Regional genetic correlations highlight relationships between neurodegenerative diseases and the immune system

Frida Lona-Durazo[1,2], Regina H. Reynolds[3,4], Sonja W. Scholz[5,6], Mina Ryten[3,4,7], Sarah A. Gagliano Taliun[1,8,*]

1. Montréal Heart Institute, Montréal, Québec, Canada
2. Université de Montréal, Montréal, Québec, Canada
3. Genetics and Genomic Medicine, Great Ormond Street Institute of Child Health, University College London, London, UK
4. Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, United States
5. Neurodegenerative Diseases Research Unit, National Institute of Neurological Disorders and Stroke, Bethesda, MD, USA
6. Department of Neurology, Johns Hopkins University Medical Center, Baltimore, MD, USA
7. NIHR Great Ormond Street Hospital Biomedical Research Centre, University College London, London, UK
8. Department of Medicine & Department of Neurosciences, Université de Montréal, Montréal, Québec, Canada

* Correspondence: sarah.gagliano-taliun@umontreal.ca

Abstract

Neurodegenerative diseases, including Alzheimer's disease (AD), Parkinson's disease (PD), Lewy body dementia (LBD) and amyotrophic lateral sclerosis (ALS), are devastating complex diseases that result in a physical and psychological burden to patients and their families. There have been significant efforts to understand the genetic basis of neurodegenerative diseases resulting in the identification of disease risk-associated variants involved in several molecular mechanisms, including those that influence immune-related pathways. Regional genetic correlations, in contrast to genome-wide correlations, between pairs of immune and neurodegenerative traits have not been comprehensively explored, but such a regional assessment could shed light on additional immune-mediated risk-associated loci. Here, we systematically assessed the potential role of the immune system in five neurodegenerative diseases, by estimating regional genetic correlations between neurodegenerative diseases and immune-cell-derived single-cell expression quantitative trait loci (sc-eQTLs), using the recently developed method of Local Analysis of [co]Variant Association (LAVA). We used the most recently published genome-wide association studies (GWASes) for five neurodegenerative diseases and publicly available sc-eQTLs derived from 982 individuals from the OneK1K Consortium, capturing aspects of the innate and adaptive immune systems. Additionally, we tested GWASes from well-established immune-mediated diseases, Crohn's disease (CD) and ulcerative colitis (UC), the immune-mediated neurodegenerative disease, multiple sclerosis (MS) and a well-powered GWAS with strong signal in the HLA region, schizophrenia (SCZ), as positive controls. Finally, we also performed regional genetic correlations between diseases and protein levels. We observed significant (FDR < 0.01) regional genetic correlations between sc-eQTLs and neurodegenerative diseases across 151 unique genes, spanning both the innate and adaptive immune systems, across most diseases tested (except for frontotemporal dementia (FTD) and LBD). Colocalization analyses on followed-up regional correlations highlighted immune-related candidate causal risk genes associated with neurodegenerative

diseases. We also observed significant regional correlations with protein levels across 156 unique proteins, across all diseases tested, except for FTD. The outcomes of this study will improve our understanding of the immune component of neurodegeneration, which can be potentially used to repurpose existing immunotherapies used in clinical care for other immune-mediated diseases, to slow the progression of neurodegenerative diseases.

Introduction

Adult-onset neurodegenerative diseases, such as Alzheimer's disease (AD), Parkinson's disease (PD), and amyotrophic lateral sclerosis (ALS), are devastating conditions affecting populations worldwide and resulting in a large physical and psychological burden to patients and their families. Neurodegeneration occurs when the cells in the nervous system deteriorate, leading to cell death caused by various pathological factors and possibly multiple biological systems. Large-scale genome-wide association studies (GWASes), which have mostly been conducted in individuals of European genetic ancestry, identified multiple risk loci associated with AD, PD, ALS, and other complex neurodegenerative diseases (Ferrari et al., 2014; Kunkle et al., 2019; Nalls et al., 2014, 2019; Schwartzentruber et al., 2021; van Rheenen et al., 2016, 2021; Wightman et al., 2021). However, there is a gap in our understanding of the mechanisms by which genetic risk factors influence the pathogenesis of neurodegenerative diseases.

Computational and experimental evidence of several complex neurodegenerative diseases suggest that the immune system is involved in disease development (Broce et al., 2018; Gagliano et al., 2016; Kim et al., 2021; Nalls et al., 2019; Schwartzentruber et al., 2021; Tansey et al., 2022) (Supplementary Figure 1). For instance, there is experimental evidence in transgenic mice suggesting an association between AD progression and accumulation of B cells and immunoglobulin deposits around Aβ plaques (Kim et al., 2021). Additionally, computational evidence has shown a genetic enrichment of AD GWAS signals across the innate and adaptive immune systems (Gagliano et al., 2016; Kim et al., 2021). Further, an experimental study uncovered the role of CD4+ T cells in brains of Lewy Body Dementia (LBD) patients and its relationship with neurodegeneration (Gate et al., 2021). In the case of ALS, the largest GWAS identified the Human Leukocyte Antigen (HLA) region as a novel disease risk locus, and an epigenome-wide association study highlighted an enrichment of Immunoglobulin E as associated with disease risk (Hop et al., 2022; van Rheenen et al., 2021). Additionally, the immune system is thought to be implicated in PD, in which the gene *LRRK2* is a shared disease risk for Crohn's disease (CD) and PD (Herrick & Tansey, 2021; Tansey et al., 2022). Further, tyrosine kinase inhibition has shown to modulate the immune response in PD (Pagan et al., 2016). These findings motivate the study of links between neurodegeneration and immune processes. The wealth of large-scale "omic" data becoming available presents a unique opportunity to apply new data-driven approaches to better understand the molecular and cellular immune-related mechanisms influencing neurodegenerative diseases, through the lens of genetics.

Pinpointing targets for neurodegenerative diseases in specific cell types involved in the immune system will be key to downstream repurposing of existing immune therapies as treatment options for certain neurodegenerative diseases. Here, we identify and investigate shared genomic loci between immune function and risk of neurogenerative disease risk using bioinformatics tools with large-scale GWAS datasets and expression quantitative trait loci (gene expression and protein). In our investigation of the role of immune-mediated pathways in neurodegenerative diseases, we not only demonstrate known relationships among genes, cell types and diseases, but also identify new potential links. Our approach pinpoints pertinent genes in a particular cell type for a particular neurodegenerative disease.

Methods

Datasets and data formatting

We obtained genome-wide association study (GWAS) datasets from publicly available repositories, or requested access to their corresponding summary statistics. We selected five GWAS datasets from European genetic ancestry case/control studies of common neurodegenerative diseases as test traits: 1) Alzheimer's disease (Schwartzentruber et al., 2021), 2) Parkinson's disease (Nalls et al., 2019), 3) Lewy body dementia (Chia et al., 2021), 4) amyotrophic lateral sclerosis (van Rheenen et al., 2021), and 5) frontotemporal dementia (Ferrari et al., 2014). We also included three GWAS datasets corresponding to case/control studies of immune-mediated diseases as control traits: 1) multiple sclerosis (IMSGC et al., 2012), 2) ulcerative colitis (de Lange et al., 2017), and 3) Crohn's disease (de Lange et al., 2017). Finally, we included as a test dataset a well-powered case/control study of schizophrenia (Pardiñas, 2018), a neuropsychiatric disorder in which there is a genome-wide association with the Human Leukocyte Antigen (HLA) region, encoding genes that play a key role in the immune system. Detailed information on the GWAS sample sizes, number of genetic variants, genomic build, and source URLs are available in Supplementary Table 1. After download, we formatted the GWAS summary statistics with R (version 4.0.2) (R Core Team, 2019) and lifted over the genomic coordinates to the Human Genome Build GRCh37 with the R package *rutils* version 0.99.2 (RHReynolds, 2022) as needed. We used the R package *SNPlocs.Hsapiens.dbSNP144.GRCh37* (Pagès, 2017) to map reference SNP IDs (rsids) to genomic coordinates or vice versa. All analyses in the present study were performed using the Digital Research Alliance of Canada compute clusters.

We obtained single-cell expression quantitative trait loci (sc-eQTLs) summary statistics from the OneK1K study (Yazar et al., 2022) by personal communication with the corresponding author. The dataset includes single-cell expression data on 1.27 million peripheral blood mononuclear cells in 982 individuals of European genetic ancestry, clustered into 14 immune cell types. To minimize the multiple testing burden, we selected a subset of these cell types for the present study. Specifically, we included the following cells from the innate and adaptive immune system: 1) classical monocytes, 2) effector memory CD4+ T cells, 3) naïve CD4+ T cells, 4) effector memory CD8+ T cells, 5) naïve CD8+ T cells, 6) naïve B cells, and 7) memory B cells. To explore an additional level of biological variation, we also obtained summary statistics of plasma protein QTLs, pQTLs (Zhang et al., 2022), corresponding to the "European American" sample, including 7,213 individuals (http://nilanjanchatterjeelab.org/pwas/).

Genome-wide genetic correlations across GWAS datasets

We estimated genome-wide genetic correlations (r_g) across GWAS trait pairs using linkage disequilibrium score regression (LDSC) (Bulik-Sullivan et al., 2015). We first formatted GWAS summary statistics for each trait using the *munge_sumstats.py* function to align the alleles and keep SNPs present in the HapMap Project Phase 3, with the MHC region removed. Next, we ran the *ldsc.py* function for each trait pair using the 1000 Genomes Project Phase 3 European super-population as the LD reference to obtain r_g estimates. We applied a Bonferroni corrected p-value threshold to account for the number of pair-wise correlations performed, and subsequently defined a significant correlation if p-value < 0.0014.

Regional genetic correlations across GWAS datasets

We estimated regional genetic correlations (r_g) across GWAS pair traits with the R package *Local Analysis of [co]Variant Association* (LAVA) (Werme et al., 2022). We used the genomic regions defined as autosomal LD blocks (N = 2,495) across autosomal chromosomes by Werme et al., which are characterized by having minimum LD across regions, a minimum of 2,500 variants included on each LD block, and with an average LD block size of 1 million bases. To define which genomic regions to test across GWAS traits, we selected LD blocks that contained at least one genome-wide significant signal in at least one GWAS trait (n = 389). We considered sample overlap across GWAS datasets in the analysis by including the pair-wise genetic covariance estimated by LDSC and further standardizing it into a correlation matrix. To estimate regional r_g, we first performed a univariate test for each trait per LD block and performed a bivariate test only for those trait pairs that had a significant univariate genetic signal (p-value < 1.28e-04, correcting for the 389 LD blocks tested). We applied a Bonferroni corrected p-value threshold to account for the number of pair-wise regional correlations performed and defined a significant correlation if p-value < 2.63e-05 (0.05/1,902).

Regional genetic correlations between GWAS and QTLs

We estimated regional r_g between GWAS and QTL datasets using LAVA (Werme et al., 2022). In the case of the regional correlations between GWAS and gene expression levels, we tested protein and non-protein coding genes harbouring at least one genome-wide significant sc-eQTL per cell type separately. We extended the tested region 100kb upstream and downstream of the start and end positions of the gene, which is where the majority of the *cis*-eQTLs are located (Võsa et al., 2021). We followed the same approach when defining the genomic regions to test between GWAS and protein levels, in which we included proteins that harboured at least one genome-wide significant pQTL, and extended the tested region +/- 100kb from the start/end gene coordinates of the respective protein. In both cases (i.e. sc-eQTLs and pQTLs) we assumed that there was no sample overlap between the GWAS and QTL datasets, which we believe is a reasonable assumption. We estimated regional r_g as described above, in which we first performed a univariate test for each trait, and then performed a bivariate test between GWAS-QTL only if both had a significant univariate genetic signal, correcting for the number of genes or proteins tested. We applied an FDR correction to the p-value threshold to account for the number of pair-wise regional correlations, separately for the analysis with sc-eQTLs and pQTLs, thus defining a significant GWAS-QTL correlation if FDR < 0.01. We chose an FDR correction, instead of a stricter Bonferroni correction, given that genic regions do not necessarily represent unique regions of linkage equilibrium (i.e. variants in one gene may be in linkage disequilibrium with variants in nearby genes as well).

***Colocalization* to follow-up on regional correlations between GWAS and gene expression levels**

We performed colocalization analysis to follow-up on the significant (FDR < 0.01) regional correlations between GWAS and gene expression levels using the tool, *coloc* (Giambartolomei et al., 2014; Wallace, 2021). For those traits where the sample minor allele frequency (MAF) was available, we checked the correlation between the MAF of the 1000 Genomes European super-population and the sample MAF, which was 0.99 in all cases. Similar to the regional correlations approach, we tested the genic region +/- 100kb from the start/end gene coordinates and assumed a maximum of one causal signal per colocalization. We tested a total of 366 GWAS-QTL pairs using the default SNP priors ($p_1 = p_2$ = 1e-04 and p_{12} = 1e-05). We considered a region to colocalize between gene expression levels and a GWAS trait, if the posterior probability (PP) of H4 ≥ 0.8, which suggests a high probability of a shared causal signal between both traits.

Gene set enrichment analysis to follow-up on regional genetic correlations between GWAS and protein levels

We performed a gene set enrichment analysis with the GENE2FUNC tool implemented in FUMA (Watanabe & Taskesen, 2017) to aid in the interpretation of the regional genetic correlations between GWAS traits and protein levels. We analysed one GWAS trait at a time and included only genes with protein levels that were significantly correlated with that GWAS trait (FDR < 0.01). We used Ensembl version 92 and included the list of 4,657 genes for which protein levels were assessed (hence the genes used in the pQTL analysis) as the background set of genes (Zhang et al., 2022). FUMA performs a hypergeometric test for gene set enrichment using gene set databases obtained from MSigDB, WikiPathways and the GWAS Catalog. Of the available datasets, we focused specifically on Gene Ontology biological processes. We set a minimum threshold of overlapping genes with gene sets of ≥ 2 and used the Benjamin-Hochberg FDR multiple testing correction method (alpha = 0.05) to define enriched gene-sets.

Validation of AD and PD signals using GWAS without proxy cases

The AD and PD GWAS used in these analyses included proxy cases (i.e., individuals who do not have the disease of interest, but have a close relative who does). The inclusion of proxy cases has been suggested as a useful means to increase case sample size, particularly for late-onset disorders such as AD and PD (Liu et al., 2017). However, there are also concerns raised on the impact of proxies on heritability and careful diagnosis, which has been examined in the context of AD (Escott-Price & Hardy, 2022). As a sensitivity analysis to verify that our results involving AD and PD are not primarily driven by possible spurious effects of the inclusion of proxy cases, we reperformed the analyses for significant findings for the regional genetic correlations with gene expression levels and with protein levels using AD and PD without proxy cases (Blauwendraat et al., 2019; Kunkle et al., 2019). Validation results are described in Supplementary Note 1.

Data and Code Availability

The source of each GWAS summary statistics file is provided in Supplementary Table 1. OneK1K sc-eQTL summary statistics that include effect sizes and standard errors were provided through personal communication with the corresponding author. pQTL summary statistics are available from http://nilanjanchatterjeelab.org/pwas/. All code generated for performing the analyses in the present study is available in the following GitHub repository: https://github.com/GaglianoTaliun-Lab/neuroimmune_genetics_project.

Results

Genome-wide overview of genetic correlations among diseases

We performed pairwise genome-wide genetic correlations (r_g) across GWAS datasets(Bulik-Sullivan et al., 2015). These GWAS datasets include five neurodegenerative diseases (AD, PD, LBD, ALS and FTD) immune-mediated diseases (MS, UC and CD) and SCZ, a neuropsychiatric disorder. Using a Bonferroni-corrected p-value = 0.0014, we identified six significant positive correlations (Figure 1), of which one was between two neurodegenerative diseases: PD and LBD (r_g = 0.65; p-value = 1e-03). We did, however, observe nominally significant correlations (p-value < 0.05) between other pairs of tested neurodegenerative diseases, except for FTD, for which there were no nominally significant correlations (Supplementary Table 2). Between

immune-mediated diseases (our control traits), we confirmed the expected significant positive correlations across these diseases (Figure 1), providing internal validation of the robustness of our approach. Finally, we saw positive significant correlations between SCZ and immune-mediated diseases (i.e. UC and CD), but not between SCZ and neurodegenerative diseases (Figure 1). All genome-wide genetic correlations results are provided in Supplementary Table 2.

Regional genetic correlations highlight pleiotropic loci implicated in neurodegenerative diseases

Reassured by the detection of known global genetic correlations and cognisant of the fact that regional correlations between two traits can be masked when assessing in a genome-wide basis (van Rheenen et al., 2019), we estimated regional genetic correlations using LAVA (Werme et al., 2022). The advantage of this tool is that it can perform correlations across multiple traits simultaneously. We performed a total of 1,902 pair-wise correlations across 389 loci with adequate univariate signal, yielding a total of 59 genomic regions (i.e. LD blocks – See Methods) with significant correlations in at least one trait pair (Bonferroni-corrected p-value threshold = 2.629e-05).

We identified significant regional correlations between various diseases and genomic loci, including loci that contain genes known to be implicated in neurodegenerative diseases. For example, we observed positive genetic correlations between AD and LBD at two genomic loci. The locus located on chromosome 2 [chr2:126754028-127895644] contains the *BIN1* gene (r_g = 0.564; p-value = 9.80e-06), whereas the locus on chromosome 19 [chr19:45040933-45893307] contains the *APOE* gene (r_g = 0.80; p-value = 1.97e-124). Both genes have been implicated in AD and LBD risk (Chia et al., 2021; Kunkle et al., 2019; Schwartzentruber et al., 2021; Seshadri et al., 2010; Wightman et al., 2021). We also observed a positive genetic correlation between PD and LBD at a locus on chromosome 4 [chr4:812416-1529267] containing *TMEM175* (r_g = 0.648; p-value = 1.49e-05). In contrast, other genomic loci containing genes that are known to be involved in more than one neurodegenerative disease did not yield significant correlations, such as the locus containing *SNCA* [chr4:90236972-91309863] between LBD and PD (r_g = 0.165; p-value = 0.130), and the locus that includes *GRN* [chr17:42348004-43460500], known to be involved in AD, PD, FTD and LBD. In the latter case of *GRN*, AD, FTD and LBD did not have sufficient univariate signal (p-value ≥ 2.63e-05) to test genetic correlations at that locus. In the case of the locus including *SNCA*, the lack of correlation may be explained by previous colocalization analyses, which have suggested that there are different regulatory causal variants implicated in PD and LBD (Chia et al., 2021).

Aside from observing significant correlations at known pleiotropic loci, we also saw significant genetic correlations between neurodegenerative and immune-mediated diseases. One such example was a genomic locus on chromosome 1 [chr1:161054077-161945442] containing, among multiple genes, *FCGR2A*. This locus was negatively correlated between PD and UC (r_g = -0.652; p-value = 1.64e-05), positively correlated between UC and CD (r_g = 0.835; p-value = 1.81e-06), and it was correlated with nominal significance between PD and CD (r_g = -0.462; p-value = 0.018) (Figure 2). The gene *FCGR2A* has been previously associated with several immune-mediated diseases, and it is a risk locus for PD (Nalls et al., 2019; Witoelar et al., 2017). Additionally, a locus [chr4:169555115-170682809] that includes six protein-coding genes (i.e. *PALLD, CBR4, NEK1, CLCN3, C4orf27* and *SH3RF1*) was significantly correlated between PD and UC (r_g = 0.525; p-value = 1.02e-05) and nominally significant between ALS and CD (r_g = 0.369; p-value = 0.044) (Figure 2). The gene *CLCN3* is a disease risk locus for PD (Nalls et al., 2019), whereas rare variants within the same gene have been associated with ALS risk (van Rheenen et al., 2021). Finally, there was a significant correlation at a locus [chr7:49632427-50894508] that includes the gene *IKZF1,*

which encodes a transcription factor of the zinc-finger DNA-binding protein family, involved in B cell activation and differentiation, between AD and MS (r_g = -0.77; p-value = 3.55e-06). This same locus was also nominally positively correlated between AD and CD (r_g = 0.489; p-value = 0.002) (Figure 2).

The Human Leukocyte Antigen (HLA) locus, a region in the genome with clear immune influences, spans multiple genomic loci that were tested in the analysis. However, we only observed one locus within this region [chr6:32208902-32454577] with significant positive correlations between neurodegenerative and immune-mediated diseases, correlated between AD and MS (r_g = 0.778; p-value = 2.01e-06). This locus contains the Major Histocompatibility Complex (MHC) class II gene *HLA-DRA*. The same genomic locus was positively correlated between SCZ and UC (r_g = 0.683; p-value = 1.47e-06). An additional locus also spanning the HLA region [chr6:32682214-32897998] was positively correlated between SCZ and UC (r_g = 0.626; p-value = 9.99e-06) and between CD and SCZ (r_g = 0.677; p-value = 2.58e-05). This locus includes the following MHC class II genes: *HLA-DQA2*, *HLA-DQB2* and *HLA-DOB*.

Gene expression levels of immune-related genes share causal signals with neurodegenerative diseases

We then moved forward to ask whether regional genetic correlations could allow us to identify immune targets for neurodegenerative diseases. To do so, we estimated regional genetic correlations between diseases and genes significantly expressed across seven immune cell types (i.e., naïve B cells, memory B cells, classical monocytes, CD4+ naïve T cells, CD8+ naïve T cells, CD4+ effector memory T cells and CD8+ effector memory T cells) from one of the largest datasets, the OneK1K dataset (Yazar et al., 2022), to assess if changes in gene expression are correlated with disease risk. We performed 1,628 pair-wise correlations across 2,553 significantly expressed genes in at least one cell type (Supplementary Table 3), which resulted in 366 significant correlations (FDR < 0.01) (Figure 3). Additionally, we followed up on the significant correlations through colocalization analyses to assess if there is a shared causal signal driving the correlation. This information provided insights about specific immune cell types and genes implicated in disease risk.

Across the tested neurodegenerative diseases, there were no expressed genes significantly correlated with FTD or LBD, which were the two GWASes with the smallest sample size. In terms of the total number of tested correlations within a disease, AD had a higher proportion of correlations with expressed genes in classical monocytes. Similarly, there was a relatively higher proportion of expressed genes significantly correlated with ALS in memory B cells (but there were only 12 significant correlations across all cell types for ALS). Finally, compared to all other tested diseases, there was a relatively higher proportion of expressed genes significantly correlated with PD in CD8+ effector memory T cells. These results provide an initial overview of how disease risk across neurodegenerative diseases may be influenced by different immune cell types.

Significant correlations between diseases and expressed genes were distributed across all autosomes except on chromosome 9, where we only observed nominally significant correlations (Supplementary Figure 2). Additionally, there were genes for which their expression was significantly correlated with a disease across more than one cell type (N = 47), whereas other expressed genes were significantly correlated with a disease in only one cell type (N = 96). For example, the expression of *BIN1* was positively correlated with AD across five immune cell types (i.e., memory B cells, CD4+ naïve and effector memory T cells and CD8+ naïve and effector memory T cells). It was also nominally correlated with LBD only in CD4+ effector memory T cells (Figure 4A). *BIN1* is ubiquitously expressed across multiple tissues, including the brain, and is implicated in AD pathogenesis, possibly through its role in neuron hyperexcitability

(Voskobiynyk et al., 2020). However, *BIN1* expression in B cells has not been associated with AD risk. Our colocalization analysis in the *BIN1* region indicate that there is no colocalization between AD risk and gene expression (H3 > 0.99). This result suggests that different variants in the locus influence either *BIN1* expression in B cells or AD risk (Figure 5), highlighting the importance of complementing significant regional genetic correlations with colocalization analyses.

We observed significant positive correlations between AD and *SCIMP* expression in naïve and memory B cells (Figure 4B). Furthermore, colocalization analysis supported the hypothesis of a single shared causal variant (H4 = 0.99 and 0.85 for naïve and memory B cells, respectively), suggesting that the expression of *SCIMP* in B cells may contribute to AD risk. *SCIMP* is a gene that has been previously associated with immune-mediated diseases, such as lupus and rheumatoid arthritis (Dozmorov et al., 2014), as well as with AD risk (Schwartzentruber et al., 2021; Wightman et al., 2021). The gene encodes a protein expressed in antigen-presenting cells, localized in the immunologic synapse, and serves as a regulator of antigen presentation (Draber et al., 2011). Overall, this result supports a role of the adaptive immune system in AD risk, specifically of B cells, mediated by the expression of genes, such as *SCIMP*.

Amongst all diseases assessed, we observed a relatively higher number of significant correlations between PD and gene expression across all cell types tested (N = 79), many of which were not correlated with other diseases (Supplementary Figure 2). After following up these significant correlations, we observed colocalization (H4 > 0.8) with three genes expressed in one or more cell types (i.e., *RAB7L1, ARSA* and *KANSL1-AS1*) (Figure 5; Supplementary Table 4).

The expression of *RAB7L1* in CD4+ naïve T cells was positively correlated with PD (r_g = 0.826; p-value = 0.001). This gene is a known risk locus for PD (Nalls et al., 2014, 2019), involved in the regulation of the T cell receptor signalling pathway. It has also been shown to interact with *LRRK2* to alter the intraneuronal sorting of proteins and the lysosomal pathway (Kuwahara et al., 2016; MacLeod et al., 2013), suggesting that the overexpression of *RAB7L1* in T cells may increase PD risk through the interaction with *LRRK2*.

The expression of *KANSL1-AS1*, an anti-sense RNA gene, was negatively correlated with PD across all adaptive immune cell types, but the correlation was the strongest for CD8+ T cells (i.e., effector memory T cells: r_g = -0.831; p-value = 1.34e-39, naïve T cells: r_g = -0.768; p-value = 1.13e-25; Supplementary Figure 2). Additionally, colocalization analysis suggested the presence of a shared causal variant at the *KANSL1-AS1* locus (Figure 5; Supplementary Table 4). The protein coding gene *KANSL1* is in the *MAPT* locus, which has been previously associated with PD (Do et al., 2011; Nalls et al., 2014, 2019; Spencer et al., 2011), but recent experimental evidence suggests that the differential expression of another gene in the *MAPT* locus, *KANSL1*, also plays a crucial role in PD risk (Soutar et al., 2022).

Of the initial 366 significant correlations observed across all tested diseases traits and cell types, 92 correlations (25.14%) implicated loci that did not encompass genome-wide significant GWAS variants (p-value ≥ 5e-08). However, 33.7% of these aforementioned loci are suggestive of association (p-value < 1e-06), whereas the remaining 66.3% loci are nominally significant (p-value < 0.05) (Supplementary Table 5). We observed colocalization with only two of these loci: 1) between AD and the expression of *FNBP4* in memory B cells (H4 = 0.913) and in CD8+ T cells (H4 = 0.85 and 0.84, for effector and naïve CD8+ T cells, respectively), and 2) between PD and the expression of *ARSA* in CD8+ effector T cells (H4 = 0.88). *FNBP4* (situated approximately 15,000 base pairs away from *CELF1*) has been previously identified as an AD risk locus (Karch et al., 2016), but in a more recent transcriptome-wide association study (TWAS) of AD, this gene was discarded in conditional analyses (Harwood et al., 2021). *ARSA* has been previously investigated

as a PD risk locus in a Chinese population, in which no significant associations were found with PD susceptibility (Pan et al., 2022). These results provide a new line of in silico evidence, suggesting that the expression of *FNBP4* and *ARSA* in adaptive immune cell types may play a role in AD and PD risk, respectively.

Regional correlations with blood protein levels provide evidence of additional mechanisms involved in disease risk

Proteins contain biologically meaningful information that cannot always be identified by solely assessing the transcriptome. For instance, as the proteome is often dysregulated by diseases, it is amenable to drug targeting and thus a better understanding of the of the proteome could aid in identifying novel treatments (Zhang et al., 2022). Therefore, we performed regional genetic correlations between diseases and protein levels in plasma using a large pQTL database (Zhang et al., 2022), with the aim of exploring an additional level of biological variation and its relation to neurodegenerative disease risk.

We performed a total of 1,863 bivariate tests between diseases and protein levels. We considered a significant regional correlation if FDR < 0.01. We observed significant correlations between protein levels and all diseases, except FTD. PD had a higher number of significant correlations, compared to other tested diseases (Supplementary Figure 3). We evaluated the concordance between the regional genetic correlations performed with gene expression levels from diverse immune cell types and regional genetic correlations performed with protein levels derived from peripheral blood samples (Supplementary Figure 4). A total of 68 unique gene/proteins were evaluated across both datasets (i.e., 68 genes with significant eQTLs also had genome-wide significant pQTLs), resulting in 35 genetic correlations that were at least nominally significant across both tests, including 24 significant correlations across both tests (FDR < 0.01). The direction of effect was consistent across 17 of the significant correlations.

To obtain a biological understanding of the significant correlations with protein levels, specifically to assess if immune-related pathways were significantly enriched, we performed a gene-set enrichment test with FUMA (Watanabe & Taskesen, 2017). We observed enrichment of several gene ontology (GO) biological processes (BP) across immune-mediated diseases, as well seven GO BP enriched for LBD (Supplementary File 1). The enriched GO BP for UC, CD and MS include several immunological processes, whereas the GO BP enriched for LBD correspond to gene-sets related to triglyceride processes. While there was a partial overlap of the GO BP among the three immune-related diseases (5.5%), there was no overlap of GO BP between LBD and other diseases (Supplementary Figure 5). Nevertheless, there were proteins harbouring significant correlations with at least one neurodegenerative disease and at least one of the immune-mediated diseases, none of which were significant in the regional genetic correlations with sc-eQTLs, but which have a function in the immune system (Figure 6).

One of the proteins harbouring a significant correlation with both a neurodegenerative and an immune-mediated disease was Fc fragment of IgG receptor IIa (*FCGR2A*), a cell surface receptor found on phagocytic cells (i.e., neutrophils, macrophages), involved in the process of clearing immune complexes (Dahlqvist et al., 2022). Protein levels of *FCGR2A* were negatively correlated with UC (r_g = -0.675; p-value = 1.44e-20) and positively correlated with PD (r_g = 0.237; p-value = 6.02e-13) (Figure 6). In line with these results, our regional genetic correlations among diseases highlighted a locus on chromosome 1 that encompasses this gene as negatively correlated between PD and UC (Figure 2A). However, there were no significant correlations between the expression of *FCGR2A* and PD or UC, even though the gene harbours genome-wide significant eQTLs in CD8+ effector memory T cells, but not in monocytes (macrophage

precursors). These findings suggest that the protein levels of *FCGR2A* have opposite risk effects in UC and PD, which may be regulated by transcriptome-independent processes. Alternatively, *FCGR2A* may be differentially expressed in another cell type (such as macrophages), which we did not assess in the current study.

The protein levels of Phospholipase C gamma 2 (*PLCG2*) were significantly correlated with UC and AD, in opposite directions (r_gUC = 0.442; p-value = 2.28e-06, and r_gAD = -0.491; p-value = 3.03e-04) (Figure 6). We did not estimate regional genetic correlations between the expression of *PLCG2* and diseases, given that this gene did not harbour genome-wide significant eQTLs in the immune cell types tested. Nonetheless, mutations in the gene *PLCG2* have been associated with dysregulation of the immune system, as well as with several dementias, in which distinct genetic variants are associated with different diseases, based on the identification of different functional point mutations across diseases (Jackson et al., 2021). For instance, the G allele of a missense variant within *PLCG2* has been shown to be protective against AD, LBD and FTD (van der Lee et al., 2019). In contrast to our in silico observations of low *PLCG2* blood protein levels correlated with high AD risk, a recent study showed an upregulation of *PLCG2* expression in post-mortem brains of late-onset AD patients and its association to inflammation in microglia (Tsai et al., 2022). These seemingly discordant findings may be explained by different effects of *PLCG2* across stages of neurodegeneration, as well as by differences across sampled tissues.

Discussion

The aim of the work was to assess the role that peripheral immune cells and related processes play in neurodegenerative diseases. We addressed this aim through orthogonal bioinformatics approaches: i) by applying regional genetic correlations to relate neurodegenerative diseases to diseases known to be driven by immune dysfunction, ii) by extending the correlation analysis by incorporating single-cell eQTLs to identify known gene-disease relationships in immune cell types, and iii) by assessing the evidence for specific genes through expression and pQTL analyses. Through our approach, we identified new links that warrant additional follow-up to better understand immune-mediated loci that may play a role in neurodegenerative diseases, such as the role of *SCIMP* expression in memory B cells as an AD risk locus, and the role of *FCGR2A* blood protein levels, correlated with PD risk.

By performing regional genetic correlations between pairs of diseases, we highlighted relationships between neurodegenerative diseases across loci encompassing known risk genes (e.g., *BIN1*, *TMEM175*, *APOE*). Additionally, we highlighted relationships between neurodegenerative and immune-mediated diseases, suggesting the presence of shared immune-related biological pathways across these diseases (e.g. *FCGR2A*, *CLCN3*, *IKZF1*). The gene *FCGR2A*, for instance, located within a locus significantly correlated between PD and UC, has been previously associated with immune-mediated diseases (Witoelar et al., 2017), and is significantly expressed in CD8+ effector T cells and memory B cells (Yazar et al., 2022). Although our regional genetic correlations with gene expression levels indicated that the expression of *FCGR2A* is correlated with neither PD nor UC risk, we observed significant correlations with the corresponding protein levels for both PD and UC. Similarly, the gene *IKZF1* is within an LD locus significantly correlated between MS and AD, and is significantly expressed in CD8+ naïve T cells (Yazar et al., 2022). However, regional genetic correlations with gene expression levels showed no evidence of significant correlation with AD (r_g = 0.113; p-value = 0.487), whereas the correlation test with MS was not performed due to lack of significant univariate signal.

The observations at these loci (i.e., *FCGR2A* and *IKZF1*) suggest that different cell types or alternative molecular mechanisms may be involved in disease risk. In fact, we observed only a partial overlap and concordance of loci evaluated in both gene expression and protein regional genetic correlations, in line with our expectations, given that the eQTL and pQTL datasets used were generated through different sources: single-cell RNA-sequencing from specific immune cell types and bulk blood tissue, respectively.

Our regional analysis of genetic correlations with gene expression levels shed light on overall differences among diseases, including the varying proportion of correlations accounted for by gene expression in immune cell types. For example, a higher proportion of significant correlations between AD and sc-eQTLs were accounted for by genes expressed in classical monocytes, including genes within and outside of the HLA region. These results are in line with previous evidence pointing at a key role of the innate immune system (i.e., microglia) in AD. However, it has also been suggested that circulating monocytes participate in the clearance of Aβ plaques that diffuse into the bloodstream (Chen et al., 2020; Xiang et al., 2015), and that monocyte-derived macrophages have a more efficacious phagocytic capacity than microglia in the brain (Malm et al., 2010; Thériault et al., 2015). Therefore, aside from the crucial role microglia play in AD, peripheral innate immune cells may be independently contributing to AD risk via changes in transcription levels.

Our gene set enrichment analysis, aimed at better understanding the significant correlations observed between diseases and protein levels, did not highlight significant immune-related biological pathways enriched for neurodegenerative diseases. However, we identified proteins across neurodegenerative and immune-related diseases for follow-up. We believe that novel pathway enrichment methods that consider gene-specific weights (i.e. weights dependent on a measure of the strength of the regional correlations) could provide an important avenue for follow-up, alongside the current methods that consider all genes as input to the analysis as having equal effects.

Through our data-driven approach, we provide fine resolution links for genomic regions to a disease in a particular cell type to better understand the etiology of neurodegenerative diseases in relation to the peripheral immune system. However, this approach is not without limitations. First, our analyses used GWAS and QTL datasets of inferred European genetic ancestry, which is a limitation stemming from the lack of diversity in GWAS cohorts (Fatumo et al., 2022). Genetic ancestry may be particularly important for immune function given different selection pressures placed by infectious diseases. Even though GWAS of neurodegenerative diseases have been performed in cohorts of other genetic ancestries (van Rheenen et al., 2021), the sample sizes needed to reach sufficient power to identify significant correlations falls short, as we observed in the case of the smaller FTD and LBD GWAS datasets. Second, it is known that there are sex differences in the incidence of some neurodegenerative and immune-mediated diseases, but the GWAS datasets used do not include sex-stratified analyses or sex chromosome data, which is a limiting factor in the identification of (i) sex-specific or sex-skewed expressed genes, or (ii) candidate immune-related genes on the sex chromosomes. Future studies that consider sex-chromosomes or sex differences may provide new insights on underlying mechanisms or cell types involved in disease pathogenesis. Third, mechanisms other than varying gene expression could be responsible for the absence of particular eQTL-disease correlations, which we were not able to capture with our approach. One such example is the absence of genome-wide significant eQTLs for *LRRK2* in any of the immune cell types tested, a gene in which missense point mutations have been associated with PD risk (Ross et al., 2011). Finally, our main analyses are based on correlations, which cannot assess causal relationship between diseases and the

molecular mechanisms assessed. Nonetheless, we have highlighted immune-related genes as clear candidates for further investigation to better understand neurodegenerative diseases.

Acknowledgements

We thank the participants from all cohorts who contributed to the study. This research was enabled in part by support provided by Calcul Québec (https://www.calculquebec.ca/) and the Digital Research Alliance of Canada (alliancecan.ca). SAGT was funded by a Junior 1 award from the Fonds de recherche du Québec - Santé (FRQS; https://frq.gouv.qc.ca) and by Operational Funds from the Institut de valorisation des données (IVADO; https://ivado.ca). SAGT and FLD acknowledge support from a Canadian Institutes of Health Research (CIHR) Project Grant (PJT 183817). SWS was supported in part by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke, National Institutes of Health (program #: ZIANS003154). MR was supported through the award of a UKRI Medical Research Council Clinician Scientist Fellowship (MRC Grant Code: MR/N008324/1).

Declaration of Interests

SWS serves on the Scientific Advisory Council of the Lewy Body Dementia Association and the Multiple System Atrophy Coalition. She receives research support from Cerevel Therapeutics.

Author Contributions

FLD, SAGT, SWS and MR conceived and designed the study. FLD analysed data and drafted the figures. RHR and FLD implemented the computer code. SAGT supervised the research. SWS and MR provided clinical insight to data interpretation. FLD and SAGT wrote the initial manuscript. All authors contributed to the critical analysis and revision of the manuscript.

Inclusion and Diversity

We support inclusive, diverse, and equitable conduct of research.

Figures

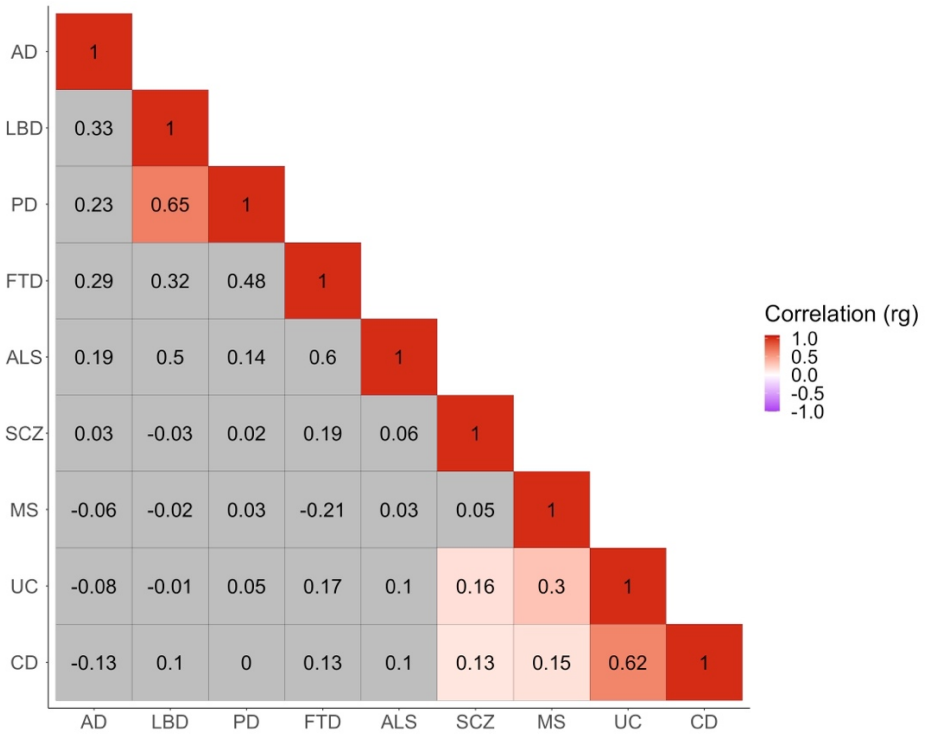

Figure 1. Genome-wide genetic correlations (r_g) across GWAS traits. Significant positive correlations (Bonferroni corrected p-value < 0.0014) are highlighted in shades of red. AD = Alzheimer's disease; LBD = Lewy body dementia; PD = Parkinson's disease; FTD = Frontotemporal dementia; ALS = Amyotrophic lateral sclerosis; SCZ = Schizophrenia; MS = Multiple sclerosis; UC = Ulcerative colitis; CD = Crohn's disease.

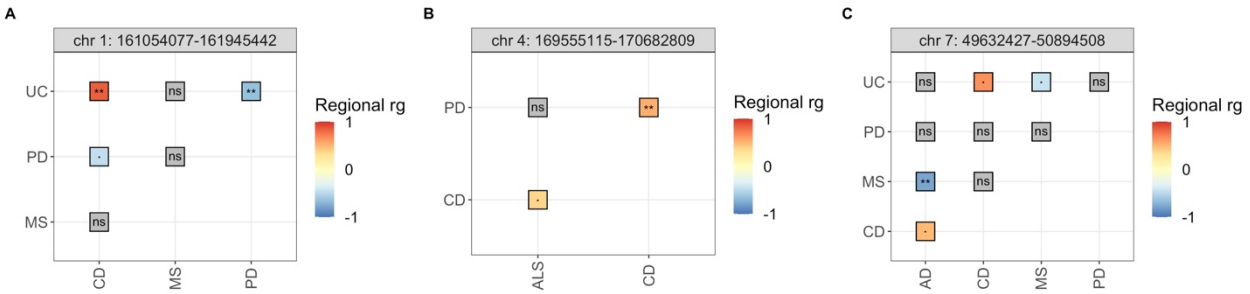

Figure 2. Significant regional genetic correlations (rg) between neurodegenerative and immune-mediated diseases across three distinct loci in chromosomes 1 (A), 4 (B) and 7 (C). ** = significant correlations (p-value < 2.629e-05); · = nominally significant correlations; ns = non-significant correlations (p-value $\geq$ 0.05). Absence of correlations means that the bivariate test was not performed (See Methods for details). CD = Crohn's disease; MS = Multiple sclerosis; UC = Ulcerative colitis; PD = Parkinson's disease.

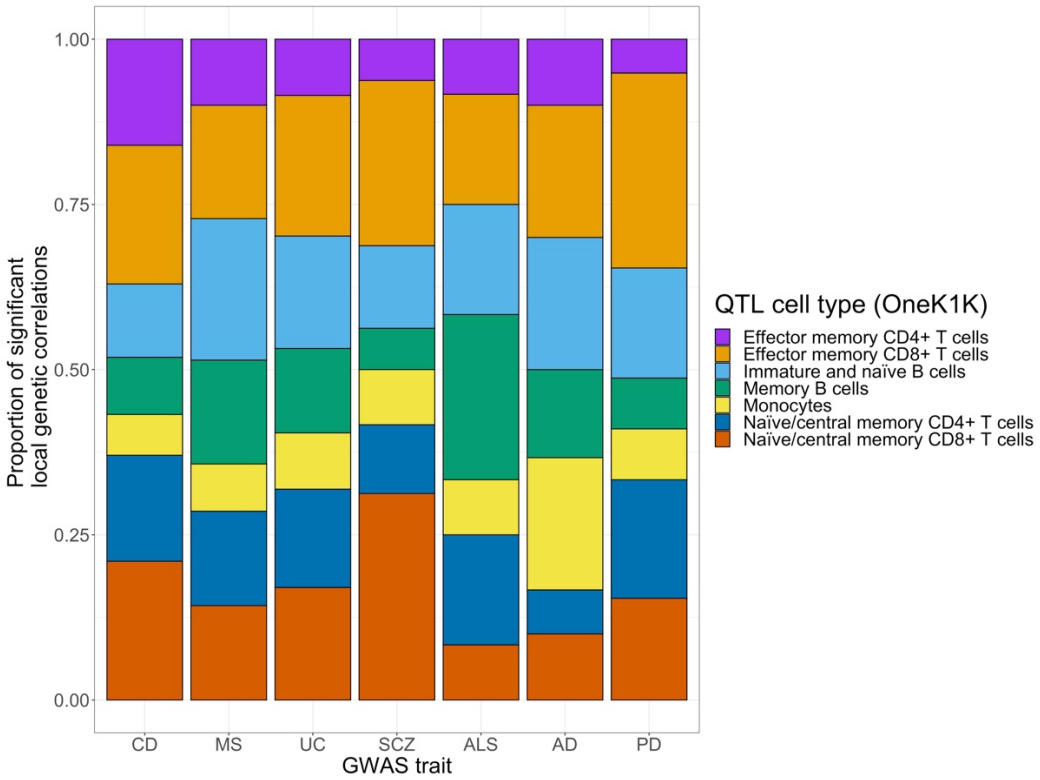

Figure 3. Proportion of significant regional genetic correlations between diseases and gene expression, across seven immune cell types. As there were no significant correlations for neither frontotemporal dementia nor Lewy body dementia, those traits are not displayed. CD = Crohn's disease; MS = Multiple sclerosis; UC = Ulcerative colitis; SCZ = Schizophrenia; ALS = Amyotrophic lateral sclerosis; AD = Alzheimer's disease; PD = Parkinson's disease.

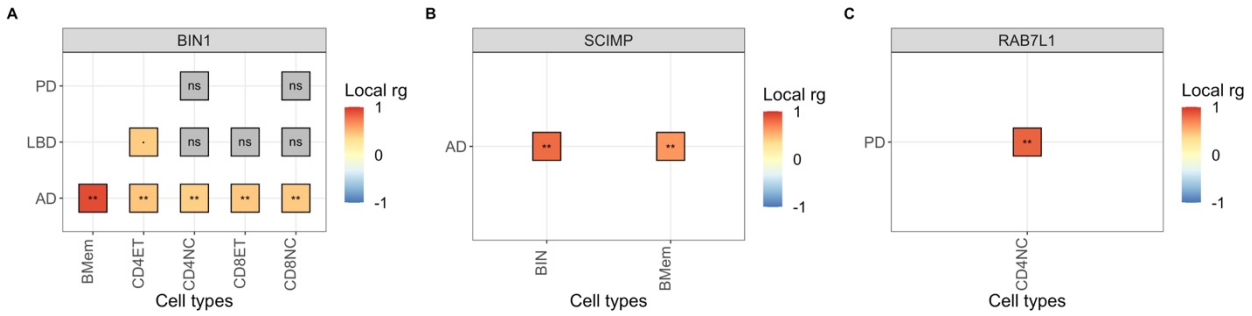

Figure 4. Examples of significant regional correlations between gene expression levels and diseases across three genes. (A) *BIN1* (chromosome 2) was significantly correlated with AD across various adaptive immune cell types. (B) *SCIMP* (chromosome 17) was significantly correlated with AD across B cells. (C) *RAB7L1* (chromosome 1) was significantly correlated with PD in naive CD4+ T cells. ** = significant correlations (FDR < 0.01); · = nominally significant correlations; ns = non-significant correlations (p-value $\geq$ 0.05). AD = Alzheimer's disease; LBD= Lewy body dementia; PD = Parkinson's disease; BMem = Memory B cells; CD4ET = CD4+ effector memory T cells; CD4NC = CD4+ naive T cells; CD8ET = CD8+ effector memory T cells; CD8NC = CD8+ naïve T cells.

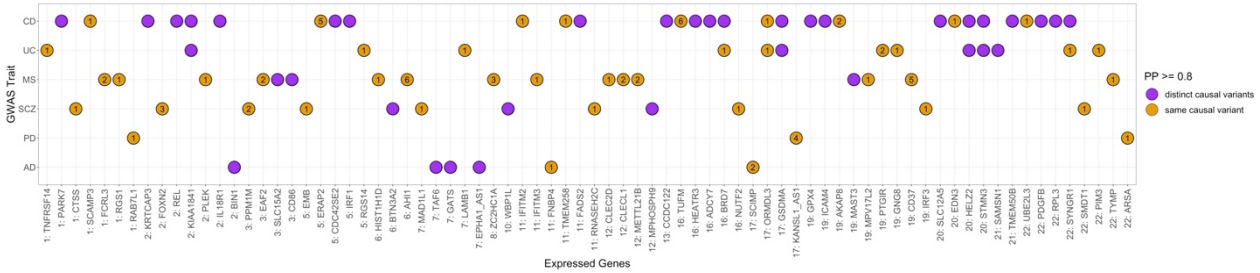

Figure 5. Colocalization results between diseases and expressed genes in at least one immune cell type. Pairs of traits that share a causal variant through colocalization analysis are highlighted in orange (Posterior Probability, PP H4 $\geq$ 0.8), whereas pairs of traits that have distinct causal variants are highlighted in purple (PP H3 $\geq$ 0.8). The number inside the colocalized signals indicates the number of cell types for which a colocalization was observed. The cell type(s) for each disease-gene pair displayed here are listed in Supplementary Table 4.

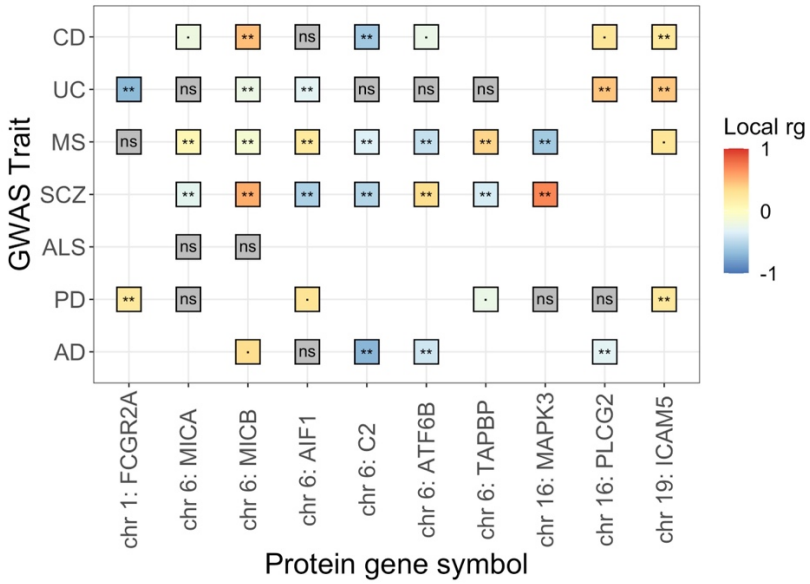

Figure 6. Regional correlations between protein levels and diseases shared between at least one test disease (i.e., AD, PD, ALS, LBD, FTD, SCZ) and at least one control disease (i.e., MS, UC, CD). ** = significant correlations (FDR < 0.01); · = nominally significant correlations; ns = non-significant correlations (p-value $\geq$ 0.05).

References

Blauwendraat, C., Heilbron, K., Vallerga, C. L., Bandres-Ciga, S., von Coelln, R., Pihlstrøm, L., Simon-Sanchez, J., Schulte, C., Sharma, M., Krohn, L., & Siitonen, A. (2019). Parkinson disease age at onset GWAS: defining heritability, genetic loci and alpha-synuclein mechanisms. *Mov Disord*, *34*(6), 866–875. https://doi.org/10.1002/mds.27659.Parkinson

Broce, I., Karch, C. M., Wen, N., Fan, C. C., Wang, Y., Tan, C. H., Kouri, N., Ross, O. A., Höglinger, G. U., Muller, U., Hardy, J., Momeni, P., Hess, C. P., Dillon, W. P., Miller, Z. A., Bonham, L. W., Rabinovici, G. D., Rosen, H. J., Schellenberg, G. D., … Sugrue, L. P. (2018). Immune-related genetic enrichment in frontotemporal dementia: An analysis of genome-wide association studies. *PLoS Medicine*, *15*(1), e1002504. https://doi.org/10.1371/journal.pmed.1002504

Bulik-Sullivan, B. K., Loh, P., Finucane, H. K., Ripke, S., Yang, J., Working, S., Consortium, G., Patterson, N., Daly, M. J., Price, A. L., & Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Publishing Group*, *47*(3). https://doi.org/10.1038/ng.3211

Chen, S. H., Tian, D. Y., Shen, Y. Y., Cheng, Y., Fan, D. Y., Sun, H. L., He, C. Y., Sun, P. Y., Bu, X. le, Zeng, F., Liu, J., Deng, J., Xu, Z. Q., Chen, Y., & Wang, Y. J. (2020). Amyloid-beta uptake by blood monocytes is reduced with ageing and Alzheimer's disease. *Translational Psychiatry*, *10*(1). https://doi.org/10.1038/s41398-020-01113-9

Chia, R., Sabir, M. S., Bandres-Ciga, S., Saez-Atienzar, S., Reynolds, R. H., Gustavsson, E., Walton, R. L., Ahmed, S., Viollet, C., Ding, J., Makarious, M. B., Diez-Fairen, M., Portley, M. K., Shah, Z., Abramzon, Y., Hernandez, D. G., Blauwendraat, C., Stone, D. J., Eicher, J., … Scholz, S. W. (2021). Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. *Nature Genetics*, *53*(3), 294–303. https://doi.org/10.1038/s41588-021-00785-3

Dahlqvist, J., Fulco, C. P., Ray, J. P., Liechti, T., de Boer, C. G., Lieb, D. J., Eisenhaure, T. M., Engreitz, J. M., Roederer, M., & Hacohen, N. (2022). Systematic identification of genomic elements that regulate FCGR2A expression and harbor variants linked with autoimmune disease. *Human Molecular Genetics*, *31*(12), 1946–1961. https://doi.org/10.1093/hmg/ddab372

de Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., Jostins, L., Rice, D. L., Gutierrez-Achury, J., Ji, S. G., Heap, G., Nimmo, E. R., Edwards, C., Henderson, P., Mowat, C., Sanderson, J., Satsangi, J., Simmons, A., Wilson, D. C., … Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics*, *49*(2), 256–261. https://doi.org/10.1038/ng.3760

Do, C. B., Tung, J. Y., Dorfman, E., Kiefer, A. K., Drabant, E. M., Francke, U., Mountain, J. L., Goldman, S. M., Tanner, C. M., Langston, J. W., Wojcicki, A., & Eriksson, N. (2011). Web-based genome-wide association study identifies two novel loci and a substantial genetic component for parkinson's disease. *PLoS Genetics*, *7*(6). https://doi.org/10.1371/journal.pgen.1002141

Dozmorov, M. G., Wren, J. D., & Alarcón-Riquelme, M. E. (2014). Epigenomic elements enriched in the promoters of autoimmunity susceptibility genes. *Epigenetics*, *9*(2). https://doi.org/10.4161/epi.27021

Draber, P., Vonkova, I., Stepanek, O., Hrdinka, M., Kucova, M., Skopcova, T., Otahal, P., Angelisova, P., Horejsi, V., Yeung, M., Weiss, A., & Brdicka, T. (2011). SCIMP, a Transmembrane Adaptor Protein Involved in Major Histocompatibility Complex Class II Signaling. *Molecular and Cellular Biology*, *31*(22), 4550–4562. https://doi.org/10.1128/mcb.05817-11

Escott-Price, V., & Hardy, J. (2022). Genome-wide association studies for Alzheimer's disease: Bigger is not always better. *Brain Communications*, *4*(3), 1–7. https://doi.org/10.1093/braincomms/fcac125

Fatumo, S., Chikowore, T., Choudhury, A., Ayub, M., Martin, A. R., & Kuchenbaecker, K. (2022). A roadmap to increase diversity in genomic studies. *Nature Medicine*, *28*(2), 243–250. https://doi.org/10.1038/s41591-021-01672-4

Ferrari, R., Hernandez, D. G., Nalls, M. A., Rohrer, J. D., Ramasamy, A., Kwok, J. B. J., Dobson-Stone, C., Brooks William S., B. S., Schofield, P. R., Halliday, G. M., Hodges, J. R., Piguet, O., Bartley, L., Thompson, E., Haan, E., Hernández, I., Ruiz, A., Boada, M., Borroni, B., … Momeni, P. (2014). Frontotemporal dementia and its subtypes: A genome-wide association study. *The Lancet Neurology*, *13*(7), 686–699. https://doi.org/10.1016/S1474-4422(14)70065-1

Gagliano, S. A., Pouget, J. G., Hardy, J., Knight, J., Barnes, M. R., Ryten, M., & Weale, M. E. (2016). Genomics implicates adaptive and innate immunity in Alzheimer's and Parkinson's diseases. *Annals of Clinical and Translational Neurology*, *3*(12), 924–933. https://doi.org/10.1002/acn3.369

Gate, D., Tapp, E., Leventhal, O., Shahid, M., Nonninger, T. J., Yang, A. C., Strempfi, K., Unger, M. S., Fehlmann, T., Oh, H., Channappa, D., Henderson, V. W., Keller, A., Aigner, L., Galasko, D. R., Davis, M. M., Poston, K. L., & Wyss-Coray, T. (2021). CD4+ T Cells Contribute to Neurodegeneration in Lewy Body Dementia. *Science*, *374*, 868–874. https://doi.org/10.1002/mds.28881

Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., & Hingorani, A. D. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genetics*, *10*(5). https://doi.org/10.1371/journal.pgen.1004383

Harwood, J. C., Leonenko, G., Sims, R., Escott-Price, V., Williams, J., & Holmans, P. (2021). Defining functional variants associated with Alzheimer's disease in the induced immune response. *Brain Communications*, *3*(2), 1–15. https://doi.org/10.1093/braincomms/fcab083

Herrick, M. K., & Tansey, M. G. (2021). Is LRRK2 the missing link between inflammatory bowel disease and Parkinson's disease? *Npj Parkinson's Disease*, *7*(1). https://doi.org/10.1038/s41531-021-00170-1

Hop, P. J., Zwamborn, R. A. J., Hannon, E., Shireby, G. L., Nabais, M. F., Walker, E. M., Rheenen, W. van, Vugt, J. J. F. A. van, Dekker, A. M., Westeneng, H., Tazelaar, G. H. P., Eijk, K. R. van, Moisse, M., Baird, D., Khleifat, A. al, Iacoangeli, A., Ticozzi, N., Ratti, A., Cooper-knock, J., … Veldink, J. H. (2022). Genome-wide study of DNA methylation shows alterations in metabolic , inflammatory , and cholesterol pathways in ALS. *Science Translational Medicine*, *14*(i), eabj0264. www.science.org/doi/10.1126/scitranslmed.abj0264

IMSGC, Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C. A., Patsopoulos, N. A., Moutsianas, L., Su, Z., Freeman, C., Hunt, S. E., Edkins, S., Gray, E., David, R., Potter, S. C., Goris, A., Band, G., Oturai, A. B., Strange, A., Comabella, M., … Lisa, F. (2012). *Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis*. *476*(7359), 214–219. https://doi.org/10.1038/nature10251.Genetic

Jackson, J. T., Mulazzani, E., Nutt, S. L., & Masters, S. L. (2021). The role of PLCγ2 in immunological disorders, cancer, and neurodegeneration. *Journal of Biological Chemistry*, *297*(2), 100905. https://doi.org/10.1016/j.jbc.2021.100905

Karch, C. M., Ezerskiy, L. A., Bertelsen, S., Goate, A. M., Albert, M. S., Albin, R. L., Apostolova, L. G., Arnold, S. E., Baldwin, C. T., Barber, R., Barmada, M. M., Barnes, L. L., Beach, T. G., Beecham, G. W., Beekly, D., Bennett, D. A., Bigio, E. H., Bird, T. D., Blacker, D., … Yu, L. (2016). Alzheimer's disease risk polymorphisms regulate gene expression in the ZCWPW1 and the CELF1 loci. *PLoS ONE*, *11*(2), 1–22. https://doi.org/10.1371/journal.pone.0148717

Kim, K., Wang, X., Ragonnaud, E., Bodogai, M., Illouz, T., DeLuca, M., McDevitt, R. A., Gusev, F., Okun, E., Rogaev, E., & Biragyn, A. (2021). Therapeutic B-cell depletion reverses progression of Alzheimer's disease. *Nature Communications*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-22479-4

Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., Boland, A., Vronskaya, M., van der Lee, S. J., Amlie-Wolf, A., Bellenguez, C., Frizatti, A., Chouraki, V., Martin, E. R., Sleegers, K., Badarinarayan, N., Jakobsdottir, J., Hamilton-Nelson, K. L., Moreno-Grau, S., … Pericak-Vance, M. A. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*, *51*(3), 414–430. https://doi.org/10.1038/s41588-019-0358-2

Kuwahara, T., Inoue, K., D'Agati, V. D., Fujimoto, T., Eguchi, T., Saha, S., Wolozin, B., Iwatsubo, T., & Abeliovich, A. (2016). LRRK2 and RAB7L1 coordinately regulate axonal morphology and lysosome integrity in diverse cellular contexts. *Scientific Reports*, *6*(June), 1–12. https://doi.org/10.1038/srep29945

Liu, J. Z., Erlich, Y., & Pickrell, J. K. (2017). Case-control association mapping by proxy using family history of disease. *Nature Genetics*, *49*(3), 325–331. https://doi.org/10.1038/ng.3766

MacLeod, D. A., Rhinn, H., Kuwahara, T., Zolin, A., di Paolo, G., MacCabe, B. D., Marder, K. S., Honig, L. S., Clark, L. N., Small, S. A., & Abeliovich, A. (2013). RAB7L1 Interacts with LRRK2 to Modify Intraneuronal Protein Sorting and Parkinson's Disease Risk. *Neuron*, *77*(3), 425–439. https://doi.org/10.1016/j.neuron.2012.11.033

Malm, T., Koistinaho, M., Muona, A., Magga, J., & Koistinaho, J. (2010). The role and therapeutic potential of monocytic cells in Alzheimer's disease. *Glia*, *58*(8), 889–900. https://doi.org/10.1002/glia.20973

Nalls, M. A., Blauwendraat, C., Vallerga, C. L., Heilbron, K., Bandres-Ciga, S., Chang, D., Tan, M., Kia, D. A., Noyce, A. J., Xue, A., Bras, J., Young, E., von Coelln, R., Simón-Sánchez, J., Schulte, C., Sharma, M., Krohn, L., Pihlstrøm, L., Siitonen, A., … Zhang, F. (2019). Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *The Lancet Neurology*, *18*(12), 1091–1102. https://doi.org/10.1016/S1474-4422(19)30320-5

Nalls, M. A., Pankratz, N., Lill, C. M., Do, C. B., Hernandez, D. G., Saad, M., Destefano, A. L., Kara, E., Bras, J., Sharma, M., Schulte, C., Keller, M. F., Arepalli, S., Letson, C., Edsall, C., Stefansson, H., Liu, X., Pliner, H., Lee, J. H., … Ansorge, O. (2014). Large-scale meta-analysis of genome-wide association

data identifies six new risk loci for Parkinson's disease. *Nature Genetics*, *46*(9), 989–993. https://doi.org/10.1038/ng.3043

Pagan, F., Hebron, M., Valadez, E. H., Torres-Yaghi, Y., Huang, X., Mills, R. R., Wilmarth, B. M., Howard, H., Dunn, C., Carlson, A., Lawler, A., Rogers, S. L., Falconer, R. A., Ahn, J., Li, Z., & Moussa, C. (2016). Nilotinib effects in Parkinson's disease and dementia with lewy bodies. *Journal of Parkinson's Disease*, *6*(3), 503–517. https://doi.org/10.3233/JPD-160867

Pagès, H. (2017). *SNPlocs.Hsapiens.dbSNP144.GRCh37: SNP locations for Homo sapiens (dbSNP Build 144)*.

Pan, H. xu, Wang, Y. ge, Zhao, Y. wen, Zeng, Q., Wang, Z., Fang, Z. huan, Zhang, Y., Zhou, X., He, R. cheng, Xu, Q., Sun, Q. ying, Tan, J. qiong, Yan, X. xiang, Li, J. chen, Tang, B. sha, & Guo, J. feng. (2022). Evaluating the role of ARSA in Chinese patients with Parkinson's disease. *Neurobiology of Aging*, *109*, 269–272. https://doi.org/10.1016/j.neurobiolaging.2021.08.008

Pardiñas, A. F. (2018). *Common schizophrenia alleles are enriched in mutation- intolerant genes and in regions under strong background selection*. *50*(4), 559–571. https://doi.org/10.1038/s41588-018-0084-1.Refining

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

RHReynolds. (2022). *Common utility functions*. https://doi.org/10.5281/zenodo.6127446

Ross, O. A., Soto-Ortolaza, A. I., Heckman, M. G., Aasly, J. O., Abahuni, N., Annesi, G., Bacon, J. A., Bardien, S., Bozi, M., Brice, A., Brighina, L., van Broeckhoven, C., Carr, J., Chartier-Harlin, M. C., Dardiotis, E., Dickson, D. W., Diehl, N. N., Elbaz, A., Ferrarese, C., … Farrer, M. J. (2011). Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: A case-control study. *The Lancet Neurology*, *10*(10), 898–908. https://doi.org/10.1016/S1474-4422(11)70175-2

Schwartzentruber, J., Cooper, S., Liu, J. Z., Barrio-Hernandez, I., Bello, E., Kumasaka, N., Young, A. M. H., Franklin, R. J. M., Johnson, T., Estrada, K., Gaffney, D. J., Beltrao, P., & Bassett, A. (2021). Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nature Genetics*, *53*(3), 392–402. https://doi.org/10.1038/s41588-020-00776-w

Seshadri, S., Fitzpatrick, A. L., Ikram, M. A., DeStefano, A. L., Gudnason, V., Boada, M., Bis, J. C., Smith, A. v., Carrasquillo, M. M., Lambert, J. C., Harold, D., Schrijvers, E. M. C., Ramirez-Lorca, R., Debette, S., Longstreth, W. T., Janssens, A. C. J. W., Pankratz, V. S., Dartigues, J. F., Hollingworth, P., … Breteler, M. M. B. (2010). Genome-wide Analysis of Genetic Loci Associated with Alzheimer Disease. *JAMA*, *303*(18), 1832–1840. http://jama.jamanetwork.com.eproxy1.lib.hku.hk/article.aspx?articleid=185849

Soutar, M. P. M., Melandri, D., O'Callaghan, B., Annuario, E., Monaghan, A. E., Welsh, N. J., D'Sa, K., Guelfi, S., Zhang, D., Pittman, A., Trabzuni, D., Verboven, A. H. A., Pan, K. S., Kia, D. A., Bictash, M., Gandhi, S., Houlden, H., Cookson, M. R., Kasri, N. N., … Plun-Favreau, H. (2022). Regulation of mitophagy by the NSL complex underlies genetic risk for Parkinson's disease at 16q11.2 and MAPT H1 loci. *Brain*. https://doi.org/10.1093/brain/awac325

Spencer, C. C. A., Plagnol, V., Strange, A., Gardner, M., Paisan-Ruiz, C., Band, G., Barker, R. A., Bellenguez, C., Bhatia, K., Blackburn, H., Blackwell, J. M., Bramon, E., Brown, M. A., Brown, M. A., Burn, D., Casas, J. P., Chinnery, P. F., Clarke, C. E., Corvin, A., … Wood, N. W. (2011). Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Human Molecular Genetics*, *20*(2), 345–353. https://doi.org/10.1093/hmg/ddq469

Tansey, M. G., Wallings, R. L., Houser, M. C., Herrick, M. K., Keating, C. E., & Joers, V. (2022). Inflammation and immune dysfunction in Parkinson disease. *Nature Reviews Immunology*, *0123456789*. https://doi.org/10.1038/s41577-022-00684-6

Thériault, P., Elali, A., & Rivest, S. (2015). The dynamics of monocytes and microglia in Alzheimer's disease. *Alzheimer's Research and Therapy*, *7*(1), 1–10. https://doi.org/10.1186/s13195-015-0125-2

Tsai, A. P., Dong, C., Lin, P. B. C., Messenger, E. J., Casali, B. T., Moutinho, M., Liu, Y., Oblak, A. L., Lamb, B. T., Landreth, G. E., Bissel, S. J., & Nho, K. (2022). PLCG2 is associated with the inflammatory response and is induced by amyloid plaques in Alzheimer's disease. *Genome Medicine*, *14*(1), 1–13. https://doi.org/10.1186/s13073-022-01022-0

van der Lee, S. J., Conway, O. J., Jansen, I., Carrasquillo, M. M., Kleineidam, L., van den Akker, E., Hernández, I., van Eijk, K. R., Stringa, N., Chen, J. A., Zettergren, A., Andlauer, T. F. M., Diez-Fairen, M., Simon-Sanchez, J., Lleó, A., Zetterberg, H., Nygaard, M., Blauwendraat, C., Savage, J. E., … Holstege, H. (2019). A nonsynonymous mutation in PLCG2 reduces the risk of Alzheimer's disease, dementia with Lewy bodies and frontotemporal dementia, and increases the likelihood of longevity. *Acta Neuropathologica*, *138*(2), 237–250. https://doi.org/10.1007/s00401-019-02026-8

van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H., & Wray, N. R. (2019). Genetic correlations of polygenic disease traits: from theory to practice. *Nature Reviews Genetics*, *20*(10), 567–581. https://doi.org/10.1038/s41576-019-0137-z

van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., van der Spek, R. A. A., Võsa, U., de Jong, S., Robinson, M. R., Yang, J., Fogh, I., van Doormaal, P. T. C., Tazelaar, G. H. P., Koppers, M., Blokhuis, A. M., Sproviero, W., Jones, A. R., Kenna, K. P., … Veldink, J. H. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature Genetics*, *48*(9), 1043–1048. https://doi.org/10.1038/ng.3622

van Rheenen, W., van der Spek, R. A. A., Bakker, M. K., van Vugt, J. J. F. A., Hop, P. J., Zwamborn, R. A. J., de Klein, N., Westra, H.-J., Bakker, O. B., Deelen, P., Shireby, G., Hannon, E., Moisse, M., Baird, D., Restuadi, R., Delzhenko, E., Dekker, A. M., Gawor, K., & Westeneng, H.-J. (2021). Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. *Nature Genetics*, *53*(December).

Võsa, U., Claringbould, A., Westra, H. J., Bonder, M. J., Deelen, P., Zeng, B., Kirsten, H., Saha, A., Kreuzhuber, R., Yazar, S., Brugge, H., Oelen, R., de Vries, D. H., van der Wijst, M. G. P., Kasela, S., Pervjakova, N., Alves, I., Favé, M. J., Agbessi, M., … Franke, L. (2021). Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nature Genetics*, *53*(9), 1300–1310. https://doi.org/10.1038/s41588-021-00913-z

Voskobiynyk, Y., Roth, J. R., Cochran, J. N., Rush, T., Carullo, N. V. N., Mesina, J. S., Waqas, M., Vollmer, R. M., Day, J. J., McMahon, L. L., & Roberson, E. D. (2020). Alzheimer's disease risk gene BIN1 induces Tau-dependent network hyperexcitability. *ELife*, *9*, 1–25. https://doi.org/10.7554/eLife.57354

Wallace, C. (2021). A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genetics*, *17*(9), 1–11. https://doi.org/10.1371/journal.pgen.1009440

Watanabe, K., & Taskesen, E. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1826), 1–10. https://doi.org/10.1038/s41467-017-01261-5

Werme, J., van der Sluis, S., Posthuma, D., & de Leeuw, C. A. (2022). An integrated framework for local genetic correlation analysis. *Nature Genetics*, *54*(3), 274–282. https://doi.org/10.1038/s41588-022-01017-y

Wightman, D. P., Jansen, I. E., Savage, J. E., Shadrin, A. A., Bahrami, S., Holland, D., Rongve, A., Børte, S., Winsvold, B. S., Drange, O. K., Martinsen, A. E., Skogholt, A. H., Willer, C., Bråthen, G., Bosnes, I., Nielsen, J. B., Fritsche, L. G., Thomas, L. F., Pedersen, L. M., … Posthuma, D. (2021). A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nature Genetics*, *53*(9), 1276–1282. https://doi.org/10.1038/s41588-021-00921-z

Witoelar, A., Jansen, I. E., Wang, Y., Desikan, R. S., Gibbs, J. R., Blauwendraat, C., Thompson, W. K., Hernandez, D. G., Djurovic, S., Schork, A. J., Bettella, F., Ellinghaus, D., Franke, A., Lie, B. A., McEvoy, L. K., Karlsen, T. H., Lesage, S., Morris, H. R., Brice, A., … Weale, M. (2017). Genome-wide pleiotropy between Parkinson disease and autoimmune diseases. *JAMA Neurology*, *74*(7), 780–792. https://doi.org/10.1001/jamaneurol.2017.0469

Xiang, Y., Bu, X. le, Liu, Y. H., Zhu, C., Shen, L. L., Jiao, S. S., Zhu, X. Y., Giunta, B., Tan, J., Song, W. H., Zhou, H. D., Zhou, X. F., & Wang, Y. J. (2015). Physiological amyloid-beta clearance in the periphery and its therapeutic potential for Alzheimer's disease. *Acta Neuropathologica*, *130*(4), 487–499. https://doi.org/10.1007/s00401-015-1477-1

Yazar, S., Alquicira-Hernandez, J., Wing, K., Senabouth, A., Gordon, M. G., Andersen, S., Lu, Q., Rowson, A., Taylor, T. R. P., Clarke, L., Maccora, K., Chen, C., Cook, A. L., Ye, C. J., Fairfax, K. A., Hewitt, A. W., & Powell, J. E. (2022). Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science*, *376*(6589). https://doi.org/10.1126/science.abf3041

Zhang, J., Dutta, D., Köttgen, A., Tin, A., Schlosser, P., Grams, M. E., Harvey, B., Yu, B., Boerwinkle, E., Coresh, J., & Chatterjee, N. (2022). Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. *Nature Genetics*, *54*(5), 593–602. https://doi.org/10.1038/s41588-022-01051-w