scientific reports

Check for updates

OPEN

Genome-wide association study indicates novel associations of annexin A13 to secretory and GAS2L2 with mucous otitis media

Argyro Bizaki-Vallaskangas[1✉], Joel Rämö[3], Eeva Sliz[2], Ilkka Kivekäs[1], Tytti Willberg[4], Elmo Saarentaus[3], Sanna Toppila-Salmi[5], Aarno Dietz[5], Teppo Haapaniemi[8,11], Vesa P. Hytönen[8,11], Sari Toivola[8], Aarno Palotie[3,6,9], Antti Mäkitie[7,10] & Johannes Kettunen[2]

To evaluate the genetics of chronic nonsuppurative otitis media (OM). We performed a genome-wide association study of 429,599 individuals included in the FinnGen study using three different case definitions: combined chronic nonsuppurative OM (7034 cases) (included serous and mucous chronic OM), mucous chronic OM (5953 cases), and secretory chronic OM (1689 cases). Individuals without otitis media were used as controls (417,745 controls). We used immunohistochemistry (IHC) of the murine middle ear to evaluate the expression of annexin A13. Four loci were significantly associated ($p < 1.7 \times 10^{-8}$) with nonsuppurative OM. Three out of the four association signals included missense variants in genes that may play a role in otitis media pathobiology. According to our subtype-specific analyses, one novel locus, located near ANXA13, was associated with secretory OM. Three loci (near TNFRSF13B, GAS2L2, and TBX1) were associated with mucous OM. Immunohistochemistry of murine middle ear samples revealed annexin A13 expression at the apical pole of the Eustachian tube epithelium as well as variable intensity of the secretory cells of the glandular structure in proximity to the Eustachian tube. We demonstrated that secretory and mucous OM have distinct and shared genetic associations. The association of GAS2L2 with ciliary epithelium function and the pathogenesis of dysfunctional mucosa in mucous OM is suggested. The abundant expression of annexin A13 in the Eustachian tube epithelium, along with its role in apical transport for the binding and transfer of phospholipids, indicates the role of annexin A13 and phospholipids in Eustachian tube dysfunction.

Keywords Finngen, Otitis media, Genome-wide association, Annexin, GAS2L2, Eustachian tube

Abbreviations

GWAS	Genome-wide association study
OM	Otitis media
ANXA13	Annexin A13
ET	Eustachian tube

[1]Department of Otolaryngology, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. [2]Biocenter Oulu and the Research Unit of Population Health, University of Oulu, Oulu, Finland. [3]Institute for Molecular Medicine Finland and the Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland. [4]Department of Otolaryngology, Turku University Hospital, Turku, Finland. [5]Department of Otolaryngology, University of Eastern Finland, Kuopio, Finland. [6]The Stanley Center for Psychiatric Research and Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Department of Otorhinolaryngology – Head and Neck Surgery, Helsinki University Hospital and University of Helsinki, Helsinki, Finland. [8]Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. [9]Analytic and Translational Genetics Unit, Department of Medicine, Department of Neurology, and Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. [10]Faculty of Medicine, Research Program in Systems Oncology, University of Helsinki, Helsinki, Finland. [11]Fimlab Laboratories, Tampere, Finland. ✉email: argyro.bizaki-vallaskangas@tuni.fi

GAS2L2	Growth arrest specific 2 like 2
TNFRSF13B	TNF receptor superfamily member 13B
TBX1	T-box transcription factor 1

Acute otitis media (AOM) is characterized by acute infectious symptoms, otalgia, and suppurative middle ear effusion (MEE)[1–5]. Nonsuppurative otitis media is an otitis media that involves transudation of fluid in the middle ear without pus formation. Acute nonsuppurative otitis media refers to the tubal pharynx, mouth, and cartilage segments; inflammatory mucosal hyperemia; swelling; and congestion after acute upper respiratory tract infection and may be accompanied by bacteria or viruses via the eustachian tube directly entering the middle ear cavity, resulting in an inflammatory reaction in the middle ear mucosa[1–4]. Middle ear effusions persisting for more than three months are considered chronic and can be categorized as serous (or mucous) based on the quality of the effusion. Chronic suppurative otitis media (CSOM) is a distinct entity from nonsuppurative mucous and secretory chronic otitis media with effusion (COME) and presents with infectious symptoms, purulent otorrhea and tympanic membrane perforation. The classification of middle ear fluid as either serous or mucous can, in some cases, be vague. However, because there is a separate diagnosis code for these two types of middle ear fluid, this method can be used to precisely characterize patients' clinical findings [1–3, and Supplement Fig. S4].

Among young children with COME, effusion is most common in mucoid individuals[6], whereas in adults, secretory COME is more common. As COME causes conductive hearing loss that may require surgical treatment, it significantly affects quality of life and increases health care costs. In childhood, a history of acute OM precedes the onset of COME. The etiology of adult-onset COME is variable and includes sinonasal and skull base lesions, chronic sinusitis, allergies, and eustachian tube dysfunction. The specific mechanisms that contribute to COME are, however, poorly understood[3].

Genetics could be used to gain a better understanding of OM pathophysiology. Based on the findings of twin studies, the heritability of OM (the fraction of phenotypic variation that can be attributed to genetic variation between individuals) has been estimated to be between 40 and 70%[7–11]. In a further three studies, otitis media was studied using genome-wide association study (GWAS) approaches[12–14]. To date, the largest study of childhood ear infections included 121,810 participants (with 46,936 cases of ear infection). In total, several genetic loci have been associated with the risk of OM[12–14].

At present, it is unclear whether secretory and mucous chronic otitis media are two phases of the same entity or if the underlying pathophysiology is distinct. Therefore, the aim of our study was to perform a separate genetic characterization of mucous and secretory OM to gain further insight into the disease process and to elaborate on the shared and distinct genetic risk factors involved. We performed three genome-wide association studies (GWASs) with 417,745 individuals from the FinnGen study using different case definitions: combined chronic nonsuppurative OM (n cases = 7034), mucous chronic OM (ICD H65.3; n cases = 5953), and secretory OM (ICD H65.2; n cases = 1689). We revealed two distinct loci and two shared loci for the two subcategories of nonsuppurative OM. Furthermore, we evaluated the role of annexin A13 using immunohistochemistry in a murine model.

Methods

FinnGen study data release 10 and ethics approval

Launched in 2017, the FinnGen study (https://www.finngen.fi/en) is a public–private research project that will, over time, combine the genome and digital health care data of 500,000 Finns. The goal of the project is to provide novel medically and therapeutically useful insights into human disease. The FinnGen study involved precompetitive collaboration between Finnish biobanks and their supporting organizations (universities and university hospitals), international pharmaceutical industry partners and a Finnish biobank cooperative (FINBB). In the present study, we included data that were released in autumn 2022. (Release 10). There was a total of 412,181 post-QC samples, 430,897 (pre-QC samples), 429,207 individuals with endpoints, and 429,861 individuals with detailed longitudinal data. In the present study, information about disease diagnoses and the procedures performed were collected from the Care Register for Health Care (THL) and the Causes of Death Register (Statistics Finland). As the population of northern and eastern Finland has expanded dramatically in isolation following a series of bottlenecks, it harbors numerous deleterious alleles at a high frequency[15].

Ethics statement, materials, and methods

Patients and control subjects in the FinnGen study provided informed written consent for biobank research, based on the Finnish Biobank Act. Alternatively, separate research cohorts collected prior to the Finnish Biobank Act coming into force (September 2013) and at the start of the FinnGen study (August 2017) were collected based on study-specific consent and later transferred to Finnish biobanks after approval by the Finnish Medicines Agency (Fimea), the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the FinnGen study is Nr HUS/990/2017.

The FinnGen study is approved by the Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019 and THL/1524/5.05/2019). Digital and population data service agencies (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), and Findata permit numbers THL/2364/14.02/2020.

The Biobank Access Decisions for FinnGen samples and the data utilized in FinnGen Data Freeze 10 include the following: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71,

BB2019_7, BB2019_8, BB2019_26, BB2020_1, BB2021_65, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, HUS/248/2020, HUS/150/2022 § 12, §13, §14, §15, §16, §17, §18, and §23; Auria Biobank AB17-5154 and amendment #1 (August 17 2020); and amendments BB_2021-0140, BB_2021-0156 (August 26 2021, Feb 2 2022), BB_2021-0169, BB_2021-0179, BB_2021-0161, AB20-5926 and amendment #1 (April 23 2020) and its modifications (Sep 22 2021), Biobank Borealis of Northern Finland_2017_101), and.

We declare that all experimental protocols were approved by a named institutional and/or licensing committee. All methods were conducted in accordance with relevant guidelines and regulations. All methods are reported in accordance with ARRIVE guidelines (https://arriveguidelines.org). The personal data and samples processed in the FinnGen study are obtained from biobanks referred to in the Biobank Act (688/2012). The DNA samples from biobanks are genotyped to provide genotype data for research purposes. Biobanks also provide other information, i.e., data collected in the context of health care / medical care and information related to blood donations. The persons who, in their biobank consent, have given their permission to be contacted in connection with possible future studies may, based on this consent, be invited to additional FinnGen studies that include the collection of self-reported health and well-being data. Participation in these studies is entirely voluntary. The processing of sensitive data is based on Article 9(2)(j) of the GDPR (processing is necessary for scientific and historical research purposes) and Article 6(1)(7) of the Data Protection Act V13.3/2023 (1050/2018) (Article 9(1) of the GDPR does not apply to the processing of data for scientific or historical research purposes or for statistical purposes). The research data for the FinnGen study is retrieved from biobanks and national registers in accordance with the relevant laws and permit processes. The disclosure of personal data from biobanks and national registers is based on the relevant laws. Section 28 of the Biobank Act also constitutes a key ground for the processing of personal data provided by biobanks for research (https://www.finngen.fi/en/code_of_conduct).

Phenotype descriptions and analysis

We began by performing a genome-wide association study (GWAS) of all instances of chronic nonsuppurative otitis media. The following codes were used as inclusion factors for the case cohort: at least one entry of the International Classification of Diseases, Tenth Revision (ICD-10) codes H65.2 or H65.3 and the International Classification of Diseases, Ninth Revision (ICD-9) codes 3812A and 3811A. Individuals with no records of chronic secretory otitis media (ICD-10 code H65.2 or ICD-9 code 3811A), chronic mucous otitis media (ICD-10 code H65.3 or ICD-9 code 3812A), nonspecific otitis codes (ICD-10 code H65.9, ICD-10 code H65.4, or ICD-9 code 3813X) or acute nonsuppurative otitis (ICD-10 code H65.1 or ICD-9 code 3810A) were included in the control cohort. A GWAS of overall nonsuppurative otitis media was performed with a case cohort of individuals with either secretory or mucous otitis media. A GWAS for chronic secretory otitis media was performed with a case cohort of individuals with chronic secretory otitis media (ICD-10 code H65.2). All patients with reported chronic secretory otitis media were included without any exclusion criteria. Finally, a case cohort of individuals with chronic mucous otitis media (ICD-10 code H65.3) was used for a GWAS of chronic mucous otitis media, without any exclusion criteria. For all three GWASs, the control cohort remained unchanged. Chronic nonsuppurative otitis media can be secretory or mucus. In clinical practice, however, the borderline can be hazy in some instances, and there is known to be a case overlap. Recognizing a risk for bias, we found 608 patients among our analyzed data with both ICD-10 codes. After excluding the overlapping cases, 1081 chronic secretory otitis media cases remained. The number of patients with only chronic mucous otitis media (excluding overlapped cases) was 5345. However, we did not exclude the 608 overlapping patients from our primary analyses, as a considerable number of patients were removed from a case cohort of 1689 individuals (ICD-10 = H65.2 code)) (see supplementary Fig. S1). Therefore, according to the "less strict" definition, there were 1689 participants in the case cohort with secretory otitis media and 5,953 participants in the case cohort with mucous otitis media. The overall number of patients with chronic secretory or mucous otitis media was 7034, which included patients with one or both diagnoses. From the total of our patients included to GWAS, the majority 72,6% of patients got that diagnosis at the age of less than 18 years old (non-adults). The rest of 23.4% of cases were diagnosed at adult age (18 years old or older). The diagnosis of mucous COME (H65.3) was present to 5953 cases and 4322 out of 5953 cases were diagnosed under the age of 18. The diagnosis of secretory COME (H65.2) was present to 1681 cases and 1309 out of 1681 cases were diagnosed at 18 years of age or older. A total number of 4670 cases of acute non-suppurative otitis media were identified from database and those cases were excluded from control cohort. This was done to reduce selection bias for control group. The control cohort included 417,745 patients. To correct for the population substructure, the outcome associations were tested using an additive model and the standard covariates sex, age, and the first ten principal components from the genetic data. Regarding to co-existing ICD codes in patients included we divided further the two study groups of secretory and mucous COME according to their age at the time of diagnosis. Some of co-existing ICD codes were equally present in both groups i.e., dental caries, otitis media, conjunctivitis. In non-adults group preferable co-existing ICD codes were acute upper respiratory infection, hypertrophy of tonsils or adenoids, pharyngitis (acute/chronic) and pharyngitis (acute/chronic). In adults group preferable co-existing ICD codes were maxillary sinusitis (acute /chronic), acute bronchitis, dyspnea, otitis externa, essential hypertension, hypercholesterolemia, type II diabetes. These results are interesting however not surprising. In clinical practice, we noticed that hypertrophy of tonsils/adenoids, pharyngitis/tonsilitis affects most commonly non-adults than adults patients. On the other hand, maxillary sinusitis is more often diagnosed to adults patients. Finally, essential hypertension, hypercholesterolemia and type II diabetes are by definition mostly diagnosed in adults patients (see Supplementary Table S4).

Genotyping, imputation, and quality control

The FinnGen study uses a biobank sample consisting of (1) prospective samples ("legacy samples") and (2) "new samples" that are collected upon request to obtain voluntary biobank consent and donate biobank samples

(usually blood). Hospital Biobank samples and Terveystalo Biobank samples are typically collected during diagnostic sampling in hospital laboratories or on hospital wards. Blood Services Biobank approval and sample collection were also obtained in connection with blood donation. Approval and sampling by the THL Biobank are typically performed in connection with the collection of research samples. In contrast, legacy samples are older sample cohorts that were collected for a specific research project before the Finnish Biobank Act came into force (September 2013). These older study cohorts were later transferred to a biobank according to the Finnish Biobank Act 13 §. FinnGen samples were genotyped with Thermo Fisher, Illumina, and Affymetrix arrays.

The chip genotype data were processed, and the QC samples were genotyped via Illumina (Illumina, Inc., San Diego, CA, USA) and Affymetrix arrays (Thermo Fisher Scientific, Santa Clara, CA, USA). Chip genotyping data produced with earlier chip platforms and reference genome builds were lifted to build version 38 (GRCh38/hg38) following the protocol described here (dx.doi.org/https://doi.org/10.17504/protocols.io.xbhfij6). The 'legacy samples' were genotyped over the years using various generations of Illumina and Affymetrix GWAS arrays.

Genotype calls were made with the GenCall and zCall algorithms for the Illumina data and the AxiomGT1 algorithm for the Affymetrix data. The 'new samples' were genotyped with the FinnGen ThermoFisher Axiom custom array at the ThermoFisher genotyping service in San Diego, CA, US. The current array (v2) contains 723,376 probesets for 664,510 markers. In addition to the core GWAS markers (approximately 500,000), it contains aboutapproximately 116,000 coding variants were enriched in Finland, > 10,000 specific markers were found for the HLA/KIR region, almost 15,000 ClinVar variants were found, approximately 4600 pharmacogenomic variants were found, and 57,000 selected markers were of special interest for partners.

All GWAS data were imputed against a Finnish population-specific whole-genome sequence (WGS) backbone. In sample wise quality control steps, individuals with ambiguous sex, high genotype missingness (> 5%), excess heterozygosity (± 4SD), or non-Finnish ancestry were excluded. This resulted in the inclusion of samples from 412,181 individuals in the association analysis. In variant wise quality control steps, variants with high missingness (> 2%), low Hardy—Weinberg equilibrium (HWE) P value ($< 1 \times 10^{-6}$), or low number of minor alleles (MAC < 3) were excluded. This resulted in the inclusion of 21,311,942 variants in the association analysis.

Genome-wide associations

To correct for the population substructure, the outcome associations were tested using an additive model and the standard covariates sex, age, and the first ten principal components from the genetic data. We performed a chi-square test of heterogeneity to explore whether the effect estimates differed significantly between the subtypes of OM.

Characterization of the associated loci

The genomic regions within the + /− 1 Mb window around the primary variant are called the associated locus. Therefore, each connected locus consisted of at least one genome-wide significant variant ($P < 5 \times 10^{-8}$) separated by at least a 1 Mb variable. According to the NHGRI-EBI catalog of human genome-wide association studies, the locus was considered novel. Candidate genes in each new locus were prioritized based on physical proximity to the index variant and the prior literature on the biological role of genes and clinical relevance. Variants with an allele frequency (AF) less than 1% were excluded.

By using bioinformatics, conservation of the residue (position) was analyzed using ConSurf with default settings. PolyPhen-2 was used to predict the pathogenicity of the variant[16]. The structural context of the target residue was analyzed using the Protein Data Bank and the AlphaFold database. The appearance of the mutation was assessed among somatic cancer mutations using ProteinPaint (see Table S1).

Immunohistochemistry

The immunohistochemistry (IHC) protocol was modified as needed, and several antibody dilutions were tested. Murine endogenous immunoglobulins were blocked with Rodent Block solution (Biocare Medical, Concord, CA, US). By using negative controls, we were able to test whether the solution had leaked and whether the mouse antibodies were specifically blocked by the primary antibody. Thus, the technical part of the optimization of the protocol was successful.

We used immunohistochemistry (IHC) of murine middle ear samples to evaluate the expression of annexin A13 in the eustachian tube and the middle ear. Samples from murine colon were used as positive controls. According to the protein atlas, there is an abundance of annexin A13 in the intestinal columnar epithelium. (https://www.proteinatlas.org/ENSG00000104537-ANXA13/tissue) Ethical permission for the use of mice for research purposes was obtained (permit number = ESAVI/23659/2018). Four adult mice (all of which were C57BL/6 wild-type mice) were euthanized by CO2 inhalation according to the guidelines for the euthanasia of rodents.

The rodents were then decapitated, and their brains were removed using blunt dissection. Middle ear specimens and temporal bones were then dissected as previously reported (http://goodrich.med.harvard.edu/uploads/3/7/7/1/37718659/cochlearwholemountstainprotocol_120420_lw.pdf). After dissection, the tissue was fixed with 4% paraformaldehyde (PFA). The specimens were decalcified with 120 mM EDTA solution for 3 days at 4 °C and then embedded in paraffin. sectioning was performed with a Leica Microtome SM2010 using a 4 µm-5 µm section thickness and 1–2 sections per slide. Paraffin sections were placed on Superfrost Plus microscope slides. Fully automated immunostaining was performed with a LabVision Autostainer instrument. Xylene was used for deparaffinization, and an ethanol series was used for rehydration of the sections. Antigen retrieval was performed at 98 °C for 30 min (98 °C PT-Module) with 10 mM Tris-1 mM EDTA buffer (+ 0.05% Tween 20, pH 9.0). TBS-0.05% Tween washing buffer was used for the washing steps. The slides were then incubated in a LabVision Autostainer at room temperature for 30 min with an unconjugated polyclonal primary antibody against annexin

A13 (ABIN7006070) diluted 1:600. Primary antibodies specific for ANXA3 were purchased from ELISA Kits & Proteins for Life Science Research (antibodies-online.com) (catalog no. ABIN7006070). Endogenous peroxidase was blocked with 3% H2O2. The Histofine SimpleStain Mouse Max Po Rabbit 414341F (Nichirei, Tokyo, Japan) was used as a detection method for the primary antibody with the diaminobenzidine (DAB) chromogen. As negative controls, we used both temporal bone and colon samples. Although the same protocol was used, we excluded the antibody for annexin A13. Finally, the slides were counterstained with hematoxylin, dehydrated, cleared, and mounted with xylene-based cover slips. Thereafter, the slides were scanned with a Nanozoomer scanner (Hamamatsu Photonics). Preparation of the slides was performed by an experienced laboratory technician. The sections were then examined with a light microscope by a pathologist and medical cell biologist.

Results

In patients with a broad definition of OM (either mucous or secretory otitis media) in the GWAS, we detected an association at four loci at genome-wide significance ($p < 1.7 \times 10^{-8}$ and AF > 1%) (Table 1). The chromosome 22 locus included the known leading variant rs1978060 near the TBX1 gene, which has previously been associated with ear infection[14,17]. In the proximity of that variant, there was a novel missense variant, rs72646967 (N (Asn) > H (His)), which, according to variant effect predictors, is tolerated, although with low confidence (CADD = 17.88)[16,18,19]. According to a recent study, the rs72646967 variant has been linked to lung function[20].

The chromosome 17 locus harbors a novel missense lead variant, rs72553883, in TNFRSF13B that has not previously been associated with ear infection even though the TNFRSF13B gene has been linked to ear infection[14]. The second locus on chromosome 17 included the missense lead variant rs3744374 (A (Ala) > (Gly)) in the GAS2L2 gene. According to a recent study, rs3744374 is protective against chronic rhinosinusitis[21]. According to the Ensembl database and the variant effect predictors Polyphen and SIFT[16,22], both rs72553883 (CADD = 8.96) and rs3744374 (CADD = 22) are predicted to be benign and well tolerated[19,23]. At the chromosome 8 locus, a novel leading variant, rs3779980, was detected near ANXA13. In addition to the lead variant, we observed a novel missense variant, rs2294013 (R (Arg) > H (His)), in ANXA13, which has genome-wide significance. Bioinformatic analysis with PolyPhen-2 predicted this mutation to be well tolerated, and the mutated residue was not highly conserved (Suppl. Table S3).

According to our GWAS of mucous otitis media, three loci were significantly associated with each other ($p < 1.7 \times 10^{-8}$) (Table 1, Fig. 1, Supplementary Fig. 3b–d). The chromosome 17 locus harbored the same two missense variants that have already been reported above (Table 1)[14,21,24]. The chromosome 22 locus near TBX1 included the same leading variant, rs1978060-A/G (OR = 1.15), that appeared in the GWAS for the combined secretory and mucous otitis media phenotype (Fig. 2, Forest plots). However, no association was detected for the locus on chromosome 8 observed in our first GWAS, which included both secretory and mucous otitis media patients (Table 1). In contrast, in the GWAS of chronic secretory otitis media, we found one locus at genome-wide significance ($p < 5 \times 10^{-8}$; Table 1; Supplementary Fig. 3a) on chromosome 8. All genome-wide significant variants in this locus were found within ANXA13. In addition to the lead variant, we also detected the missense variant

Chr:Pos (hg38)	Rsid	EAF	EA	NEA	Gene	Phenotype	OR	OR 95% CI	P value
8:123,697,261 intron	rs3779980	0.530	G	A	*ANXA13*	Nonsuppurative	0.90	0.867–0.929	6.24×10^{-9}
		0.530				Mucous chronic otitis media	0.93	0.898–1.060	1.13×10^{-4}
		0.530				Secretory chronic otitis media	0.77	0.724–1.174	1.45×10^{-13}
17:16,940,415 missense	rs72553883	0.024	G	T	*TNFRSF13B*	Nonsuppurative	1.42	1.28–1.57	6.76×10^{-11}
		0.025				Mucous chronic otitis media	1.47	1.31–1.63	9.45×10^{-12}
		0.025				Secretory chronic otitis media	1.27	1.15–1.73	9.12×10^{-05}
17:35,745,536 missense	rs3744374	0.230	G	A	*GAS2L2*	Nonsuppurative	0.90	0.87–0.93	6.01×10^{-8}
		0.230				Mucous chronic otitis media	0.86	0.82–0.90	5.7×10^{-10}
		0.230				Secretory chronic otitis media	1.00	0.92–1.09	9.80×10^{-01}
22:19,762,002 intron	rs1978060	0.640	A	G	*TBX1*	Nonsuppurative	1.14	1.10–1.18	4.55×10^{-13}
		0.640				Mucous chronic otitis media	1.15	1.10–1.19	1.3×10^{-11}
		0.640				Secretory chronic otitis media	1.11	1.04–1.17	4×10^{-03}

Table 1. Loci associated with nonsuppurative otitis media and its subtypes according to three GWASs. The protein-coding gene nearest to the leading variant in the locus was the nearest protein-coding gene, which was the number of genes within 1 MB of the Functional Mapping and Annotation locus. Chr, chromosome; Pos, position; EA, effect allele; NEA, noneffect allele; OR odds ratio; CI, confidence interval; EAF, effect allele frequency ;study groups : (a) total chronic nonsuppurative otitis media cases (both secretory and mucous otitis media cases together): N = 7034 patients, (b) mucous chronic otitis media, N = 5953 and (c) secretory chronic otitis media, N = 1689.

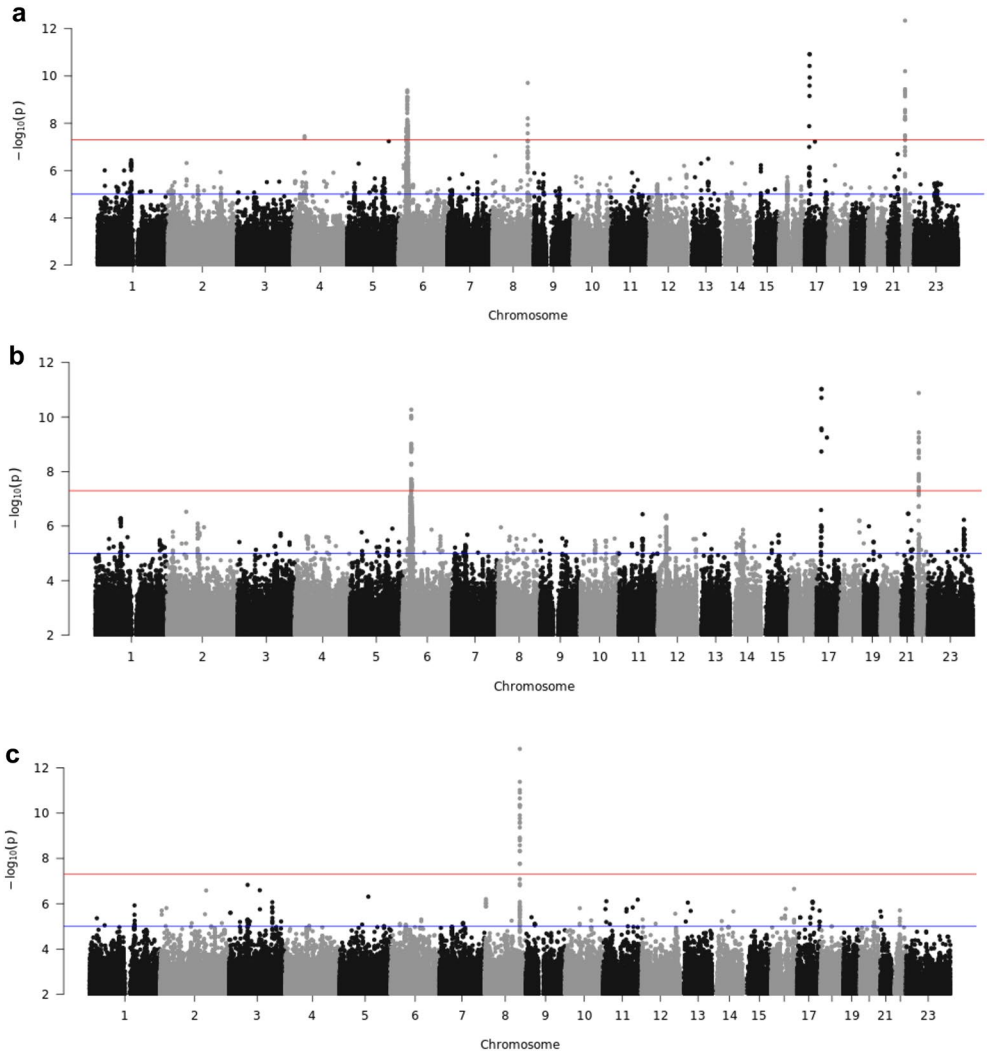

Figure 1. Manhattan plots of the GWAS results. Analysis of chronic nonsuppurative otitis media cases (both secretory and mucous otitis media cases together): N = 7034 patients (R10) (**a**), mucous chronic otitis media (**b**), N = 5953 and secretory chronic otitis media (**c**), N = 1689.

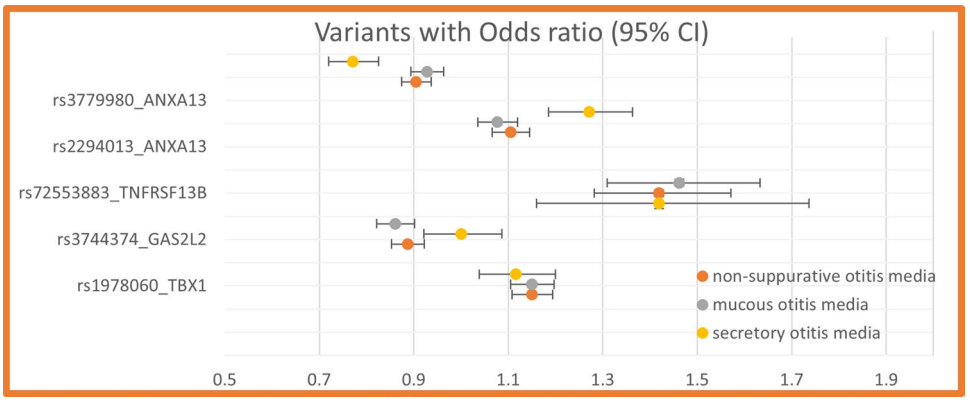

Figure 2. Forest plots of the odds ratios and confidence intervals for lead variants and missense variants for combined OM, mucous OM, and secretory OM.

rs2294013 (R (Arg) > H (His)) in ANXA13 at the genome level. The missense rs2294013 variant is predicted to be benign based on the variant effect predictors PolyPhen and SIFT, with a CADD of 12.3[16]. For all the referred variants, the GWAS analyses and Manhattan plots are available in Fig. 1, and locus zooms for all the reported variants are available in supplementary Fig. S3a–d.

We further compared the effect sizes of index variants from any of the three GWASs to assess heterogeneity based on phenotype definitions (Table 1). The effect on chronic secretory otitis media was significantly greater for the chromosome 8 locus and ANXA13 for both lead variant rs3779980 (OR = 0.77) and missense variant rs2294013 (OR = 1.28) than for mucous otitis media and cross-phenotype meta-analysis without overlapping confidence intervals. In contrast, the effect on chronic mucous otitis media was greater for the GAS2L2 gene variant rs3744374 (OR = 0.90) than for chronic secretory otitis media, without overlapping confidence intervals (Fig. 2. Forest plots).

The critical value for chi-square for degrees of freedom (df) 1 and P = 0.05 is 5.991. Thus, the data for the ANXA13 (P = 0.005) and GAS2L2 (P = 0.05) variants revealed that the variability across effect sizes exceeded what would be expected based on sampling error. However, for the TNFRSF13B (P = 0.95) and TBX1 (P = 0.9) variants, the variability across effect sizes did not exceed what would be expected based on sampling error.

According to our GWAS for both mucous and secretory otitis media (Fig. 1a) and for patients with only mucous otitis media (Fig. 1b), there was a strong locus on chromosome 6. This locus was in the human leukocyte antigen (HLA)/major histocompatibility complex (MHC) region and was not studied further or analyzed at that time.

Phenome-wide association studies of the lead variants in the FinnGen study

We evaluated the co-associations of the lead variants in the FinnGen study. Among the variants associated with chronic mucous otitis media, the missense variant rs72553883 in TNFRSF13B was associated most strongly with chronic mucous otitis media in the primary GWAS. Furthermore, it was strongly associated with multiple diseases, including chronic diseases of the tonsils and adenoids, nonsuppurative otitis media, eustachian salpingitis, and tonsillectomy ($p < 1.7 \times 10^{-8}$) (Supplementary Table S2). The missense variant rs3744374 near GAS2L2 was significantly associated with chronic sinusitis, diseases of the middle ear and mastoid, nonsuppurative otitis media, and other diseases of the upper respiratory tract ($p < 1.7 \times 10^{-8}$). The variant rs1978060 near TBX1 was significantly associated with suppurative and nonsuppurative otitis media, nasal polyps, and diseases of the middle ear and mastoid ($p < 1.7 \times 10^{-8}$). The variant chromosome 8 locus and the reported variants near ANXA13 were associated with nonsuppurative otitis media, even though the association was clearly stronger with the more specific phenotype of secretory otitis media.

Immunohistochemistry

To further assess the potential role of annexin A13 in otitis media pathophysiology, we conducted immunohistochemistry experiments that revealed the expression of annexin A13 at the apical pole of the epithelium of the eustachian tube as well as variable intensity of annexin A13 in the secretory cells of the glandular structure in proximity to the eustachian tube (Figs. 3 and 4, Supplementary Fig. S2). Expression of annexin A13 was seen at the apical pole of the epithelium of the nasopharynx too. No significant staining was found in the middle ear's epithelium or ossicles. Very strong staining for annexin A13 was found in Zymbal's gland. This is an auditory sebaceous gland ventral to the orifice of the external ear. Humans do not have a Zymbal's gland, although the human external auditory meatus dermis contains both sebaceous glands and coiled tubular ceruminous glands that secrete wax. As previously reported in the literature[25], the expression of annexin A13 was intense in the intestinal epithelium (Fig. 5). Furthermore, negative controls stay clearly negative.

Cross referencing with UKBB results

To further validate our findings, we investigated how the variants were associated with similar outcomes in publicly available UKBB data. The UKBB associations for the variants in Table 1 are shown in Supplementary Table S2. In brief, the ANXA13-leading variant was associated with otitis media and eustachian tube disorders. In

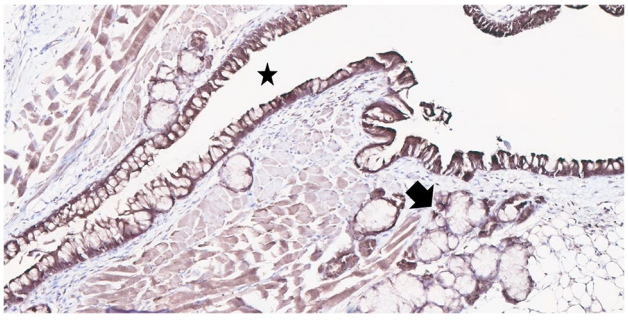

Figure 3. Immunohistochemistry for annexin A13 in murine eustachian tubes. Immunohistochemistry revealed localization of annexin A13 (brown) in the epithelium of the eustachian tube (Black star) as well as in the epithelial mucous glands (arrow) (magnification, × 20).

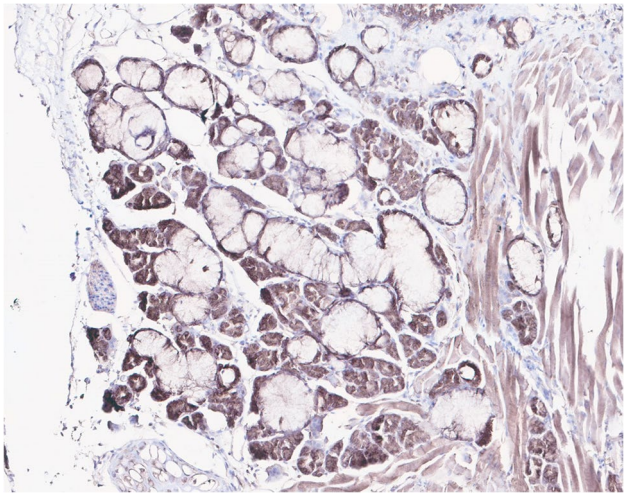

Figure 4. Immunohistochemistry for annexin A13 in the eustachian tube and upper airway. Redundant expression of annexin A13 in the glandular epithelium of the eustachian tube in proximity to the nasopharynx (magnification × 40).

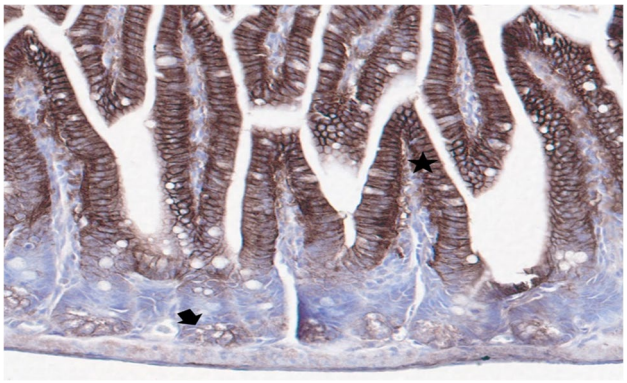

Figure 5. Immunohistochemistry for annexin A13 in the intestinal epithelium. Annexin A13 was strongly expressed in the absorptive columnar epithelium (black star) of the villous processes (small intestine, magnification × 40). Neither the lamina propria nor the bottom of the crypts were stained with annexin A13 (light blue). However, Paneth cells (black arrow) showed particularly clear expression of annexin A13.

particular, the otitis media and eustachian disorder phenotypes correspond quite well with our chronic secretory otitis media phenotype. Additionally, the TBX1 leading variant was associated with the nasal polyp phenotype. For both the ANXA13 and TBX1 variants, the allele frequency (AF) was the same for both the Finngen and the UKBB data.

Discussion

This is the first description of a subtype GWAS for chronic nonsuppurative otitis media. Here, we found four loci that were significantly associated ($p < 1.7 \times 10^{-8}$) with overall nonsuppurative OM. Moreover, three out of the four association signals included missense variants in genes with a plausible role in otitis media pathobiology. According to our subtype-specific analyses, one novel locus near ANXA13 was associated with the secretory subtype, and three loci (near TNFRSF13B, GAS2L2, and TBX1) were associated with the mucous subtype. Interestingly, the association of ANXA13 was replicated in the UKBB data.

Immunohistochemistry revealed the expression of annexin A13 at the apical pole of the Eustachian tube epithelium as well as at variable intensities in the secretory cells of the glandular structure in proximity to the Eustachian tube. We discussed the potential mechanisms by which proteins encoded by the highlighted genes could lead to OM susceptibility.

There are two variants each for ANXA13 and TBX1; however, based on the odds ratios, their direction of effects is opposite to each other for both genes. Regarding the ANXA13 locus, rs3779980 was the lead variant, whereas for the missense variant rs2994013, the r2 value was between 0.6 and 0.8, which indicated that these two variants were not in complete linkage disequilibrium (LD) but were highly correlated. Regarding the TBX1

variants, rs1978960 was the lead variant, and for the missense variant rs72646967, the r2 was between 0.2 and 0.4. This means that they are not in linkage disequilibrium.

In this study we focused on studying only chronic OME and not acute OME. It is true that in most cases of chronic OME, symptoms subside before the age of six years. However, it has been advocated that patients with chronic OME during childhood may still have a partially dysfunctional eustachian tube as adults. On the other hand, our cohorts revealed that secretory COME is mostly seen among adults and not in childhood. Additionally, even as a mostly transient, benign, and relatively mild condition, COME causes disability to the affected subjects and remarkable costs and stress to the closet family. The early childhood is critical phase for language and speech development and COME causes conductive hearing loss which consequently will negatively affect the language and speech development.

Annexin-13 (ANXA13)

Annexin A13 is considered the original progenitor of the 11 other members of the vertebrate annexins. The specific function of this gene is not yet fully understood. However, it is a calcium dependent, phospholipid-binding protein and an apical vesicle that may play a crucial role in the lipid raft-mediated delivery of apical proteins[26]. Earlier work has shown that ANXA13 may also play a role in epithelial cyst formation and ciliary function[27].

The eustachian tube connects the middle ear to the nasopharynx, and it is important that it is patent so that the middle ear ventilates. Previously, it was observed that, similar to the lung, there is a surfactant in the Eustachian tube epithelium that consists partially of phospholipid proteins. Based on earlier research[28], a phospholipid layer is needed to make the surface of the epithelium of the Eustachian tube hydrophobic. This process is crucial for opposing the strongly adhesive nature of proteins and easing the patency of the eustachian tube.

Together with our findings that link the expression of ANXA13 to the eustachian tube, these findings suggest for the first time that ANXA13 plays an important role in the function of the eustachian tube and the middle ear[8,26,29,30]. Furthermore, according to the Human Protein Atlas[10], cytoplasmic expression of annexin-13 has been detected in the gastrointestinal tract, bile ducts, gallbladder, renal tubules, and fallopian tubes.

Unfortunately, there is quite an amount of non-specific staining of the mucosa and further studies are needed to confirm those preliminary encouraging results. According to pathologist who assessed the results of IHC for annexin A13, we should not let the dyeing become too dull, as the proper dyeing will not occur. The considerable amount of non-specific staining of the mucosa to due to annexin A13 protein's broad-spectrum appearance. To increase the quality of the results, a special protocol was following for tissue fixation and embedding to preserve structural integrity. Sectioning was done at consistent thicknesses, and we ensured that tissue sections are selected at regular intervals and randomly oriented to avoid bias. However, we did not quantify the results by any stereological methods.

A query of our reported leading variants in the UKBB revealed the strong association of the variant rs3779980 with otitis media and eustachian tube disorders[31]. Interestingly, the allele frequency (AF) and odds ratio (OR) were similar to those in our GWAS in patients with secretory otitis media (Fig. 2. Forest plots). Although TNFRSF13B and GAS2L2 have previously been linked to childhood ear infection, specific missense variants (rs72553883 and rs3744374) were not significantly associated with any phenotype in the UKBB cohort. Finally, the variant rs1978060 was associated with nasal polyps ($P = 5.80 \times 10^{-7}$) in the UKBB cohort. (Supplementary Table S2).

TNF receptor superfamily member 13B (TNFRSF13B)

The protein encoded by this gene is a lymphocyte-specific member of the tumor necrosis factor (TNF) receptor superfamily. It interacts with calcium modulators and cyclophilin ligands (CAMLs). The protein induces the activation of the transcription factors NFAT, AP1, and NF-kappa-B and plays a crucial role in humoral immunity by interacting with a TNF ligand. TNFRSF13B mediates the calcineurin-dependent activation of NF-AT as well as the activation of NF-kappa-B and AP-1. Additionally, it stimulates B- and T-cell functions and regulates humoral immunity[24]. The missense variant rs72553883, which was significantly associated with mucous otitis our media according to analyses, has been reported in the literature in multiple individuals diagnosed with common variable immunodeficiency[32–34]. This variant is detected in clinically asymptomatic relatives as well as in the Finnish European population, with an overall allele frequency of 2.4% (603/24990 alleles, including eight homozygotes) in the Genome Aggregation Database. These findings suggest that these variants may act as susceptibility or disease-associated variants with other genetic and/or environmental factors[35]. According to Clinvar[36,37] and Ensembl[22], the missense variant rs72553883 is most likely pathogenic. Other variants in TNFRSF13B (rs11649804, rs34557412) have previously been linked to tonsillectomy (rs34557412) and childhood ear infection[14,21,38]. In line with these findings, we observed coassociations with sinusitis and chronic diseases of the tonsils and adenoids in the FinnGen study[21]. Functional characterization of the rs72553883 variant in mice and human B cells revealed disrupted NF-κB signaling and significantly impaired B-cell function[39,40].

Growth arrest specific 2 like 2 (GAS2L2)

Growth arrest specific 2 like 2 (GAS2L2) is a protein-coding gene that has previously been associated with primary ciliary dyskinesia (PCD) by impairing cilia orientation and mucociliary clearance[41,42]. The gene ontology (GO) annotations related to these genes included the regulation of microtubule dynamics and stability via interactions with microtubule plus-end tracking proteins, such as MAPRE1[43]. GAS2L2 regulates ciliary orientation and performance in cells found in the airway[40]. The rs3744374 variant has been recently linked to chronic rhinosinusitis and other sinonasal and pharyngeal diseases[21], although based on previous studies, it is predicted to be benign[16,44]. There is aA mouse model (C57BL/6 N-Gas2l2tm1a(KOMP)Wtsi/Wtsi) with an abnormal auditory tube and middle ear was used[32].

T-box transcription factor 1 (TBX1)

Loss of function of the TBX1 gene causes dysregulation of neural competence in otocyst regions, which is linked to the formation of either mechanosensory or structural sensory organ epithelia[44–46]. The TBX1 gene is also needed for inner ear morphogenesis[47]. The leading variant, rs1978060, which was associated with mucous otitis media in this FinnGen study, was also associated with nasal polyps ($p = 9 \times 10{-}09$). In an earlier GWAS in the Finnish population, the variant rs1978060 was reported to have a discordant association with chronic rhinosinusitis[21].

According to a GWAS meta-analysis of Japan Biobank data with data from the UKBB and Finngen, rs1978060 is associated with nasal polyps[12,17,31]. In previous studies, GWAS data from research participants of European ancestry revealed the association of rs1978060 with childhood ear infection and myringotomy[14,31,38]. Childhood ear infections are strongly influenced by the shape of the eustachian tube. Murine models in which the TBX1 gene was knocked out did not exhibit any eustachian tube abnormalities but did reveal abnormal development of the ear (outer, middle, or inner ear) and hearing (decreased hearing or even deafness)[48]. These findings show the crucial role of TBX1 in ear development.

Potential impact of knowledge on loci

By identifying loci associated with various conditions, GWAS provides a powerful tool for revealing the genetic background of health and disease. Loci discovered through GWAS can provide evidence for previously unknown biological pathways involved in disease processes, introducing new targets for therapeutic intervention. The association of GWAS loci to specific genes and understanding their functions can reveal how genetic variations influence phenotypic traits. GWAS findings can provide hints for innovating epigenetic research, exploring how genetic variants influence gene expression through epigenetic modifications. The pathophysiology of COME is undoubtedly multifunctional[49]. It is possible to tailor GWAS that assesses interactions between specific genetic variants and environmental factors with impact on related COME traits. Combining GWAS with other omics data (e.g., transcriptomics, proteomics) can provide a more comprehensive view of the genetic architecture of diseases. As a next phase of our study, long-term studies tracking genetic and phenotypic changes over time will be essential for understanding the dynamic nature of genetic influences on health.

Limitations of the study

Because the populations of northern and eastern Finland have expanded dramatically and in isolation following a series of bottlenecks, these populations harbor numerous deleterious alleles at relatively high frequencies compared to those of other European countries. This increase in allele frequency aids the statistical power to detect certain disease alleles. In contrast, the population history has also removed some genetic variants or lowered their frequency, making it much more difficult to find those variants using the Finnish population[18].

Although the number of cases in our study was lower than usual in complex trait GWASs, we were able to detect loci associated with the risk of chronic nonsuppurative otitis media and its subtypes. The four lead variants in the discovered loci explained 0.4% to 0.8% of the trait variation, which means 0.1% to 0.2% per locus. This value is greater than that typically found for common complex traits.

We acknowledge the presence of selection bias in our GWAS analyses especially since in Finngen database and Finnish biobanks there are more adults patients compared to children. In addition to that, we are aware of the selection bias since our study cohorts was made according to ICD-10 codes (H65.2 for secretory otitis media and H65.3 for mucous otitis media. It is true that ICD code especially regarding to H65.2 and H65.3 codes, relies heavily on subjective impression. Knowing the standards of clinical practice in Finland we can assume that most of the cases got their ICD-10 code based on microscopical otoscopy performed by an otolaryngologist. However, diagnosis may be modified later in some cases. Those patients' ICD-10 codes are available, and their genome has been analyzed. Since we do not have access directly to those patients' medical records it is impossible to answer how precisely diagnosis was placed for each case.

Our strategy to ameliorate selection bias is to use large population-based cohorts that are more representative of the general population.

Conclusions

We revealed distinct genetic associations of ANXA13 and GASL2L2 with secretory OM and mucous OM. In contrast, variants of TBX1 and TNFRSF13B were associated with both subtypes of nonsuppurative otitis media. Immunohistochemistry (IHC) of annexin A13 in the ear in a murine model confirmed abundant expression of annexin A13 in the glandular epithelium of the eustachian tube, supporting the role of annexin A13 in the function of the eustachian tube via regulation of phospholipid metabolism. However, further studies are needed to explore the potential pathophysiological role of annexin A13 in chronic secretory otitis media and the potential association of surfactant with annexin A13.

Data availability

FinnGen results are subjected to one year embargo and, after that, available to the larger scientific community. The datasets generated and/or analysed during the current study are available in the Finngen repository (Pheweb browser or through data download), https://www.finngen.fi/en/access_results; https://r10.finngen.fi/ and https://finngen.gitbook.io/documentation/data-download.

Received: 24 March 2024; Accepted: 29 July 2024
Published online: 07 August 2024

References

1. Williamson, I. G., Dunleavey, J., Bain, J. & Robinson, D. The natural history of otitis media with effusion–a three-year study of the incidence and prevalence of abnormal tympanograms in four South West Hampshire infant and first schools. *J. Laryngol. Otol.* **108**(11), 930–934. https://doi.org/10.1017/s0022215100128567 (1994).
2. Midgley, E. J., Dewey, C., Pryce, K. & Maw, A. R. The frequency of otitis media with effusion in British preschool children: A guide for treatment. ALSPAC Study Team. *Clin Otolaryngol. Allied Sci.* **25**(6), 485–91. https://doi.org/10.1046/j.1365-2273.2000.00360.x (2000).
3. Rosenfeld, R. M. *et al.* Clinical practice guideline: Otitis media with effusion (update). *Otolaryngol. Head Neck Surg.* **154**(1 Suppl), S1–S41. https://doi.org/10.1177/0194599815623467 (2016).
4. Rye, M. S., Blackwell, J. M. & Jamieson, S. E. Genetic susceptibility to otitis media in childhood. *Laryngoscope* **122**(3), 665–675. https://doi.org/10.1002/lary.22506 (2012).
5. Schilder, A. G. *et al.* Otitis media. *Nat. Rev. Dis. Primers* **2**(1), 16063. https://doi.org/10.1038/nrdp.2016.63 (2016).
6. Duah, V. *et al.* Younger patients with COME are more likely to have mucoid middle ear fluid containing mucin MUC5B. *Int. J. Pediatr. Otorhinolaryngol.* **90**, 133–137. https://doi.org/10.1016/j.ijporl.2016.09.009 (2016).
7. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era–concepts and misconceptions. *Nat. Rev. Genet.* **9**(4), 255–266. https://doi.org/10.1038/nrg2322 (2008).
8. Casselbrant, M. L. *et al.* The heritability of otitis media: A twin and triplet study. *JAMA* **282**(22), 2125–2130. https://doi.org/10.1001/jama.282.22.2125 (1999).
9. Casselbrant, M. L. *et al.* The genetic component of middle ear disease in the first 5 years of life. *Arch. Otolaryngol. Head Neck Surg.* **130**(3), 273–278. https://doi.org/10.1001/archotol.130.3.273 (2004).
10. Kvaerner, K. J. *et al.* Distribution and heritability of recurrent ear infections. *Ann. Otol. Rhinol. Laryngol.* **106**(8), 624–632. https://doi.org/10.1177/000348949710600802 (1997).
11. Kvestad, E. *et al.* Otitis media: genetic factors and sex differences. *Twin Res.* **7**(3), 239–244. https://doi.org/10.1375/136905204774200514 (2004).
12. Rye, M. S. *et al.* Genome-wide association study to identify the genetic determinants of otitis media susceptibility in childhood. *PLoS One* **7**(10), e48215. https://doi.org/10.1371/journal.pone.0048215 (2012).
13. Jiang, L. *et al.* Genome-wide association analyses of common infections in a large practice-based biobank. *BMC Genom.* **23**(1), 672. https://doi.org/10.1186/s12864-022-08888-9 (2022).
14. Tian, C. *et al.* Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat. Commun.* **8**(1), 599. https://doi.org/10.1038/s41467-017-00257-5 (2017).
15. Kurki, M. I. *et al.* FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**(7944), 508–518. https://doi.org/10.1038/s41586-022-05473-8 (2023).
16. Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* https://doi.org/10.1002/0471142905.hg0720s76 (2013).
17. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**(10), 1415–1424. https://doi.org/10.1038/s41588-021-00931-x (2021).
18. Locke AE et al. Exome sequencing of Finnish isolates enhances rare-variant association power. Nature. 2019;572(7769):323–328. https://doi.org/10.1038/s41586-019-1457-z. Epub 2019 Jul 31. Erratum in: Nature. 2019;575(7783): E4.
19. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**(3), 310–315. https://doi.org/10.1038/ng.2892 (2014).
20. Shrine N. et al. Multiancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. Nat Genet. 2023;55(3):410–422. https://doi.org/10.1038/s41588-023-01314-0. Epub 2023 Mar 13. Erratum in: Nat Genet. 2023;55(10):1778–1779. PMID: 36914875; PMCID: PMC10011137.
21. Saarentaus, E. C. *et al.* Inflammatory and infectious upper respiratory diseases associate with 41 genomic loci and type 2 inflammation. *Nat. Commun.* **14**(1), 83. https://doi.org/10.1038/s41467-022-33626-w (2023).
22. Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**(1), D988–D995. https://doi.org/10.1093/nar/gkab1049 (2022).
23. Schubach, M., Maass, T., Nazaretyan, L., Röner, S. & Kircher, M. CADD v1.7: Using protein language models, regulatory CNNs and other nucleotide-level scores to improve genome-wide variant predictions. *Nucleic Acids Res.* **52**(D1), D1143–D1154. https://doi.org/10.1093/nar/gkad989 (2024).
24. Vuckovic, D. *et al.* The polygenic and monogenic basis of blood traits and diseases. *Cell* **182**(5), 1214-1231.e11. https://doi.org/10.1016/j.cell.2020.08.008 (2020).
25. Thul, P. J. & Lindskog, C. The human protein atlas: A spatial map of the human proteome. *Protein Sci.* **27**(1), 233–244. https://doi.org/10.1002/pro.3307 (2018).
26. Lizarbe, M. A., Barrasa, J. I., Olmo, N., Gavilanes, F. & Turnay, J. Annexin-phospholipid interactions. Functional implications. *Int. J. Mol. Sci.* **14**(2), 2652–2683. https://doi.org/10.3390/ijms14022652 (2013).
27. Torkko, J. M., Manninen, A., Schuck, S. & Simons, K. Depletion of apical transport proteins perturbs epithelial cyst formation and ciliogenesis. *J. Cell Sci.* **121**(Pt 8), 1193–1203. https://doi.org/10.1242/jcs.015495 (2008).
28. Hills, B. A. Surface tension versus straight adhesion in the Eustachian tube. *Otol. Neurotol.* **23**(4), 620–621. https://doi.org/10.1097/00129492-200207000-00040 (2002).
29. Lecat, S. *et al.* Different properties of two isoforms of annexin XIII in MDCK cells. *J. Cell Sci.* **113**(Pt 14), 2607–2618. https://doi.org/10.1242/jcs.113.14.2607 (2000).
30. Lafont, F. *et al.* Annexin XIIIb associates with lipid microdomains to function in apical delivery. *J. Cell Biol.* **142**(6), 1413–1427. https://doi.org/10.1083/jcb.142.6.1413 (1998).
31. Zhou et al. Efficiently Controlling for Case—Control Imbalance and Sample Relatedness in Large-Scale Genetic Association Studies.
32. Castigli, E. *et al.* TACI is mutant in common variable immunodeficiency and IgA deficiency. *Nat. Genet.* **37**(8), 829–834. https://doi.org/10.1038/ng1601 (2005).
33. Salzer, U. *et al.* Mutations in TNFRSF13B encoding TACI are associated with common variable immunodeficiency in humans. *Nat. Genet.* **37**(8), 820–828. https://doi.org/10.1038/ng1600 (2005).
34. Martinez-Gallo, M. *et al.* TACI mutations and impaired B-cell function in subjects with CVID and healthy heterozygotes. *J. Allergy Clin. Immunol.* **131**(2), 468–476. https://doi.org/10.1016/j.jaci.2012.10.029 (2013).
35. Dong, X. *et al.* Phenotypic and clinical heterogeneity associated with monoallelic TNFRSF13B-A181E mutations in common variable immunodeficiency. *Hum. Immunol.* **71**(5), 505–511. https://doi.org/10.1016/j.humimm.2010.02.002 (2010).
36. Landrum, M. J. *et al.* ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**(D1), D1062–D1067. https://doi.org/10.1093/nar/gkx1153 (2018).
37. Landrum, M. J. *et al.* improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**(D1), D1062–D1067. https://doi.org/10.1093/nar/gkx1153 (2018).
38. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection, and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48(7):709–17. https://doi.org/10.1038/ng.3570. Epub 2016 May 16. Erratum in: Nat. Genet. 2016;48(10):1296.

39. Lee, J. J. *et al.* The murine equivalent of the A181E TACI mutation associated with common variable immunodeficiency severely impairs B-cell function. *Blood* **114**(11), 2254–2262. https://doi.org/10.1182/blood-2008-11-189720 (2009).
40. Fried, A. J. *et al.* Functional analysis of transmembrane activator and calcium-modulating cyclophilin ligand interactor (TACI) mutations associated with common variable immunodeficiency. *J. Allergy Clin. Immunol.* **128**(1), 226-228.e1. https://doi.org/10.1016/j.jaci.2011.01.048 (2011).
41. Feng, G. *et al.* A novel homozygous variant in GAS2L2 in two sisters with primary ciliary dyskinesia. *Intern. Med.* **61**(18), 2765–2769. https://doi.org/10.2169/internalmedicine.8884-21 (2022).
42. Bustamante-Marin, X. M. *et al.* Lack of GAS2L2 causes PCD by impairing cilia orientation and mucociliary clearance. *Am. J. Hum. Genet.* **104**(2), 229–245. https://doi.org/10.1016/j.ajhg.2018.12.009 (2019).
43. Stroud, M. J. *et al.* GAS2-like proteins mediate communication between microtubules and actin through interactions with end-binding proteins. *J. Cell Sci.* **127**(Pt 12), 2672–82. https://doi.org/10.1242/jcs.140558 (2014).
44. Ng, P. C. & Henikoff, S. Predicting deleterious amino acid substitutions. *Genome Res.* **11**(5), 863–874. https://doi.org/10.1101/gr.176601 (2001).
45. Lim, D. J. Functional morphology of the mucosa of the middle ear and Eustachian tube. *Ann. Otol. Rhinol. Laryngol.* **85**(2 Suppl 25 Pt 2), 36–43. https://doi.org/10.1177/00034894760850S209 (1976).
46. Raft, S., Nowotschin, S., Liao, J. & Morrow, B. E. Suppression of neural fate and control of inner ear morphogenesis by Tbx1. *Development* **131**(8), 1801–1812. https://doi.org/10.1242/dev.01067 (2004).
47. Vitelli, F. *et al.* TBX1 is required for inner ear morphogenesis. *Hum. Mol. Genet.* **12**(16), 2041–2048. https://doi.org/10.1093/hmg/ddg216 (2003).
48. Tian, C. & Johnson, K. R. TBX1 is required for normal stria vascularis and semicircular canal development. *Dev. Biol.* **457**(1), 91–103. https://doi.org/10.1016/j.ydbio.2019.09.013 (2020).
49. Wojas, O. *et al.* Co-occurrence of otitis media with effusion and another environment-dependent disease (selected allergic conditions). *Postepy Dermatol. Alergol.* **41**(1), 78–84. https://doi.org/10.5114/ada.2023.135602 (2024).

Acknowledgements

We would like to acknowledge the participants and investigators of the FinnGen study (see supplement). The FinnGen project was funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie, Inc.; AstraZeneca, UK; Biogen MA, Inc.; Bristol Myers Squibb (and Celgene Corporation & Celgene International II Sàrl); Genentech, Inc.; Merck Sharp & Dohme LCC; Pfizer, Inc.; GlaxoSmithKline Intellectual Property Development, Ltd.; Sanofi US Services, Inc.; Maze Therapeutics, Inc.; Janssen Biotech, Inc.; Novartis AG; and Boehringer Ingelheim International GmbH. The following biobanks are acknowledged for delivering biobank samples to the FinnGen project: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki/fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta), and Terveystalo Biobank (www.tervealo.com All Finnish Biobanks are members of the BBMRI.fi infrastructure (www.bbmri.fi). The Finnish Biobank Cooperative FINBB (https://finbb.fi/) is the coordinator of BBMRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious® services (https://site.fingenious.fi/en/) managed by FINBB.

Author contributions

Argyro Bizaki-Vallaskangas contributed to the conception of the work , prepared analyses of the data, wrote manuscript and prepared the figures and tables based on the results Joel Rämö contributed with interpretation of data and preparation of the manuscript Eeva Sliz prepared analyses of the data and interpreted the results of GWAS analysis Ilkka Kivekäs contributed with reviewing of the literature and with interpretation of data Tytti Willberg contributed to the preparation of the manuscript and the review of literature Elmo Saarentaus contributed with interpretation of data Sanna Toppila-Salmi and Aarno Dietz interpreted the results of GWAS analysis and contributed to the design of the study Teppo Haapaniemi reviewed and interpreted the results of immunohistochemistry Vesa P. Hytönen reviewed and interpreted the results of GWAS analysis. Sari Toivola performed the immunohistochemistry Antti Mäkitie, Aarno Palotie and Johannes Kettunen made substantial contributions to the conception or design of the work All authors reviewed the manuscript.

Funding

This project was funding by the Finnish ORL-HNS Foundation (Korvatautien tutkimussäätiö).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-68781-1.

Correspondence and requests for materials should be addressed to A.B.-V.