Genetic variability in both the adaptive and innate immune systems contribute to Alzheimer's and Parkinson's disease risk

Running head: Immune system and neurodegenerative disorders

Authors:
Sarah A Gagliano, PhD [1,2,3]
Jennie G Pouget, PhD [2,3]
John Hardy, PhD [4]
Michael R Barnes, PhD [5,*]
Jo Knight, PhD [6,*]
Mina Ryten, MBPhD [4,*]
Michael E Weale, PhD [1,*]

Affiliations:
1: King's College London, Dept. of Medical & Molecular Genetics 8th Floor, Tower Wing, Guy's Hospital, London SE1 9RT, United Kingdom
2: Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, 250 College Street, Toronto, ON, M5T 1R8, Canada
3: Institute of Medical Science, University of Toronto, 1 King's College Circle, Room 2374, Toronto, ON, M5S 1A8, Canada
4: Institute of Neurology, University College London, Queen Square, London WC1N 3BG, United Kingdom
5: William Harvey Research Institute Barts and The London School of Medicine and Dentistry, Queen Mary University of London, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ, United Kingdom
6: Faculty of Health and Medicine and Data Science Institute, Furness College, Lancaster University, Lancaster, LA1 4YG, United Kingdom
* Equal authorship

Corresponding author: Sarah A Gagliano
gsarah6221@gmail.com

Abstract

Neurodegenerative disorders are devastating diseases with a worldwide health-care burden. Studies have demonstrated enrichment of disease-associated genetic variants with functional genomic annotations. Determining associated cell-types is important to understand pathogenicity.

We obtained GWAS summary statistics from Parkinson's disease (PD), Alzheimer's disease (AD), amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS), and frontotemporal dementia (FTD). We applied stratified LD score regression to determine if functional categories are enriched for heritability.

There was little enrichment of brain annotations, but annotations from both the innate and adaptive immune systems were enriched for MS (as expected), AD, and PD, in decreasing order of statistical significance.

Introduction

Neurodegenerative disorders, such as Alzheimer's (AD) and Parkinson's disease (PD), are personally devastating, and a burden on health-care systems worldwide. Previous studies have demonstrated enrichment of disease-associated variants (for numerous diseases) with functional genomic annotations, including DNase I hypersensitive sites, transcription factor binding sites, histone modifications, and expression quantitative trait loci (eQTLs).[1,2] These annotations vary depending on cell/tissue-type. Given the many ways in which complex disorders arise, and for human brain disorders, the well-recognized cellular heterogeneity of the brain, pinpointing cell-types of interest is important to further understand pathogenicity.

Recently there has been much progress in identifying genetic variants associated with neurodegenerative disorders. The latest PD meta-analysis brought the total number of established PD loci to 26.[3] In the latest AD meta-analysis 19 loci in addition to the well-established APOE locus were pinpointed.[4] Despite progress in identifying genetic hits in these neurodegenerative diseases, the underlying processes or cell-types leading to pathology remain uncertain. Efforts to obtain brain samples (the seemingly most obvious tissue for neurodegenerative disorders) for eQTL analyses are ongoing.[5–7] However, it is not as easy to obtain large numbers of post-mortem human brains. Characterization of eQTLs and other DNA regulatory elements in blood is a complimentary approach that allows large-scale sample collection, helping to combat noise in gene expression data.

With regard to the relevant tissue for neurodegenerative disorders, there is growing evidence that in addition to the brain, the immune system also plays a role in these disorders. There is little doubt that MS is an immune-mediated disorder[8], and so this disease serves as a positive control with regard to expected enrichment in heritability for annotations from immune cells. For AD, Yokoyama et al.[9] showed that eight variants were associated with both AD and immune-mediated diseases, and there is further evidence from animal models.[10] For PD, the role of the immune system has been suggested through pathway analysis[11,12] and animal models.[13]

Finucane et al.[14] introduced stratified LD score regression as a method for partitioning heritability from genome-wide association study (GWAS) summary statistics while accounting for markers in linkage disequilibrium (LD) using a reference population. They partitioned heritability for a variety of tissue/cell-types for 17 GWASs, none of which were neurodegenerative disorders. We applied this methodology to three neurodegenerative disorders to test for enrichment of heritability using Finucane et al.'s[14] cell-type group annotations, and additional annotations from brain and immune cells as well as previously published sets of brain and immune genes.

Methods

As a first step towards identifying the biological mechanisms underpinning our neurological phenotype-associated genes, we first performed an Ingenuity Pathway Analysis (IPA) (www.ingenuity.com) to identify pathway enrichment among genes associated with different neurological traits for canonical pathways, and diseases and biological functions. Data from the different phenotypes were integrated and subjected to network analysis via IPA to identify pathway enrichment. Cancer-related functions were removed from the disease and biological function results, due to their over-representation in the database.
For each disorder we included SNPs with a p-value $<5x10^{-4}$, and excluded SNPs in the MHC region.

We obtained GWAS summary statistics for three neurodegenerative disorders: Parkinson's disease (PD),[3] Alzheimer's disease (AD),[4] amyotrophic lateral sclerosis (ALS).[15] We used multiple sclerosis (MS)[16] as a positive control, a disease affecting the brain with known immune etiology. All studies were conducted in European populations, and are summarized in **Table 1**. For AD, which is a two-stage study, we only used data from the first stage. We also obtained summary statistics for frontotemporal dementia (FTD),[17] but the mean chi-square was not large enough to accurately infer heritability using LD score regression. We also thought to include Huntington's disease, since there are other genetic modifiers in addition to the primary locus in the *HTT* gene, but the sample size was only modest, and thus the dataset was not suitable for the method, which requires a large sample size for reasonable power.[14]

First we estimated pairwise genetic correlations among the four disorders using cross-trait LD score regression.[18] Next, we applied stratified LD score regression to determine if various functional categories (cell-type groups, annotations at the tissue/cell level for brain or immune cells, and sets of brain and immune gene lists) are enriched for heritability. LD score regression exploits the expected relationships between true association signals and local LD around them to correct out systematic biases and arrive at unbiased estimates of genetic heritability within a given set of SNPs (here stratified according to their functional category).[14] Following Finucane et al.[14], we added annotations individually to the baseline model; we used HapMap Project Phase 3 SNPs for the regression and 1000 Genomes Project European

population SNPs for the reference panel; and we only partitioned the heritability of SNPs with minor allele frequency >5%.

We first investigated the grouped cell-type annotations provided by Finucane et al.[14] (i.e. the union of histone marks for 10 categories including central nervous system (CNS), cardiovascular, immune/hematopoietic, and liver). For this analysis we corrected for multiple testing of four GWASs across 10 cell-type groups (4 x 10 = 40 hypotheses tested), resulting in a Bonferroni significance threshold of $p= 1.2 \times 10^{-3}$.

In addition to cell-type group analysis, we also looked at the individual tissue/cell level, assessing enrichment of additional annotations for various brain tissues (13 annotations) and immune cells (38 annotations). From brain tissue we assessed eQTLs derived from brain regions from the UK Brain Expression Consortium and the GTEx Consortium, and histone marks and DNase I hypersensitive sites data from the Roadmap Epigenomics Consortium. In contrast to Finucane et al.'s[14] cell-type specific analysis approach, we took a union of histone marks and DNase I hypersensitive data within each brain region to limit the multiple testing burden. This processing resulted in one annotation per brain region (10 annotations). We took a union of the GTEx eQTLs for the brain tissues, and also a union of the UKBEC eQTLs for the brain tissues (taking the union of microarray and RNA-sequencing data separately), resulting in three annotations. Details are in **Supplementary Table 1**.

With immune cells we assessed the histone marks previously described.,[14] and also the histone marks and DNase I hypersensitive data from the Roadmap Epigenomics Consortium for immune and blood cells.[19] We took the union for each cell-type as described above, resulting in 20 annotations from Finucane et al.[14] and 14 annotations from Roadmap. Promoter capture hiC array express data in CD34 from GM12878 (reference: E-MTAB-2323) was also used.[20] The data for the prey and bait were analyzed separately for interactions between captured promoter and captured promoter interactions and for captured promoter and all other regions, which resulted in four annotations.

For the tissue analysis we corrected for multiple testing of four GWASs across 51 (13 brain + 38 immune) annotations (4 x 51 = 204 hypotheses tested), resulting in a Bonferroni significance threshold of $p= 2.4 \times 10^{-4}$. Note that there are correlations among the immune annotations and correlations among the brain annotations, making our Bonferroni correction somewhat conservative.

We also looked at sets of genes with known brain and immune function for enrichment of heritability of neurodegenerative disorders. We used a brain gene list of 973 genes previously described by Raychaudhuri et al.[21], and an immune gene list of 2,635 genes previously described by Pouget et al.[22] Brain genes were defined as those fulfilling any of the following criteria: preferential expression in the brain compared to other tissues, "neural-activity" annotation in Panther, "learning" annotation in Ingenuity, and "synapse" annotation in Gene Ontology. Immune genes

were defined as those with an "immune response" annotation in at least three of the following databases: Kyoto Encyclopedia of Genes and Genomes, Gene Ontology, Ingenuity, and Immunology Database and Analysis Portal. SNPs were annotated to genes using a 50 kb window, and a baseline list of all genes using this 50 kb window was included in the model as previously described.[22]

Results

We performed an IPA pathway enrichment analysis for canonical pathways, and diseases and biological functions on the combined results for AD, ALS, FTD, MS and PD (**Supplementary Table 2**). Canonical pathway showed evidence for enrichment of immune functions across all five phenotypes, particularly in the MS and PD (e.g. dendritic cell maturation, role of macrophages, fibroblasts, and endothelial cells in RA), MS and AD (e.g. B-cell receptor signaling), and MS, PD and ALS (e.g. NFKB signaling). Disease and biofunction analysis (**Supplementary Table 2**, second tab) showed enrichment of leukocyte signaling in MS.

There is limited evidence of pairwise genetic correlation among the four neurodegenerative disorders using cross-trait LD score regression. The lack of an AD-PD pairwise correlation has already been reported, as well as between AD-MS and PD-MS.[23] We also found no statistically significant evidence for genetic correlation between ALS-AD (0.2, p= 0.08), ALS-PD (-0.08, p=0.01), and ALS-MS (-0.04, p=0.7).

For the cell-type group analysis, the most significant enrichment was seen for the immune/hematopoietic category for MS (10.1, $p= 3.8 \times 10^{-13}$), confirming the recognized role of the immune system in this immune-mediated disorder. This category was also significantly enriched for heritability of AD (5.5, $p= 2.4 \times 10^{-7}$), in addition to liver (10.5, $p= 1.1 \times 10^{-5}$). For PD and ALS, there were no significantly enriched functional categories (**Fig 1**).

At the tissue level, none of the enrichments were significant for the brain annotations. However, from the Roadmap data, the inferior temporal region was nominally significant in AD (4.9, $p= 6.6 \times 10^{-4}$).

At the cell level, for the immune annotations assessed (which were from outside the brain), from both the innate and adaptive immune systems, nearly all came up as significantly enriched for MS heritability, several came up as significant for AD and some for PD. There was no enrichment of heritability for ALS (**Supplementary Table 1, Fig 2**). The most significant enrichment for MS was one of the primary T helper cells annotations (17 cells PMA-I stimulated) (21.5, $p= 7.3 \times 10^{-21}$). The most significant enrichment for AD was primary T cells from cord blood (9.6, $p= 3.0 \times 10^{-7}$). Cells from both the innate (e.g. CD14, CD15 and CD34) and adaptive immune systems came up as significant for AD (following a similar pattern to MS). Only two annotations passed the multiple testing threshold for PD: primary T helper cells PMA-I stimulated and primary T regulatory cells from peripheral blood (5.2 and 5.4,

respectively, p= 0.0002 for both), but several other immune annotations were suggestive.

Our gene list analysis provided complimentary results. As expected, the immune gene list was enriched for heritability in MS (1.6, $p= 4.6x10^{-14}$). We have previously reported the enrichment of this immune gene list in the same MS dataset, using an earlier version of LDSC.[22] The immune gene list was also enriched for heritability in AD (5.2, $p= 4.8 \times 10^{-4}$), and the effects in PD and ALS were suggestive but would not survive multiple testing correction (4.5, p= 0.02 and 2.5, p= 0.03, respectively). The brain gene list was not significantly enriched in any of the neurodegenerative disorders assessed (among the other three disorders enrichment ranges from 0.9 to 1.9, p >0.04 for all three) (**Supplementary Table 1**).

Discussion

From multiple lines of evidence we show that there is a significant contribution of variants exhibiting functional marks in immune cells to the heritability of three neurodegenerative disorders, namely MS, AD and PD. Annotations from immune cells are most significantly enriched for the heritability of MS. Immune annotations are also consistently enriched but to a lesser degree for AD, and some cell-specific immune annotations (T-cells) were significantly enriched for PD. A lack of results from the ALS dataset could be attributed to this dataset being smaller than the other datasets investigated (**Table 1**). These results suggest that immune modulation may be a treatment target for such diseases.

We note that if we correct for the 17 GWASs assessed in Finucane et al.[14] as well as the four GWASs we assessed here for the 10 cell-type groups ((17+4) x 10 = 210 hypotheses tested), both the immune/hematopoietic and liver categories remain significant for AD.

The role of the immune system in AD pathogenicity has been previously shown[9,10] and previous pathway analysis of the AD GWAS we assessed here showed enrichment in immune-related pathways.[24] Our findings further support the role of immune variation in AD susceptibility. Interestingly, using LD score regression, AD was found to be not significantly correlated with a variety of immune disorders.[18] This lack of correlation could be because when considering the entire genome the signal coming from the correlated loci between the disorders is diluted, or the immune variants involved in AD are different from those involved in immune disorders.

Consistent with previous applications of the LD-score method, we included the annotations separately into the regression model. This means that that enrichments can be due to correlation with other cell-types. However, including the annotations simultaneously results in a decrease in power.[14]

Our results do not provide statistically significant evidence that variants overlapping with functional annotations from the brain contribute excessively to the

heritability of neurodegenerative disorders. The brain doubtless plays an important role in the genetic aetiology of these disorders. The lack of brain annotation enrichment could be due to data being based on few samples for the brain. This analysis should be revisited as brain annotation information improves.

In summary, our results suggest a significant contribution of variants that exhibit functional marks in immune cells to the heritability of three neurodegenerative disorders, namely MS, AD and PD.

Acknowledgements
We would like to thank Andy Singleton and Mike Nalls for access to the Parkinson's disease summary statistics. We would also like to thank Isabella Fogh and John Powell for providing access to the ALS summary statistics. The SLAGEN (Italian Consortium for the Genetics of ALS) and ALSGEN (International Consortium of Amyotrophic Lateral Sclerosis Genetics) Consortia contributed ALS/controls samples. The multiple sclerosis summary statistics were accessed through Immunobase (https://www.immunobase.org/).

Author Contributions
S.A.G., M.R., M.R.B., J.K., and M.E.W. are responsible for concept and study design.
S.A.G. analyzed data, and J.G.P. analyzed data for the gene list enrichment.
S.A.G. drafted the manuscript and figure, with input from all the authors.

References

1. Maurano MT, Haugen E, Sandstrom R, et al. Large-scale identification of sequence variants influencing human transcription factor occupancy in vivo. Nat. Genet. 2015;

2. Farh KK-H, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518(7539):337–343.

3. Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat. Genet. 2014;46(9):989–993.

4. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45(12):1452–8.

5. The GTEx Consortium. The genotype-tissue expression (GTEx) project. Nat Genet 2013;45(6):580–585.

6. Ardlie KG, Deluca DS, Segrè AV, et al. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science 2015;348(6235):648–660.

7. Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat. Neurosci. 2014;17(10):1418–1428.

8. Steinman L. Multiple sclerosis: a two-stage disease. Nat. Immunol. 2001;2(9):762–764.

9. Yokoyama JS, Wang Y, Schork AJ, et al. Association Between Genetic Traits for Immune-Mediated Diseases and Alzheimer Disease. JAMA Neurol. 2016;

10. Wang Y, Cella M, Mallinson K, et al. TREM2 Lipid Sensing Sustains the Microglial Response in an Alzheimer's Disease Model. Cell 2015;160(6):1061–1071.

11. Holmans P, Moskvina V, Jones L, et al. A pathway-based analysis provides additional support for an immune-related genetic susceptibility to Parkinson's disease. Hum. Mol. Genet. 2013;22(5):1039–1049.

12. Jones L, Holmans PA, Hamshere ML, et al. Genetic Evidence Implicates the Immune System and Cholesterol Metabolism in the Aetiology of Alzheimer's Disease. PLOS ONE 2010;5(11):e13950.

13. Kannarkat GT, Boss JM, Tansey MG. The Role of Innate and Adaptive Immunity in Parkinson's Disease. J. Park. Dis. 2013;3(4):493–514.

14. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 2015;47(11):1228–1235.

15. Fogh I, Ratti A, Gellera C, et al. A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. Hum. Mol. Genet. 2014;23(8):2220–2231.

16. International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011;476(7359):214–219.

17. Ferrari R, Hernandez DG, Nalls MA, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. Lancet Neurol. 2014;13(7):686–699.

18. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. Nat. Genet. 2015;47(11):1236–1241.

19. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. Nature 2015;518(7539):317–330.

20. Mifsud B, Tavares-Cadete F, Young AN, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat. Genet. 2015;47(6):598–606.

21. Raychaudhuri S, Korn JM, McCarroll SA, et al. Accurately assessing the risk of schizophrenia conferred by rare copy-number variation affecting genes with brain function. PLoS Genet. 2010;6(9):e1001097.

22. Pouget JG, Gonçalves VF, Consortium SWG of the PG, et al. Genome-Wide Association Studies Suggest Limited Immune Gene Enrichment in Schizophrenia Compared to 5 Autoimmune Diseases. Schizophr. Bull. 2016;sbw059.

23. Anttila V, Bulik-Sullivan B, Finucane HK, et al. Analysis of shared heritability in common disorders of the brain [Internet]. bioRxiv 2016;Available from: http://biorxiv.org/content/early/2016/04/16/048991.article-info

24. International Genomics of Alzheimer's Disease Consortium (IGAP). Convergent genetic and expression data implicate immunity in Alzheimer's disease. Alzheimers Dement. J. Alzheimers Assoc. 2015;11(6):658–671.

Figure Legends

Fig 1. Enrichment of cell-type groups as used in Finucane et al. 2015. The black dashed lines at $-\log_{10}(P) = 2.9$ is the cutoff for Bonferroni significance.

Fig 2. Enrichment of immune cell annotations. The black dashed lines at $-\log_{10}(P) = 3.6$ is the cutoff for Bonferroni significance. The list of annotations can be found in the immune section of Supplementary Table 1.

Tables

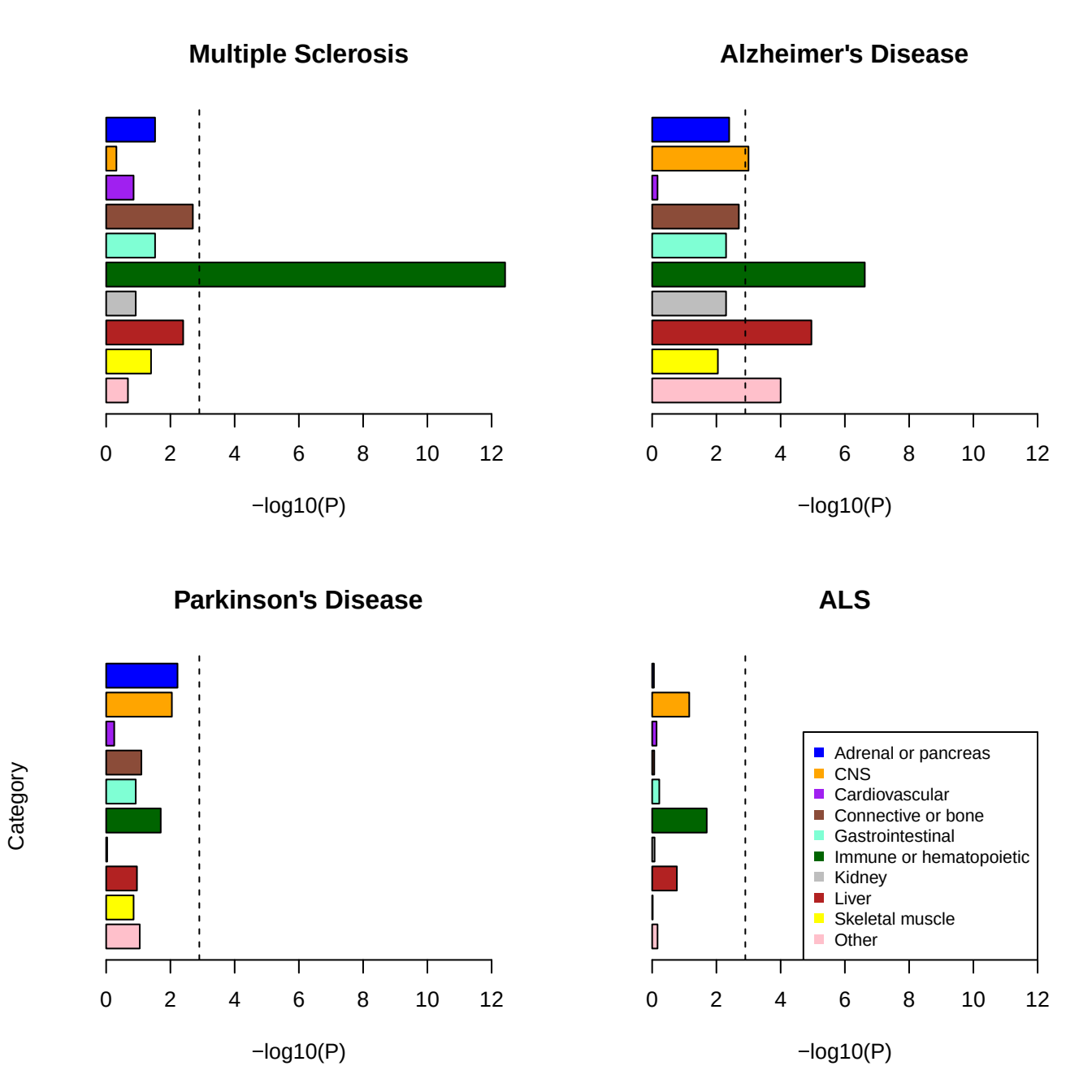
Table 1. Description of the GWASs summary statistics

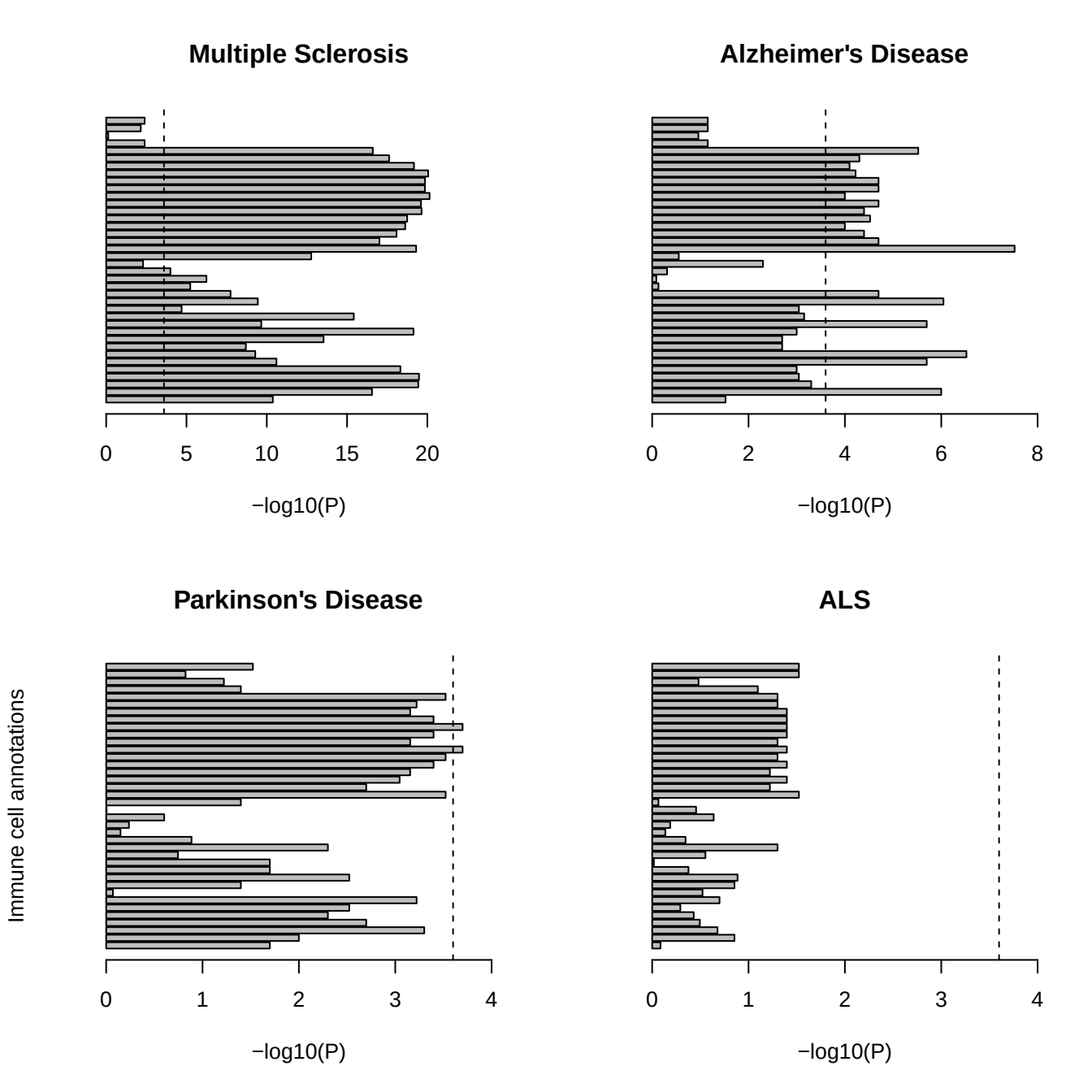
Neurodegenerative disorder	PMID	Cases	Controls	Cohorts
Parkinson's disease	25064009	13,708	95,282	15
Alzheimer's disease	24162737	17,008	37,154	19
Amyotrophic lateral sclerosis	24256812	7,177	8,393	8
Multiple sclerosis	21833088	9,772	17,376	23
Frontotemporal dementia	24943344	2,154	4,308	3

Supplementary Information

Supplementary Table 1. Annotation enrichment results

Supplementary Table 2. Pathway analysis results


Multiple Sclerosis
Alzheimer's Disease
Parkinson's Disease
ALS
0
2
4
6
8
10
12
−log10(P)
Category
Adrenal or pancreas
CNS
Cardiovascular
Connective or bone
Gastrointestinal
Immune or hematopoietic
Kidney
Liver
Skeletal muscle
Other


Multiple Sclerosis
0
5
10
15
20
−log10(P)
Alzheimer's Disease
0
2
4
6
8
−log10(P)
Parkinson's Disease
Immune cell annotations
0
1
2
3
4
−log10(P)
ALS
0
1
2
3
4
−log10(P)