Our study reveals a burden of rare variants in SG, including *ANK2*, *AKAP9 and TSC2* in patients with severe tinnitus and predicts the involvement of *ANK2* in the cytoskeleton re-organization in the axon initial segment. We thus propose that axonal branching is a likely mechanism to enhance connectivity in auditory and non-auditory brain regions ultimately leading to severe tinnitus.

Data sharing

Clinical and genomic data from each individual are protected according to EU regulation on data protection. Aggregated, anonymized genomic datasets will be shared upon request according to consortium agreement regulations defined by H2020 MSCA-ITN-2016–722046, the H2020-SC1-2019-848261, and the GNP-182 GENDER-Net Co-Plus Fund.

Acknowledgments

This project is a part of European School of Interdisciplinary Tinnitus (ESIT) research and Sana Amanat is a PhD student in Biomedicine Program at the University of Granada.-We would like to thank all patients for their participation in this study. We thank Pablo Roman-Naranjo and Marisa Flook for their assistance with DNA extraction; Alba Escalera-Balsera for her help during sample preparation for Sanger sequencing; Niklas Edvall and Natalia Trpchevska for assistance with the STOP cohort.

Funding sources

This study has been funded by H2020 MSCA-ITN-2016–722046, the H2020-SC1-2019-848261, and the GNP-182 GENDER-Net Co-Plus Fund (JALE and CRC). The project leading to these results has received funding from “la Caixa” Foundation (ID 100010434), under agreement LCF/PR/DE18/52010002 (JALE). This project is a part of European School of Interdisciplinary Tinnitus (ESIT) research and Sana Amanat is a student in Biomedicine Program at the University of Granada. CRC received additional funding from Svenska Läkaresällskapet (SLS-779681), Hörselforskningsfonden (503).The data handling for STOP and SweGen cohorts were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2018-05973. The time provided by JS was funded by the University of Nottingham, Nottingham Research Fellowship. PM was supported by the BMBF Treat-ION grant (01GM1907) and the DFG Research Unit FOR2715 (FNR INTER/DFG/17/11583046).

Declaration of interests

C.R.C. is supported by the UK National Institute for Health Research (NIHR) Biomedical Research Centre but the views expressed herein are his own and do not represent those of NIHR nor the UK Department of Health and Social Care. The other authors declare no competing interest.

Author contributions

JALE conceived the study design and developed the scientific arguments. SA and AGM performed the bioinformatics analyses for MD cohort, AGM performed bioinformatics analyses on Swedish tinnitus cohort. SA performed statistical analysis for MD and Swedish tinnitus cohorts. JS contributed to the analysis of mouse gene expression datasets. JALE, PPC, JMES, IA, ASV, ABC recruited extreme phenotype patients from the MD cohort. BC and

CRC collected data for STOP cohort. PM performed bioinformatic analyses in the epilepsy cohort. JALE, SA, AGM draft the manuscript and all authors have read and approved the final draft.

References

[1] S. Vanneste, K. Joos, D. De Ridder Prefrontal cortex based sex differences in tinnitus perception: same tinnitus intensity, same tinnitus distress, different mood. PLoS One 2012;7. doi:10.1371/journal.pone.0031182.
[2] A. Gilles, G. Camp, P. Van de Heyning, E. Fransen A pilot genome-wide association study identifies potential metabolic pathways involved in tinnitus. Front Neurosci 2017;11:1–10.
[3] J.A. Lopez-Escamez, S. Amanat Heritability and genetics contribution to tinnitus. Otolaryngol Clin 2020;53:501–513. doi:10.1016/j.otc.2020.03.003.
[4] A. Gallego-Martinez, T. Requena, P. Roman-Naranjo, J.A. Lopez-Escamez Excess of rare missence variants in hearing loss genes in sporadic Meniere disease. Front Genet 2019;10. doi:10.3389/fgene.2019.00076.
[5] P. Perez-Carpena, J.A. Lopez-escamez Current understanding and clinical management of Meniere 's disease. Semin Neurol 2020;40:138–150.doi:10.1055/s-0039-3402065.
[6] M.A. Ueberfuhr, L. Wiegrebe, E. Krause, R. Gürkov, M. Drexl Tinnitus in normal-hearing participants after exposure to intense low-frequency sound and in Ménière's disease patients. Front Neurol 2017;7:1–11.
[7] C. Herraiz, M.C. Tapia, G. Plaza Tinnitus and Ménière's disease: characteristics and prognosis in a tinnitus clinic sample. Eur Arch Oto Rhino Laryngol 2006;263:504–509.
[8] B. Vona, I. Nanda, W. Shehata-Dieler, T. Haaf Genetics of tinnitus: still in its infancy. Front. Neurosci. 2017. doi:10.3389/fnins.2017.00236.
[9] S. Amanat, A. Gallego-Martinez, J.A. Lopez-Escamez Genetic inheritance and its contribution to tinnitus. Curr Top Behav Neurosci 2020. doi:10.1007/ 7854_2020_155.
[10] J.L. Perez-Gracia, M.G. Ruiz-Ilundain, I. Garcia-Ribas, E.M. Carrasco The role of extreme phenotype selection studies in the identification of clinically relevant genotypes in cancer research. Cancer 2002;95:1605–1610.
[11] G.M. Peloso, D.J. Rader, S. Gabriel, S. Kathiresan, M.J. Daly, B.M. Neale Phenotypic extremes in rare variant study designs. Eur J Hum Genet 2016;24:924–930.
[12] S. Amanat, T. Requena, J.A. Lopez-Escamez A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders. Genes 2020;11:987.
[13] J.A. Lopez-Escamez, J. Carey, W.H. Chung, J.A. Goebel, M. Magnusson, M. Mandalà Diagnostic criteria for Menière's disease. J Vestib Res Equilib Orientat 2015;25:1–7.
[14] C. Herráiz, G. Plaza, M.C. Tapia Evaluación de la incapacidad en pacientes con acúfenos. Acta Otorrinolaringol Esp 2001;52:534–538.
[15] C.W. Newman, S.A. Sandridge, G.P. Jacobson Psychometric adequacy of theTinnitus Handicap Inventory (THI) for evaluating treatment outcome. J Am Acad Audiol 1998;9:153–160.
[16] P. Perez-Carpena, M. Martinez-Martinez, R.A. Martínez Carranza, A. Batuecas-Caletrio, J.A. Lopez-Escamez A tinnitus symphony in 100 patients with Meniere's disease. Clin Otolaryngol 2019;44:1176–1180.
[17] C.R. Cederroth, A. Lugo, N.K. Edvall, A. Lazar, J. Bulla, I. Uhlen Association between hyperacusis and tinnitus. J Clin Med 2020 Jul;289(8):2412.doi:10.3390/jcm90824122020
18] C. Almqvist, H.A. Paul, W.F. Leif, G. Erik LifeGene — a large prospective population-based study of global relevance. Eur J Epidemiol 2011 Jan;26(1):67–77. doi:10.1007/s10654-010-9521-x.
[19] P. May, S. Girard, M. Harrer, D.R. Bobbili, J. Schubert, S. Wolking Rare coding variants in genes encoding GABA a receptors in genetic generalized epilepsies: an exome-based case-control study. Lancet Neurol 2018;17:699–708.
[20] A.J. Szczepek, L. Frejo, B. Vona, N. Trpchevska, C.R. Cederroth, H. Caria Recommendations on collecting and storing samples for genetic studies in hearing and tinnitus research. Ear Hear. Mar/Apr 2019;40(2):219–226.doi:10.1097/AUD.0000000000006142019.
[21] Depristo M.A., Banks E., Poplin R., Garimella K. V., Maguire J.R., Hartl C. Technical
reports a framework for variation discovery and genotyping using next-generation DNA sequencing data. 2011; 43. doi:10.1038/ng.806.
[22] M.X. Li, H.S. Gui, J.S.H. Kwan, S.Y. Bao, P.C. Sham A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. Nucleic Acids Res 2012;40. doi:10.1093/nar/gkr1257.
[23] K. Wang, M. Li, H. Hakonarson ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010 Sep;38(16):e164. doi:10.1093/nar/gkq603.
[24] M. Pirooznia, T. Wang, D. Avramopoulos, D. Valle, G. Thomas, R.L.D.B. HuganirSynaptome An ontology-based knowledgebase for synaptic genes. Bioinformatics 2012;28(6):897–899. doi:10.1093/bioinformatics/bts040.
[25] H. Azaiez, K.T. Booth, S.S. Ephraim, B. Crone, E.A. Black-ziegelbein, R.J. Marini Genomic landscape and mutational signatures of deafness-associated genes. Am J Hum Genet 2018;103:484–497.
[26] K.J. Karczewski, L.C. Francioli, G. Tiao, B.B. Cummings, J. Alföldi, Q. Wang The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434–443. doi:10.1038/s41586-020-2308-7.
[27] M. Peña-Chilet, G. Roldán, J. Perez-Florido, F.M. Ortuño, R. Carmona, V. Aquino CSVS, a crowdsourcing database of the Spanish population genetic variability. Nucleic Acids Res 2020. doi:10.1093/nar/gkaa794.

[28] Richards S., Aziz N., Bale S., Bick D., Das S. ACMG Standards and guidelinesstandards and guidelines for the interpretation of sequence variants : a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. 2015: 1-20.
[29] W Liu, Y Xie, J Ma, X Luo, P Nie, Z Zuo, et al. an illustrator for the presentation and visualization of biological sequences. Bioinformatics 2015 Oct;1531(20):3359–3361. doi:10.1093/bioinformatics/btv362.
[30] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–15550.
[31] E.S. Lein, M.J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard Genome-wide atlas of gene expression in the adult mouse brain. Nature 2007 Jan;445(7124):168–176. doi:10.1038/nature054532007; 445.
[32] A. Ameur, J. Dahlberg, P. Olason, et al. SweGen: a whole-genome data resource of genetic variability in a cross-section of the Swedish population. Eur J Hum Genet 2017;25:1253–1260. https://doi.org/10.1038/ejhg.2017.130.
[33] R. Walsh, F. Mazzarotto, N. Whiffin, et al. Quantitative approaches to variant classification increase the yield and precision of genetic testing in Mendelian diseases: the case of hypertrophic cardiomyopathy. Genome Med 2019;11:5. https://doi.org/10.1186/s13073-019-0616-z.
[34] H. Staining Allen cell types databases. Allen Inst. Brain Sci. 2015;88:1–9.
[35] M.J. Guitton Tinnitus: pathology of synaptic plasticity at the cellular and system levels. Front System Neurosci 2012;6:1–7. https://doi.org/10.3389/fnsys.2012.00012.
[36] Schaette R, Kempter R. Development of tinnitus-related neuronal hyperactivity through homeostatic plasticity after hearing loss: a computational model. Eur J Neurosci. 2006 Jun;23(11):3124-38. doi: 10.1111/j.1460-9568.2006.04774.x.PMID: 16820003.
[37] P J Mohler, A O Gramolini, V Bennett The ankyrin-B C-terminal domain determines activity of ankyrin-B/G chimeras in rescue of abnormal inositol 1,4,5-trisphosphate and ryanodine receptor distribution in ankyrin-B (-/-) neonatal cardiomyocytes. J Biol Chem 2002 Mar;22277(12):10599–10607.doi:10.1074/jbc.M110958200.
[38] YangR Walder-ChristensenKK KimN WuD LorenzoDN BadeaA et al. *ANK2* autism mutation targeting giant ankyrin-B promotes axon branching and ectopic connectivityProc Natl Acad Sci U S A2019 Jul 116301526215271 10.1073/pnas.1904348116
[39] H. Kuba, Y. Oichi, H. Ohmori Presynaptic activity regulates Na+ channel distribution at the axon initial segment. Nature 2010;465:1075–1078.
[40] SwayneLA MurphyNP AsuriS ChenL XuX McIntoshS et al. Novel Variant in the *ANK2* Membrane-Binding Domain Is Associated With Ankyrin-B Syndrome and Structural Heart Disease in a First Nations Population With a High Rate of Long QT SyndromeCirc Cardiovasc Genet2017 Jan 101e00153710.1161/CIRCGENETICS.116.001537
[41] C.R. Cederroth, M. Pirouzifard, N. Trpchevska, E. Idrizbegovic, B. Canlon, J. Sundquist Association of genetic vs environmental factors in Swedish adoptees with clinically significant tinnitus. JAMA Otolaryngol Head Neck Surg 2019;145:222–229.
[42] R E Clifford, A X Maihofer, M B Stein, A F Ryan, C M Nievergelt Novel Risk Loci in Tinnitus and Causal Inference With Neuropsychiatric Disorders Among Adults of European Ancestry. JAMA Otolaryngol Head Neck Surg 2020 Sep;24146(11):1015–1025. doi:10.1001/jamaoto.2020.2920.
[43] W Schlee, N Weisz, O Bertrand, T Hartmann, T Elbert Using auditory steady state responses to outline the functional connectivity in the tinnitus brain. PLoSOne 2008;3(11):e3720. doi:10.1371/journal.pone.0003720.
[44] P.L. Hedley, P. Jørgensen, S. Schlamowitz, R. Wangari, J. Moolman-Smook, P.A. Brink The genetic basis of long QT and short QT syndromes: a mutation update. Hum Mutat 2009;30:1486–1511.
[45] L. Chen, R.S. Kass A-kinase anchoring protein 9 and IKs channel regulation. J Cardiovasc Pharmacol 2011;58:459–461.
[46] U. Ozcan, L. Ozcan, E. Yilmaz, K. Düvel, M. Sahin, B.D. Manning Loss of the tuberous sclerosis complex tumor suppressors triggers the unfolded protein response to regulate insulin signaling and apoptosis. Mol Cell 2008;29:541–551.
[47] N Trpchevska Journal of Clinical Medicine 2020. doi:10.3390/jcm9123812.
[48] Maas Genetics in Medicine 2017. doi:10.1038/gim.2017.4

Figures

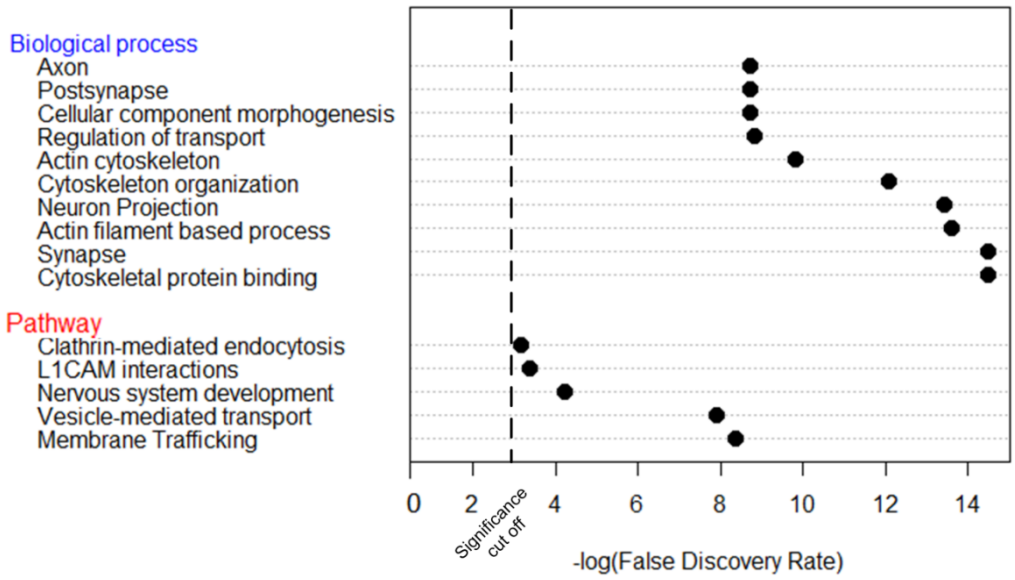


Fig. 1 Gene ontology (GO) and gene-set enrichment analysis

This analyses were performed using 55 genes obtained in the gene burden analysis of MD-EP and GSEA tool was used to obtained molecular pathways and biological processes.

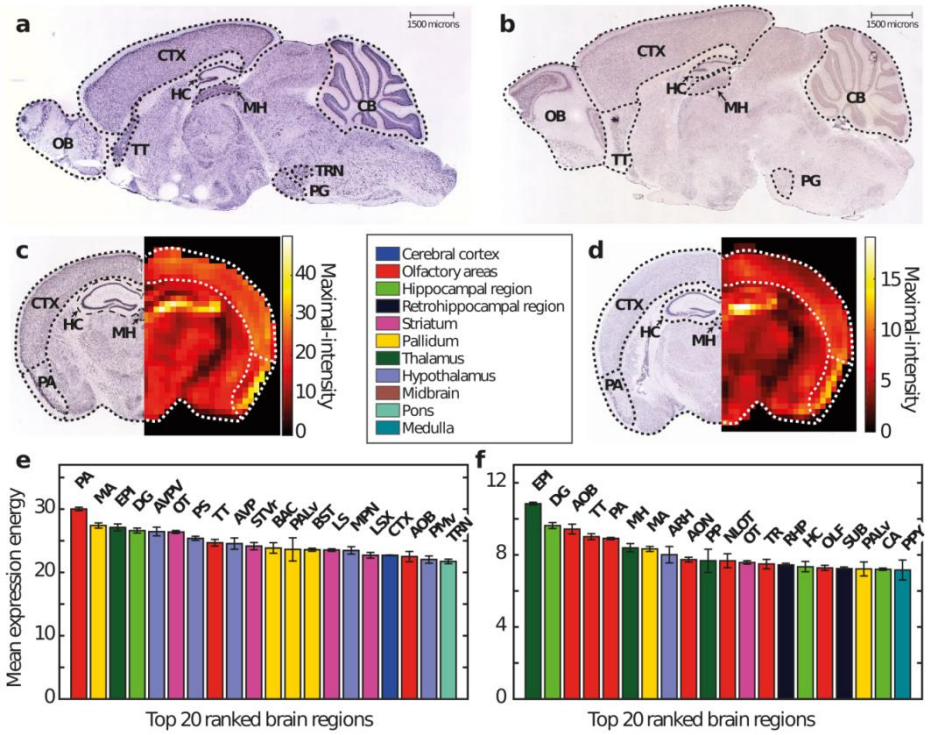


Fig. 2 Brain wide expression profiles of Ank2 and Tsc2 in the mouse brain taken from in-situ hybridization data from the Allen Brain Atlas data set (http://www.brain-map.org).

Sagittal sections of expression in the adult mouse (P56) brain for both a) Ank2 and B) Tsc2. Strong expression for both genes is found in a number of brain regions, including: CTX = Cortex, HC = Hippocampus, CB = Cerebellum, MH = Medial habenula, TT = Taenia tecta. Coronal sections (c and d, left panels) were taken from Allen Brain Atlas in a pre-rendered to fit an annotated format (c and d, right panels) allowing easy identification expression in different brain regions. This was used to identify brain regions demonstrating strongest expression (e and f, see text for details). PA = piriform area, MA = magnocellular nucleus, EPI = epithalamus, DG = dentate gyrus, AVPV = anteroventral periventricular nucleus, OT = olfactory

tubercle, PS = parastrial nucleus, TT = tenia tecta, AVP = anteroventral preoptic nucleus, STVr = Striatum, ventral region, BAC = bed nucleus of the anterior commissure, PALv = pallidum, ventral region, BST = bed nuclei of the stria terminalis, LS = lateral septal nucleus, MPN = medial preoptic nucleus, LSX = lateral septal complex, CTX = cortex, AOB = accessory olfactory bulb, PMv = ventral premammillary nucleus, TRN = tegmental reticular nucleus, PP = peripeduncular nucleus, NLOT = nucleus of the lateral olfactory tract, ARH = arcuate hypothalamic nucleus, PP = peripenducular nucleus, TR = postpiriform transition area, RHP = retrohippocampal region, PPY = Parapyramidal nucleus.

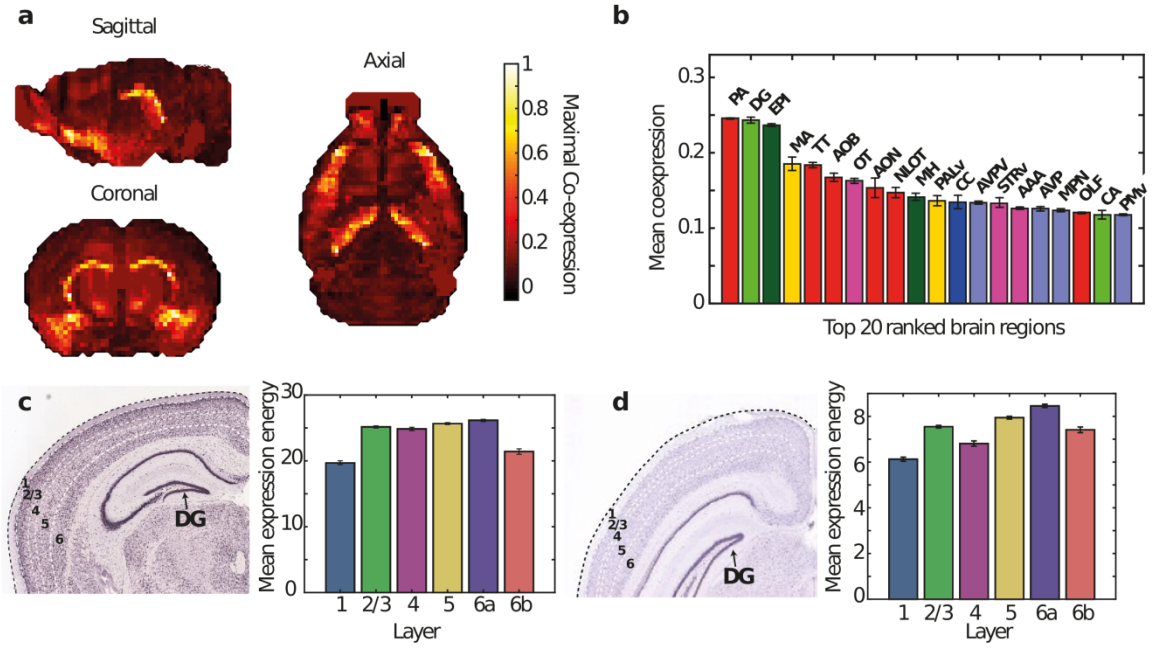


Fig. 3 Significant brain wide co-expression of Ank2 and Tsc2 was found in the mouse brain (see text).

a) Mean across-section (sagittal, coronal and axial planes) co-expression of Ank2 and Tsc2. Colorbar indicates strong coexpression in yellow. b) Mean coexpression in 209 brain regions was calculated and ranked revealing the top 20 brain regions where Ank2 and Tsc2 were co-expressed. Coronal sections revealing layer specific Ank2 expression in the cortex (c, left panel), mean expression for each cortical layer (across the whole brain) was calculated (c, right panel) revealing strongest expression in layers 2/3, 5 and 6a. Coronal sections revealing layer specific Tsc2 expression in the cortex (d, left panel), mean expression for each cortical layer (across the whole brain) was calculated (d, right panel) revealing strongest expression in layers 2/3, 5 and 6a.

Tables

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	
PRUNE2	*9*	6.02(3.60-10.07)	0.83	1.44E-08	5.89(3.52-9.85)	0.83	2.75E-08	5
AKAP9	*6*	12.32(5.48-27.68)	0.92	2.2E-06	13.89(6.17-31.27)	0.93	4.04E-07	6
SORBS1	*6*	10.93(4.87-24.55)	0.91	1.31E-05	11.52(5.12-25.93)	0.91	6.57E-06	7
ITGAX	*5*	73.02(29.68-179.66)	0.99	<1.00E-15	61.68(24.88-152.94)	0.98	<1.00E-15	14
ANK2	*4*	18.30(6.78-49.40)	0.95	1.80E-05	19.95(7.36-54.08)	0.95	7.55E-06	2
KIF20B	*4*	7.76(3.45-17.49)	0.87	1.42E-03	8.43(3.74-19.01)	0.88	5.27E-04	1
TSC2	*4*	63.73(23.35-173.96)	0.98	8.38E-13	53.56(19.47-147.30)	0.98	2.35E-11	2
SPHK2	*4*	5.47(2.25-13.28)	0.82	*NS*	5.51(2.27-13.39)	0.82	*NS*	8
SYNPO	*4*	74.87(27.35-204.94)	0.99	<1.00E-15	78.43(28.21-218.03)	0.99	<1.00E-15	32
LRPPRC	*4*	49.75(18.29-135.32)	0.98	3.73E-11	73.20(26.39-203.03)	0.99	4.19E-13	2
XYLT1	4	2.00(0.74-5.38)	0.50	*NS*	2.09(0.78-5.61)	0.52	*NS*	1
ALCAM	*3*	8.22(2.62-25.81)	0.88	*NS*	9.48(3.01-29.81)	0.89	*NS*	2
CDH13	*3*	12.15(4.50-32.85)	0.92	1.60E-03	13.09(4.83-35.46)	0.92	8.05E-04	33
DOCK7	*3*	52.57(16.54-167.10)	0.98	3.53E-08	70.09(21.61-227.27)	0.99	2.71E-09	2
BIN1	*3*	56.70(17.82-180.42)	0.98	1.53E-08	73.20(22.54-237.74)	0.99	1.72E-09	49
FLII	*3*	26.77(8.48-84.44)	0.96	3.87E-05	33.95(10.66-108.12)	0.97	4.64E-06	2
HSPA4L	*3*	17.41(5.53-54.77)	0.94	1.95E-03	16.71(5.29-52.75)	0.94	2.98E-03	1
IQSEC1	*3*	32.85(10.39-103.82)	0.97	5.12E-06	30.49(9.59-96.94)	0.97	1.32E-05	2
IQSEC3	*3*	4.16(1.33-13.04)	0.76	*NS*	4.47(1.43-14.03)	0.78	*NS*	1
LLGL1	*3*	27.29(10.06-74.03)	0.96	1.57E-07	25.08(9.21-68.31)	0.96	5.53E-07	1
MADD	*3*	128.54(39.58-417.46)	0.99	1.26E-12	73.20(22.54-237.74)	0.99	1.72E-09	49
MBP	*3*	170.13(51.78-559.02)	0.99	<1.00E-15	82.35(25.24-268.68)	0.99	4.99E-10	1
MPRIP	*3*	82.62(25.77-264.88)	0.99	2.10E-10	78.43(24.09-255.39)	0.99	8.32E-10	49
NRCAM	*3*	69.68(21.82-222.56)	0.99	1.49E-09	50.67(15.78-162.74)	0.98	8.07E-08	49

TRAP1	*3*	13.72(4.37-43.13)	0.93	1.39E-02	11.27(3.58-35.47)	0.91	*NS*	2
VCAN	3	90.37(28.13-290.35)	0.99	7.41E-11	76.61(23.55-249.22)	0.99	1.07E-09	2
MYO18A	3	204.11(72.08-577.95)	0.99	<1.00E-15	169.90(58.69-491.88)	0.99	<1.00E-15	33
MYO5A	3	13.72(4.37-43.13)	0.93	1.39E-02	16.97(5.37-53.37)	0.94	2.62E-03	1
PPP1R9A	*2*	171.57(51.97-566.41)	0.99	<1.00E-15	127.78(38.15-427.99)	0.99	7.12E-12	2
CCDC22	*2*	7.38(2.72-20.04)	0.86	*NS*	8.37(3.08-22.75)	0.88	*NS*	1
EPX	*2*	8.17(3.01-22.19)	0.88	*NS*	8.69(3.20-23.64)	0.88	4.28E-02	1

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by number

Table 1: **List of synaptic genes showing enrichment of missense variants in Spanish patients with Meniere disease (MD) ar**
reference datasets (Non-Finnish European from gnomAD.v2 or gnomAD.v3, Spanish from CSVS) were used to compar
MD cohort. Listed genes were significant when they were compared against CSVS refere

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	
Non selected Tinnitus (N=97)								
ANK2	8	3.20(1.92-5.33)	0.69	6.25E-05	3.28(1.97-5.47)	0.70	4.08E-05	2.8
MYO18A	5	5.99(2.84-12.64)	0.83	1.34E-05	5.94(2.81-12.57)	0.83	1.61E-05	6.6
MADD	4	4.99(2.23-11.17)	0.80	3.78E-04	4.78(2.13-10.73)	0.79	6.04E-04	3.8
KIF20B	4	4.99(2.06-12.06)	0.80	1.45E-03	4.89(2.02-11.87)	0.80	1.77E-03	3.7
MPRIP	3	35.36(11-113.70)	0.97	6.53E-09	28.77(8.82-93.82)	0.97	7.60E-08	15.5
MBP	2	12.95(3.18-52.76)	0.92	7.04E-04	35.34(8.20-152.23)	0.97	3.43E-06	10.3
NRCAM	2	47.15(11.13-199.77)	0.98	3.37E-07	111.91(22.52-556.20)	0.99	1.62E-08	20.7
Severe tinnitus (N=34)								
AKAP9	3	4.93(2.03-12)	0.80	1.29E-03	5.80(2.38-14.12)	0.83	3.20E-04	3.3
TSC2	2	13.92(4.41-43.93)	0.93	1.4E-05	10.97(3.47-34.66)	0.91	9.02E-05	12.8
ANK2	2	11.51(3.65-36.28)	0.91	6.06E-05	13.20(4.17-41.78)	0.92	2.26E-05	4.7

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are c

Table 2: **Swedish Tinnitus replication cohort, Synaptic genes showing an enrichment of missense rare variants in Swedish p**
diagnosis of MD

Supplementary appendix

This appendix formed part of the original submission.

SUPPLEMENTARY DATA

Table of contents

Supplementary Tables

Supplementary Table1 Clinical profile of 30 Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

Supplementary Table 2 Clinical profile of 29 Spanish patients with Meniere disease and tinnitus almost extreme phenotype (MD-AEP)

Supplementary Table 3 Clinical profile of 97 Swedish individuals with tinnitus from STOP cohort

Supplementary Table 4 Synaptic genes showing enrichment of missense variants in Spanish patients with MD and tinnitus almost extreme phenotype (MD- AEP). Three reference datasets (Non-Finnish European from gnomAD.v2 or gnomAD.v3, Spanish from CSVS) were used to compare allelic frequencies for each gene in the MD cohort. Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 5 Synaptic genes showing enrichment of synonymous variants in Spanish patients with MD and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 6 Synaptic genes showing enrichment of 5´UTR variants in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 7 Synaptic genes showing enrichment of synonymous variants in Spanish patients with MD and tinnitus almost extreme phenotype (MD-AEP). Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 8 Rare missense variants found in the gene burden analysis for *ANK2* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus.

Supplementary Table 9 Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with MD and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 10 Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with MD and tinnitus almost extreme phenotype (MD-EP). Listed genes were significant when they were compared against CSVS reference dataset

Supplementary Table 11 Synaptic genes showing enrichment of indels in Spanish patients with MD and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Supplementary Table 12 Synaptic genes showing enrichment of indels in Spanish patients with MD and tinnitus almost extreme phenotype (MD-AEP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Supplementary Table 13 Rare missense variants found in the gene burden analysis for *AKAP9* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus

Supplementary Table 14 Rare missense variants found in the gene burden analysis for *TSC2* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus

Supplementary Table 15 List of synaptic genes showing enrichment of indels in Swedish tinnitus cohort compared with SweGen and gnomAD Non-Finnish European reference datasets

Supplementary Table 16 List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) when they were compared against CSVS reference dataset

Supplementary Table 17 List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients with MD and tinnitus almost extreme phenotype (MD-AEP) when they were compared against CSVS reference dataset

Supplementary Table 18 Gene Ontology analysis showing the list of genes found in Reactome pathways and GO biological processes

Supplementary Figures

Supplementary Fig.1 Flowchart for filtering and prioritization of variants associated with tinnitus in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP). Single variant and gene burden analyses were performed in a set of 1886 synaptic genes selected for EP and almost extreme phenotypes (AEP) for tinnitus. Individuals with MD and no persistent tinnitus were used as an internal control to filter variants associated with MD

Supplementary Fig.2 Flowchart for variant analysis according to the type of variant, location in coding or non-coding regions and effect on the protein. SVA, single variant analysis, GBA, gene burden analysis

Supplementary Fig.3 Distribution of rare variants across *ANK2*, *AKAP9* and *TSC2* genes found in the gene burden analysis in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish tinnitus cohorts. Each octagon/circle indicates the position of the involved amino acid in the protein sequence

Supplementary Fig.4 Sanger sequencing of rare variants in *ANK2* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP)

Supplementary Fig.5 Sanger sequencing of rare variants in *TSC2* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP)

Supplementary Fig.6 Single-cell RNA-seq data demonstrating trimmed mean gene expression for Ank2 and Tsc2 from cortical neurons (data from the Allen Brain Institute Cell Types Database). Ank2 expression data in human (A) and mouse (C) and Tsc2 expression in human (B) and mouse (D) cortex and hippocampus. Using the Allen Brain institute taxonomy, neurons were grouped into two classes: excitatory neurons (Exc, red lines) and inhibitory neurons (Inhib, blue lines). For each population the proportion of cells in that population was plotted against the trimmed mean expression. Excitatory EP-neurons consistently displayed higher expression for each gene across species.

Supplementary Notes

Supplementary Note 1 Gene expression analysis using the Allen Brain Atlas

Supplementary Note 2 *ANK2* and *TSC2* gene expression profile in the mouse brain

Supplementary References

Supplementary Tables

Supplementary Table 1 Clinical profile of 30 Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

No.	Sex	Age at onset	THI score	Familial MD	Type of migraine	High blood pressure	Type 2 diabetes	Diagnosis of autoimmune disease	HADS score	Others
1	X	34	76	Yes	MA	-	-	-	31	-
2	X	33	98	-	MA	-	-	-	34	Arthrosis
3	X	40	76	-	MA	Yes	-	-	21	-
4	X	31	82	-	MA	-	-	-	28	-
5	X	41	78	-	-	-	-	-	27	Arthrosis
6	X	22	76	-	MO	Yes	-	-	23	-
7	X	56	82	-	-	-	-	-	20	-
8	X	14	80	-	-	-	-	-	34	Asthma
9	X	25	76	-	MA	Yes	Yes	-	NA	Anxiety Arthrosis
10	X	20	86	Yes	-	-	Yes	-	27	ArthrosisAsthma
11	X	55	86	-	-	Yes	-	-	NA	Arthrosis
12	X	35	82	-	-	NA	NA	-	NA	-
13	X	38	76	-	-	NA	NA	NA	NA	-
14	X	50	88	-	-	-	-	-	17	-
15	X	37	96	-	-	-	-	-	31	-
16	X	40	82	Yes	-	-	Yes	-	NA	-
17	X	31	80	-	-	-	-	-	NA	-
18	X	33	90	Yes	MA	-	-	Hypothyroidism	NA	Asthma, seasonal allergy
19	X	22	95	-	MO	-	-	-	24	Asthma, seasonal allergy
20	Y	29	78	-	-	Yes	-	-	30	-
21	Y	35	88	-	-	-	-	Psoriasis	16	-
22	Y	25	84	-	-	-	-	-	NA	Asthma, seasonal allergy
23	Y	57	90	-	-	-	-	-	21	-
24	Y	36	82	-	-	-	-	-	23	-
25	Y	31	94	Yes	-	Yes	-	-	10	_
26	Y	49	82	-	-	-	-	-	26	-
27	X	20	96	-	MO	-	-	-	15	-
28	X	46	88	-	-	-	-	-	18	-
29	X	54	76	-	-	Yes	-	-	28	-
30	X	40	90	Yes	-	Yes	-	-	32	-

THI= Tinnitus handicap inventory, MA=migraine with aura, MO= migraine without aura, HADS=Hospital anxiety and depression scale,

NA= data not available

Supplementary Table 2 Clinical profile of 29 Spanish patients with Meniere disease and tinnitus almost extreme phenotype (MD-AEP)

No.	Sex	Age at onset	THI score	Familial MD	Type of migraine	High blood pressure	Type 2 diabetes	Diagnosis of autoimmune disease	HADS score	Others
1	X	40	74	-	MO	Yes	-	Psoriasis	14	-
2	Y	50	68	-	-	-	-	Spondylitis, ulcerative colitis	20	Anxiety, depression
3	X	31	64	-	-	-	-	Vitiligo	NA	Artrosis
4	Y	53	72	-	-	Yes	-	-	NA	-
5	X	45	60	-	-	Yes	-	-	4	Headache
6	X	22	70	-	-	-	-	-	NA	-
7	X	31	70	-	-	Yes	Yes	Hypothryrodism Rheumatoid arthritis	NA	-
8	X	33	70	-	MA	Yes	-	Antiphospholipid syndrome	18	-
9	X	39	60	-	-	-	-	-	17	-
10	Y	58	60	-	MO	-	-	-	18	-
11	X	29	66	-	MA	Yes	-	-	NA	Arthrosis
12	Y	39	72	-	-	-	-	-	NA	-
13	X	57	70	-	NA	NA	NA	NA	NA	-
14	X	33	62	-	-	-	-	Celico	13	-
15	X	42	68	Yes	-	Yes	-	-	12	-
16	X	42	72	Yes	-	-	-	-	8	-
17	X	41	66	Yes	-	Yes	Yes	-	25	-
18	X	15	62	Yes	-	-	-	-	11	-
19	X	48	64	-	-	-	-	-	NA	Arthrosis
20	X	51	72	-	-	Yes	-	-	NA	-
21	Y	24	58	-	-	-	-	-	NA	-
22	X	39	72	-	MA	-	-	Spondylitis	NA	-
23	Y	33	56	-	-	-	-	Psoriasis	2	-
24	X	45	56	-	-	-	-	Spondylitis	NA	Arthrosis
25	Y	47	74	-	MO	Yes	Yes	-	27	-
26	X	56	74	-	-	-	-	-	22	-
27	Y	30	74	-	-	-	-	-	26	-
28	Y	21	74	-	-	-	-	-	NA	-
29	Y	48	74	Yes	-	-	-	-	10	-

THI= Tinnitus handicap inventory, MA=migraine with aura, MO= migraine without aura, HADS=Hospital anxiety and depression scale, NA= data not available

Supplementary Table 3 Clinical profile of 97 Swedish individuals with tinnitus from STOP cohort

No.	Age	Sex	THI score	Hearing disorder	Headaches	Vertigo	TMJ	Neck pain	Other pain	Under treatment for psychiatric disorder
1	53	Y	14	Yes	-	-	-	-	Yes	-
2	54	Y	18	Yes	-	-	-	-	-	-
3	69	X	24	NA	Yes	NA	-	Yes	Yes	-
4	51	Y	24	Yes	-	-	-	-	-	-
5	68	X	28	Yes	-	-	-	-	-	-
6	50	X	28	NA	-	NA	NA	NA	-	-
7	43	X	28	NA	Yes	Yes	NA	Yes	-	Yes
8	63	X	30	Yes	Yes	-	-	-	Yes	-
9	35	Y	30	Yes	-	Yes	Yes	Yes	-	-
10	37	Y	32	Yes	-	Yes	-	-	-	-
11	52	Y	32	Yes	-	-	-	-	-	-
12	39	X	32	Yes	-	-	-	Yes	Yes	-
13	60	X	32	NA	Yes	Yes	NA	Yes	-	-
14	77	X	34	Yes	-	-	-	-	-	-
15	53	Y	34	-	Yes	-	-	-	-	-
16	41	Y	34	Yes	-	-	-	-	-	Yes
17	47	Y	34	Yes	-	-	-	-	-	-
18	47	X	36	Yes	-	Yes	-	-	-	-
19	39	X	36	Yes	Yes	Yes	Yes	Yes	Yes	-
20	41	Y	36	-	-	-	-	-	-	Yes
21	37	Y	36	Yes	-	-	-	NA	-	-
22	45	X	38	Yes	Yes	-	-	Yes	Yes	-
23	48	X	38	NA	Yes	Yes	-	Yes	Yes	Yes
24	34	X	38	-	Yes	Yes	NA	Yes	Yes	Yes
25	66	X	40	Yes	-	-	-	-	-	-
26	54	Y	40	Yes	-	-	-	-	Yes	-
27	70	Y	40	Yes	-	-	-	-	-	-
28	77	Y	40	Yes	-	-	-	-	-	-
29	48	X	42	-	Yes	Yes	Yes	Yes	Yes	-
30	48	X	42	Yes	-	-	-	-	-	-
31	39	X	42	Yes	-	-	-	-	-	Yes
32	33	Y	42	-	Yes	Yes	Yes	Yes	Yes	-
33	49	Y	42	Yes	-	-	-	Yes	Yes	-
34	32	X	42	Yes	Yes	-	-	Yes	Yes	-
35	51	Y	42	NA	-	-	Yes	-	-	Yes
36	59	X	44	Yes	-	-	-	-	-	-
37	49	X	44	NA	Yes	Yes	Yes	Yes	Yes	-
38	72	X	46	Yes	-	Yes	-	-	-	-
39	52	X	48	Yes	-	Yes	Yes	Yes	Yes	-
40	35	X	48	NA	Yes	-	-	Yes	-	-
41	33	Y	48	NA	-	-	Yes	-	Yes	-
42	52	X	48	-	-	-	-	Yes	-	-
43	58	Y	48	Yes	-	-	-	-	Yes	-
44	30	X	50	Yes	Yes	NA	Yes	Yes	Yes	-
45	48	X	50	NA	Yes	Yes	Yes	Yes	NA	-
46	32	Y	50	-	-	-	Yes	-	-	-
47	33	X	50	Yes	Yes	Yes	-	-	-	-
48	54	X	52	NA	-	Yes	NA	Yes	Yes	Yes
49	77	X	52	NA	-	Yes	-	-	-	-
50	63	X	52	Yes	-	NA	Yes	-	Yes	Yes
51	37	Y	52	Yes	-	-	-	Yes	-	-
52	52	Y	52	Yes	Yes	Yes	-	-	Yes	-
53	77	Y	52	Yes	-	-	-	NA	-	-
54	79	X	52	Yes	-	-	-	-	-	-
55	56	Y	54	Yes	-	Yes	-	-	Yes	-
56	58	X	54	NA	-	Yes	-	-	-	-
57	60	X	54	NA	Yes	Yes	Yes	Yes	Yes	-
58	50	Y	54	NA	-	-	-	-	Yes	-
59	72	X	54	Yes	-	-	-	Yes	Yes	-
60	43	X	56	-	Yes	Yes	-	-	-	-
61	57	X	56	Yes	-	NA	NA	Yes	-	-

62	30	Y	56	Yes	-	-	-	-	-	-
63	37	Y	56	Yes	-	Yes	-	Yes	-	-
64	53	X	58	Yes	Yes	Yes	-	-	NA	Yes
65	35	Y	58	Yes	-	-	-	-	-	Yes
66	52	X	58	NA	NA	NA	-	-	-	-
67	29	Y	58	Yes	Yes	Yes	NA	-	Yes	-
68	52	Y	60	Yes	Yes	-	Yes	-	-	-
69	78	X	60	Yes	Yes	Yes	-	-	Yes	-
70	48	X	60	Yes	-	-	-	-	-	-
71	52	Y	60	Yes	Yes	Yes	Yes	Yes	Yes	-
72	52	Y	62	Yes	-	-	-	Yes	Yes	-
73	40	Y	64	-	-	-	Yes	Yes	-	-
74	39	X	66	Yes	Yes	Yes	Yes	Yes	-	-
75	37	X	68	Yes	-	Yes	Yes	-	-	-
76	32	Y	68	NA	-	-	-	-	-	-
77	66	X	68	NA	-	Yes	-	Yes	Yes	-
78	35	Y	68	Yes	-	-	-	-	-	-
79	50	Y	70	Yes	Yes	-	Yes	Yes	Yes	-
80	37	X	70	-	NA	-	Yes	Yes	-	Yes
81	67	Y	70	Yes	-	-	-	-	-	-
82	31	X	72	NA	Yes	Yes	-	Yes	-	-
83	58	X	72	-	Yes	-	Yes	Yes	Yes	Yes
84	53	X	72	NA	Yes	Yes	Yes	Yes	Yes	-
85	56	X	72	Yes	-	Yes	-	-	Yes	-
86	50	Y	76	Yes	Yes	-	-	Yes	Yes	-
87	40	X	78	Yes	Yes	Yes	Yes	-	Yes	Yes
88	32	X	78	NA	Yes	Yes	-	Yes	Yes	-
89	61	Y	82	-	-	-	Yes	Yes	Yes	Yes
90	52	X	82	Yes	-	NA	Yes	Yes	Yes	-
91	30	X	82	Yes	-	-	-	-	-	-
92	37	Y	86	-	-	Yes	-	-	-	-
93	35	Y	88	-	Yes	-	-	-	-	-
94	33	X	88	Yes	-	NA	-	-	-	-
95	29	Y	90	Yes	Yes	Yes	Yes	Yes	Yes	Yes
96	27	X	92	Yes	-	Yes	-	Yes	Yes	-
97	55	X	96	Yes	-	-	-	-	-	-

THI= Tinnitus handicap inventory, TMJ=Temporo-mandibular joint dysfunction, NA= data not available

Supplementary Table 4 Synaptic genes showing enrichment of missense variants in Spanish patients with MD and tinnitus almost
reference datasets (Non-Finnish European from gnomAD.v2 or gnomAD.v3, Spanish from CSVS) were used to compare allelic freq
Listed genes were significant when they were compared against CSVS reference dataset

Gene	#Variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
ITGAX	5	65.65(28.87-149.28)	0.98	<1.00E-15	65.79(28.66-151.06)	0.98	<1.00E-15
KIAA1549	5	10.18(4.19-24.69)	0.90	5.46E-04	11.82(4.86-28.76)	0.92	9.62E-05
GOLGB1	4	41.14(15.16-111.66)	0.98	5.60E-10	41.32(15.11-113.01)	0.98	7.87E-10
GPR158	4	13.80(5.12-37.20)	0.93	4.06E-04	14.70(5.44-39.75)	0.93	2.24E-04
WFS1	4	10.77(4-29.02)	0.91	4.90E-03	12.51(4.63-33.80)	0.92	1.19E-03
WNK1	4	21.10(7.81-57.01)	0.95	3.40E-06	17.88(6.60-48.43)	0.94	2.63E-05
PPFIA1	4	92.82(33.78-255.07)	0.99	<1.00E-15	85.78(30.79-238.97)	0.99	<1.00E-15
RIN1	4	35(12.91-94.86)	0.97	5.22E-09	30.50(11.20-83.04)	0.97	4.25E-08
TAOK2	4	27.05(10-73.16)	0.96	1.57E-07	21.95(8.09-59.54)	0.95	2.46E-06
CAD	3	6.07(1.94-19.05)	0.84	*NS*	5.86(1.87-18.41)	0.83	*NS*
FASN	3	7.09(2.26-22.26)	0.86	*NS*	7.19(2.29-22.60)	0.86	*NS*
KIF5A	3	27.83(8.82-87.84)	0.96	2.65E-05	27.71(8.73-87.96)	0.96	3.27E-05
ANK1	3	30.85(9.77-97.44)	0.97	9.63E-06	25.06(7.90-79.44)	0.96	8.39E-05
LMO7	3	35.35(13.01-96.10)	0.97	5.24E-09	47.13(17.14-129.56)	0.98	1.54E-10
MYO1C	3	12.78(4.73-34.56)	0.92	9.65E-04	11.92(4.40-32.29)	0.92	2.06E-03
PLXNA2	3	99.12(35.92-273.54)	0.99	<1.00E-15	120.34(42.48-340.89)	0.99	<1.00E-15
PTPRS	3	16.95(5.39-53.32)	0.94	2.46E-03	17.75(5.62-56.06)	0.94	1.80E-03
RYR2	3	153.52(46.99-501.54)	0.99	<1.00E-15	106.56(32.32-351.29)	0.99	3.22E-11
SPHK2	3	6.09(2.50-14.84)	0.84	*NS*	6.35(2.60-15.50)	0.84	*NS*
TRAP1	3	42.75(13.49-135.49)	0.98	3.31E-07	32.77(10.30-104.30)	0.97	6.50E-06

UNC13A	3	18.75(5.96-59.03)	0.95	1.03E-03	11.16(3.55-35.15)	0.91	*NS*
ST14	2	60.38(18.87-193.23)	0.98	9.15E-09	61.41(18.94-199.08)	0.98	1.28E-08
CRMP1	2	3.34(1.75-6.39)	0.70	*NS*	3.56(1.86-6.81)	0.72	*NS*

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 5 Synaptic genes showing enrichment of synonymous variants in Spanish patients with MD and tinnitus e
were significant when they were compared against CSVS reference dataset

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
SYNJ2	5	3.44(1.70-6.95)	0.71	*NS*	3.39(1.68-6.85)	0.70	*NS*
BRSK2	5	6.02(2.48-14.58)	0.83	1.34E-01	6.60(2.72-16.01)	0.85	*NS*
TLN1	5	4.42(1.82-10.71)	0.77	*NS*	4.44(1.83-10.76)	0.77	*NS*
PC	4	233.72(82.15-664.89)	0.99	<1.00E-15	141.70(49.63-404.55)	0.99	<1.00E-15
CRIP2	4	46.17(16.99-125.48)	0.98	1.09E-10	42.63(15.58-116.70)	0.98	5.26E-10
RIMBP2	4	26.73(10.96-65.20)	0.96	9.71E-10	25.51(10.42-62.48)	0.96	2.58E-09
FGD4	4	32.25(13.21-78.78)	0.97	4.65E-11	26.36(10.76-64.60)	0.96	1.56E-09
HGS	4	7.60(2.82-20.45)	0.87	*NS*	7.87(2.92-21.22)	0.87	*NS*
IARS	4	35.69(13.17-96.75)	0.97	4.00E-09	35.70(13.08-97.42)	0.97	5.58E-09
IQSEC2	3	22.66(8.36-61.39)	0.96	1.59E-06	25.37(9.31-69.11)	0.96	4.81E-07
AP3D1	3	525.88(145.47-1901.09)	0.9	<1.00E-15	366.07(98.29-1363.47)	0.99	<1.00E-15
MYH14	3	10.58(3.37-33.22)	0.91	*NS*	9.40(2.99-29.56)	0.89	*NS*
HTT	3	74.15(23.19-237.13)	0.99	7.31E-10	8.45(2.69-26.58)	0.88	*NS*
SBF1	3	11.28(3.59-35.43)	0.91	*NS*	11(3.50-34.64)	0.91	*NS*
DPP3	3	12.53(3.99-39.37)	0.92	2.83E-02	11.11(3.53-34.99)	0.91	*NS*
SYNM	3	152.22(46.56-497.67)	0.99	<1.00E-15	106.27(32.19-350.79)	0.99	3.56E-11
UNC13A	3	15.37(4.89-48.32)	0.93	5.56E-03	13.32(4.23-41.99)	0.92	1.85E-02
KIAA1217	2	9.96(3.67-27.05)	0.90	1.23E-02	10.12(3.72-27.55)	0.90	1.10E-02
LLGL1	2	97.21(30.06-314.36)	0.99	3.98E-11	237.33(67.31-836.75)	0.99	<1.00E-15
SLC25A3	2	25.56(8.07-81.01)	0.96	6.84E-05	23.21(7.29-73.87)	0.96	1.91E-04

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 6 Synaptic genes showing enrichment of 5´UTR variants in Spanish patients with Meniere disease and tin
genes were significant when they were compared against CSVS reference datase

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[CSV OR(
TUBB3	5	62.28(23.33-166.26)	0.98	4.19E-13	59.67(24.09-147.84)	0.98	<1.00E-15	20.56(7.48
ACTG1	4	261.61(65.04-1052.23)	0.99	8.38E-12	399.35(126.26-1263.12)	0.99	<1.00E-15	32.90(9.84
DPYSL2	4	7.18(3.99-12.93)	0.86	9.20E-08	6.85(3.83-12.26)	0.85	1.74E-07	6.34(3.46
RTN4	3	49.04(14.16-169.79)	0.98	1.53E-06	39.21(12.28-125.20)	0.97	1.11E-06	24.67(6.49
TRIO	3	31.38(9.39-104.87)	0.97	4.10E-05	23.69(7.48-75.05)	0.96	1.41E-04	49.36(10.97
MYO1D	2	43.95(12.77-151.23)	0.98	3.71E-06	34.23(10.70-109.55)	0.97	4.95E-06	24.87(6.52
PDE4D	2	15.75(6.27-39.53)	0.94	8.18E-06	11.41(4.64-28.05)	0.91	2.15E-04	10.85(4.15
RPLP1	2	27.27(8.19-90.77)	0.96	1.34E-04	41(12.77-131.64)	0.98	8.28E-07	24.87(6.52
SEPT2	2	16.13(4.96-52.48)	0.94	7.25E-03	12.80(4.04-40.54)	0.92	2.73E-02	16.57(4.62

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 7 Synaptic genes showing enrichment of synonymous variants in Spanish patients with MD and tinnitus al
genes were significant when they were compared against CSVS reference datas

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	
PTPRS	6	5.19(2.31-11.65)	0.81	*NS*	5.11(2.27-11.47)	0.80	*NS*	
KALRN	5	38.66(15.84-94.39)	0.97	2.09E-12	44.39(18.03-109.29)	0.98	4.19E-13	
AGAP1	4	14.96(6.15-36.41)	0.93	4.69E-06	14.43(5.92-35.20)	0.93	8.32E-06	
PIP5K1C	4	130.87(47.19-362.89)	0.99	<1.0E-15	137.78(48.42-392.07)	0.99	<1.00E-15	1
TSC2	4	34.69(12.80-94.03)	0.97	5.88E-09	49.41(18.01-135.58)	0.98	6.91E-11	1
PKP4	4	5.41(2.40-12.17)	0.82	*NS*	5.24(2.33-11.81)	0.81	*NS*	
ANK2	4	133.05(47.96-369.12)	0.99	<1.00E-15	108.25(38.50-304.37)	0.99	<1.00E-15	1
REV3L	4	70.64(25.84-193.09)	0.99	0<1.00E-15	101.03(36.04-283.24)	0.99	<1.00E-15	1
SHANK3	4	8.79(3.62-21.37)	0.89	3.02E-03	8.52(3.50-20.75)	0.88	4.39E-03	1
FARP1	3	105.03(32.58-338.63)	0.99	1.26E-11	106.56(32.32-351.29)	0.99	3.22E-11	2
AKAP9	3	99.78(30.99-321.24)	0.99	2.26E-11	136.40(40.80-456.01)	0.99	2.51E-12	2
PLXNA3	3	126.19(50.49-315.40)	0.99	<1.00E-15	112.74(44.45-285.97)	0.99	<1.00E-15	86
PRUNE2	3	19.24(6.11-60.56)	0.95	8.20E-04	19.14(6.06-60.51)	0.95	9.39E-04	51
RIMBP2	3	16.85(5.36-53.01)	0.94	2.58E-03	15.92(5.04-50.24)	0.94	4.46E-03	1
ROCK1	3	171.07(52.11-561.52)	0.99	<1.00E-15	227.34(65.22-792.44)	0.99	<1.00E-15	2
VCPIP1	3	88.04(27.43-282.52)	0.99	9.80E-11	126.29(37.96-420.24)	0.99	5.86E-12	51
SCN1A	2	14.85(5.46-40.40)	0.93	2.39E-04	12.51(4.59-34.09)	0.92	1.48E-03	1
SLC2A1	2	54.87(19.98-150.70)	0.98	1.47E-11	42.04(15.24-115.97)	0.98	9.70E-10	
KEL	2	33.97(12.43-92.83)	0.97	1.19E-08	34(12.36-93.47)	0.97	1.56E-08	1
CASK	2	183.03(55.34-605.39)	0.99	<1.00E-15	181.06(52.84-620.41)	0.99	4.19E-13	2
CAD	2	54.87(19.98-150.70)	0.98	1.47E-11	79.76(28.48-223.43)	0.99	<1.00E-15	34

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 8 Rare missense variants found in the gene burden analysis for *ANK2* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus.

Spanish MD-EP cohort (N=30)

Pos	Exon	rsID	MAF (MD-EP)	MAF NFE		MAF. CSVS	CADD	ACMG	Amino acid change
				gnomAD.v2	gnomAD.v3				
4:114262911:A>G	33	-	0.0167	-	-	-	23.2	VUS (PS4,PM2,BP1)	I1321V
4:114277102:T>G	38	-	0.0167	-	-	-	24.1	VUS (PS4,PM2,PP3,BP1)	L2443R
4:114294509:G>C	45	-	0.0167	-	-	-	25.7	VUS (PS4,PM2,PP3,BP1)	Q3921H
4:114294537:G>A	45	rs45454496	0.0167	0.0037	0.0034	0.003	25.4	Benign (PP3,PP5,BS1,BS2,BP1,BP6)	E3931K

Swedish tinnitus cohort (N=97)

Pos	Exon	rsID	MAF (Swedish cohort)	MAF NFE		MAF. SweGen	CADD	ACMG	Amino acid change
				gnomAD.v2	gnomAD.v3				
4:114275980:G>A	38	rs149645600	0.0052	0.0013	0.00113	0.001	23.1	Benign (BS1,BS2,BP4,BP1,BP4,BP6)	R2069H
4:114276906:G>A	38	rs141191319	0.0206	0.0039	0.0034	0.0095	7.91	Benign (BS1,BS2,BP4,BP1,BP4)	E2378K
4:114277914:G>A	38	rs753223319	0.0052	2.65E-05	-	-	12.84	Benign (PS4,BS1,BS2,BP4,BP1,BP4)	V2714I
4:114278016:C>A	38	rs764914059	0.0052	8.95E-06	-	-	9.27	Likely benign (PS4,PM2,BP1,BP4)	H2748N
4:114278128:C>T	38	rs145895389	0.0052	0.003	0.0031	0.0035	3.78	Benign (BS1,BS2,BP4,BP1,BP4,BP6)	S2785L
4:114279628:T>C	38	rs36210417	0.0258	0.0107	0.0110	0.0065	25	Benign (PP3,BS1,BS2,BP1,BP6)	I3285T
4:114294462:C>T	45	rs121912706	0.0052	0.0017	0.0016	0.003	35	Benign (PP3,PP5,BS1,BS2,BP1)	R3906W
4:114294537:G>A	45	rs45454496	0.0052	0.0037	0.0034	0.004	25.4	Benign (PP3,PP5,BS1,BS2,BP1)	E3931K

Swedish severe tinnitus cohort (N=34)

Pos	Exon	rsID	MAF (Swedish cohort)	MAF NFE		MAF. SweGen	CADD	ACMG	Amino acid change
				gnomAD.v2	gnomAD.v3				
4:114276906:G>A	38	rs141191319	0.0294	0.0039	0.0034	0.0095	7.91	Benign (BS1,BS2,BP4,BP1,BP4)	E2378K
4:114278016:C>A	38	rs764914059	0.0147	8.95E-06	-	-	9.27	Likely benign (PS4,PM2,BP1,BP4)	H2748N

VUS= Variant of uncertain significance

Supplementary Table 9 Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with MD a
Listed genes were significant when they were compared against CSVS reference da

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	
ANK2	4	18.30(6.78-49.40)	0.95	1.80E-05	19.95(7.36-54.08)	0.95	7.55E-06	21.9
SPTB	4	23.36(8.64-63.12)	0.96	9.86E-07	25.23(9.29-68.54)	0.96	4.60E-07	10.9
ARHGAP23	3	11.28(3.59-35.43)	0.91	*NS*	7.57(2.41-23.80)	0.87	*NS*	24.6
BIN1	3	56.70(17.83-180.42)	0.98	1.53E-08	73.20(22.54-237.74)	0.99	1.72E-09	49.36
FLII	3	26.77(8.48-84.44)	0.96	3.87E-05	33.95(10.66-108.12)	0.97	4.64E-06	24.6
TRAP1	3	13.72(4.37-43.13)	0.93	1.39E-02	11.27(3.58-35.47)	0.91	*NS*	24.6
TSC2	3	50.73(15.97-161.15)	0.98	5.22E-08	42.77(13.37-136.83)	0.98	4.60E-07	24.6
CCDC22	2	7.38(2.72-20.04)	0.86	*NS*	8.37(3.08-22.75)	0.88	*NS*	14.0
CDH13	2	19.69(6.22-62.27)	0.95	7.45E-04	21.28(6.69-67.65)	0.95	4.17E-04	49.77
FASN	2	8.94(2.83-28.18)	0.89	*NS*	7.20(2.28-22.73)	0.86	*NS*	16.5
MYO18A	2	153.51(46.73-504.26)	0.99	1.00E-15	127.78(38.15-427.99)	0.99	7.12E-12	24.8

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 10 Synaptic genes showing enrichment of missense variants with CADD≥20 in Spanish patients with MD an
AEP). Listed genes were significant when they were compared against CSVS reference

Gene	#variants	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	
DMD	3	8.04(3.30-19.61)	0.88	8.55E-03	7.49(3.07-18.28)	0.87	1.84E-02	28.
GOLGB1	3	30.85(9.77-97.44)	0.97	9.63E-06	30.99(9.75-98.52)	0.97	1.12E-05	25.5
MYO1C	3	12.78(4.73-34.56)	0.92	9.65E-04	11.92(4.40-32.29)	0.92	2.06E-03	13.6
PPFI A1	3	74.83(23.40-239.27)	0.99	6.46E-10	66.85(20.67-216.27)	0.99	4.31E-09	51.09
PTPRS	3	16.95(5.39-53.32)	0.94	2.46E-03	17.75(5.62-56.06)	0.94	1.80E-03	51.09
RYR2	3	153.52(46.99-501.54)	0.99	1.00E-15	106.56(32.32-351.29)	0.99	3.22E-11	17.0
TRAP1	3	42.75(13.49-135.49)	0.98	3.31E-07	32.77(10.30-104.30)	0.97	6.50E-06	17.0
CRMP1	2	3.34(1.75-6.39)	0.70	*NS*	3.56(1.86-6.81)	0.72	*NS*	5.2
OGDHL	2	105.95(32.70-343.29)	0.99	1.42E-11	64.89(19.98-210.70)	0.98	7.17E-09	25.7
ST14	2	60.38(18.87-193.23)	0.98	9.15E-09	61.41(18.94-199.08)	0.98	1.28E-08	17.1

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 11 Synaptic genes showing enrichment of indels in Spanish patients with MD and tinnitus extreme phenotype (MD-EP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
GPSM1	6	8.90(4.37-18.15)	0.89	3.33E-06	6.08(3.01-12.29)	0.84	9.14E-04
SGTA	6	4.91(2.52-9.58)	0.80	5.65E-03	5.11(2.63-9.92)	0.80	2.77E-03
CACNA2D1	5	13.88(5.61-34.33)	0.93	2.37E-05	24.32(9.95-59.43)	0.96	4.77E-09
TSC2	5	11.30(4.59-27.84)	0.91	2.58E-04	10.21(4.20-27.82)	0.90	5.57E-04
AAK1	4	9.50(3.48-25.95)	0.89	2.15E-02	14.01(5.18-37.87)	0.93	3.73E-04
MRAS	4	17.25(6.23-47.77)	0.94	7.95E-05	12.07(4.47-32.60)	0.92	1.71E-03
SIPA1L1	4	11.94(5.22-27.32)	0.92	8.08E-06	11.07(4.90-25)	0.91	1.39E-05
WASL	4	86.34(27.93-266.86)	0.99	1.84E-11	56.76(20.63-156.15)	0.98	9.63E-12
PIP4K2A	3	26.29(7.97-86.71)	0.96	1.49E-04	20.18(6.38-63.82)	0.95	5.93E-04
PLXNA2	3	16.59(5.98-45.99)	0.94	1.28E-04	14.34(5.29-38.90)	0.93	3.17E-04
GSK3B	3	26.29(7.97-86.71)	0.96	1.49E-04	22.78(7.20-72.15)	0.96	2.00E-04
FARSA	3	23.37(7.13-76.64)	0.96	3.74E-04	18.20(5.76-57.51)	0.95	1.45E-03
AKAP9	3	33.66(10.06-112.57)	0.97	2.15E-05	55.21(17.17-177.49)	0.98	3.16E-08
ANXA11	3	33.66(10.06-112.57)	0.97	2.15E-05	30.45(9.58-96.78)	0.97	1.32E-05
SH3PXD2A	3	17.90(5.51-58.07)	0.94	2.95E-03	13.22(4.19-41.67)	0.92	1.97E-02
STK32C	3	22.14(6.77-72.43)	0.95	5.71E-04	18.58(5.88-58.73)	0.95	1.21E-03
AP1G1	2	62.35(13.32-291.88)	0.98	2.91E-04	61.99(14.77-260.12)	0.98	3.21E-05
AP2A2	2	40.08(9-178.42)	0.98	2.40E-03	22.64(5.52-92.88)	0.96	2.79E-02
ATF7IP	2	13.25(4.10-42.82)	0.92	2.98E-02	13.64(4.30-43.22)	0.93	1.70E-02
BAZ1B	2	56.11(12.15-259.06)	0.98	4.65E-04	52.34(12.54-218.44)	0.98	1.06E-04
CORO1C	2	52.90(21.71-128.88)	0.98	<1.00E-15	103.26(43.93-242.75)	0.99	1.00E-15
DNM3	2	46.76(10.34-211.37)	0.98	1.11E-03	33.17(8.04-136.91)	0.97	2.44E-03
HSPA12A	2	112.24(21.55-584.75)	0.99	3.91E-05	196.34(43.43-887.54)	0.99	1.30E-08
ICA1	2	35.06(7.97-154.32)	0.97	4.79E-03	22.64(5.52-92.88)	0.96	2.79E-02
NDRG2	2	70.15(14.73-334.07)	0.99	1.77E-04	61.99(14.77-260.12)	0.98	3.21E-05
RGS8	2	10.33(4.15-25.72)	0.90	9.85E-04	8.86(3.60-21.79)	0.89	3.82E-03
SNAP47	2	47.18(13.70-162.48)	0.98	1.90E-06	26.20(8.22-83.52)	0.96	6.35E-05
SNX5	2	28.05(6.48-121.45)	0.96	1.56E-02	43.62(10.50-181.12)	0.98	3.81E-04
TAOK2	2	62.35(13.32-291.88)	0.98	2.91E-04	94.23(22.05-402.69)	0.99	1.61E-06
TRAP1	2	280.64(39.18-2010.18)	0.99	3.78E-05	235.61(51.03-1087.74)	0.99	4.87E-09
TUBB2A	2	190.53(53.02-684.73)	0.99	1.68E-12	92.27(32.80-259.57)	0.99	<1.00E-15

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by number of indels

Supplementary Table 12 Synaptic genes showing enrichment of indels in Spanish patients with MD and tinnitus almost extreme phenotype (MD-AEP). Listed genes were significant when they were compared against gnomAD Non-Finnish European reference dataset

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p
RYR2	6	6.30(2.95-13.42)	0.84	3.62E-03	5.34(2.52-11.33)	0.81	2.39E-02
TSC2	4	39.86(14.04-113.21)	0.97	8.49E-09	36.07(13.232-98.33)	0.97	4.58E-09
ANK2	4	9.81(3.59-26.79)	0.90	1.58E-02	11.18(4.13-30.24)	0.91	3.73E-03
SFXN5	3	71.77(20.53-250.91)	0.99	4.14E-08	88.62(27.31-287.58)	0.99	1.55E-10
BRSK1	3	119.64(32.02-446.97)	0.99	2.11E-09	54.44(16.99-174.50)	0.98	3.28E-08
ENAH	3	153.82(39.34-601.44)	0.99	8.54E-10	71.74(22.24-231.37)	0.99	1.61E-09
EPB41L2	3	28.32(8.63-92.90)	0.96	6.55E-05	21.82(6.89-69.09)	0.95	3.01E-04
GRIP1	3	89.72(25.02-321.70)	0.99	9.66E-09	113.00(34.52-369.91)	0.99	1.05E-11
HSPA9	3	215.36(50.93-910.60)	0.99	5.29E-10	188.34(56.02-633.25)	0.99	<1.00E-15
JUP	3	8.79(3.21-24.07)	0.89	4.38E-02	10.23(3.77-27.77)	0.90	9.43E-03
VPS8	3	23.91(7.34-77.92)	0.96	2.63E-04	29.14(9.18-92.54)	0.97	2.01E-05
AARS	2	29.89(6.96-128.45)	0.97	9.31E-03	65.50(15.65-274.08)	0.98	1.94E-05
ACTR2	2	65.25(14.25-298.75)	0.98	1.39E-04	66.95(15.99-280.35)	0.99	1.65E-05
ANXA6	2	358.95(49.99-2577.53)	0.99	9.32E-06	376.71(78.85-1799.65)	0.99	1.99E-10
ATP2A1	2	47.84(10.78-212.39)	0.98	6.89E-04	34.62(8.39-142.88)	0.97	1.80E-03
ATP6V1C2	2	119.64(23.81-601.07)	0.99	1.18E-05	502.29(99.98-2523.52)	0.99	8.12E-11
BASP1	2	71.77(15.50-332.42)	0.99	8.77E-05	47.07(11.34-195.35)	0.98	2.13E-04
BIN1	2	39.86(9.11-174.44)	0.97	1.87E-03	42.43(10.24-175.72)	0.98	4.44E-04
CSNK2A1	2	27.59(6.45-118.06)	0.96	1.46E-02	35.86(8.68-148.07)	0.97	1.42E-03
DAAM1	2	24.73(5.81-105.28)	0.96	2.67E-02	25.74(6.26-105.80)	0.96	1.26E-02
DNM1L	2	23.91(5.63-101.61)	0.96	3.22E-02	23.16(5.64-95.10)	0.96	2.45E-02
DOCK9	2	79.75(16.98-374.52)	0.99	5.42E-05	65.50(15.65-274.08)	0.98	1.94E-05
DPYSL2	2	57.31(16.65-197.29)	0.98	2.58E-07	54.42(16.87-175.52)	0.98	4.24E-08
EIF3C	2	119.64(23.81-601.07)	0.99	1.18E-05	158.60(36.38-691.39)	0.99	2.89E-08
NRXN3	3	9.81(3.58-26.90)	0.90	1.70E-02	8.89(3.28-24.12)	0.89	3.36E-02
PDE10A	3	12.95(4.04-41.51)	0.92	3.08E-02	11.81(3.74-37.27)	0.92	4.78E-02
EPS15L1	2	44.85(10.16-198.04)	0.98	9.78E-04	57.94(13.89-241.63)	0.98	4.77E-05
GAPVD1	2	12.31(4.44-34.10)	0.92	2.61E-03	18.26(6.66-50.08)	0.95	3.13E-05
HIBCH	2	37.76(8.66-164.63)	0.97	2.52E-03	42.43(10.24-175.72)	0.98	4.44E-04
HPCAL1	2	239.29(39.49-1450.11)	0.99	4.79E-06	94.16(22.22-399.07)	0.99	1.30E-06
ITSN2	2	35.11(10.53-117.05)	0.97	1.31E-05	21.95(6.89-69.99)	0.95	3.34E-04
MCCC1	2	51.26(11.18-228.96)	0.98	4.77E-04	35.86(8.68-148.07)	0.97	1.42E-03
MYO5A	2	239.29(39.49-1450.11)	0.99	4.79E-06	111.60(26.13-476.64)	0.99	3.68E-07
NOMO1	2	47.84(10.78-212.39)	0.98	6.89E-04	27.38(6.65-112.62)	0.96	8.50E-03
PARP1	2	65.25(14.25-298.75)	0.98	1.39E-04	38.13(9.22-157.61)	0.97	9.33E-04
PDHB	2	358.95(49.99-2577.53)	0.99	9.32E-06	1506.91(209.86-10820.36)	0.99	6.50E-10
PFN2	2	143.57(27.48-750.14)	0.99	7.39E-06	215.25(48.19-961.45)	0.99	3.76E-09
POR	2	20.52(6.29-66.96)	0.95	1.04E-03	18.56(5.83-59.09)	0.95	1.45E-03
PYGB	2	358.95(49.99-2577.53)	0.99	9.32E-06	251.13(55.38-1138.89)	0.99	1.48E-09
RAB5B	2	33.36(11.73-94.94)	0.97	9.24E-08	35.42(12.85-97.64)	0.97	1.01E-08
REV3L	2	119.64(23.81-601.07)	0.99	1.18E-05	115.89(27.08-495.90)	0.99	2.78E-07
RPL30	2	102.54(21-500.67)	0.99	1.97E-05	79.29(18.83-333.84)	0.99	4.69E-06

SAE1	2	179.47(34.44-992.81)	0.99	5.16E-06	273.96(59.83-1254.41)	0.99	9.04E-10
SH3GL3	2	90.76(25.16-327.42)	0.99	1.07E-08	101.61(30.98-333.30)	0.99	4.61E-11
SORBS2	2	37.76(8.66-164.63)	0.97	2.52E-03	42.43(10.24-175.72)	0.98	4.44E-04
SYNPO	2	43.55(12.90-146.97)	0.98	2.25E-06	33.85(10.57-108.37)	0.97	5.65E-06
TCP11L1	2	15.10(4.67-48.86)	0.93	1.11E-02	21.34(6.69-68.01)	0.95	4.30E-04
VPS41	2	143.57(27.48-750.14)	0.99	7.39E-06	111.60(26.13-476.64)	0.99	3.68E-07

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by number of indels

Supplementary Table 13 Rare missense variants found in the gene burden analysis for *AKAP9* gene in Spanish patients with MD and tinnitus extreme phenotype (MD-EP) and Swedish patients with severe tinnitus

Spanish MD-EP tinnitus cohort (N=30)

Pos	Exon	rsID	MAF (MD-EP)	MAF NFE		MAF. CSVS	CADD	ACMG	Amino acid change
				gnomAD.v2	gnomAD.v3				
7:91622303:G>C	5	rs144888041	0.0167	0.0026	0.0030	0.008	20.1	Benign (PS4,BS1,BS2,BP1,BP4)	E170D
7:91631849:A>G	8	rs746429266	0.0167	0	0	0.001	17.79	Benign (PS4,BS1,BS2,BP1,BP4)	K873R
7:91643610:G>A	10	rs139965373	0.0167	0.0004	0.0004	0.001	25	Benign (PS4,BS1,BS2,BP1,BP4)	A1194T
7:91670121:G>A	18	rs148146011	0.0167	0.0003	0.0001	0.002	22.8	Benign (PS4,BS1,BS2,BP1,BP4,BP6)	R1609K
7:91700267:T>C	28	rs76177450	0.0167	0.0049	0.0038	0.003	16.42	Benign (PS4,BS1,BS2,BP1,BP4,BP6)	S2186P
7:91732039:G>C	46	rs143306820	0.0167	4.48E-05	3.10E-05	-	24.4	Benign (PS4,PP3,BS1,BS2,BP1)	M3743I
7:91574215:CT/C-	-	rs1309343726	0.0179	0.0006	0.0006	-	-	Benign(PS4,BS1,BS2,BP4)	c.48+3755delT
7:91659313:AT/A-	-	rs779223487	0.0179	0.0002	0.0001	-	-	(Benign PS4,BS1,BS2,BP4)	c.4245+13delT
7:91706410:CT/C-	-	rs370936884	0.0179	0.0009	0.0003	-	-	(Benign PS4,BS1,BS2,BP4)	c.6765+106delT

Swedish tinnitus cohort (N=34)

Pos	Exon	rsID	MAF (Swedish cohort)	MAF NFE		MAF. SweGen	CADD	ACMG	Amino acid change
				gnomAD.v2	gnomAD.v3				
7:91603115:C>T	2	rs35669569	0.0441	0.0133	0.011	0.0205	0.009	Benign (BS1,BS2,BP1,BP4,BP6)	H47Y
7:91712609:A>C	33	rs144875383	0.0147	0.0019	0.0018	0.002	0.211	Benign (BS1,BS2,BP1,BP4)	K2762N
7:91727526:G>A	43	-	0.0147	-	-	-	32	Pathogenic (PS4,PM2,PP3,BP1)	E3571K

ACMG=American college of medical genetics and genomics, VUS= Variant of uncertain significance

Supplementary Table 14 Rare missense variants found in the gene burden analysis for *TSC2* gene in Spanish patients w
(MD-EP) and Swedish patients with severe tinnitus

Spanish MD-EP tinnitus cohort (N=30)

Pos	Exon	rsID	MAF (MD-EP)	MAF NFE		MAF. CSVS	CADD	ACMG
				gnomAD.v2	gnomAD.v3			
16:2110765:C>T	11	rs150195368	0.0167	0.0006	0.0009	-	23.8	Likely benign (PS4,PP,BS2,BP
16:2129140:C>T	27	-	0.0167	-	-	-	21.7	Likely pathogenic (PS4,PM2,PP2,PP3)
16:2133726:C>T	33	rs45517320	0.0167	6.09E-05	7.74E-05	0.001	14.01	VUS (PS4,PM5,PP2,BS1,BS2,BP4,B
16:2138096:C>T	40	rs45517391	0.0167	0.0004	0.0003	0.002	23.2	Likely pathogenic (PS4,PM1,PP2,PP3,BS2)
16:2114151:C/+TG	-	rs754285275	0.0179	0.0003	0.0002	-	-	Likely pathogenic (PS4,PM2,B
16:2123243: G/+T	-	rs141745833	0.0179	0.0003	0.0007	-	-	Benign (PS4,BA1,BP4)
16:2127041:C/+TA	-	rs200120767	0.0179	0.0018	0.0023	-	-	Likely pathogenic (PS4,M2,P
16:2130492: G/+T	-	rs112025110	0.0179	0.0048	0.0051	-	-	Benign (BA1,BP4)
16:2136476:CA/C-	-	rs142421783	0.0179	0.0006	0.0006	-	-	Benign (PS4,BA1BP4,BP6)

Swedish tinnitus cohort(N=34)

Pos	Exon	rsID	MAF (Swedish cohort)	MAF NFE		MAF. SweGen	CADD	ACMG
				gnomAD.v2	gnomAD.v3			
16:2129044:C>T	27	rs137854410	0.0147	3.60E-05	4.64E-05	-	33	VUS (PS4,PP2,PP3,BS2,BP6
16:2138546:G>A	42	rs45517419	0.0294	0.0032	0.0040	0.0035	1.823	Benign (PP2,BS1,BS2,BP4,BP

VUS= Variant of uncertain significance

Supplementary Table 15 List of synaptic genes showing enrichment of indels in Swedish tinnitus cohort compared with SweGen a
reference datasets

Gene	#indels	gnomAD.v2 OR(CI)	EF	Corrected p	gnomAD.v3 OR(CI)	EF	Corrected p	
Non selected Tinnitus (N=97)								
APC	4	2.58(2.08-3.20)	0.61	<1.00E-15	1.93(1.56-2.39)	0.48	2.58E-06	1.7
CLASP2	4	3.13(1.92-5.10)	0.68	9.48E-03	2.82(1.74-4.56)	0.64	4.92E-02	11.1
Severe tinnitus(N=34)								
AGL	4	6.93(4.01-11.95)	0.86	6.6E-09	7.91(4.61-13.58)	0.87	1.11E-10	5.3
APC	4	3.48(2.52-4.81)	0.71	9.67E-11	2.61(1.89-3.61)	0.62	1.08E-05	2.3
CLASP2	4	5.51(2.99-10.13)	0.82	8.00E-05	5.06(2.77-9.27)	0.80	2.75E-04	16.8
PC	4	4.22(3.07-5.81)	0.76	<1.00E-15	2.99(2.17-4.10)	0.67	2.67E-08	4.3
ACACA	3	172.27(92.86-319.59)	0.99	<1.00E-15	4.13(2.54-6.70)	0.76	1.88E-05	4.0
APPL2	2	5.16(3.41-7.80)	0.81	1.42E-11	5.28(3.50-7.96)	0.81	4.19E-12	5.6

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are

Supplementary Table 16 List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients (MD-EP) when they were compared against CSVS reference dataset

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[CS OR
ADGRV1	16	1.74(1.24-2.44)	0.42	*NS*	1.70(1.21-2.38)	0.41	*NS*	2.18(1.
USH1G	6	19.56(8.69-44.01)	0.95	1.01E-10	22.63(10.02-51.12)	0.96	9.65E-12	7.16(3.0
ILDR1	3	24.09(7.64-75.93)	0.96	8.50E-06	28.14(8.86-89.37)	0.96	2.29E-06	16.44(4.
MYO3A	3	3.40(1.40-8.28)	0.71	*NS*	3.48(1.43-8.49)	0.71	*NS*	5.71(2.2
OTOA	3	96.40(29.95-310.25)	0.99	2.80E-12	102.95(31.24-339.27)	0.99	3.98E-12	16.44(4.
PCDH15	3	13.99(4.45-43.97)	0.93	9.61E-04	14.37(4.56-45.32)	0.93	8.22E-04	12.33(3.
NARS2	2	116.85(27.72-492.54)	0.99	1.34E-08	122.01(28-531.73)	0.99	2.41E-08	32.90(5.9
CACNA1D	2	37.43(9.13-153.43)	0.97	7.37E-05	26.13(6.36-107.46)	0.96	9.24E-04	32.90(5.9
CDC14A	2	5.69(1.40-23.06)	0.82	*NS*	6.48(1.60-26.33)	0.85	*NS*	32.90(5.9

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 17 List of hearing loss genes showing an enrichment of missense rare variants in Spanish patients with MD
(MD-AEP) when they were compared against CSVS reference dataset

Gene	#Variants	[gnomAD.v2] OR(CI)	EF	Corrected p	[gnomAD.v3] OR(CI)	EF	Corrected p	[G O
TRIOBP	13	2.09(1.35-3.23)	0.52	*NS*	2.19(1.42-3.39)	0.54	*NS*	2.42(1
DSPP	12	2.26(1.43-3.57)	0.56	*NS*	1.52(0.96-2.40)	0.34	*NS*	2.92(1
DFNB31	3	35.21(11.13-111.34)	0.97	2.04E-07	22.87(7.22-72.43)	0.96	1.56E-05	25.54(6
DIAPH1	2	58.69(14.21-242.34)	0.98	2.77E-06	43.71(10.52-181.58)	0.98	3.05E-05	34.05(6.

OR (CI)= odds ratio (95% confidence interval), EF=etiological fraction, p-corrected values obtained after Bonferroni correction, genes are ordered by

Supplementary Table 18 Gene Ontology analysis showing the list of genes found in Reactome pathways and GO biol

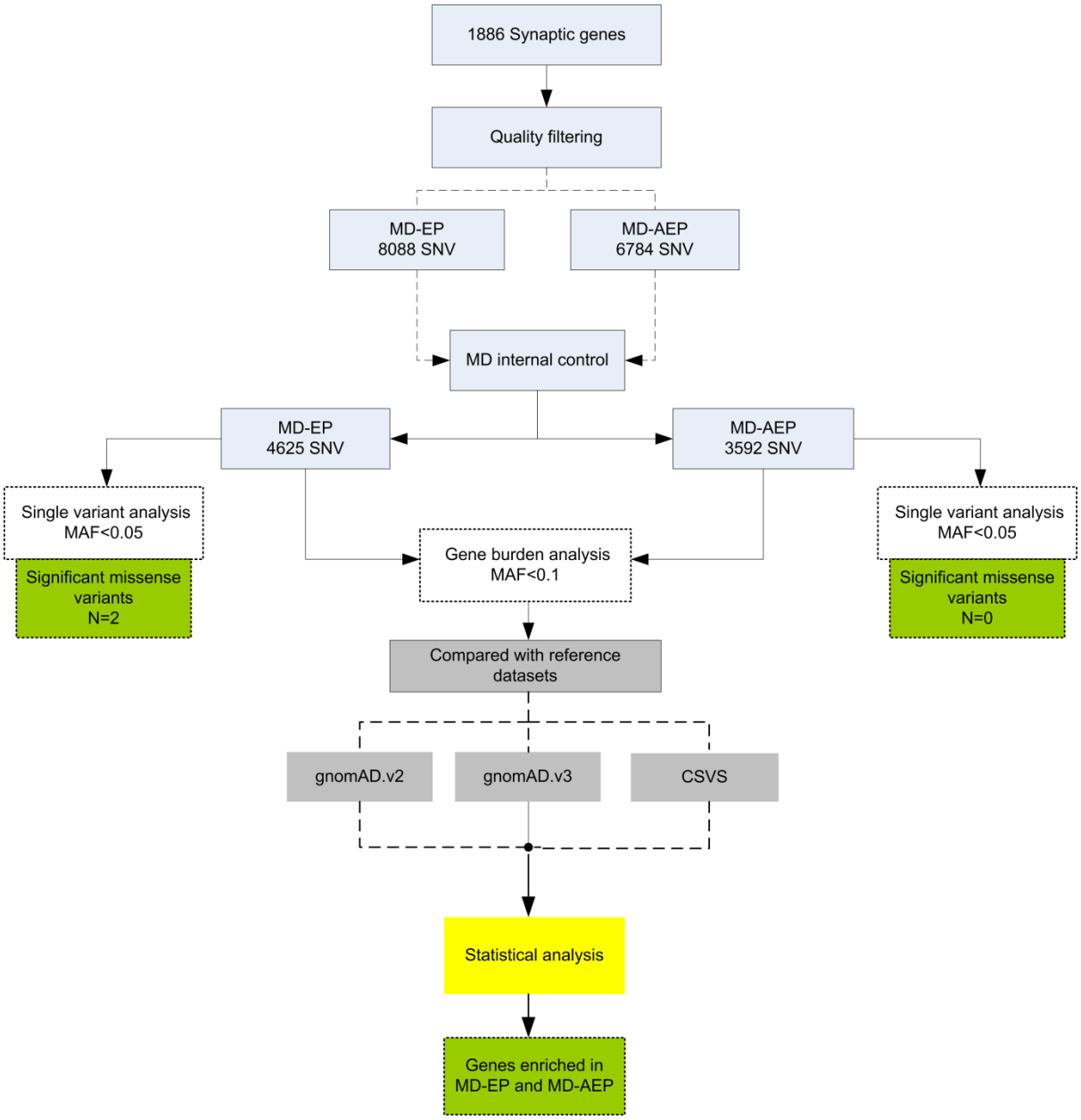
Pathway	**Total genes (N)**	**Candidate genes (N)**	**Gene(s) name**	**P-valu**
Membrane trafficking	629	13	*AP2A2,TUBB2A,DNM3,ANK2,WASL,TSC2,AP1G1,MYO5A,BIN 1,AAK1,KIF20B,SNX5,MADD*	1.47E-
Vesicle mediated transport	724	13	*AP2A2,TUBB2A,DNM3,ANK2,WASL,TSC2,AP1G1,MYO5A,BIN 1,AAK1,KIF20B,SNX5,MADD*	8.52E-
Nervous system development	580	9	*AP2A2,TUBB2A,DNM3,ANK2,WASL,,NRCAM,GSK3B,PLXNA2, MBP*	6.59E-
L1cam interactions	121	5	*AP2A2,TUBB2A,DNM3,ANK2,WASL,NRCAM*	5.90E-
Clathrin mediated endocytosis	145	5	*AP2A2,DNM3,WASL,BIN1,AAK1*	1.44E-
Biological process	**Total genes (N)**	**Candidate genes (N)**	**Gene(s) name**	**P-valu**
Cytoskeletal protein binding	979	20	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C,GSK3 B,DNM3,NRCAM,LLGL1,TAOK2,WASL,SORBS1,FLII,MPRIP,KIF 20B,AP1G1,SNX5,LRPPRC*	4.49E-
Synapse	1357	22	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C,GSK3 B,DNM3,NRCAM, AKAP9,SIPA1L1,IQSEC1,AAK1,CDH13,SNAP47,MBP,RGS8,TSC 2,ICA1,MADD,SGTA*	6.48E-
Actin filament based process	804	18	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C,LLGL 1,TAOK2,WASL,SORBS1,FLII,MPRIP,AKAP9,SIPA1L1,IQSEC1,M RAS, CACNA2D1*	7.71E-
Neuron projection	1366	21	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2, GSK3B,DNM3,NRCAM,LLGL1,TAOK2,KIF20B, AKAP9, SIPA1L1, AAK1,CDH13,SNAP47,MBP,RGS8,DOCK7,NDRG2,VCAN*	1.48E-
Cytoskeleton organization	1396	20	*PPP1R9A,MYO5A,BIN1,SYNPO,ANK2,MYO18A,CORO1C,GSK3 B,LLGL1,TAOK2,WASL,SORBS1,FLII,MPRIP, AKAP9,SIPA1L1,IQSEC,MARS,DOCK7,TUBB2A*	4.11E-
Actin cytoskeleton	503	13	*PPP1R9A,MYO5A,BIN1,SYNPO,*	8.77E-

			MYO18A,CORO1C,LLGL1,TAOK2,WASL,SORBS1,FLII,MPRIP, SH3PXD2A	
Regulation of transport	1856	19	*PPP1R9A,MYO5A,BIN1, ANK2,MYO18A, GSK3B,DNM3, LLGL1, WASL,SORBS1, KIF20B,AP1G1,SNX5,AKAP9,AAK1,CDH13,TSC2,ICA1, CACNA2D1*	1.07E-
Cellular component morphogenesis	800	14	*PPP1R9A,ANK2, CORO1C,GSK3B,DNM3,NRCAM,LLGL1,TAOK2,WASL, FLII,KIF20B, SIPA1L1,DOCK7, PLXNA2*	1.67E-
Postsynapse	640	13	*PPP1R9A,MYO5A,SYNPO,ANK2,MYO18A, GSK3B,DNM3,NRCAM, AKAP9,SIPA1L1,IQSEC1, SNAP47,TSC2*	1.83E-
Axon	643	13	*PPP1R9A,MYO5A,BIN1, GSK3B,DNM3,NRCAM,LLGL1,TAOK2,KIF20B,AAK1,MBP,DOCK 7, NDRG2*	1.94E-

FDR=False discovery rate

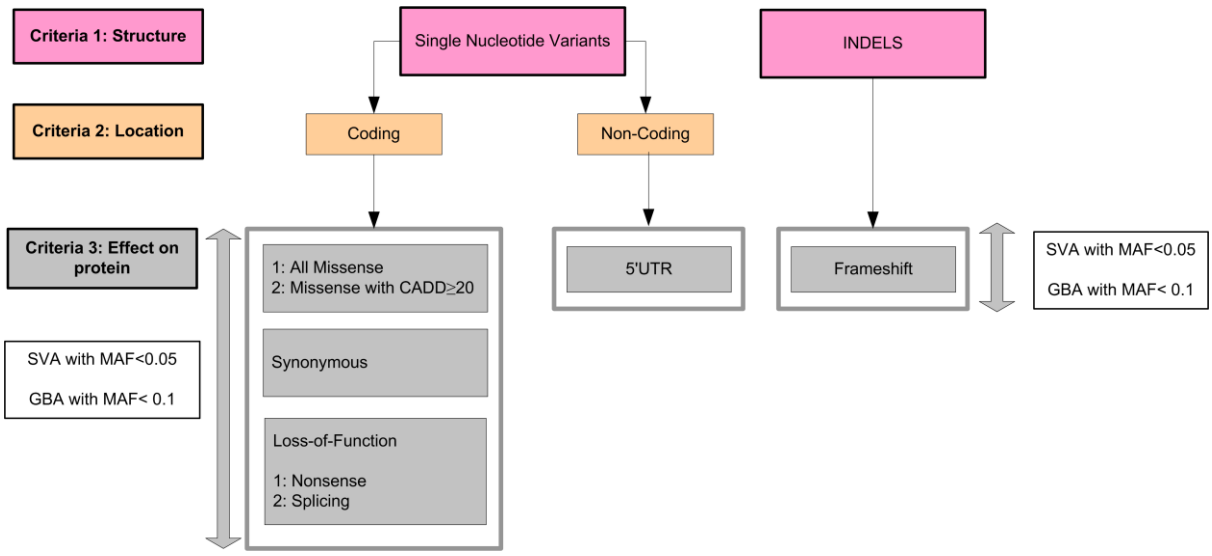
Supplementary Figures

Supplementary Fig. 1 Flowchart for filtering and prioritization of variants associated with tinnitus in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP). Single variant and gene burden analyses were performed in a set of 1886 synaptic genes selected for EP and almost extreme phenotypes (AEP) for tinnitus. Individuals with MD and no persistent tinnitus were used as an internal control to filter variants associated with MD



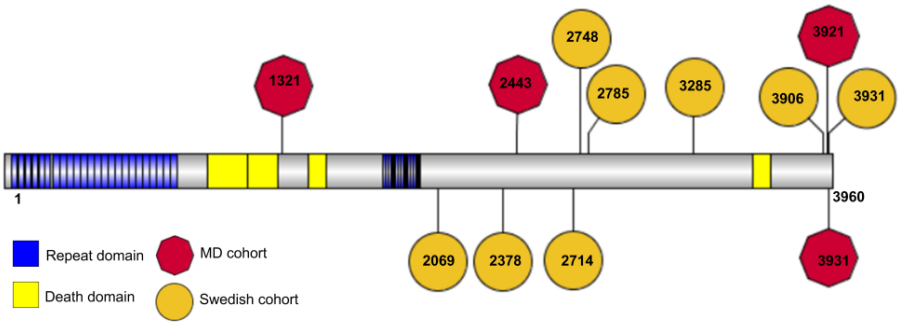
MD-EP= Meniere disease and tinnitus extreme phenotype, MD-AEP= Meniere disease and tinnitus almost extreme phenotype, CSVS=Collaborative Spanish Variant Server

Supplementary Fig. 2 Flowchart for variant analysis according to the type of variant, location in coding or non-coding regions and effect on the protein. SVA, single variant analysis, GBA, gene burden analysis.

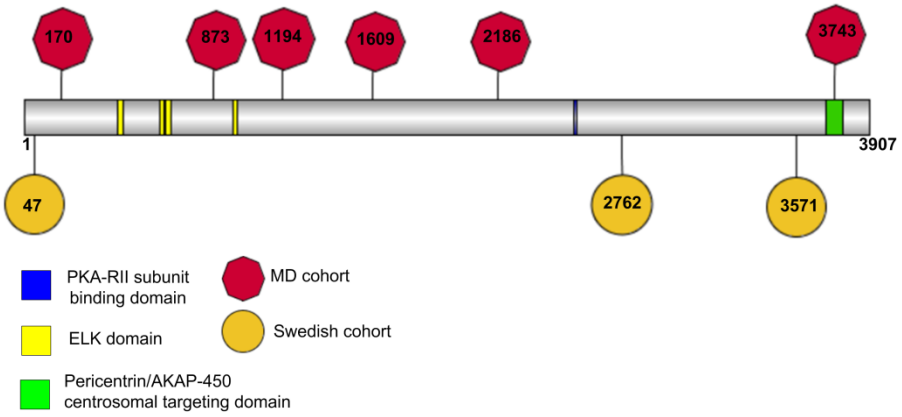


Supplementary Fig. 3 Distribution of rare variants across *ANK2*, *AKAP9* and *TSC2* genes found in the gene burden analysis in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP) and Swedish tinnitus cohorts. Each octagon/circle indicates the position of the involved amino acid in the protein sequence

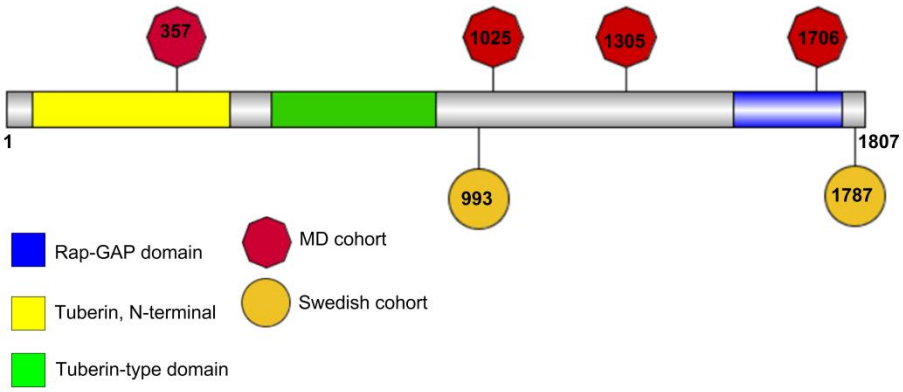
ANK2



AKAP9

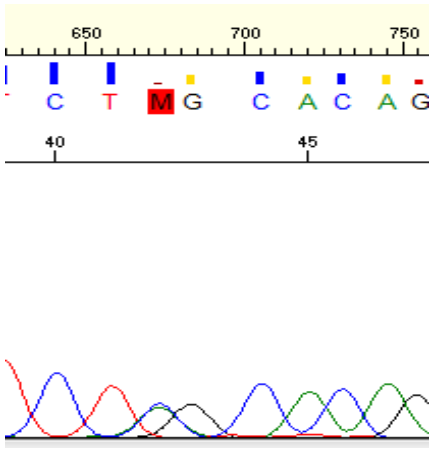


TSC2

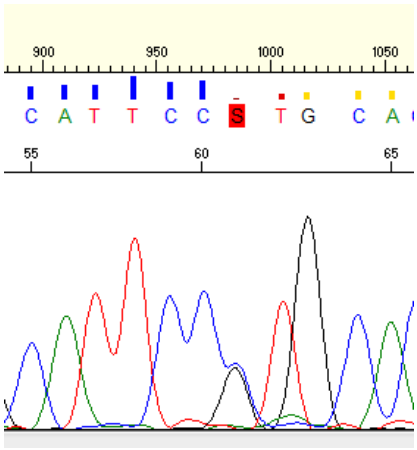


Supplementary Fig.4 Sanger sequencing of rare variants in *ANK2* gene in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

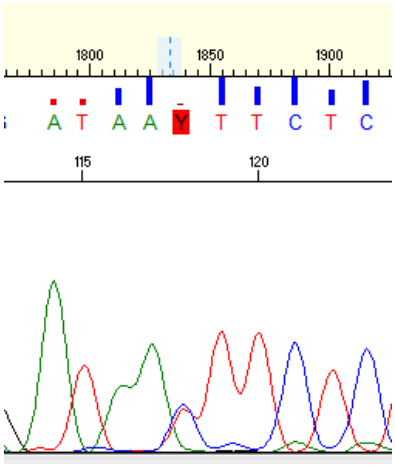
Case#8 (4:114277102 T>G)



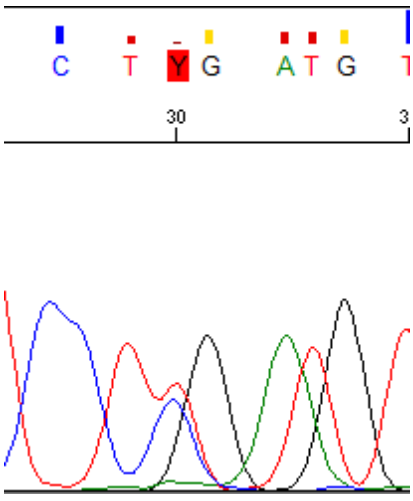
Case#8 (4:114294509 G>C)



Case#13(4:114262911 A>G)

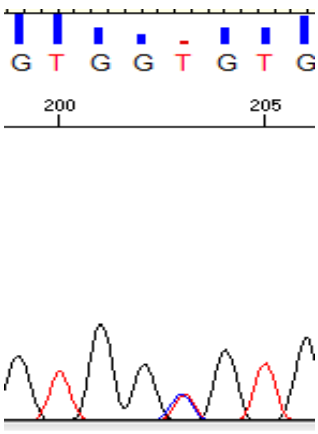


Case#6(4:1 114294537 G>A)

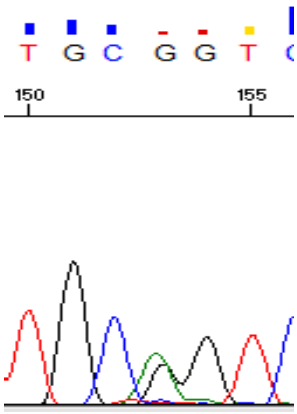


Supplementary Fig.5 Sanger sequencing of rare variants in *TSC2* gene in Spanish patients with Meniere disease and tinnitus extreme phenotype (MD-EP)

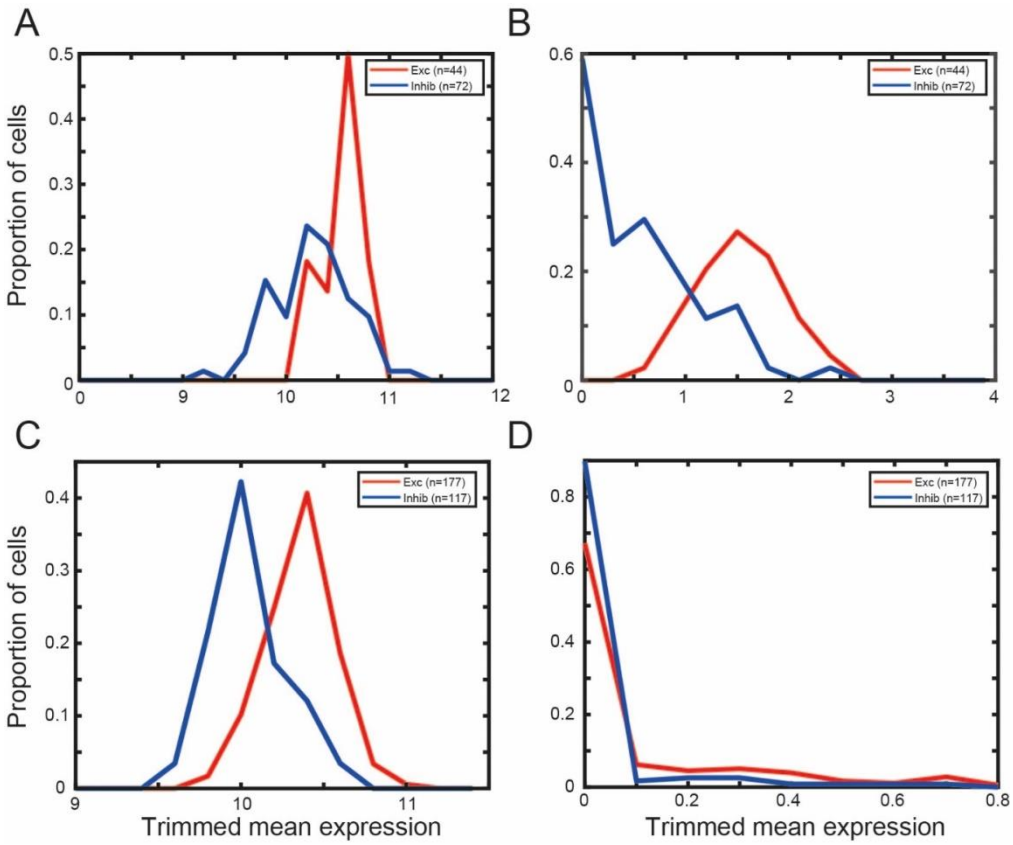
Case#19(16: 2110765 C>T)



Case#12(16: 2138096 C>T)



Supplementary Fig. 6 Single-cell RNA-seq data demonstrating trimmed mean gene expression for Ank2 and Tsc2 from cortical neurons (data from the Allen Brain Institute Cell Types Database). Ank2 expression data in human (A) and mouse (C) and Tsc2 expression in human (B) and mouse (D) cortex and hippocampus. Using the Allen Brain institute taxonomy, neurons were grouped into two classes: excitatory neurons (Exc, red lines) and inhibitory neurons (Inhib, blue lines). For each population the proportion of cells in that population was plotted against the trimmed mean expression. Excitatory neurons consistently displayed higher expression for each gene across species.



Supplementary Notes

Supplementary Note 1: Gene expression analysis using the Allen Brain Atlas

Antisense expression data were available in coronal and sagittal sectioned brains for both Ank2 and Tsc2 (4 mice: Ank2, sagittal section = Exp 68844707, coronal section = Exp 71924087 and Tsc2, sagittal section = Exp 70919985, coronal section = Exp 1431). Sagittal sections, for both brains, were visually examined and areas of high expression were noted. These regions were then confirmed via inspection of coronal sections and a good correspondence was found (every highlighted region was confirmed). To quantify these findings we obtained these data in a Matlab format[1]. ISH data for 4,104 genes were downloaded as mouse brain-wide expression profiles partitioned into 49,742 cubic voxels of 200 micron size [2,3]. In this format *expression energy* of a given gene, *g*, is a weighted sum of the greyscale-intensity of the pixels within a voxel:

$$E(v,g) = \frac{\sum_{p \epsilon v} M(p) I(p)}{\sum_{p \epsilon v} 1}$$

Where *p* denotes a given pixel, *v* a given voxel, *I(p)* the intensity within a given pixel and *M(p)* is a Boolean mask that equals 1 if the gene is expressed at pixel *p* or a 0 otherwise. Coronal data came pre-annotated to allow allocation of each voxel to a given brain region [2,3]. Mean expression energy for each brain region was simply the mean of the expression energy for all voxels annotated to fall within this region, likewise for standard deviations and counts used for calculation of the standard error of mean. The raw annotated voxel data were also transferred into SPSS to allow statistical testing of the variation in expression data. Voxels were treated as independent samples of expression within a given brain region and a Kruskal-Wallis test performed to determine if expression within different regions differed statistically.

Expression of genes were also contrasted in order determine if visually observed strong correspondence of brain wide expression of Tsc2 and Ank2 was statistically significant. First, for each pair of genes within the set of 4,104 genes a co-expression value was calculated:

$$CoExpr(g,g') = \frac{\sum_{v=1}^{V} E(v,g) E(v,g')}{\sum_{u=1}^{V} E(u,g)^2 \sum_{w=1}^{V} E(w,g')^2}$$

Where *V* is the total number of voxels. The coexpression values from all gene pairs were used to estimate the probability distribution of a given coexpression value being obtained. This distribution was then used to determine if the coexpression of Ank2 and Tsc2 was statistically significant.

Supplementary Note 2: *ANK2* and *TSC2* gene expression profile in the mouse brain

These ISH data are available pre-rendered into a three-dimensional annotated reference volume [2] of 200 micron voxels containing the maximal-intensity value within each [1]. This allowed quantitative analysis of gene expression across 209 registered brain regions (e.g. see Supplementary figure 3C-D). *Ank2* ISH data revealed significant variations in expression across brain regions (voxels were grouped by brain region and compared, Kruskal-Wallis, $p<0.001$). Mean expression was ranked across brain regions confirming the strongest expression in a number of regions (Supplementary figure 3E), including: olfactory areas (piriform area, tenia tecta and accessory olfactory bulb), the pallidum (ventral regions, particularly: the magnocellular nucleus and caudal regions, particularly: the bed nucleus of the anterior commissure and the bed nuclei of the stria terminalis), the epithalamus (particularly: medial habenula), the hippocampus (dentate gyrus and the pyramidal layers of Ammon's horn, i.e. CA) the hypothalamus (periventricular regions, particularly: anteroventral periventricular nucleus, parastrial nucleus, anteroventral preoptic nucleus and medial preoptic nucleus, and the hypothalamic medial zone, particularly: the ventral premammillary nucleus), the striatum (ventral regions, particularly: the olfactory tubercle and the lateral septal complex, particularly the lateral septal nucleus) and the cortex. In addition, strong expression was observed in pontine gray and tegmental reticular nucleus. Similarly, *Tsc2* showed significant variations in expression (Kruskal-Wallis, $p<0.001$) with the strongest expression in: olfactory areas (accessory olfactory bulb, tenia tecta, piriform area, the nucleus of the lateral olfactory tract, anterior olfactory nucleus and the postpiriform transition area), the pallidum (ventral regions, particularly: the magnocellular nucleus), the thalamus (epithalamus, particularly: medial habenula, and the peripeduncular nucleus), the hippocampus (dentate gyrus and the pyramidal layers of Ammon's horn, i.e. CA) the hypothalamus (arcuate hypothalamic nucleus), the striatum (the olfactory tubercle) and the medulla (parapyramidal nucleus).In order to identify the brain regions with high coexpression brain wide expression profiles were normalised and multiplied (see

methods) allowing a brain wide coexpression map .This revealed the strongest coexpression in: olfactory areas (piriform area, tenia tecta, accessory olfactory bulb, anterior olfactory nucleus and the nucleus of the lateral olfactory tract), the hippocampus (dentate gyrus and the pyramidal layers of ammon's horn), the epithalamus (particularly: medial habenula), the pallidum (ventral regions, particularly: the magnocellular nucleus), the striatum (ventral regions, particularly: the olfactory tubercle and anterior amygdalar area), the cerebral cortex (analysis of cortical layers revealed strong expression for both in layers 2/3, 5 and 6a but not in layers 1 and 4) and the hypothalamus (periventricular regions, particularly: anteroventral periventricular nucleus, anteroventral preoptic nucleus and medial preoptic nucleus, and the hypothalamic medial zone, particularly: the ventral premammillary nucleus).

Supplemental References

1 Grange P, Bohland JW, Hawrylycz M, Mitra PP. Brain Gene Expression Analysis: a MATLAB toolbox for the analysis of brain-wide gene-expression data. 2012.http://arxiv.org/abs/1211.6177.

2 Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A *et al.* Genome-wide atlas of gene expression in the adult mouse brain. 2007; **445**. doi:10.1038/nature05453.

3 Goldowitz D. Allen Reference Atlas. A Digital Color Brain Atlas of the C57BL/6J Male Mouse - by H. W. Dong. *Genes, Brain Behav* 2010; **9**: 128–128.