

Genetic overlap between Alzheimer's disease and immune-mediated diseases: An atlas of shared genetic determinants and biological convergence

Brisa Fernandes

brisasf@gmail.com

University of Texas Health Center at Houston <https://orcid.org/0000-0002-3797-7582>

Nitesh Enduru

<https://orcid.org/0000-0002-0255-706X>

Brisa Fernandes

University of Texas Health Center at Houston

Shahram Bahrami

Yulin Dai

The University of Texas Health Science Center at Houston <https://orcid.org/0000-0003-1874-7893>

Ole Andreassen

Oslo University Hospital & Institute of Clinical Medicine, University of Oslo <https://orcid.org/0000-0002-4461-3568>

Zhongming Zhao


The University of Texas Health Science Center at Houston <https://orcid.org/0000-0002-3477-0914>

Article

Keywords: Alzheimer's disease, immune-mediated diseases, genome-wide association studies, pleiotropy, Mendelian randomization, genetic correlation

Posted Date: September 28th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3346282/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: OAA is a consultant to Cortechs.ai, and has received speaker's honoraria from Lundbeck, Janssen and Sunovion. NE, BSF, SB, YD, and ZZ have nothing to disclose.

Version of Record: A version of this preprint was published at Molecular Psychiatry on March 18th, 2024. See the published version at <https://doi.org/10.1038/s41380-024-02510-y>.

Abstract

The occurrence of immune disease comorbidities in Alzheimer's disease (AD) has been observed in both epidemiological and molecular studies, suggesting a neuroinflammatory basis in AD. However, their shared genetic components have not been systematically studied. Here, we composed an atlas of the shared genetic associations between 11 immune-mediated diseases and AD by analyzing genome-wide association studies (GWAS) summary statistics. Our results unveiled a significant genetic overlap between AD and 11 individual immune-mediated diseases despite negligible genetic correlations, suggesting a complex shared genetic architecture distributed across the genome. The shared loci between AD and immune-mediated diseases implicated several genes, including *GRAMD1B*, *FUT2*, *ADAMTS4*, *HBEGF*, *WNT3*, *TSPAN14*, *DHODH*, *ABCB9* and *TNIP1*, all of which are protein-coding genes and thus potential drug targets. Top biological pathways enriched with these identified shared genes were related to the immune system and cell adhesion. In addition, *in silico* single-cell analyses showed enrichment of immune and brain cells, including neurons and microglia. In summary, our results suggest a genetic relationship between AD and the 11 immune-mediated diseases, pinpointing the existence of a shared however non-causal genetic basis. These identified protein-coding genes have the potential to serve as a novel path to therapeutic interventions for both AD and immune-mediated diseases and their comorbidities.

Introduction

Alzheimer's disease (AD) is a neurodegenerative brain disease that affects cognition, behavior and social skills¹. Twin studies have shown AD to be a highly heritable disease². As the older population is increasing, AD is becoming an emerging global public health challenge in our century, with an estimated 6.2 million Americans aged 65 and older living with AD. This number is projected to increase to 13.8 million by 2060¹. An increasing body of evidence indicates that the pathogenesis of AD is not confined to the neurons but instead that neuroinflammation plays a significant role in the disease, with an interplay between the brain and the immune system³. In particular, the microglia, the brain-resident immune cells that contribute to the regulation of the brain immune response, is considered to have a pivotal role in the pathophysiology of AD, as a core alteration in AD^{4,5} or at least as a particular endophenotype in what is thought to be a heterogeneous disease⁶. Other studies pinpoint to alterations in peripheral immune cells in AD and a role for innate immune cells in the development of AD^{7,8}.

Chronic inflammation, which is often associated with immune-mediated diseases, has been implicated in the development and progression of AD. The comorbidities between immune diseases, showing immune system dysregulation, and AD have been widely reported. Epidemiological studies show that the overall incidence of AD is higher among individuals with inflammatory bowel disease, other inflammatory polyarthropathies and systematic connective tissue disorders, psoriasis, rheumatoid arthritis (RA), or multiple sclerosis compared to the age and sex-matched comparison groups free from inflammatory diseases⁹⁻¹¹. A previous study investigated the pleiotropy between immune-mediated diseases and AD¹². However, it is still unclear to what extent a multitude of immune-mediated diseases and AD share genetic underpinnings, and now larger and more powerful GWAS data are available.

In this study, we aimed to comprehensively investigate the genetic relationship, if any, between immune-mediated diseases and AD. We hypothesized that the genetic determinants contributing to AD overlap with those contributing to these immune-mediated diseases. To achieve this goal, we probed the shared genetic architecture between AD and a plethora of immune-mediated diseases. Specifically, we collected the genome-wide association studies (GWAS) summary statistics data from AD and 11 immune-mediated diseases: allergic rhinitis (AR), asthma, atopic dermatitis (AtD), celiac disease (CeD), Crohn's disease (CD), hypothyroidism, primary sclerosing cholangitis (PSC), RA, systemic lupus erythematosus (SLE), ulcerative colitis (UC), and vitiligo. We identified the genes linked to the associated single nucleotide polymorphisms (SNPs) for each of the 11 immune-mediated diseases and AD. Moreover, we characterized the shared loci and their implications on genes, tissues, cell types, and biological functions between immune-mediated diseases and AD. Lastly, we ascertained their possible causal relationship, shedding light on the intricate interplay between these immune-mediated diseases and AD.

Methods

GWAS data sets used

For AD, we obtained GWAS summary statistics data from Wightman *et al.*, which included only diagnosed AD individuals (n = 398,058)¹³. We used GWAS data with AD cases only, excluding AD proxies from UK Biobank, and such data was made available by the authors upon request. GWAS summary statistics data for AR (n = 289,307)¹⁴, asthma (n = 385,822)¹⁴, atopic dermatitis (AtD) (n = 40,835)¹⁵, celiac disease (CeD) (n = 15,283)¹⁶, Crohn's disease (CD) (n = 40,266)¹⁷, hypothyroidism (n = 244,890)¹⁴, primary sclerosing cholangitis (PSC) (n = 309,154)¹⁸, rheumatoid arthritis (RA) (n = 58,284)¹⁹, systemic lupus erythematosus (SLE) (n = 23,210)²⁰, ulcerative colitis (UC) (n = 45,975)¹⁷, and vitiligo (n = 44,266)²¹ were obtained. Moreover, GWAS for hypertrophic cardiomyopathy (HCM)²² (n = 309,154) was obtained to serve as negative control for AD, considering that HCM is limited to heart pathology. All individuals were of European ancestry (EA). All summary statistics were built based on human reference genome GRCh37. There was no sample overlap between the AD GWAS and those immune-mediated diseases or the HCM GWAS datasets. Full details regarding the sample size of cases and controls for each GWAS is given in **Table S1**.

Linkage disequilibrium score regression between AD and immune-mediated diseases

We calculated the genetic correlation between AD and 11 immune-mediated diseases using the cross-trait linkage disequilibrium score regression (LDSR) method²³. Pre-estimated linkage disequilibrium (LD) scores provided by LDSC developers were obtained from the 1,000 Genomes Project (1kGP) European reference population; we preprocessed the summary statistics using LDSC `munge_sumstats.py` and used the odds ratio (OR) as the signed summary statistic. We calculated the genetic correlation by employing HapMap3 SNPs only with LD reference panel SNPs to minimize potential bias due to differences in LD structure.

Genetic overlap between AD and immune-mediated diseases

To complement the genetic correlation analysis, we used the statistical tool MiXeR (<https://github.com/precimed/mixer>). MiXeR quantifies polygenic overlap irrespective of genetic correlation using GWAS summary statistics, and is agnostic to effect directions²⁴. This method estimates the total number of both shared and trait-specific causal variants (i.e., variants with nonzero additive genetic effects on a given trait). We applied MiXeR to verify genetic overlap between AD and each immune-mediated disease. To evaluate model fit (the ability of the MiXeR model to predict the actual GWAS data), we constructed modeled vs. actual conditional quantile-quantile (Q-Q) plots, log-likelihood plots, and calculated the Akaike information criterion (AIC). A negative AIC value indicates that the MiXeR model cannot be adequately differentiated from a scenario of maximum possible overlap and a scenario of minimum overlap, while a positive AIC value indicates adequate fit for the MiXeR analyses. The major histocompatibility complex (MHC) region was excluded from the MiXeR analyses for all the considered traits.

Conditional Q-Q plots to assess pleiotropic enrichment

We created conditional Q-Q plots to assess the cross-trait genetic enrichment, conditioning AD on the 11 immune-mediated diseases and one negative control, and we did this analysis inversely as well. The genetic enrichment can be referred to as a leftward shift in the Q-Q curve, indicating a considerable number of SNPs with their p -values greater than or equal to a given threshold. Q-Q plot is used to visualize the quantile distribution of the statistics for the observed result with respect to the expected quantile distribution under the null hypothesis, where the theoretical distribution is uniform on the interval [0,1]. In this case, enrichment is considered to occur when an increased proportion of SNPs associated with one phenotype or disease (AD) is observed as a function of the strength of the association with a second phenotype or disease, and enrichment is considered to occur in case the degree of the shift of the Q-Q curve to the left from the expected null line for the first disease is dependent on the magnitude of the association with the second disease²⁵.

Conditional and conjunctive false discovery rate analysis

Conditional false discovery rate (CondFDR) analysis was used to detect SNPs associated with AD given their association with each immune-mediated disease, and *vice versa*. CondFDR is defined as “the posterior probability that a given SNP is null (i.e., no SNPs associated with the first trait) for the primary trait (i.e., AD) given that the p -values for both traits are as small or smaller than the observed p -values”²⁵. ConjFDR is defined as “the posterior probability that a given SNP is null for either phenotype or both simultaneously, given the p -values for both traits are as small or smaller than the observed p -values”^{25,26}. The enrichment verified in the conditional Q-Q plots is shown for each individual SNP according to the FDR. In this project, we used the conjFDR analysis with the threshold value 0.05 to identify the shared loci between AD and each of the 11 immune-mediated diseases. To examine the chromosomal location of SNPs in combination with AD and these immune-mediated diseases, we constructed a Manhattan plot using the bioinfokit package²⁷.

Local genetic correlation analysis to assess local genetic correlation

We used Local Analysis of covariant Association (LAVA) to further calculate the local genetic correlation between AD and immune-mediated diseases across the pleiotropic SNPs identified from conjFDR analysis. Using the pre-defined 2,495 genomic regions (2,492 after excluding the *APOE* region) provided by the LAVA developers, which minimize the local LD in those blocks²⁸, the pleiotropic SNPs from conjFDR results were mapped and grouped into LAVA genomic regions. The genomic partition was performed using the 1kGP European individuals. To identify the presence of genetic signals, we initially performed a univariate test to estimate local heritability across each phenotype with a p -value < 0.05 being considered as a significant locus to evaluate for the bivariate local genetic correlation. For the loci with significant heritability in both disorders, we calculated the bivariate local genetic correlation. We adjusted the p -values using Bonferroni correction between AD and each immune-mediated disease based on the number of genomic loci tested in local univariate and bivariate analyses.

Enrichment and assessment of SNP novelty

SNPs within the MHC region (defined as chr6: 25,119,106 – 33,854,733), the 8p23 inversion (chr8:7,200,000–12,500,000), and apolipoprotein E (*APOE*) gene (chr19:44,000,000–47,000,000) were excluded from the analyses due to known association to AD^{29,30} or a very complex LD structure^{31,32}. For the conjFDR analysis, we used a conjFDR p -value < 0.05 as statistically significant. A GWAS p -value ($p_{AD \& ID}$) > 5×10^{-8} cutoff in the original GWAS was used to identify the shared SNPs between two phenotypes as novel. In other words, to be considered ‘novel’, a SNP identified using the conjFDR should not have been found statistically significant in the original GWAS. For the local genetic correlation, we used a p -value < 0.05 as cutoff in the univariate and bivariate test with Bonferroni corrected p -value for AD and each immune-mediated disease. To check the novelty of the SNPs and loci identified in our study, we examined previously reported GWAS associations in the National Human Genome Research Institute (NHGRI)-EBI GWAS Catalog³³. We further identified if the gene was novel to any of our traits based on the traits reported in the NHGRI-EBI GWAS Catalog³³ and previous AD studies based on conjFDR analysis^{25,34-36}.

Definition of genomic loci

Independent genomic loci were identified based on the FUMA protocol, an online tool for functional mapping of genetic variants (<http://fuma.ctglab.nl/>)³⁷. SNPs with conjFDR < 0.1 were identified as independently significant and LD r^2 < 0.6 with each other. A subset of these independent SNPs (LD r^2 < 0.1) was then considered as the lead SNPs. The boundaries for genomic loci were then determined by identifying all candidate SNPs which were in LD ($r^2 \geq 0.6$) with the lead SNP. Then, loci were combined if they were separated by less than 250 kb. These separate regions, containing all of these candidate SNPs, were treated as a single independent genomic locus. All LD information was calculated based on the EA reference panel from 1kGP³⁸.

Functional annotation and tissue and cell-specificity analyses

We annotated the genes to the lead SNPs based on positional mapping using ANNOtate VARIation (ANNOVAR)³⁹ in FUMA. We mapped all the candidate SNPs with conjFDR value < 0.1 and having an LD $r^2 \geq 0.6$ with one of the independently significant SNPs using ANNOVAR. In addition, these SNPs were annotated with Combined Annotation Dependent Depletion (CADD) scores⁴⁰ to predict how certain the SNP effect is on protein structure or function and possible contribution to genetic disease. Similarly, we used RegulomeDB (RDB version 1.1) scores⁴¹ to predict the likelihood of regulatory functionality, and

chromatin states, to predict transcription and also regulatory effects from chromatin states at the SNP locus. RegulomeDB (version 1.1) is a database that scores SNP's functionality based upon its presence in a DNase hypersensitive site or transcription factor binding site. RDB ranks the SNPs with a score from 1 to 6. The SNP with strong evidence of being a regulatory variant is given the score of 1 and one with the least evidence is scored as 6⁴¹. SNPs with RDB scores of 1 (1a – 1f) were queried for known eQTLs in brain (GTEx v8, Braineac, CMC, PsychENCODE, xQTLServer), blood (GTEx v8, Blood eQTL, BIOS QTL, eQTLGen,) and immune system tissues (GTEx v8, DICE, scRNA eQTLs) along with other lead SNPs. An RDB score of 1 suggests the presence of the variant in multiple data regions such as combination of eQTL + TF binding + matched TF motif + matched DNase Footprint + DNase peak (for more details about the RDB score, see <https://regulomedb.org/regulome-help/>).

In addition, we ranked the tissues implicated by the SNPs using our in-house method DeepFun⁴². The DeepFun web server (<https://bioinfo.uth.edu/deepfun/>) is a convolutional neural network framework that was trained on ~ 8,000 chromatin profiles of 225 tissues or cell types from ENCODE and Roadmap projects. We imported all independent SNPs shared between AD and immune-mediated diseases to query the pretrained model on the web server and compute the SNP Activity Difference (SAD, ranging from - 1 to 1) between SNP alleles, indicating the accessibility or binding probability difference between reference and alternative alleles. We took the mean of absolute SAD of each SNP for each of the 50 unique tissues and cell lines and counted the number of tissues or cell types with the largest absolute SAD for each of the considered SNPs. The tissue and cell lines with most counts were defined as the most related tissue and cell lines to those independent SNPs shared between AD and immune-mediated diseases. We used all the shared SNPs between AD and immune-mediated diseases as the imputed SNPs. Then, we used Web-based Cell-type Specific Enrichment Analysis of Genes (WebCSEA, available at <https://bioinfo.uth.edu/webcsea/>)⁴³ to assess the tissue and cell type enrichment of all the shared SNPs. WebCSEA curated a total of 111 single cell RNA-seq (scRNA-seq) panels of human tissues and 1,355 tissue-cell types (TCs) from 61 different general tissues across 11 human organ systems and uses decoding tissue-specificity (deTS) algorithm to measure the enrichment for each tissue-cell type⁴³. In the WebCSEA analyses, a nominal p -value $< 1 \times 10^{-3}$ was considered statistically significant.

Finally, we used FUMA to analyse the Gene Ontology (GO) terms in three domains (biological processes, molecular functions, and cellular components terms) and biological pathways from KEGG (Kyoto encyclopedia of genes and genomes), Panther BD, Reactome, and Wikipathways. An FDR corrected by Benjamini and Hochberg (BH) procedure⁴⁴ < 0.05 was considered significant for the GO and pathway analyses.

Causal relationship between AD and immune-mediated diseases

We evaluated the potential causal relationship between AD and immune-mediated diseases using Mendelian randomization (MR) analysis. MR uses SNPs as instrument variables (IV) to assess causality and these IVs are defined by three assumptions⁴⁵. First, the selected IVs are significantly associated (p -value_{GWAS} $< 5 \times 10^{-8}$) with the exposure variable. Second, the IVs are independent between the exposure and the outcome. Third, the effect of IVs on the outcome must precede through the exposure. We used the two-sample MR (2SMR) method (<https://mrcieu.github.io/TwoSampleMR/articles/introduction.html>). Initially, independent ($r^2 < 0.001$) genome-wide significant SNPs (p -value_{GWAS} $< 5 \times 10^{-8}$) associated with the exposure (immune-mediated diseases) were considered as IVs and assessed against outcome variables (AD). We used the five MR methods [MR Egger, Weighted median, Inverse variance weighted (IVW), Simple mode, and Weighted mode] implemented in the 2SMR R package⁴⁶.

Results

The general workflow of our study is illustrated in Fig. 1. We developed a pipeline to perform the global and local genetic correlations between AD and 11 immune-mediated diseases and to quantify the shared loci among them using MiXeR. We used the conditional/comjunctive FDR to identify the shared loci and, finally, performed bi-directional MR to assess the causal relationship between AD and the 11 immune-mediated diseases.

Genetic correlation and genetic overlap beyond genetic correlation

We computed the genetic correlation based on the LDSR method. In general, there was no significant genetic correlation between AD and AR ($r_g = 0.043$, standard error (SE) = 0.08, $p = 0.6$), AD and asthma ($r_g = 0.005$, SE = 0.06, $p = 0.9$), AD and AtD ($r_g = -0.023$, SE = 0.11, $p = 0.8$), AD and CeD ($r_g = -0.011$, SE = 0.12, $p = 0.9$), AD and CD ($r_g = 0.0006$, SE = 0.02, $p = 0.9$), AD and hypothyroidism ($r_g = 0.05$, SE = 0.07, $p = 0.5$), AD and PSC ($r_g = -0.09$, SE = 0.07, $p = 0.2$), AD and RA ($r_g = -0.12$, SE = 0.07, $p = 0.1$), AD and SLE ($r_g = -0.05$, SE = 0.1, $p = 0.6$), AD and UC ($r_g = 0.03$, SE = 0.02, $p = 0.2$), and AD and vitiligo ($r_g = 0.03$, SE = 0.08, $p = 0.7$). In addition, no genetic correlation was found between AD and the negative control HCM ($r_g = 0.14$, SE = 0.18, $p = 0.42$) (Table S2).

To capture mixtures of effect directions across shared genetic variants, we performed MiXeR analysis with AD and immune-mediated diseases to verify genetic overlap beyond genetic correlation by determining the number of overlapping variants between trait pairs. The observed small genetic overlap in the Venn diagrams with nonexistent genetic correlations suggests a balanced mixture of concordant and discordant genetic effects across shared loci between AD and immune-mediated diseases, with adequate quality of model fit as suggested by positive AIC values. The predicted number of shared causal loci between AD and immune-mediated diseases varied from 6 (between AD and PSC; AD and SLE) to 40 (AD and AR). There was complete genetic overlap between AD and the negative control HCM. However, the AIC values were all negative, indicating a poor model fit and, thus, this particular analysis is unreliable (Fig. 2, Figure S1, and Table S3).

Pleiotropic enrichment

The enrichment observed in QQ-plots can be translated to FDR for each SNP. The genetic enrichment can be identified from the Q-Q plot with a leftward shift in the Q-Q curve, indicating a considerable number of SNPs with p -value greater than or equal to a given threshold. There was an enrichment of associations with AD given increasing SNP associations with each immune-mediated disease and conversely in the conditional Q-Q-plots, with a leftward shift of decreasing

values of empirical $-\log_{10}(q_{AD})$, showing polygenic overlap between AD and each immune-mediated disease across common genetic variants. No enrichment was verified between AD and the negative control HCM (Fig. 3, **Figure S1**).

Shared and novel loci for AD and immune-mediated diseases

We employed conjFDR to ascertain SNPs mutually associated with AD and each immune-mediated disease; in order to determine the allelic direction of effects in the diseases, we used the z-scores from the original GWAS. Considering a conjFDR < 0.05, we identified 76 shared genomic loci (**Table S4**), of which 6 loci were jointly associated with more than one pair of diseases. Five of the 6 multi-trait shared loci were identified as novel loci for AD (**Table S5**). After excluding these 6 overlapping SNPs, we identified 70 unique shared genomic loci (71 unique SNPs) between AD and immune-mediated diseases, out of which 38 genomic loci were novel for AD only and 25 genomic loci were novel for immune-mediated diseases only (Tables 1 and S4); 17 genomic loci were identified as novel to both AD and immune-mediated diseases (Table 2). We then constructed conjFDR Manhattan plots to visualize all the novel and all shared loci between these diseases (**Figure S2**), and, based on the z-scores, we evaluated the directionality of allelic effects in the loci shared between AD and the immune-mediated diseases. All the shared loci had a mixed direction of effect, except for AtD and AD, which had complete concordance effect (Tables 1 and S4). The shared loci (1q23.3, 7p15.1, 22q13.2, 1q32.1, 4p14, 6q23.3) and their respective genes (*ADAMTS4*, *JAZF1*, *Y_RNA*, *C1orf106*, *AC195454.1*, and *RP11-95M15.1*) were mapped to more than one immune-mediated disease. No shared loci were identified between AD and the negative control HCM (**Figures S2 and 4 and Table S5**).

Table 1
Summary of total loci shared between Alzheimer's disease (AD) and immune-mediated diseases (ID). AD = Alzheimer's disease. ID = immune-mediated disease.

Immune-mediated disease (ID)	# loci	# novel AD loci	Concordance effect (%) for AD novel loci	# novel ID loci	Concordance effect (%) for novel ID loci	# novel loci for AD & ID	Concordance effect (%) for novel loci
Allergic rhinitis	4	3	66.7	2	100	2	100
Asthma	8	2	50	3	33.3	1	100
Atopic dermatitis	3	1	100	2	100	1	100
Celiac disease	3	1	100	1	100	0	NA
Crohn's disease	22	11	36.4	6	16.7	5	20
Hypothyroidism	3	2	50	0	0	0	NA
Primary sclerosing cholangitis	6	2	50	3	66.7	0	NA
Rheumatoid arthritis	7	4	25	1	100	1	100
Systemic lupus erythematosus	9	5	20	3	33.3	3	33.3
Ulcerative colitis	6	4	50	3	33.3	3	33.3
Vitiligo	5	3	0	1	0	1	0
Overall	76	38	49.8	25	53	17	60.8

Table 2
Novel loci associated with AD and Immune-mediated diseases based on conjFDR analysis.

Trait1	Trait2	Locus	Chr	SNP	Pos	Nearest Gene	Distance	Function	A1/A2	zscore_Trait1	zscore_Trait2	cor
AD	AR	1	7	rs17171172	37872650	GPR141, EPDR1*	0:00	intronic	T/C	3.9	4.1	1.9
AD	AR	2	9	rs10993704	93597359	SYK	0	intronic	T/C	3.7	3.8	4.7
AD	Asthma	3	5	rs2028258	110553497	AC010468.3	8451	intergenic	A/G	4.3	3.9	2.4
AD	AtD	4	8	rs897150	126620703	RP11-136012.2	23086	intergenic	A/G	4.3	4.4	8.5
AD	CD	5	5	rs2963769	114632469	CCDC112*	0	intronic	T/C	3.7	-3.8	3.6
AD	CD	6	8	rs2923447	42457590	RP11-503E24.3	29511	intergenic	T/G	3.6	-3.7	4.7
AD	CD	7	10	rs9665686	82302047	SH2DB4	0	Intronic	T/C	3.7	3.4	4.7
AD	CD	8	14	rs45441198	104010198	TRMT61A	0	UTR3	T/C	4.0	-4.8	1.3
AD	CD	9	16	rs3764310	72073122	DHODH	0	intronic	T/C	-3.8	4.6	2.7
AD	RA	10	11	rs12790721	123385425	LINC01059	12142	intergenic	T/C	3.8	3.5	4.3
AD	SLE	11	2	rs62168048	135152287	MGAT5 ^A	0	intronic	A/G	-4.1	4.1	1.2
AD	SLE	12	4	rs16890310	14378658	AC006296.1	17952	intergenic	A/G	-3.7	4.1	4.6
AD	SLE	13	10	rs4948365	61574694	CCDC6	0	intronic	T/C	3.7	3.8	4.2
AD	UC	14	1	rs148784768	20242739	OTUD3 ^A	22669	intergenic	A/G	3.9	-5.1	3.7
AD	UC	15	11	rs117197442	100142344	CNTN5	0	intronic	T/C	-3.8	4.2	3.9
AD	UC	16	16	rs11644964	23962293	PRKCB ^A	0	intronic	A/G	3.9	4.2	3.6
AD	Vitiligo	17	5	rs62380789	141050575	ARAP3	0	intronic	A/C	4.0	-4.4	2.5

Local genetic correlation analysis across the LAVA predefined genomic loci and conjFDR results

Regional genetic correlations may be masked when r_g is assessed on a genome-wide level. To investigate whether there are any genomic loci with pronounced genetic correlations despite negligible genome-wide r_g , we estimated regional genetic correlations using LAVA. Regional genetic correlations are a better tool to capture genetic associations with mixed effect directions. That is, a pair of traits may exhibit no global genetic correlations as a result of an equal number of positive and negative (opposite effect directions) local genetic correlations of similar magnitude. We initially performed LAVA on the pre-defined 2,492 genomic loci to calculate the pairwise local genetic correlation across AD and 11 immune-mediated diseases. We used a nominal p -value < 0.05 to detect the univariate signals and used the Bonferroni corrected p -value between AD and each of the 11 immune-mediated diseases for bivariate analysis. The LAVA results were summarized in **Table S6**. The overall average concordant effect in univariate analysis was 7%, and 46% of these loci had 95% confidence intervals (CIs) for the variance that included 1. The overall average concordant effect in bivariate analysis was 53%, and 81% of these loci had 95% CIs for the variance that included 1. The overall average concordant effect in Bonferroni adjusted bivariate analysis was 56%, and 87% of these loci had 95% CIs for the variance that included 1. Overall 1,765 unique loci found to univariately heritable (**Table S8**) and 745 unique loci found significant for bivariate analysis at p -value < 0.05 for AD and 11 immune-mediated diseases (**Table S9**) and 220 unique loci are shared between more than one immune-mediated disease with AD (**Table S10**). After adjusting for Bonferroni correction between AD and each of the 11 immune-mediated disease, we found 39 genomic loci shared between AD and 11 immune-mediated diseases (**Table S11**).

Based on the conjFDR analysis results, we further performed local genetic correlation using LAVA on the 17 novel and 59 previously reported shared loci (6 of these loci are shared between multiple pair of traits, see **Table S5**) (**Tables S4 and S7**) identified between AD and immune-mediated diseases. Among the 17 novel loci, 11 were significantly heritable at univariate p -value < 0.05 , and 3 were bivariate significant at p -value < 0.05 and 5 of those loci had 95% confidence

intervals (CIs) for the variance that included 1 (**Table S12**). After adjusting for Bonferroni correction, 2 out of 3 loci were statistically significant for correlation. Thirty-four of the 59 loci had significant heritability for both diseases, and 7 of those loci had a bivariate p -value < 0.05 ; 7 of those loci had 95% CIs for the variance that included 1, suggesting a shared signal between the diseases and that the local genetic signal of those phenotypes is completely shared even though they are not statistically significant (**Tables S7 and S12**). After adjustment using Bonferroni correction, 6 out of 7 loci were statistically significant for correlation. Overall, 8 out of 76 loci had a Bonferroni adjusted bivariate p -value < 0.05 . These results provide further evidence of pleiotropy with mixed effect direction between AD and immune-mediated diseases with 3 positively correlated and 5 negatively correlated at Bonferroni adjusted bivariate p -value < 0.05 . We found only one genomic locus (chr 2:64696090–65938002) shared between LAVA and ConjFDR results.

Exploration of shared biological mechanisms: Functional annotation and tissue and cell-specificity analyses

Functional annotation for all the lead SNPs at conjFDR < 0.05 within the loci shared between AD and immune-mediated diseases showed that the majority of the lead SNPs are intronic or intergenic (**Table S4**). Two lead SNPs were reported to have a CADD score > 12.37 , suggesting deleteriousness (rs12790721 and rs602662, *GRAMD1B* and *FUT2*, respectively)⁴⁰. Seven lead SNPs reported a low RegulomeDB scores of 1d or 1f, suggesting regulatory functionality (rs4233366, rs739954, rs2074404, rs10748526, rs3764310, rs4275659 and rs4958435, nearest genes *ADAMTS4*, *HBEGF*, *WNT3*, *TSPAN14*, *DHODH*, *ABCB9* and *TNIP1*, respectively)⁴¹ (**Table S13**). These SNPs were associated with eQTL functionality in different brain, blood, and immune system regions. For a few examples, we found eQTL functionality for genes *NDUFS2*, *PCP4L*, *PCDHA8*, *SRA1*, *PCDHA10*, *VTRNA1-3*, *PCDHA3*, *PCDHA6*, *PCDHB15*, *WDR55*, *CYSTM1* and 739 unique genes (**Table S14**). Functional annotation for all SNPs associated with lead SNPs at conjFDR < 0.1 and $r^2 > 0.6$ were extracted from GWAS summary statistics, between AD and immune-mediated diseases. The majority of the candidate SNPs are either intronic or intergenic (**Table S15**). In **Figures S2 and 4**, the shared genes between AD and each specific immune-mediated disease are shown.

Functional annotation of the lead SNPs using positional, eQTL and chromatin interaction mapping was performed to understand the biological mechanism affecting AD and immune-mediated diseases. Based on the three gene mapping techniques, we identified 142 unique credible genes mapped to the lead SNPs between AD and immune-mediated diseases (**Table S16**). The credible genes were mapped to the selected eQTLs available in FUMA. Out of the 70 unique shared loci (71 unique SNPs) identified using the conjFDR method, 50 loci were identified as eQTLs covering 750 unique genes across the different brain, blood, and immune system tissues in the GTEx database (**Table S16**).

To understand the tissues that these SNP loci tend to manifest their impacts, we leveraged the in-house method DeepFun and identified that three out of the top-ranked five tissues were immune and brain tissues (thymus, cerebellum, and frontal cortex) (**Figure S3**). We mapped these SNP loci to the proximity gene using the online tool WebCSEA, and found enrichment in the lymphatic system. There was a signal in the nervous, cardiovascular, and endocrine systems (**Figure S4**); the top enriched cell types were mostly related to the immune and nervous systems (neutrophil, macrophage, monocytes, innate lymphoid cell, plasma cell, excitatory neuron, purkinje cell, microglia, dendritic and glial cells) (**Figure S5**) (nominal significance p -value $< 1 \times 10^{-3}$). Finally, the top biological pathways were related to cell adhesion and the immune system (**Table S17**).

Lack of significant causal association observed between AD and immune-mediated diseases

We evaluated the potential causal relationship between immune-mediated diseases and AD using five MR methods (MR Egger, weighted median, inverse variance weighted (IVW), simple mode, and weighted mode) implemented in the 2SMR package. There was no sufficient evidence for causal relationship between immune-mediated diseases and AD in all these tests, except for asthma which had a significant causal relationship with AD based on weighted median and weighted mode methods (**Table S18**).

Discussion

In this study, we performed a comprehensive assessment of the shared genetic architecture between AD and immune-mediated diseases by analyzing large-scale GWAS summary data using different but complementary genetic approaches in order to shed light into their shared underlying molecular biology mechanisms. The findings show extensive genetic overlap between AD and immune-mediated diseases regardless of genetic correlation. We identified several shared and novel loci using the conjFDR approach and identified several pathways and immunological signatures enriched in the brain and immune system.

In our analyses, LDSR did not show significant genome-wide genetic correlation between AD and each immune-mediated trait. Still, the evidence of significant genetic overlap verified by MiXeR in tandem with nonsignificant genetic correlation reflects shared genetic etiologies with mixed effect directions, a suggestion which is corroborated by the local genetic correlation performed using LAVA; this balanced mixture of concordant and discordant shared loci distributed across the genome indicates that some genetic variants can increase the risk of one disorder while decreasing the risk of the other. In addition, all MR analyses, with the exception of perhaps asthma, gave us no support for a significant causal relationship between AD and immune-mediated diseases, indicating that pleiotropic and common biological pathways may be a better explanation for their association. A similar study showed no significant causal relationship between AD and immune-mediated diseases with the exception of multiple sclerosis and Sjogren's syndrome⁴⁷. In our study, a total of 70 unique shared genomic loci were identified between AD and immune-mediated diseases by the conjFDR analyses employing GWAS data, being enriched in biological pathways related to cell adhesion and the immune system. They were mostly enriched in the lymphatic system, with a signal being seen in the central nervous system; the top enriched cell types were also related to the immune and nervous systems, including neurons and microglia.

Two SNPs were suggestive of deleteriousness, rs12790721 and rs646327. The first is an intron variant and eQTL of *GRAMD1B* (11q24.1) in blood, and the second is an exonic variant of *FUT2* (19q13.33). Both genes are protein-coding genes, for the proteins aster-B and galactoside alpha-(1, 2)-fucosyltransferase 2, respectively. *GRAMD1B* is expressed in the brain, in astrocytes, oligodendrocytes precursor cells, and in inhibitory and excitatory neurons, mainly

maintaining synaptic function, and is also expressed in the immune system. It is predicted to be involved in cellular response to cholesterol and cholesterol homeostasis. Aster-B, as a novel regulator of mitochondrial cholesterol and fatty acid transport, and excess mitochondrial cholesterol, has also been a biomarker in AD. Because abnormal mitochondrial cholesterol is a common phenotype in AD, Aster-B has been suggested as a target for the development of therapeutics for AD⁴⁸. Regarding *FUT2*, fucosylated host glycoproteins or glycolipids mediate interaction with intestinal microbiota, influencing its composition^{49–51}. There is evidence suggesting a link between gut microbiota dysbiosis and the pathogenesis of AD, with the gut microbiota modulating neuroinflammation indirectly by impacting microglia and having effects on synaptic neurotransmission dysfunction, with a cross-talk between peripheral and central inflammation through microbiota-mediated microglial alterations^{52, 53}.

In addition to deleteriousness, seven lead SNPs were considered to have important regulatory functionality, mapped to the genes *ADAMTS4* (UTR3 region), *HBEGF* (intergenic), *WNT3* (intronic), *TSPAN14* (intronic), *DHODH* (intronic), *ABCB9* (intronic) and *TNIP1* (intronic), all of each are protein-coding genes. For instance, the protein coded by *ADAMTS4* is expressed in the cytoplasm in several tissues, being most abundant in the central nervous system. Its RNA is mostly expressed in oligodendrocytes, and is related to myelination. The enzyme coded by *ADAMTS4* is responsible for amyloid deposition in AD^{54, 55}. *HBEGF* has been associated with AD, with its overexpression resulting in increased APP protein level⁵⁶. It is expressed in the cortex and hippocampus, mostly in neurons and astrocytes, and also in the immune system. Genetic deletion of *HBEGF* cause cognitive dysfunction, which is reversed by NMDA receptor agonists⁵⁷. Heparin-binding epidermal growth factor-like growth factor (HB-EGF), coded by *HBEGF*, also restores neurogenesis in the hippocampus of aged mice⁵⁸. It also contributes to the proliferation of glial cells and to the survival of dopaminergic neurons⁵⁹. *DHODH*, it is detected in the immune system and is related to mitochondrial function and cellular homeostasis and also to alterations in reactive oxygen species levels⁶⁰. *TNIP1* is related to inflammatory response, including neuroinflammation and microglia activation, in addition to regulating nuclear factor kappa-B activation⁶¹, being related to AD⁶². Finally, in the loci to eQTL analyses, several SNPs were found associated with eQTL functionality in brain tissue and in the immune system. For instance, *NDUFS2* gene is a protein-coding gene that is associated with mitochondrial complex I alterations. A transcriptome-wide association study of AD using brain tissue found that *NDUFS2* is one of the putative causal genes in AD^{63, 64}.

Six genes were found to be shared between AD and at least two immune-related diseases: *ADAMTS4*, *JAZF1*, *Y_RNA*, *C1orf106*, *AC195454.1*, and *RP11-95M15.1*. As discussed above, *ADAMTS4* is responsible for amyloid deposition in AD^{54, 55}. *JAZF1* is a protein-coding gene which functions as a transcriptional repressor; it is involved in glucose and lipid metabolisms⁶⁵. *Y_RNA* is a small non-coding RNA gene. It has been implicated in cellular processes such as DNA replication and RNA quality control. *Y_RNA* has been found in extracellular vesicles (EV) from multiple different cell lines, and EV-associated Y-RNA may be involved in a range of immune-related processes, including inflammation and immune suppression, being regulated in immune cells by Toll-like receptor (TLR) signaling⁶⁶. Dysregulation of Y-RNA also has been found to cause alternative splicing in neurons of individuals with AD⁶⁷. *C1orf106*, also known as *CALHM6*, encodes the protein calcium homeostasis modulator, which may play an important role in several immune inflammatory responses. It is upregulated by interferon-gamma and tumor necrosis factor alpha⁶⁸. *AC195454.1* is a long non-coding gene with unknown functional category. It is overexpressed in the brain and has been implicated in SLE^{69, 70}. Finally, *RP11-95M15.1*, also a long non-coding gene, has been associated with PSC⁷¹.

One of the major strengths of our study is the employment of diverse albeit complementary statistical genetic approaches, enabling an extensive analysis of the genetic associations between AD and immune-mediated diseases. In addition, the fact that we included generally well-powered GWAS suggests that our results are mostly not due to small sample sizes. Another strength is that we included only diagnosed cases of AD and not by proxy subjects. Having said that, our study has inherent limitations that should be considered alongside the present findings. The analyses were restricted to participants of European ancestry; thus, our findings may not be generalizable to populations of other ancestries. In addition, since we used the most comprehensive GWAS data available, replication analyses in independent samples were not possible. Also, AD is not a highly polygenic disease, which might have underestimate the number of variants found in the MiXeR analyses. Further, the AIC values from the MiXeR analysis for the negative control HCM were all negative, showing poor fit and that the results of this particular analysis are unreliable. Finally, the gene mapping strategy was based on statistical analyses, and should be validated with experimental studies.

In summary, our study provides genetic insights into the observed epidemiological relationship between AD and immune-mediated diseases, exposing shared genetic susceptibility extensively distributed across the genome. Our results support a significant genetic association between AD and immune-mediated diseases with mixed effect directions, with some genes increasing the risk of one disease but decreasing the risk of the other. Furthermore, we pinpoint shared loci and genes between AD and immune-mediated diseases that have the potential to be intended for further research, such as *GRAMD1B*, *FUT2*, *ADAMTS4*, *HBEGF*, *WNT3*, *TSPAN14*, *DHODH*, *ABCB9* and *TNIP1*, all of which are protein-coding, and thus show promise as potential drug targets, and showed deleteriousness or regulatory functionality. We also highlight the significance of the immune system as a shared mechanism, albeit a non-causal one, in AD and immune-mediated diseases, providing further support for the immunoinflammatory hypothesis in AD. Overall, our results provide an atlas of the shared genetic architecture of AD and immune-mediated diseases, and their long-noticed comorbidity association, with implications for personalized interventions for the prevention and treatment of AD and immune-mediated diseases.

Declarations

Acknowledgements

We thank those investigators who generated and shared the genetic datasets used in the present study.

Funding

This research was partially supported by National Institutes of Health grants awarded to Z.Z. (R01LM012806, R01LM012806-07S1, R03AG077191, and U01AG079847). We thanked the resource support from Cancer Prevention and Research Institute of Texas (CPRIT RP180734). OAA is supported by the Research Council of Norway (#324499, #324252, #223273), South-East Norway Health Authority (2022-073), KG Jebsen Stiftelsen, Norwegian Health Association (#22731), and EU Horizon 2020 (# 847776). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of interest disclosures

OAA is a consultant to Cortechs.ai, and has received speaker's honoraria from Lundbeck, Janssen and Sunovion. NE, BSF, SB, YD, and ZZ have nothing to disclose.

References

1. 2021Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 2021. 17(3): p. 327–406.
2. Gatz, M., et al., *Role of Genes and Environments for Explaining Alzheimer Disease*. *Archives of General Psychiatry*, 2006. 63(2).
3. Heneka, M.T., et al., *Neuroinflammation in Alzheimer's disease*. *Lancet Neurol*, 2015. 14(4): p. 388–405.
4. Jorfi, M., A. Maaser-Hecker, and R.E. Tanzi, *The neuroimmune axis of Alzheimer's disease*. *Genome Med*, 2023. 15(1): p. 6.
5. Podlesny-Drabiniok, A., E. Marcora, and A.M. Goate, *Microglial Phagocytosis: A Disease-Associated Process Emerging from Alzheimer's Disease Genetics*. *Trends Neurosci*, 2020. 43(12): p. 965–979.
6. Olah, M., et al., *Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease*. *Nature Communications*, 2020. 11(1).
7. Bettcher, B.M., et al., *Peripheral and central immune system crosstalk in Alzheimer disease - a research prospectus*. *Nat Rev Neurol*, 2021. 17(11): p. 689–701.
8. Stephenson, J., et al., *Inflammation in CNS neurodegenerative diseases*. *Immunology*, 2018. 154(2): p. 204–219.
9. Huang, J., et al., *Inflammatory Diseases, Inflammatory Biomarkers, and Alzheimer Disease: An Observational Analysis and Mendelian Randomization*. *Neurology*, 2023. 100(6): p. e568-e581.
10. Liu, L., et al., *Association Between Psoriasis and Dementia: Current Evidence*. *Front Aging Neurosci*, 2020. 12: p. 570992.
11. Zhang, B., et al., *Inflammatory bowel disease is associated with higher dementia risk: a nationwide longitudinal study*. *Gut*, 2021. 70(1): p. 85–91.
12. Yokoyama, J.S., et al., *Association Between Genetic Traits for Immune-Mediated Diseases and Alzheimer Disease*. *JAMA Neurol*, 2016. 73(6): p. 691–7.
13. Wightman, D.P., et al., *A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease*. *Nat Genet*, 2021. 53(9): p. 1276–1282.
14. Watanabe, K., et al., *A global overview of pleiotropy and genetic architecture in complex traits*. *Nat Genet*, 2019. 51(9): p. 1339–1348.
15. Paternoster, L., et al., *Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis*. *Nat Genet*, 2015. 47(12): p. 1449–1456.
16. Dubois, P.C., et al., *Multiple common variants for celiac disease influencing immune gene expression*. *Nat Genet*, 2010. 42(4): p. 295–302.
17. de Lange, K.M., et al., *Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease*. *Nat Genet*, 2017. 49(2): p. 256–261.
18. Ji, S.-G., et al., *Genome-wide association study of primary sclerosing cholangitis identifies new risk loci and quantifies the genetic relationship with inflammatory bowel disease*. *Nature Genetics*, 2016. 49(2): p. 269–273.
19. Okada, Y., et al., *Genetics of rheumatoid arthritis contributes to biology and drug discovery*. *Nature*, 2014. 506(7488): p. 376–81.
20. Bentham, J., et al., *Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus*. *Nat Genet*, 2015. 47(12): p. 1457–1464.
21. Jin, Y., et al., *Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants*. *Nat Genet*, 2016. 48(11): p. 1418–1424.
22. Kurki, M.I., et al., 2022.
23. Bulik-Sullivan, B.K., et al., *LD Score regression distinguishes confounding from polygenicity in genome-wide association studies*. *Nat Genet*, 2015. 47(3): p. 291–5.
24. Frei, O., et al., *Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation*. *Nature Communications*, 2019. 10(1).
25. Drange, O.K., et al., *Genetic Overlap Between Alzheimer's Disease and Bipolar Disorder Implicates the MARK2 and VAC14 Genes*. *Front Neurosci*, 2019. 13: p. 220.
26. Andreassen, O.A., et al., *Abundant genetic overlap between blood lipids and immune-mediated diseases indicates shared molecular genetic mechanisms*. *PLoS One*, 2015. 10(4): p. e0123057.
27. Bedre, R. and *reneshbedre/bioinfokit: Bioinformatics data analysis and visualization toolkit*. 2022.
28. Werme, J., et al., *An integrated framework for local genetic correlation analysis*. *Nat Genet*, 2022. 54(3): p. 274–282.

29. Lambert, J.C., et al., *Implication of the immune system in Alzheimer's disease: evidence from genome-wide pathway analysis*. J Alzheimers Dis, 2010. 20(4): p. 1107–18.
30. Scheltens, P., et al., Alzheimer's disease. Lancet, 2016. 388(10043): p. 505–17.
31. de Bakker, P.I. and S. Raychaudhuri, *Interrogating the major histocompatibility complex with high-throughput genomics*. Hum Mol Genet, 2012. 21(R1): p. R29-36.
32. Jia, P., et al., *A bias-reducing pathway enrichment analysis of genome-wide association data confirmed association of the MHC region with schizophrenia*. J Med Genet, 2012. 49(2): p. 96–103.
33. Buniello, A., et al., *The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019*. Nucleic Acids Res, 2019. 47(D1): p. D1005-D1012.
34. Yokoyama, J.S., et al., *Association Between Genetic Traits for Immune-Mediated Diseases and Alzheimer Disease*. JAMA Neurology, 2016. 73(6).
35. Monereo-Sanchez, J., et al., *Genetic Overlap Between Alzheimer's Disease and Depression Mapped Onto the Brain*. Front Neurosci, 2021. 15: p. 653130.
36. Fominykh, V., et al., *Shared genetic loci between Alzheimer's disease and multiple sclerosis: Crossroads between neurodegeneration and immune system*. Neurobiol Dis, 2023. 183: p. 106174.
37. Watanabe, K., et al., *Functional mapping and annotation of genetic associations with FUMA*. Nat Commun, 2017. 8(1): p. 1826.
38. Clarke, L., et al., *The international Genome sample resource (IGSR): A worldwide collection of genome variation incorporating the 1000 Genomes Project data*. Nucleic Acids Res, 2017. 45(D1): p. D854-D859.
39. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data*. Nucleic Acids Res, 2010. 38(16): p. e164.
40. Rentzsch, P., et al., *CADD: predicting the deleteriousness of variants throughout the human genome*. Nucleic Acids Res, 2019. 47(D1): p. D886-D894.
41. Boyle, A.P., et al., *Annotation of functional variation in personal genomes using RegulomeDB*. Genome Res, 2012. 22(9): p. 1790–7.
42. Pei, G., et al., *DeepFun: a deep learning sequence-based model to decipher non-coding variant effect in a tissue- and cell type-specific manner*. Nucleic Acids Research, 2021. 49(W1): p. W131-W139.
43. Dai, Y., et al., *WebCSEA: web-based cell-type-specific enrichment analysis of genes*. Nucleic Acids Research, 2022. 50(W1): p. W782-W790.
44. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. Journal of the Royal Statistical Society: Series B (Methodological), 1995. 57(1): p. 289–300.
45. Davies, N.M., M.V. Holmes, and G. Davey Smith, *Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians*. BMJ, 2018. 362: p. k601.
46. Hemani, G., et al., *The MR-Base platform supports systematic causal inference across the human phenome*. eLife, 2018. 7.
47. Yeung, C.H.C., S.L. Au Yeung, and C.M. Schooling, *Association of autoimmune diseases with Alzheimer's disease: A mendelian randomization study*. Journal of Psychiatric Research, 2022. 155: p. 550–558.
48. Andersen, J.P., et al., *Aster-B coordinates with Arf1 to regulate mitochondrial cholesterol transport*. Mol Metab, 2020. 42: p. 101055.
49. Rausch, P., et al., *Colonic mucosa-associated microbiota is influenced by an interaction of Crohn disease and FUT2 (Secretor) genotype*. Proc Natl Acad Sci U S A, 2011. 108(47): p. 19030–5.
50. Wacklin, P., et al., *Secretor genotype (FUT2 gene) is strongly associated with the composition of Bifidobacteria in the human intestine*. PLoS One, 2011. 6(5): p. e20113.
51. Wacklin, P., et al., *Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status*. PLoS One, 2014. 9(4): p. e94863.
52. Wu, S., et al., *Roles and Mechanisms of Gut Microbiota in Patients With Alzheimer's Disease*. Frontiers in Aging Neuroscience, 2021. 13.
53. Bairamian, D., et al., *Microbiota in neuroinflammation and synaptic dysfunction: a focus on Alzheimer's disease*. Molecular Neurodegeneration, 2022. 17(1).
54. Matsuzaki, M., et al., *ADAMTS4 is involved in the production of the Alzheimer disease amyloid biomarker APP669-711*. Molecular Psychiatry, 2023. 28(4): p. 1802–1812.
55. Walter, S., et al., *The metalloprotease ADAMTS4 generates N-truncated Aβ4-x species and marks oligodendrocytes as a source of amyloidogenic peptides in Alzheimer's disease*. Acta Neuropathologica, 2018. 137(2): p. 239–257.
56. da Rocha, J.F., et al., *APP Binds to the EGFR Ligands HB-EGF and EGF, Acting Synergistically with EGF to Promote ERK Signaling and Neuritogenesis*. Molecular Neurobiology, 2020. 58(2): p. 668–688.
57. Sasaki, K., et al., *NMDA receptor agonists reverse impaired psychomotor and cognitive functions associated with hippocampal Hbegf-deficiency in mice*. Molecular Brain, 2015. 8(1).
58. Jin, K., et al., *Neurogenesis and aging: FGF-2 and HB-EGF restore neurogenesis in hippocampus and subventricular zone of aged mice*. Aging Cell, 2003. 2(3): p. 175–183.
59. Oyagi, A. and H. Hara, *Essential Roles of Heparin-Binding Epidermal Growth Factor-Like Growth Factor in the Brain*. CNS Neuroscience & Therapeutics, 2012. 18(10): p. 803–810.
60. Scialò, F., D.J. Fernández-Ayala, and A. Sanz, *Role of Mitochondrial Reverse Electron Transport in ROS Signaling: Potential Roles in Health and Disease*. Frontiers in Physiology, 2017. 8.
61. Adelaja, A. and A. Hoffmann, *Signaling Crosstalk Mechanisms That May Fine-Tune Pathogen-Responsive NFκB*. Frontiers in Immunology, 2019. 10.

62. Jones, S.V. and I. Kounatidis, *Nuclear Factor-Kappa B and Alzheimer Disease, Unifying Genetic and Environmental Risk Factors from Cell to Humans*. Frontiers in Immunology, 2017. 8.
63. Sun, Y., et al., *A transcriptome-wide association study of Alzheimer's disease using prediction models of relevant tissues identifies novel candidate susceptibility genes*. Genome Medicine, 2021. 13(1).
64. Wang, X.-L. and L. Li, *Cell type-specific potential pathogenic genes and functional pathways in Alzheimer's Disease*. BMC Neurology, 2021. 21(1).
65. Nakajima, T., et al., *TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4*. Nucleic Acids Res, 2004. 32(14): p. 4194–204.
66. Driedonks, T.A.P. and E.N.M. Nolte-'t Hoen, *Circulating Y-RNAs in Extracellular Vesicles and Ribonucleoprotein Complexes; Implications for the Immune System*. Frontiers in Immunology, 2019. 9.
67. Scheckel, C., et al., *Regulatory consequences of neuronal ELAV-like protein binding to coding and non-coding RNAs in human brain*. eLife, 2016. 5.
68. Malik, U., et al., *Structural and functional annotation of human FAM26F: A multifaceted protein having a critical role in the immune system*. Gene, 2017. 597: p. 66–75.
69. Owen, K.A., A.C. Grammer, and P.E. Lipsky, *Deconvoluting the heterogeneity of SLE: The contribution of ancestry*. Journal of Allergy and Clinical Immunology, 2022. 149(1): p. 12–23.
70. Yin, X., et al., *Meta-analysis of 208370 East Asians identifies 113 susceptibility loci for systemic lupus erythematosus*. Annals of the Rheumatic Diseases, 2021. 80(5): p. 632–640.
71. Hitomi, Y. and M. Nakamura, *The Genetics of Primary Biliary Cholangitis: A GWAS and Post-GWAS Update*. Genes, 2023. 14(2).

Figures

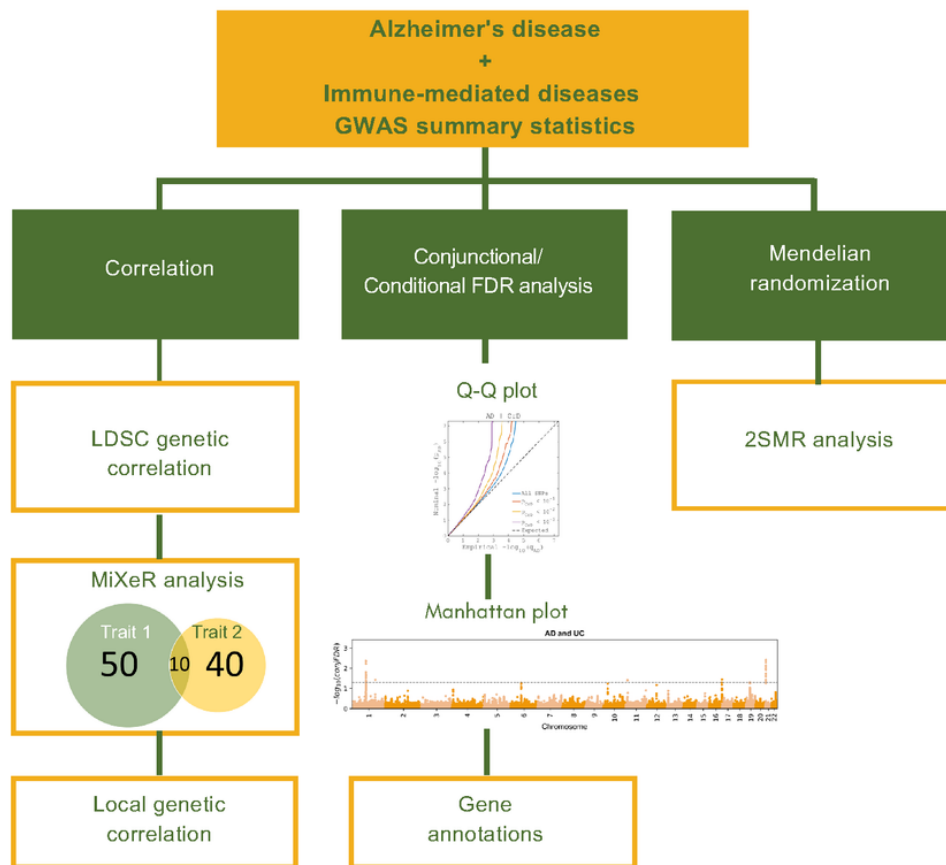


Figure 1

Workflow for the study of AD and 11 immune-mediated diseases.

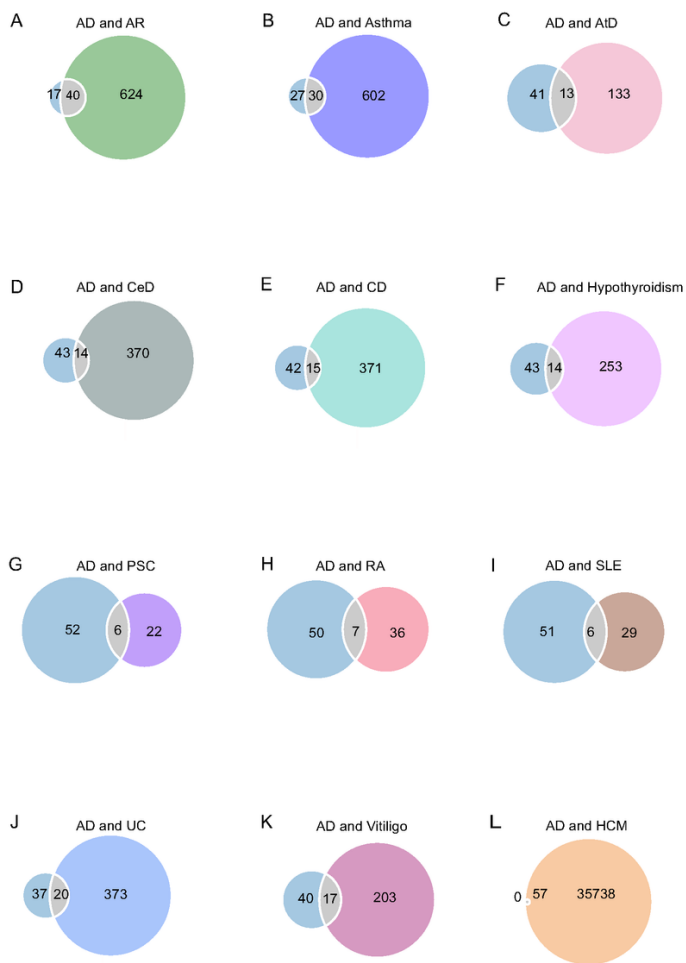


Figure 2

Genetic overlap between AD and 11 immune-mediated diseases using MiXeR. Venn diagrams represent the genetic overlap, regardless of direction, between AD and each immune-mediated disease. The size of the circles represents the polygenicity of the phenotype. Numeric values represent the estimated number of variants. Abbreviations: AD: Alzheimer's disease; AR: allergic rhinitis; AtD: atopic dermatitis; CeD: celiac disease; CD: Crohn's disease; HCM: hypertrophic cardiomyopathy; PSC: primary sclerosing cholangitis; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; and UC: ulcerative colitis.

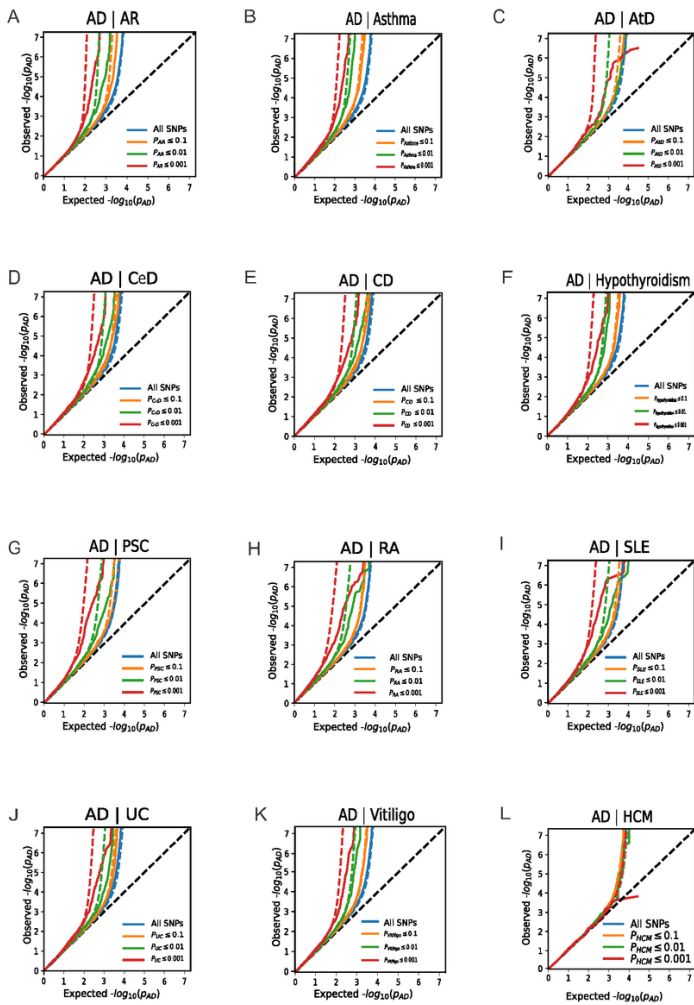


Figure 3

Conditional quantile-quantile plots of nominal vs. empirical $-\log_{10} p$ -values in AD below the standard genome-wide association study threshold of p -value $< 5 \times 10^{-8}$ as a function of significance of association with immune-mediated diseases at the level of $-\log_{10} p$ -values of 1, 2, or 3, corresponding to p -value = 0.1, p -value = 0.01, and p -value = 0.001, respectively. Significant association was found between AD and all immune-mediated diseases. No significant association was found between AD and HCM (negative control). The null hypothesis is indicated by the dashed lines. Abbreviations: AD: Alzheimer's disease; AR: allergic rhinitis; AtD: atopic dermatitis; CeD: celiac disease; CD: Crohn's disease; HCM: hypertrophic cardiomyopathy; PSC: primary sclerosing cholangitis; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; and UC: ulcerative colitis.

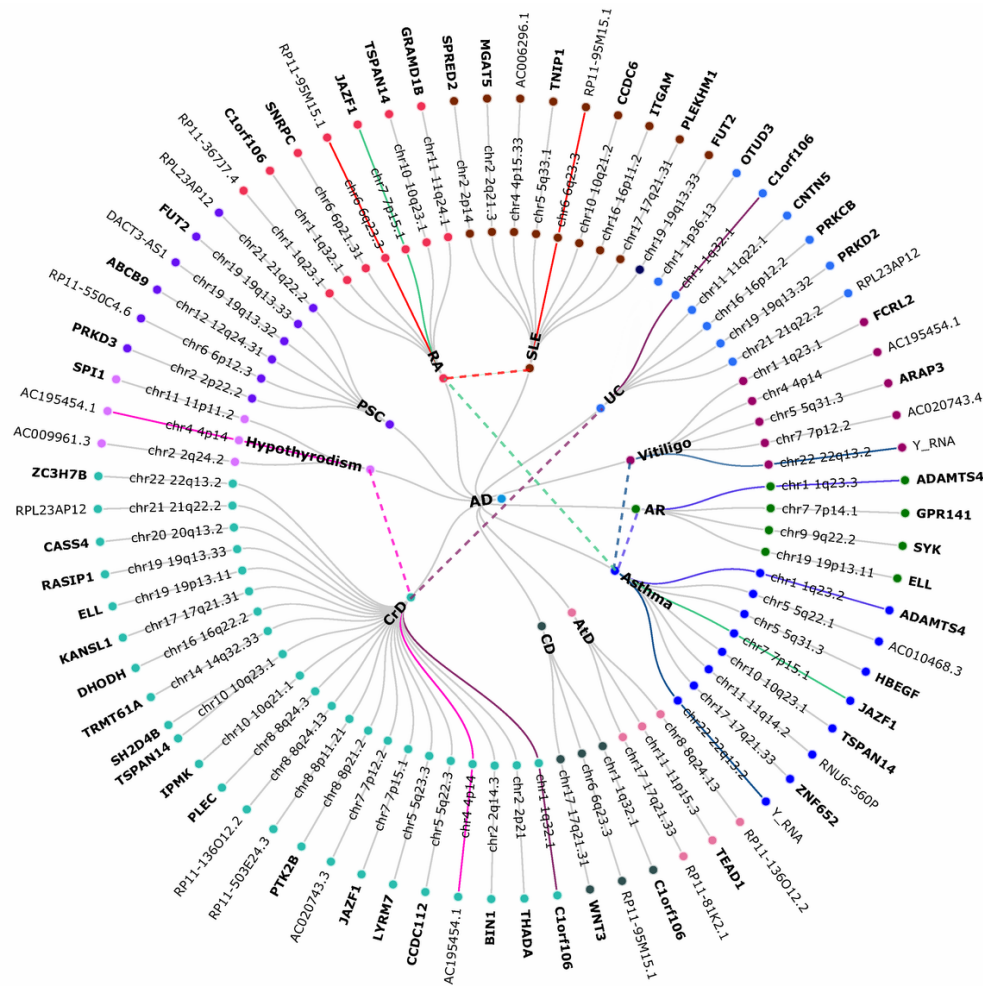


Figure 4

A circular dendrogram showing the shared genes between AD (center circle) and each of eleven specific immune-mediated diseases (first circle), resulting in eleven pairs. A total of 70 unique shared loci were identified across eleven trait pairs, which mapped to 43 distinct protein-coding genes and 16 distinct non-protein coding genes (second and third circles) identified using ANNOVAR. The dashed lines denote genes shared between AD and more than one immune-mediated disease. Genes in bold font suggest protein-coding genes. Abbreviations: ANNOVAR: analysis ANNOtate VARIation (ANNOVAR); AD: Alzheimer's disease; AR: allergic rhinitis; AtD: atopic dermatitis; CeD: celiac disease; CD: Crohn's disease; HCM: hypertrophic cardiomyopathy; PSC: primary sclerosing cholangitis; RA: rheumatoid arthritis; SLE: systemic lupus erythematosus; and UC: ulcerative colitis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xlsx](#)
- [SupplementaryMaterial.pdf](#)