

**Genome-wide association, Mendelian Randomization and polygenic risk score studies  
converge on a role of  $\beta$ -amyloid and *APOE* locus in Parkinson disease**

Laura Ibanez<sup>1,2</sup>, Jorge A. Bahena<sup>1,2</sup>, Chengran Yang<sup>1,2</sup>, Umber Dube<sup>1,2</sup>, Fabiana G. Farias<sup>1,2</sup>, John P. Budde<sup>1,2</sup>, Kristy Bergmann<sup>1,2</sup>, Carol Brenner-Webster<sup>1,2</sup>, John C. Morris<sup>3,4,5</sup>, Richard J. Perrin<sup>3,4,5,6</sup>, Nigel Cairns<sup>3,4,5,6</sup>, John O'Donnell<sup>4</sup>, Ignacio Álvarez<sup>7,8</sup>, Monica Diez-Fairen<sup>7,8</sup>, Miquel Aguilar<sup>7,8</sup>, Rebecca Miller<sup>4</sup>, Albert A. Davis<sup>3,4</sup>, Pau Pastor<sup>7,8</sup>, Paul Kotzbauer<sup>3,4</sup>, Meghan C. Campbell<sup>4,9</sup>, Joel S. Perlmutter<sup>3,4,9</sup>, Herve Rhinn<sup>10</sup>, Oscar Harari<sup>1,2,3,5</sup>, Carlos Cruchaga<sup>1,2,3,5,11</sup>, and Bruno A. Benitez<sup>1,2\*</sup>

**Author Affiliations:**

<sup>1</sup> Department of Psychiatry, Washington University, St. Louis, MO 63110

<sup>2</sup> NeuroGenomics and Informatics. Washington University, St. Louis, MO 63110

<sup>3</sup> Hope Center for Neurologic Disorders, Washington University, St. Louis, MO 63110

<sup>4</sup> Department of Neurology, Washington University, St Louis, MO, 63110

<sup>5</sup> The Charles F. and Joanne Knight Alzheimer Disease Research Center, Washington University School of Medicine, St Louis, MO, 63110.

<sup>6</sup> Department of Pathology and Immunology, Washington University, St. Louis, MO 63110

<sup>7</sup> Memory Unit, Department of Neurology, University Hospital Mutua de Terrassa, University of Barcelona, Terrassa, Barcelona, Spain.

<sup>8</sup> Fundació per a la Recerca Biomèdica i Social Mútua de Terrassa, University of Barcelona, Terrassa, Barcelona, Spain.

<sup>9</sup> Departments of Neuroscience and Radiology, Programs in Physical Therapy and Occupational Therapy, Washington University, St. Louis, MO, 63110

<sup>10</sup> Department of Bioinformatics, Alector, INC, San Francisco, CA, 94080

<sup>11</sup> Department of Genetics, Washington University School of Medicine, St Louis, MO, 63110

\*To whom correspondence should be addressed:

Bruno A. Benitez, MD

Department of Psychiatry

Washington University School of Medicine

BJC Institute of Health. Box 8134

425 S. Euclid Ave. St. Louis, Missouri, 63110.

E-mail: babenitezv@wustl.edu

**Keywords:** Parkinson Disease, Genetics, Biomarkers, Alpha-Synuclein, Amyloid Beta, Tau, APOE, polygenic risk scores

## Abstract

Alpha-Synuclein ( $\alpha$ -Syn) is the main protein component of Lewy bodies (LB), the pathological hallmark of Parkinson's disease (PD). Cerebrospinal fluid (CSF) levels of  $\alpha$ -Syn are not currently used as a clinical biomarker but may be a proxy for pathological  $\alpha$ -Syn accumulation in the brain. Therefore, identifying genetic modifiers of CSF  $\alpha$ -Syn levels could provide insights into the underlying molecular mechanisms leading to PD. However, genetic modifiers of CSF  $\alpha$ -Syn levels remain unknown. CSF levels of amyloid beta<sub>1-42</sub> (A $\beta$ 42), total tau (t-tau), and phosphorylated tau<sub>181</sub> (p-tau<sub>181</sub>) are standard biomarkers for the diagnosis of Alzheimer disease (AD); its use as quantitative traits in genetic studies have provided novel insights into AD pathophysiology. A systematic study of the genomic architecture of CSF biomarkers in PD has not been conducted. Here, genome-wide association studies (GWAS) were performed using CSF biomarker levels as quantitative traits in four PD cases and control cohorts (combined N=1,960). CSF biomarker ( $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub>) levels were significantly lower in PD cases compared with controls. An SNP, proxy for *APOE*  $\epsilon$ 4, was associated with CSF A $\beta$ 42 levels (effect=-0.5,  $p=9.2\times 10^{-19}$ ). Several genome-wide suggestive loci associated with CSF  $\alpha$ -Syn, t-tau, or p-tau<sub>181</sub> were found. Polygenic risk scores (PRS) were constructed using the latest PD risk meta-analysis (49,731 PD cases and 784,343 controls) and the largest CSF biomarkers GWAS (N=3,146). PRS calculated using META-PD were associated with PD status in the four cohorts included in the present study ( $p=2.2\times 10^{-16}$ ). A highly significant correlation (Nagelkerke pseudo- $R^2=2.29\%$ ;  $p=2.5\times 10^{-11}$ ) of the genomic architecture between CSF A $\beta$ 42 and PD risk was also found. Higher PRS scores were associated with lower CSF A $\beta$ 42 levels ( $p=7.3\times 10^{-4}$ ). Two-sample Mendelian Randomization (MR) approach revealed that CSF A $\beta$ 42 plays a role in PD risk ( $p=1.4\times 10^{-5}$ ) and age at onset ( $p=7.6\times 10^{-6}$ ), an effect mainly mediated by variants in the *APOE* locus. Subsequently, the *APOE*  $\epsilon$ 4 allele was associated with significantly lower levels of CSF A $\beta$ 42 ( $p=3.8\times 10^{-6}$ ), higher mean cortical binding potentials (cortical binding of Pittsburgh compound B PET) ( $p=5.8\times 10^{-8}$ ) and higher Braak A $\beta$  score ( $p=4.4\times 10^{-4}$ ) in PD participants. Together these results from high-throughput and hypothesis-free approaches (GWAS, PRS and MR) converge on a genetic link between PD with CSF A $\beta$ 42 and *APOE*.

## Introduction

Parkinson's disease (PD) is a neurodegenerative disease characterized by rest tremor, rigidity, bradykinesia, and postural instability.<sup>1</sup> PD is the most common neurodegenerative movement disorder, affecting more than six million people worldwide, with its prevalence projected to double in the next several decades.<sup>2</sup> Currently, the diagnosis of PD relies almost entirely on history and clinical examination, as there are no reliable biomarkers for PD. A modest but significant decrease (~10% to 15%) in cerebrospinal fluid (CSF) alpha-synuclein ( $\alpha$ -Syn) levels in PD cases compared with controls<sup>3</sup> and a correlation with disease progression support the role of CSF  $\alpha$ -Syn as a potential PD biomarker.<sup>4-6</sup> Aggregated and phosphorylated  $\alpha$ -Syn is the main protein component of neuronal Lewy bodies (LB) and neurites, the pathological hallmark of Lewy body diseases (LBD). Gene dosage effect of the *SNCA* gene, which encodes  $\alpha$ -Syn, correlates with CSF  $\alpha$ -Syn levels and a more severe PD phenotype.<sup>7,8</sup> Additionally, common variants in the *SNCA* promoter are among the top GWAS signals for PD<sup>9</sup>, suggesting that genetic control of CSF  $\alpha$ -Syn level plays a role in PD phenotype variability. CSF  $\alpha$ -Syn is not currently used as a clinical biomarker<sup>5,10</sup>, but has potential as a proxy for pathological brain  $\alpha$ -Syn accumulation.<sup>11</sup> Therefore, identifying genetic modifiers of CSF  $\alpha$ -Syn levels could provide insight into PD pathogenesis. To date, genetic modifiers of CSF  $\alpha$ -Syn remain unknown.

The  $\alpha$ -Syn accumulation in specific brain regions defines different subtypes of LBD, including neocortical (nLBD); limbic (lLBD); and brainstem or amygdala (bLBD). However, pure  $\alpha$ -syn pathology is only found in 45% (bLBD), 32% (lLBD) and 19% (nLBD). Concomitant proteinopathies (presence of A $\beta$ , tau, and TDP-43) are common findings in LBD. A $\beta$  and tau pathology is present in up to 80% and 53% in cases of nLBD, respectively.<sup>12</sup> LBD with concomitant Alzheimer disease (AD) pathology exhibit a faster cognitive decline.<sup>13</sup> CSF levels of A $\beta$ 42, tau and p-tau are used as proxies of A $\beta$  and tau pathology in the brain.<sup>14</sup> A correlation between decreased A $\beta$ 42 CSF levels and Braak stages of AD neuropathology was found in neuropathologically confirmed LBD.<sup>14</sup> In cross-sectional studies, PD cases exhibit lower CSF levels of A $\beta$ 42 compared to age and gender-matched healthy controls.<sup>15,16</sup> CSF levels of A $\beta$ 42 and total tau (t-tau) change with PD cognitive decline progression.<sup>16</sup> Decreased CSF A $\beta$ 42 levels predict the development of dementia in PD patients.<sup>17,18</sup> These results suggest that dementia associated CSF biomarker profiles in PD behave similarly to AD. Genome-wide association

studies (GWAS) using CSF A $\beta$ 42, t-tau, and p-tau<sub>181</sub> levels as quantitative traits have identified genes involved in AD pathogenesis<sup>19</sup>. However, a systematic study of the role of genetic modifiers of dementia CSF biomarkers in PD has not been thoroughly evaluated yet.

This study aimed to uncover genetic modifiers of  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub> CSF levels in PD patients by performing a large (N=1,960) GWAS meta-analysis of CSF biomarkers in PD cohorts. PRS and MR analyses were integrated with the latest PD risk meta-GWAS and CSF biomarker summary statistics to examine the causal relationship between CSF biomarkers and PD risk. This is the first comprehensive analysis of CSF biomarkers using GWAS, PRS, and MR in PD.

## **Material and Methods**

### **Study design**

The goal of this study was to identify common genetic variants and genes associated with CSF  $\alpha$ -Syn, A $\beta$ 42, tau, and p-tau<sub>181</sub> in PD. A three-stage GWAS was used: discovery, replication, and meta-analyses. The discovery phase included 729 individuals from the Protein and Imaging Biomarkers in Parkinson's disease study (PIB-PD) at the Washington University Movement Disorder Center<sup>20</sup> (n=103) and the Knight ADRC<sup>19</sup> (n=626). The replication phase included 1231 independent CSF samples obtained from PD cases and healthy elderly individuals from three additional studies (PPMI, ADNI, and Spain). Meta-analyses were performed using a fixed-effects model. Genetic loci that passed the multiple test correction for GWAS ( $P < 5 \times 10^{-8}$ ) were functionally annotated using bioinformatics tools to identify variants and genes driving the GWAS signal. Polygenic risk scores were used to test the correlation between CSF biomarkers and PD risk genetic architecture. Instrumental variables were selected from summary statistics of CSF biomarkers, and Mendelian Randomization methods were applied to test causality with PD risk. PD genetic architecture was obtained from the latest PD risk meta-analysis summary statistics (META-PD).<sup>9</sup>

The Institutional Review Boards of all participating institutions approved the study, and this research was carried out in accordance with the recommended protocols. Written informed consent was obtained from participants or their family members.

## Cohorts/Datasets

This cross-sectional multicenter study was performed using 1,960 samples from non-Hispanic white (NHW) individuals from four cohorts: Washington University in Saint Louis (WUSTL–N=729), the Parkinson’s Progression Markers Initiative (PPMI–N=785), the University Hospital Mutua Terrassa (Spain–N=130) and the Alzheimer Disease Neuroimaging Initiative (ADNI–N=316). Cohorts included 700 clinically diagnosed PD cases, 564 controls, and 386 clinically-diagnosed AD cases. The remaining N=310 individuals do not exhibit symptoms of neurodegenerative disease (Table 1 and Table S1). PD clinical diagnoses were based on the UK Brain Bank criteria.<sup>21</sup> Clinical, biomarker, and genetic data from the PPMI and ADNI were obtained from the PPMI ([www.ppmi-info.org](http://www.ppmi-info.org)) and ADNI database repository (<http://adni.loni.usc.edu/>), respectively and accessed most recently on 1 April 2019. Demographic characteristics of some of those cohorts have been published previously.<sup>22-24</sup> PPMI is a prospective study with ongoing recruitment. CSF samples were obtained at baseline (N=510), six months (N=385), and yearly after enrolment (N<sub>1stYear</sub>=428, N<sub>2ndYear</sub>=404 and N<sub>3rdYear</sub>=320). CSF  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub> were available for all the mentioned time points in the PPMI cohort.

## Biomarker Measurements

$\alpha$ -Syn in CSF was measured in 107 samples from the WUSTL cohort<sup>20</sup> and the entire PPMI cohort, using a commercial ELISA kit (Covance, Dedham, MA).<sup>25</sup> The additional samples (N=622) from WUSTL were quantified using the *SOMAScan* platform (See below). A $\beta$ 42, t-tau, and p-tau<sub>181</sub> were quantified using the INNOTEST assay (WUSTL cohort) and xMAP-Luminex with INNOBIA AlzBio3 (PPMI cohort). The immunoassay platform from Roche Elecsys cobas e 601 was used in the ADNI cohort to quantify all four biomarkers. ELISA assays from Euroimmun (Germany) were used in the Spanish cohort to measure the CSF levels of  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub>. The  $\alpha$ -Syn levels were normalized by log<sub>10</sub> transformation. A $\beta$ 42, t-tau, and p-tau<sub>181</sub> values were normalized by the Z score transformation. Individuals with biomarker levels outside three standard deviations of the mean were removed from the analysis (Table 1).

## PIB Imaging

PIB acquisition and analysis were performed according to published methods.<sup>26</sup> PET imaging was performed using a Siemens 962 HR ECAT EXACT PET scanner (CTI, Knoxville, TN).

Approximately 10-15 mCi of the radiotracer (specific activity  $\geq 1,200\text{Ci}/\text{mmol}$ ) was injected via an antecubital vein, and a 60-minute, three-dimensional (septa retracted) dynamic PET scan was collected in 53 frames (25 $\times$ 5sec, 9 $\times$ 20sec, 10 $\times$ 60sec and 9 $\times$ 300sec). Emission data were corrected for scattering, randoms, attenuation and dead time. Image reconstruction produced images with a final resolution of 6mm full-width half-maximum at the center of the field of view. Frame alignment was corrected for head motion and co-registered to each person's T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) MR scan.<sup>27</sup> For quantitative analyses, three-dimensional regions of interest (prefrontal cortex, gyrus rectus, lateral temporal cortex, precuneus, occipital lobe, caudate nucleus, brainstem, and cerebellum) were created by a blinded observer for each subject based on the individual's MRI scans, with boundaries defined as previously described.<sup>28</sup> Binding potentials ( $\text{BP}_{\text{ND}}$ ) were calculated using Logan graphical analysis, with the cerebellum as the reference tissue input function.<sup>28,29</sup> Mean cortical binding potentials (MCBP) were calculated for each subject as the average of all cortical regions except the occipital lobe. Positron emission tomography (PET) with [11C]-Pittsburgh Compound B (PiB) has identified PD patients with elevated cortical PiB binding corresponding to A $\beta$  plaque deposition at autopsy.<sup>30</sup> PiB PET is a reliable method for identifying A $\beta$  deposition in patients with Lewy body disorders.<sup>26</sup>

### **Neuropathologic analysis**

The neuropathological analysis was done at Washington University in Saint Louis, as previously reported.<sup>18</sup> Briefly, brains were fixed in 10% neutral buffered formalin for two weeks. Paraffin-embedded sections were cut at six  $\mu\text{m}$ . Blocks were taken from the frontal, temporal, parietal, and occipital lobes (thalamus, striatum, including the nucleus basalis of Meynert, amygdala, hippocampus, midbrain, pons, medulla oblongata) and the cervical spinal cord. Histologic stains included hematoxylin-eosin and a modified Bielschowsky silver impregnation. Immunohistochemical analysis was performed using the following antibody: A $\beta$  (10D5, Elan Pharmaceuticals). The AD pathologic changes were rated using an amyloid plaque stage (range, 0 to A-C)<sup>31</sup> and diffuse and neuritic plaques were also assessed, and cases were classified according to the neuropathologic criteria of Khachaturian<sup>32</sup>, the Consortium to Establish a Registry for Alzheimer Disease (CERAD)<sup>33</sup> and NIA-Reagan.<sup>34</sup>

## **Genotyping**

All cohorts, except PPMI, were genotyped using the GSA Illumina platform. Genotyping quality control and imputation were performed using SHAPEIT<sup>35</sup> and IMPUTE2<sup>36</sup>. SNPs with a call rate lower than 98% and autosomal SNPs that were not in Hardy-Weinberg equilibrium ( $P < 1 \times 10^{-6}$ ) were excluded from downstream analyses. The X chromosome SNPs were used to determine sex based on heterozygosity rates. Samples in which the genetically inferred sex was discordant with the reported sex were removed. Whole-genome sequence data from the PPMI cohort was merged with imputed genotyping data; only variants present in both files were included in further analysis. Pairwise genome-wide estimates of proportion identity-by-descent tested the presence of unexpected duplicates and cryptically related samples ( $P_{ihat} > 0.30$ ). Unexpected duplicates were removed; the sample with a higher genotyping rate in the merged file was kept for those cryptically related samples. Finally, principal components were calculated using HapMap as an anchor. Only samples with European descent, an overall call rate higher than 95%, and variants with minor allele frequency (MAF) greater than 5% were included in the analyses.

## **Single Variant Analysis**

The three-stage single variant analysis was performed due to differences in time and platform for biomarker quantification. PLINK1.9<sup>37,38</sup> was used to perform the analysis of each cohort independently. A linear model corrected by sex, age, and the first two principal components was used. Results from WUSTL, PPMI, ADNI, and Spanish cohort analyses were meta-analyzed using METAL.<sup>39</sup> For the  $\alpha$ -Syn analyses, the WUSTL cohort was divided into two samples quantified using ELISA based method and samples quantified using the SomaScan platform.

## **Analysis of Variance**

To test the amount of variance explained by specific loci, the genome-wide complex traits analysis (GCTA)<sup>40</sup> software was used. GCTA estimates the degree to which the phenotypic variance is explained by the provided variants for a complex trait by fitting the effect of these SNPs as random effects in a linear mixed model. It relies on substantial and homogeneous populations.

## **Multi-Tissue Analysis**

The levels of  $\alpha$ -Syn were measured in CSF, plasma, and brain (parietal cortex) using an aptamer-based approach (SomaScan platform). After stringent quality control, CSF (n=835), plasma (n=529), and brain (n=380) samples were included in the downstream analyses (Table S2). The protein level was 10-based log-transformed to approximate the normal distribution and used as phenotype for the subsequent GWAS. The single variant analysis was used in each tissue independently using PLINK1.9<sup>37</sup>. A multi-tissue analysis using the multi-trait analysis of GWAS (MTAG)<sup>41</sup> was performed to increase the power of detecting a no tissue-specific protein quantitative trait loci (pQTL) for  $\alpha$ -Syn. MTAG calculates the trait-specific effect estimate for each tissue separately and then performs the meta-analysis while accounting for possible sample overlap. Measurements of A $\beta$ 42, t-tau, and p-tau<sub>181</sub> were not available in different tissues.

### **Polygenic Risk Score (PRS)**

The PRSice2 software<sup>42,43</sup> was used to test if the genetic architecture of PD risk (using the summary statistics from the last PD meta-analysis<sup>9</sup>) was correlated with that of  $\alpha$ -Syn CSF levels or with the levels of the other biomarkers in CSF (A $\beta$ 42, t-tau, and p-tau<sub>181</sub>) (omics.wustl.edu).<sup>44</sup> Briefly, the approach constructs a PRS by summing all trait-associated alleles in a target sample, weighted by the effect size of each allele in a base GWAS. SNPs in linkage disequilibrium (LD) are grouped together to avoid extra weight into a single marker. PRSs are computed at various p-value thresholds using the base GWAS. The optimal threshold is considered the one that explains the maximum variance in the target sample. The association was tested using the default parameters and nine p-value cutoffs:  $5.0 \times 10^{-08}$ ,  $5.0 \times 10^{-06}$ ,  $1.0 \times 10^{-05}$ ,  $1.0 \times 10^{-03}$ , 0.01, 0.05, 0.2, 0.5 and 1. Longitudinal measures of CSF  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub> were available for the PPMI cohort. A simplified PRS (detailed below) was used to test if the genetic architecture of PD risk was predictive of biomarker level progression by means of a mixed model. The PD risk PRS using sentinel SNPs from the PD meta-analysis<sup>9</sup> was modeled using the method described by the Psychiatric Consortium<sup>45</sup> and ourselves.<sup>46-48</sup> Briefly, only genetic variants corresponding to the top hit on each GWAS locus (also known as sentinel SNP) available in the dataset with a minimum call rate of 85% were included in the PRS. If not possible, a proxy with  $R^2 > 0.90$  was used. The weight of each variant was calculated using the binary logarithm transformation of the reported Odd ratios (ORs). The final PRS is the sum of the weighted values for the alternate allele of all the sentinel SNPs.<sup>46,47</sup>



## Mendelian Randomization

The latest meta-GWAS of PD genetic risk was used<sup>9</sup>. Briefly, Nalls et al. performed a fixed-effects meta-analysis of 17 datasets from European ancestry individuals, analyzing 7.8 million SNPs in 37,688 diagnosed cases, 18,618 UK Biobank proxy-cases (individuals who do not have PD but have a first-degree relative who does), and 1.4 million controls. Summary statistics from the largest GWAS of CSF A $\beta$ 42, t-tau, and p-tau<sub>181</sub> were also used.<sup>44</sup> Deming et al. performed a one-stage GWAS for 3,146 individuals of European ancestry across nine independent studies.<sup>44</sup> None of these cohorts included PD affected individuals for each biomarker (A $\beta$ 42, t-tau, and p-tau<sub>181</sub>). Finally, the summary statistics of  $\alpha$ -Syn CSF levels generated in the current study were used. There was no overlap between CSF biomarker datasets and PD risk datasets.

Two-sample Mendelian randomization to estimate causal effects using the Wald ratio for single variants and an inverse-variance-weighted (IVW) fixed-effects meta-analysis for an overall estimate<sup>49</sup> was performed using the R package “MendelianRandomization”<sup>50</sup> (version 0.4.1). To account for potential violations of the assumptions underlying the IVW analysis, a sensitivity analysis using MR-Egger regression and the weighted median estimator<sup>49</sup> were conducted. Robust regression to downplay the contribution to the causal estimate of instrumental variables with heterogeneous ratio estimates were also performed.<sup>51,52</sup> Instrumental variables for each GWAS were obtained by clumping each GWAS summary statistics based on a significance threshold of  $1 \times 10^{-5}$  using PLINK1.9.<sup>37</sup> Instrumental variables were restricted to those that are uncorrelated (in linkage equilibrium) by setting the --clump-r2 flag to 0.0 and the --clump-kb flag to 1000 (1 Mb). Heterogeneity (i.e., instrument strength) was tested using the  $I^2$  statistic.<sup>53</sup>

## Results

### Association of CSF biomarkers with disease status

A generalized linear model (CSF biomarker levels~Age+Sex+PD<sub>Status</sub>) including PD cases (N=700) and controls (N=189) from two independent datasets (WUSTL and PPMI –Table S1) in which  $\alpha$ -Syn levels were measured with the same platform revealed that all CSF biomarker levels were significantly lower in PD cases compared with healthy-controls ( $\alpha$ -Syn:  $\beta_{PD}=-0.05$ ,  $p=2.10 \times 10^{-04}$ ; A $\beta$ 42:  $\beta_{PD}=-0.34$ ,  $p=4.38 \times 10^{-05}$ ; t-tau:  $\beta_{PD}=-0.23$ ,  $p=4.58 \times 10^{-03}$ ; and p-tau<sub>181</sub>:  $\beta_{PD}=-0.25$ ,  $p=2.46 \times 10^{-03}$  – Figure 1). All associations passed multiple test correction ( $p < 0.013$ ).

Using a longitudinal model adjusted by age at lumbar puncture, sex, and the first two principal components, significant changes over time were found for CSF A $\beta$ 42 ( $p=0.01$ ) but not for  $\alpha$ -Syn, t-tau, or p-tau<sub>181</sub> in the PPMI cohort ( $N=785$ ). These results demonstrate that CSF AD biomarkers are associated with PD status.

### **No significant loci were identified for CSF $\alpha$ -Syn, t-tau or p-tau<sub>181</sub> in PD cohorts**

Within each cohort, a linear regression testing the additive genetic model of each single nucleotide polymorphism (SNP) for association with CSF protein levels using age, gender, and two principal component factors for population stratification as covariates did not reveal any genome-wide significant loci associated with CSF  $\alpha$ -Syn. Although several suggestive loci ( $p$  value  $<10^{-6}$  to  $10^{-8}$ ) were identified in these analyses (Table S3 and Figure S1), none of them passed a genome-wide significant threshold when cohorts were combined in the meta-analysis (Figure 2A, Table 2, and Table S3).

Joint analysis for CSF  $\alpha$ -Syn levels stratifying by PD cases ( $N=700$ ), PD cases and controls ( $N=889$ ), AD cases only ( $N=386$ ), AD cases and controls ( $N=575$ ) and controls only ( $N=189$ ) were also performed. None of these analyses revealed any genome-wide significant locus, suggesting that these sample sizes might be underpowered to uncover the genetic modifiers of CSF  $\alpha$ -Syn.

For t-tau, individual cohort analyses revealed four genome-wide significant loci and more than 20 loci that were suggestive (Table S5 and Figure S3). However, none of them remained significant in the fixed effect meta-analyses (Figure 2B, Tables 2 and S5). For p-tau<sub>181</sub>, individual cohort analyses revealed three genome-wide significant loci for p-tau<sub>181</sub> and several suggestive signals (Table S6 and Figure S4). However, none achieved significance in the fixed effect meta-analyses (Figure 1C, Tables 2 and S6).

### **Multi-tissue $\alpha$ -Syn genetic GWAS analyses**

In a subgroup of samples,  $\alpha$ -Syn levels were measured in plasma ( $N=529$ ), brain ( $N=380$ ), and CSF ( $N=835$ ) using the SomaScan platform (Table S2). Single variant analysis was performed in each tissue separately (Figure S2A to S2C). The multi-tissue analysis was performed using MTAG.<sup>41</sup> This analysis increases the statistical power of a GWAS by increasing the sample size and removing the variance specific to each tissue, thereby increasing the chances of finding genetic modifiers for  $\alpha$ -Syn. Although some suggestive loci were observed in chromosomes 3 and 13 (Table S4 and Figure S2D) within genomic regions enriched with long intergenic non-protein

coding (LINC) genes (Figure S2E and F), no genome-wide significant locus was identified. These results suggest that the power boost of using MTAG is not enough to unveil the genetic architecture of  $\alpha$ -Syn.

### ***APOE* locus is associated with A $\beta$ 42 CSF levels in PD cohorts**

A proxy SNP for APOE  $\epsilon$ 4, rs769449, was associated at a genome-wide level with CSF levels of A $\beta$ 42 in the WUSTL (effect=-0.56,  $p=4.15\times 10^{-19}$ ), and ADNI cohorts (effect=-0.73,  $p=1.25\times 10^{-15}$ ). A suggestive association was found in PPMI (effect=-0.43,  $p=3.09\times 10^{-07}$ ). There was no association in the Spanish cohort ( $p>0.05$ ). Additional suggestive loci were also identified (Table S7 and Figure S5). The *APOE* locus (effect=-0.57,  $p=4.46\times 10^{-43}$ ) and a locus in the *HLA* region (effect=0.23,  $p=2.88\times 10^{-08}$ ) remained significant in the meta-analysis (Figure 2D to 2F). When the cohorts containing only PD cases and controls were analyzed jointly (WUSTL and PPMI – N=700 cases and 189 controls), the *APOE* locus remained GWAS significant (effect=-0.50,  $p=9.25\times 10^{-19}$ ); the *HLA* region did not pass multiple test correction (effect=0.22,  $p=3.58\times 10^{-04}$ ).

In the combined analysis of all cohorts (N=1,960), the *APOE* locus accounted for 36.2% of the CSF A $\beta$ 42 levels variance ( $p=2.35\times 10^{-03}$ ). This estimation should be interpreted with caution due to the sample size. Overall, these results revealed a strong and highly significant association between *APOE* locus and lower CSF A $\beta$ 42 levels in PD participants.

### **Significant correlation of genomic architecture of PD risk and CSF A $\beta$ 42**

PRS at different p-value thresholds were used to test if the genetic variants associated with neurodegeneration biomarkers were associated with the genomic architecture of PD risk.<sup>9</sup> PRS calculated using the latest meta-analysis of PD risk (META-PD) were associated with PD status in the four cohorts included in the present study ( $p<2.20\times 10^{-16}$ ). No correlation was observed between the genetic architecture of PD risk and that of CSF  $\alpha$ -Syn, t-tau, or p-tau<sub>181</sub> (Figure 3).

In contrast, the genetic architecture of A $\beta$ 42 was correlated with PD risk, with the best fit when collapsing independent SNPs with p-value<0.01 ( $p=2.50\times 10^{-11}$ ) with a correlation coefficient ( $R^2$ ) of 2.29%. The correlation remained significant ( $p=4.78\times 10^{-08}$ ) with an  $R^2$  of 2.36% after collapsing independent SNPs with p value<0.01 in PD cases and controls only. In PD patients with both GWAS and CSF biomarker data, the CSF levels of each biomarker was analyzed by quartiles of the PRS calculated from META-PD risk. A significant difference ( $p=7.30\times 10^{-04}$ ) was found among the top and the bottom quartiles; higher PRS values exhibit lower levels of CSF A $\beta$ 42 (Figure S6).

No association between PD PRS and longitudinal changes of  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub> was found. These results indicate that PD risk and A $\beta$ 42 CSF levels share genomic architecture.

### **Mendelian Randomization suggest a link between CSF A $\beta$ 42 and PD**

Robust regression with the MR-Egger method found no association for t-tau (effect=-0.33; p=0.25) or p-tau<sub>181</sub> (effect=-0.10; p=0.79), but revealed a trend towards significance for  $\alpha$ -Syn (effect=-1.40; p=0.06), and a significant causal effect for A $\beta$ 42 on PD risk (effect=0.43; p=1.44 $\times$ 10<sup>-05</sup>) (Table 3 and S8; Figure 4A to 4C and S7). When each cohort included in the PD meta-analysis was tested separately, A $\beta$ 42 showed a causal effect in Nalls 2014<sup>54</sup> and Nalls 2019<sup>9</sup> (p=1.54 $\times$ 10<sup>-07</sup> and 8.74 $\times$ 10<sup>-05</sup> respectively), but not in Chang 2017<sup>55</sup> (Table S8). Additionally, a significant causal effect for A $\beta$ 42 on PD age-at-onset was found using the data from Blauwendraat, 2019<sup>56</sup> (effect=7.75; p=7.65 $\times$ 10<sup>-06</sup> – Table S8). A leave-one-out sensitivity analysis on A $\beta$ 42 revealed that the proxy SNP for *APOE*  $\epsilon$ 4, rs769449 is the main driver of the causal effect of A $\beta$ 42 on PD risk. Other SNPs contribute in a smaller proportion to the causal effect (Figure 4D). The MR analysis removing the rs769449 decreased the I<sup>2</sup> statistic (I<sup>2</sup>=0.00%) and increased the p-value to not-significant levels while removing any of the other variants maintained a high I<sup>2</sup> (I<sup>2</sup>>95%) and a significant p-value. Altogether these results suggest a causal role of variants on the *APOE* locus and CSF A $\beta$ 42 on PD. All the MR methods for  $\alpha$ -Syn, t-tau and p-tau<sub>181</sub> and A $\beta$ 42 did not produce significant results when using the META-PD summary statistics or de individual datasets (Table S8 and Figure S7).

### ***APOE* $\epsilon$ 4 is associated with A $\beta$ deposition in brains of PD individuals**

CSF A $\beta$ 42 and *APOE* genotype data were available for a total of 134 participants, including healthy controls (N=26) and PD participants (N=108). No difference in the *APOE*  $\epsilon$ 4 frequency was found between PD cases (0.14%) compared to controls (0.11%). However, the CSF A $\beta$ 42 levels were significantly different between controls (p=3.00 $\times$ 10<sup>-02</sup>) and PD cases (p=3.80 $\times$ 10<sup>-06</sup>) when stratifying by the presence of *APOE*  $\epsilon$ 4 allele (Figure 5A), as we have previously reported with a smaller sample size.<sup>20</sup>

PET PiB analysis revealed a MCBP that increased with age-at-onset (r=0.20, p=3.00 $\times$ 10<sup>-02</sup>) and number of *APOE*  $\epsilon$ 4 alleles (r=0.22, p=8.00 $\times$ 10<sup>-03</sup>) (Figure 5A), but decreased CSF A $\beta$ 42 (r=-0.55, p=3.33 $\times$ 10<sup>-12</sup>) in 108 PD patients. A linear regression model, indicates that CSF A $\beta$ 42 and *APOE*

$\epsilon 4$ , explain 48% of the variance of MCBP. *APOE*  $\epsilon 4$  is also significantly associated with MCBP ( $\beta=0.14$ ,  $p=1.40 \times 10^{-06}$ ) in an analysis that included 200 participants including sex and age as covariates. *APOE*  $\epsilon 4$  and age at onset (AAO) explain 20 % of the MCBP variance in this larger cohort. The presence of *APOE*  $\epsilon 4$  did not affect the MCBP in healthy controls ( $p=0.19$ ). However, PD patients carrying *APOE*  $\epsilon 4$  exhibit significantly ( $p=5.80 \times 10^{-08}$ ) higher levels of MCBP than non-carriers (Figure 5B).

Neuropathological data and *APOE* genotype were available from 92 PD cases. Individuals carrying an *APOE*  $\epsilon 4$  allele had significantly ( $p=4.40 \times 10^{-04}$ ) higher Braak A $\beta$  stage (Figure 5C). *APOE*  $\epsilon 4$  correlated with Braak A $\beta$  stage ( $r=0.33$ ,  $p=1.00 \times 10^{-03}$ ) and diffuse plaques ( $r=0.42$ ,  $p=5.00 \times 10^{-03}$ ), but not with neuritic plaques ( $r=0.42$ ,  $p=0.12$ ). The best multiple linear regression model for the Braak A $\beta$  stage, which included AAO and *APOE*  $\epsilon 4$ , explained 42% of the variance of the Braak A $\beta$  stage. All these results suggest that *APOE*  $\epsilon 4$  drives the A $\beta$  deposition in PD participants.

## Discussion

Here, CSF  $\alpha$ -Syn, A $\beta 42$ , t-tau, and p-tau<sub>181</sub> levels were significantly lower in PD cases compared with controls, as we have previously reported with smaller sample size.<sup>20</sup> GWAS were performed using CSF biomarker levels as quantitative traits in a large combined PD cases and healthy controls cohort (N=1,960). Several genome-wide suggestive loci were associated with CSF  $\alpha$ -Syn, t-tau, or p-tau<sub>181</sub> but with the current sample size, no signal was below the GWAS significant threshold. An SNP proxy for *APOE*  $\epsilon 4$  was associated with CSF A $\beta 42$  levels in the GWAS. PRS constructed using the latest PD risk (META-PD) and the largest CSF biomarkers GWAS were used to calculate the genetic correlation between PD risk and CSF biomarker level. The PRS calculated using META-PD were associated with PD status in the four cohorts included in the present study and a highly significant correlation of the genomic architecture between CSF A $\beta 42$  and PD risk was also found. Individuals with higher PRS scores exhibit lower CSF A $\beta 42$  levels. Two-sample MR analysis revealed that CSF A $\beta 42$  plays a role in PD risk and age at onset, an effect mainly mediated by variants in the *APOE* locus. Using a subset of participants from the WUSTL cohort with additional clinical and neuropathological data, we found that the *APOE*  $\epsilon 4$  allele was associated with significantly lower levels of CSF A $\beta 42$ , higher cortical binding of Pittsburgh compound B PET and higher Braak A $\beta$  score.

This is the first comprehensive analysis of CSF  $\alpha$ -Syn and AD biomarkers using GWAS, polygenic risk scores, and MR in PD. We found significantly lower levels of CSF  $\alpha$ -Syn in PD cases compared to controls in a cross-sectional analysis but no significant differences in the longitudinal study (PPMI). CSF  $\alpha$ -syn, as measured with ELISA-based assays, is not a clinically useful diagnostic marker for PD, and utility as an outcome measure for clinical trials or progression is still controversial<sup>5,10</sup>. CSF biomarkers in AD used as quantitative endophenotypes have provided insights into AD pathophysiology.<sup>19</sup> Here, we used a large CSF  $\alpha$ -Syn cohort (N=1,920) to identify genetic modifiers. However, we did not find any genomic locus associated with CSF  $\alpha$ -Syn levels. Recently, a GWAS on CSF  $\alpha$ -Syn (N=209) reported a genome-wide significant locus<sup>57</sup> (rs7072338). In the present meta-analyses (N=1,960), the p-value for rs7072338 was not significant (0.99). In the same cohort (ADNI dataset) used by Zhong et al.<sup>57</sup>, we found a nominal association for this SNP ( $p=0.50\times 10^{-3}$ ). No correlation was found between the genetic architecture of PD risk with cross-sectional or longitudinal CSF  $\alpha$ -Syn levels, consistent with what we have previously reported.<sup>47</sup> Using MR methods, we found a trend towards significance for the association between the CSF  $\alpha$ -Syn levels and the risk of developing PD. However, sensitivity analyses showed limited power due to the small number of variants included in the analyses.

*MAPT* is one of the most consistently replicated loci associated with PD risk in individuals of European descent.<sup>9</sup> Accumulation of t-tau and p-tau<sub>181</sub> is found in PD patients<sup>58</sup>. However, it is not clear if tau accumulation plays a role in PD dementia.<sup>18</sup> The accumulation of t-tau and p-tau<sub>181</sub> plays a role in parkinsonism of progressive supranuclear palsy and corticobasal degeneration. However, tau pathology is not generally associated with the movement disorder aspects of idiopathic PD. Our results did not reveal any genome-wide significant loci associated with t-tau or p-tau<sub>181</sub> CSF levels in PD cohorts. A suggestive signal was found in the same genomic locus (3q26) previously reported for t-tau and p-tau<sub>181</sub> in AD cohorts.<sup>19</sup> The same direction but half of the effect size found in PD cohorts compare to previously reported in AD cohorts. In PD cohorts, a suggestive association with the same direction (effect=0.21;  $p=1.98\times 10^{-6}$ ) was also found between the *APOE* locus and CSF p-tau<sub>181</sub>.<sup>19</sup> One possibility is that our sample could be underpowered to detect a small effect size. No association was found between CSF t-tau or p-tau<sub>181</sub> genetic architectures with PD risk. We have previously reported a statistical trend for the association of PD risk with the levels of p-tau<sub>181</sub>.<sup>47</sup> We did not replicate this trend with the current study with a

larger sample size. MR analyses did not support a causal relationship between CSF t-tau or p-tau<sub>181</sub> and PD.

We found a genome-wide significant association between the *APOE* and CSF A $\beta$ 42 levels in two cohorts (WUSTL and ADNI), and the association was borderline significant ( $p=3.09\times 10^{-07}$ ) in the PPMI cohort, as previously reported.<sup>59</sup> *APOE* locus remained significant in the meta-analysis that included four cohorts. The association of the *APOE* locus with CSF A $\beta$ 42 levels has been previously reported in AD case-control cohorts.<sup>60</sup> We also found that in PD cohorts, the direction of the effect and the p-value of the association with *APOE* were similar to what was previously reported in AD but with higher effect size (-0.57 in PD compared to -0.10 in AD).<sup>44</sup> There are differences in the strength of the association across the PD cohorts used in this study. The WUSTL cohort includes PD patients with a longer disease duration than those included in the PPMI cohort and is likely to be enriched in PD patients with cognitive decline.<sup>61</sup> In contrast, the PPMI cohort includes PD patients recently diagnosed with PD and excludes PD patients with dementia.<sup>61</sup> The *APOE* locus is the most significant locus associated with sporadic DLB,<sup>62,63</sup> PD age at onset<sup>56</sup> and cognitive decline in PD.<sup>61</sup> However, *APOE* was not found to be associated with PD in the latest PD risk GWAS.<sup>9</sup> Here, we found that the genetic architecture of CSF A $\beta$ 42 levels was correlated with PD risk. PD patients with higher PRS for PD risk exhibit lower levels of CSF A $\beta$ 42, suggesting that similar genes or pathways predispose individuals to brain accumulation of A $\beta$  and risk for developing PD. Low levels of CSF A $\beta$ 42 predict cognitive decline in PD.<sup>64</sup> Further studies, ideally with neuropathological information, stratifying PD individuals by dementia or cognitive status, are needed to clarify the role of *APOE* on PD risk or PD dementia.

MR analyses suggest that A $\beta$ 42 could play a causal role in PD. Our MR results consistently identified a causal correlation between the *APOE* locus, CSF A $\beta$ , and PD risk in Nalls, 2014<sup>54</sup> and Nalls, 2019<sup>9</sup> but not in Chang, 2017<sup>55</sup>. PD patients and controls were classified and ascertained differently among the cohorts with Nalls, 2014<sup>54</sup> containing the most clinically-diagnosed PD participants, Chang, 2017<sup>55</sup> using self-reported PD individuals, and Nalls, 2019<sup>9</sup> including clinically-diagnosed PD plus UKB-proxy PD participants. Our MR results showed a difference in the strength of the association, which may be due to different criteria for case-control classification. In summary, the results from our high-throughput and hypothesis-free approaches (GWAS, PRS and MR) detected the existence of PD genetic risk linked to CSF A $\beta$ 42 and *APOE* locus. Further

evidence was provided by showing that PD patients carrying *APOE*  $\epsilon 4$  allele presented significantly lower levels of CSF A $\beta$ 42 ( $p=3.8\times 10^{-06}$ ), higher MCBP (PiB) ( $p=5.80\times 10^{-08}$ ) and higher Braak A $\beta$  score ( $p=4.40\times 10^{-04}$ ). A synergistic relationship between  $\alpha$ -Syn and A $\beta$  pathology has been reported in AD, PD and LBD brains.<sup>65</sup> The presence of A $\beta$  plaques exacerbates propagation of  $\alpha$ -Syn pathology in mouse models.<sup>66</sup> *APOE*  $\epsilon 4$  in AD patients drives the production of A $\beta$  and the accumulation of A $\beta$  fibrils.<sup>67</sup> *APOE*  $\epsilon 4$  markedly exacerbates tau-mediated neurodegeneration in a mouse model of tauopathy.<sup>68</sup> The role of *APOE* in human synucleinopathies is more complex. In LBD patients, the *APOE*  $\epsilon 4$  effect on  $\alpha$ -Syn pathology could be dependent on concurrent A $\beta$  and/or tau pathology.<sup>12</sup> However, *APOE*  $\epsilon 4$  also promotes  $\alpha$ -Syn pathology independently.<sup>69,70</sup> We recently showed that *APOE*  $\epsilon 4$  aggravated  $\alpha$ -Syn phosphorylation, worsened motor impairment, and increased neuroinflammation and neurodegeneration in different mouse models.<sup>71</sup> The present results suggest a genetic link between *APOE* and CSF A $\beta$ 42 and A $\beta$  brain deposition in PD.

### **Limitations**

This study uses the largest sample size reported to date and yet found no significant genetic modifier of CSF  $\alpha$ -Syn levels. However, this study had the power to detect strong genetic associations (effect=0.57,  $p=4.46\times 10^{-43}$ ), comparable to the association of the *APOE* locus with A $\beta$  CSF values. It is possible that the current sample size is not sufficiently powered to detect signals with a smaller effect. Here, we found lower levels of CSF  $\alpha$ -Syn in PD patients, which aligns with previous reports. However, neither PRS nor MR analysis revealed evidence of the causal link of CSF  $\alpha$ -Syn with PD risk. It has been reported that  $\alpha$ -Syn aggregation is neither necessary nor sufficient for neurodegeneration or clinical parkinsonism.<sup>72,73</sup> The cohorts used in this study rely on clinical diagnosis rather than neuropathological confirmation, which precludes analyses of correlation between CSF  $\alpha$ -Syn levels and pathologic brain accumulation of brain  $\alpha$ -Syn. Thus, it is possible that CSF  $\alpha$ -Syn levels used here may not be reflecting aggregation in the brain. It is not entirely understood how  $\alpha$ -Syn gets into CSF, but its levels in CSF may reflect changes in the balance between  $\alpha$ -Syn synthesis, secretion, solubility, or degradation pathways.<sup>10</sup> Consequently, factors that influence CSF  $\alpha$ -Syn level variability remain poorly understood. Factors that may have contributed to the lack of power to replicate or detect genetic modifiers of CSF  $\alpha$ -Syn include participant characteristics (PD subtypes, misdiagnosis, comorbidities,



medications, disease duration), preanalytical factors (blood contamination at LP), and different assays (measuring various abnormal pathological or normal forms of  $\alpha$ -Syn)<sup>74</sup>. A unified and standardized platform, such as  $\alpha$ -Syn seeding aggregation assays, is necessary to maximize the power of future genetic analyses of CSF  $\alpha$ -Syn levels.<sup>75</sup>

## Conclusions

PD is a heterogeneous disorder with different identifiable clinical-pathological subtypes based on symptom severity and predominance. It is conceivable that more homogeneous PD subtypes could be defined using biomarker-driven, clinical-molecular phenotyping approaches. Our results have implications for designing anti- $\alpha$ -Syn therapeutic trials in which both *APOE* genotype and CSF A $\beta$ 42 levels could be highly informative. Markers of AD pathology (CSF A $\beta$ 42 and PiB) during life may have important prognostic indications in PD to guide clinical trials for homogeneous patient selection. This study, with 1,960 samples with CSF  $\alpha$ -Syn levels, shows that the genomic architecture of  $\alpha$ -Syn is complex and not correlated with the genomic landscape of PD risk. Additional studies with larger sample sizes and standardized methods to quantify  $\alpha$ -Syn in both CSF and brain are needed to uncover genetic modifiers of  $\alpha$ -Syn levels. Our results using high-throughput and hypothesis-free approaches detected the existence of PD genetic risk linked to CSF A $\beta$ 42 and *APOE* locus. These findings were further validated by strong significant associations of *APOE*  $\epsilon$ 4 with A $\beta$  deposition in cortical regions of living and postmortem PD patients.

## Acknowledgments

We thank all the participants and their families, as well as the many institutions and their staff.

## Data availability Statement

All data is available in the Center for Neurogenomics and informatics (NGI) website (<https://neurogenomics.wustl.edu/>). The summary statistics for all the analyses can be easily explored in the Online Neurodegenerative Trait Integrative Multi-Omics Explorer (ONTIME) (<https://omics.wustl.edu>) and the Charles F. and Joanne Knight Alzheimer Disease Research Center (<https://knightadrc.wustl.edu/research/resourcerequest.htm>).

## Funding

This work was supported by grants from the National Institutes of Health (R01AG044546, P01AG003991, RF1AG053303, R01AG058501, U01AG058922, RF1AG058501, R01AG057777, NS097437, NS075321, NS097799 and NS07532), the Alzheimer Association (NIRG-11-200110, BAND-14-338165, AARG-16-441560, and BFG-15-362540), the American Parkinson Disease Association (APDA), the Greater St. Louis Chapter of the APDA, the Jo Riney Fund, the Barnes Jewish Hospital Foundation (Elliot Stein Family Fund and Parkinson Disease Research Fund), and the Paula and Roger Riney Fund. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine. The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276.

### **Competing Interests**

CC receives research support from Biogen, Eisai, Alector, and Paragon. The funders of the study had no role in the collection, analysis, or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. CC is a member of the advisory board of Vivid genetics, Halia Therapeutics, and ADx Healthcare.

**Table 1:** Summary Demographics for the individuals with CSF measurements available

|                                       | <b>All</b>            | <b>PD Cases</b>       | <b>Controls</b>       | <b>AD Cases</b>         | <b>Non-neurodegenerative<sup>#</sup></b> |
|---------------------------------------|-----------------------|-----------------------|-----------------------|-------------------------|--|
| N                                     | 1,960                 | 700                   | 564                   | 386                     | 310                                      |
| Age (years, 95% CI)                   | 69.3<br>(62.0-75.0)   | 66.2<br>(59.6-73.3)   | 70.0<br>(64.9-74.3)   | 75.0<br>(70.0-80.0)     | 64.0<br>(55.0-73.0)                      |
| Males (N, %)                          | 1,107<br>(56.5%)      | 435<br>(62.1%)        | 297<br>(52.7%)        | 222<br>(57.5%)          | 153<br>(49.4%)                           |
| Alpha-Synuclein<br>(ZScore(pg/mL))    | -0.02<br>(-0.67-0.65) | -0.03<br>(-0.70-0.60) | 0.02<br>(-0.71-0.71)  | 0.14<br>(-0.57-0.80)    | -0.23<br>(-0.76-0.47)                    |
| Amyloid Beta<br>(ZScore(pg/mL))       | -0.20<br>(-0.72-0.64) | -0.20<br>(-0.72-0.45) | 0.11<br>(-0.52-1.04)  | -0.63<br>(-1.00- -0.25) | 0.03<br>(-0.49-0.90)                     |
| Total Tau<br>(ZScore(pg/mL))          | -0.27<br>(-0.66-0.40) | -0.27<br>(-0.70-0.39) | -0.34<br>(-0.67-0.28) | 0.22<br>(-0.41-1.00)    | -0.50<br>(-0.73- -0.08)                  |
| Phosphorylated Tau<br>(ZScore(pg/mL)) | -0.25<br>(-0.69-0.42) | -0.30<br>(-0.72-0.35) | -0.27<br>(-0.70-0.35) | 0.22<br>(-0.42-0.95)    | -0.48<br>(-0.81- -0.04)                  |

\* Concentration values have been standardized using ZScore for comparison purposes. Non-transformed values cannot be compared because there are several measuring methods: SomaScan platform, INNOTEST assay, xMAP-Luminex with INNOBIA AlzBio3, Roche Elecsys cobas e 601 and Euroimmun.

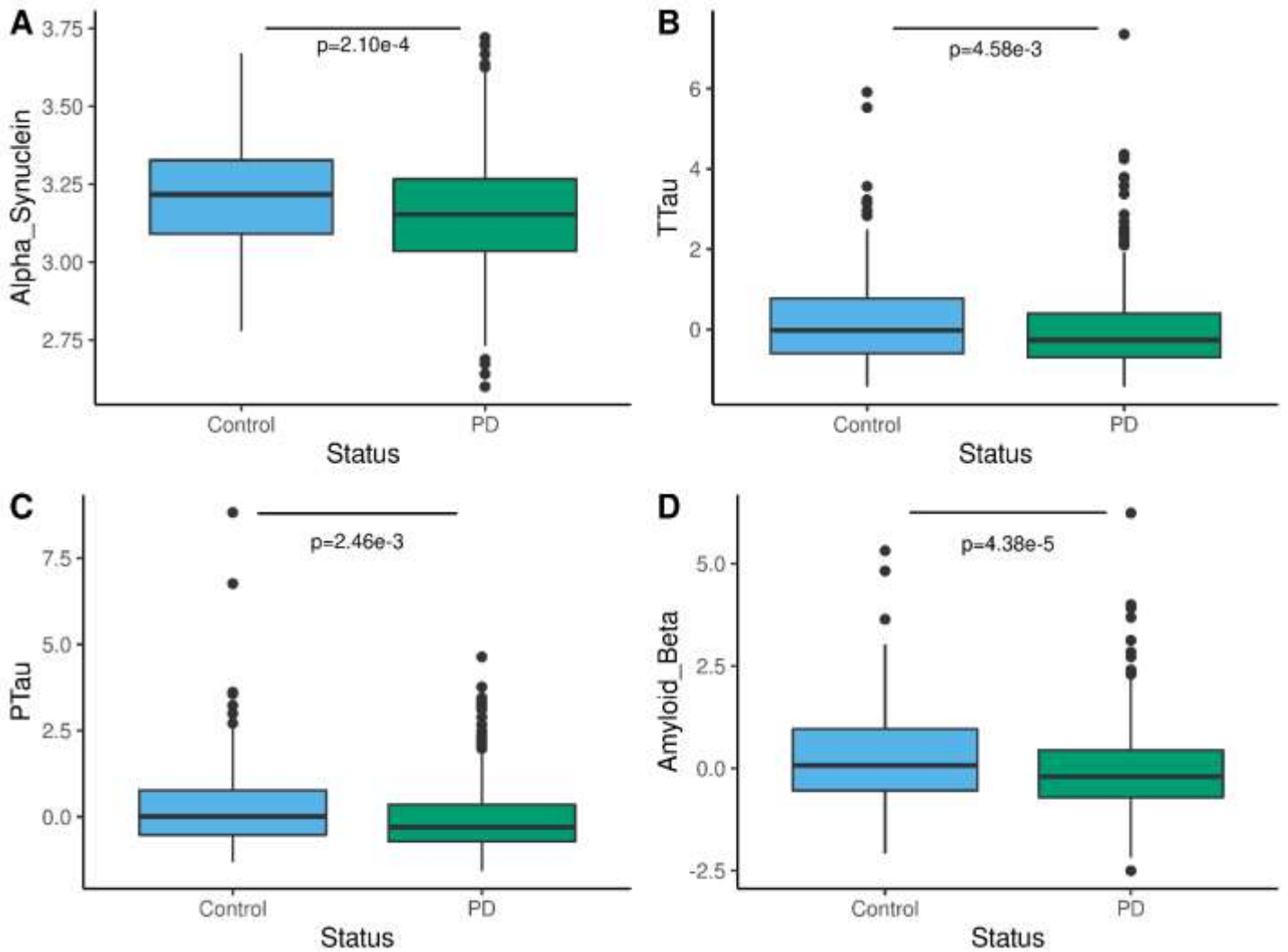
<sup>#</sup> Includes individuals classified as not being AD, PD or having dementia but neither controls such as essential tremor

**Table 2.** Significant and suggestive sentinel variants in the meta-analysis for GWA of CSF  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau levels.

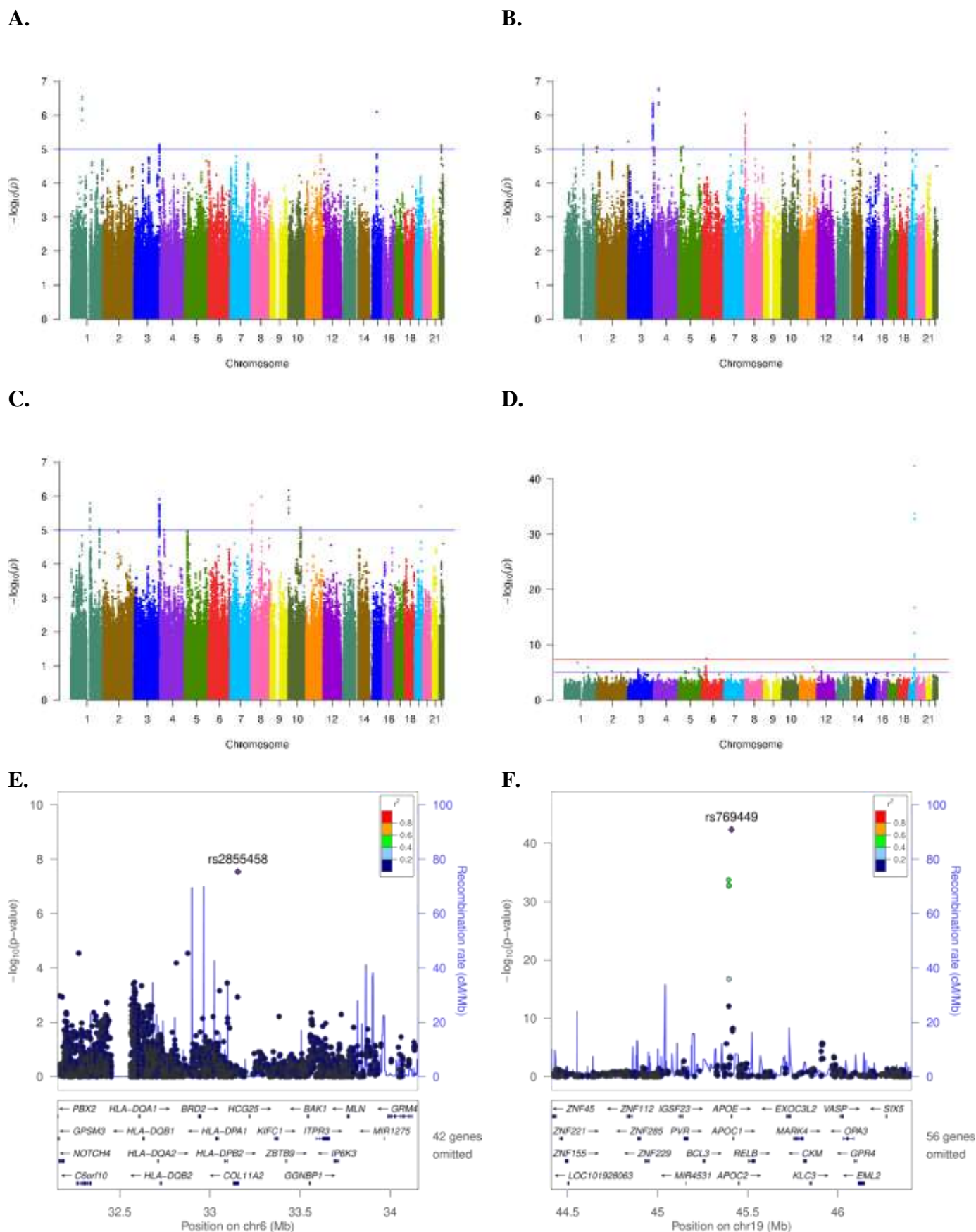
| SNP                       | Location     | Effect Allele | MAF   | Effect | STD Error | P-Value                |
|---------------------------|--------------|---------------|-------|--------|-----------|------------------------|
| <b>Alpha Synuclein</b>    |              |               |       |        |           |                        |
| rs74924741                | 1:86091232   | T             | 0.061 | -0.051 | 0.010     | $7.20 \times 10^{-07}$ |
| rs905124                  | 3:190657360  | A             | 0.360 | 0.023  | 0.005     | $7.22 \times 10^{-06}$ |
| rs76163673                | 15:55519643  | T             | 0.082 | 0.045  | 0.009     | $7.85 \times 10^{-07}$ |
| rs4820059                 | 22:32346468  | A             | 0.366 | 0.023  | 0.005     | $7.60 \times 10^{-06}$ |
| <b>Total Tau</b>          |              |               |       |        |           |                        |
| rs517201                  | 1:146528706  | T             | 0.238 | 0.170  | 0.038     | $7.40 \times 10^{-06}$ |
| rs57133061                | 2:476531     | A             | 0.053 | -0.314 | 0.070     | $8.46 \times 10^{-06}$ |
| rs2055568                 | 2:242393574  | A             | 0.715 | -0.161 | 0.036     | $5.88 \times 10^{-06}$ |
| rs13066848                | 3:190668204  | T             | 0.643 | -0.170 | 0.034     | $4.46 \times 10^{-07}$ |
| rs2101094                 | 4:1235643    | A             | 0.106 | 0.232  | 0.052     | $8.95 \times 10^{-06}$ |
| rs28656190                | 4:36521271   | A             | 0.055 | 0.385  | 0.074     | $1.58 \times 10^{-07}$ |
| rs76987944                | 5:19423843   | T             | 0.833 | 0.189  | 0.043     | $8.71 \times 10^{-06}$ |
| rs2453299                 | 5:36031870   | A             | 0.137 | 0.212  | 0.048     | $8.35 \times 10^{-06}$ |
| rs7830049                 | 8:3640771    | A             | 0.933 | -0.314 | 0.064     | $8.92 \times 10^{-07}$ |
| rs55803311                | 10:90753135  | A             | 0.117 | 0.229  | 0.051     | $7.16 \times 10^{-06}$ |
| rs11232401                | 11:80633016  | T             | 0.927 | -0.284 | 0.063     | $6.17 \times 10^{-06}$ |
| rs12888726                | 14:27302413  | A             | 0.506 | -0.143 | 0.032     | $9.28 \times 10^{-06}$ |
| rs7160073                 | 14:65450630  | A             | 0.359 | 0.149  | 0.033     | $8.55 \times 10^{-06}$ |
| rs799103                  | 14:82705630  | A             | 0.816 | -0.188 | 0.042     | $6.81 \times 10^{-06}$ |
| rs236012                  | 16:71305579  | A             | 0.235 | 0.180  | 0.039     | $3.14 \times 10^{-06}$ |
| <b>Phosphorylated Tau</b> |              |               |       |        |           |                        |
| rs517201                  | 1:146528706  | T             | 0.238 | 0.182  | 0.038     | $1.62 \times 10^{-06}$ |
| rs6659896                 | 1:219365421  | T             | 0.141 | 0.201  | 0.045     | $9.25 \times 10^{-06}$ |
| rs34911650                | 3:190670658  | C             | 0.642 | -0.163 | 0.034     | $1.21 \times 10^{-06}$ |
| rs7686226                 | 4:31156405   | A             | 0.558 | -0.145 | 0.033     | $9.76 \times 10^{-06}$ |
| rs35487809                | 8:79343275   | A             | 0.934 | -0.324 | 0.066     | $1.03 \times 10^{-06}$ |
| rs2502076                 | 10:2989837   | T             | 0.318 | 0.176  | 0.035     | $6.60 \times 10^{-07}$ |
| rs55803311                | 10:90753135  | A             | 0.117 | 0.227  | 0.051     | $8.22 \times 10^{-06}$ |
| rs769449                  | 19:45410002  | A             | 0.164 | 0.213  | 0.045     | $1.98 \times 10^{-06}$ |
| <b>Amyloid Beta</b>       |              |               |       |        |           |                        |
| rs7368467                 | 1:181249490  | A             | 0.145 | 0.216  | 0.045     | $1.39 \times 10^{-06}$ |
| rs3920740                 | 1:98233221   | A             | 0.357 | 0.292  | 0.056     | $1.64 \times 10^{-07}$ |
| rs2708959                 | 2:113670415  | A             | 0.238 | 0.207  | 0.046     | $5.91 \times 10^{-06}$ |
| rs116413594               | 2:237102380  | A             | 0.155 | 0.255  | 0.057     | $8.44 \times 10^{-06}$ |
| rs4993942                 | 3:76733503   | T             | 0.514 | -0.152 | 0.032     | $2.76 \times 10^{-06}$ |
| rs2174433                 | 5:118582779  | C             | 0.363 | 0.269  | 0.056     | $1.77 \times 10^{-06}$ |
| rs35978892                | 5:156025784  | A             | 0.947 | 0.340  | 0.072     | $2.40 \times 10^{-06}$ |
| rs13172707                | 5:53146524   | T             | 0.210 | 0.175  | 0.039     | $6.66 \times 10^{-06}$ |
| rs3752417                 | 6:26045905   | C             | 0.230 | 0.212  | 0.043     | $1.02 \times 10^{-06}$ |
| rs2855458                 | 6:33154698   | C             | 0.239 | 0.231  | 0.042     | $2.88 \times 10^{-08}$ |
| rs17435959                | 11:102651313 | C             | 0.276 | 0.243  | 0.050     | $1.11 \times 10^{-06}$ |
| rs4466925                 | 12:34380058  | T             | 0.167 | -0.195 | 0.043     | $5.95 \times 10^{-06}$ |
| rs769449                  | 19:45410002  | A             | 0.168 | -0.570 | 0.041     | $4.46 \times 10^{-43}$ |

**Table 3.** Mendelian randomization results for the causal role of  $\alpha$ -Syn, A $\beta$ 42, tau, and t-tau in PD using the robust regression MR-Egger method with robust regression

| <b>Biomarker</b>   | <b>PD Risk<sup>9</sup></b> |                        | <b>PD Age at Onset<sup>56</sup></b> |                        |
|--------------------|----------------------------|------------------------|-------------------------------------|------------------------|
|                    | <b>Effect</b>              | <b>P-Value</b>         | <b>Effect</b>                       | <b>P-Value</b>         |
| Alpha-Synuclein    | -1.389                     | 0.064                  | -11.018                             | 0.835                  |
| Amyloid-Beta       | 0.430                      | 1.44x10 <sup>-05</sup> | 7.746                               | 7.65x10 <sup>-06</sup> |
| Total Tau          | -0.338                     | 0.246                  | 11.276                              | 0.069                  |
| Phosphorylated Tau | -0.096                     | 0.785                  | -0.298                              | 0.912                  |



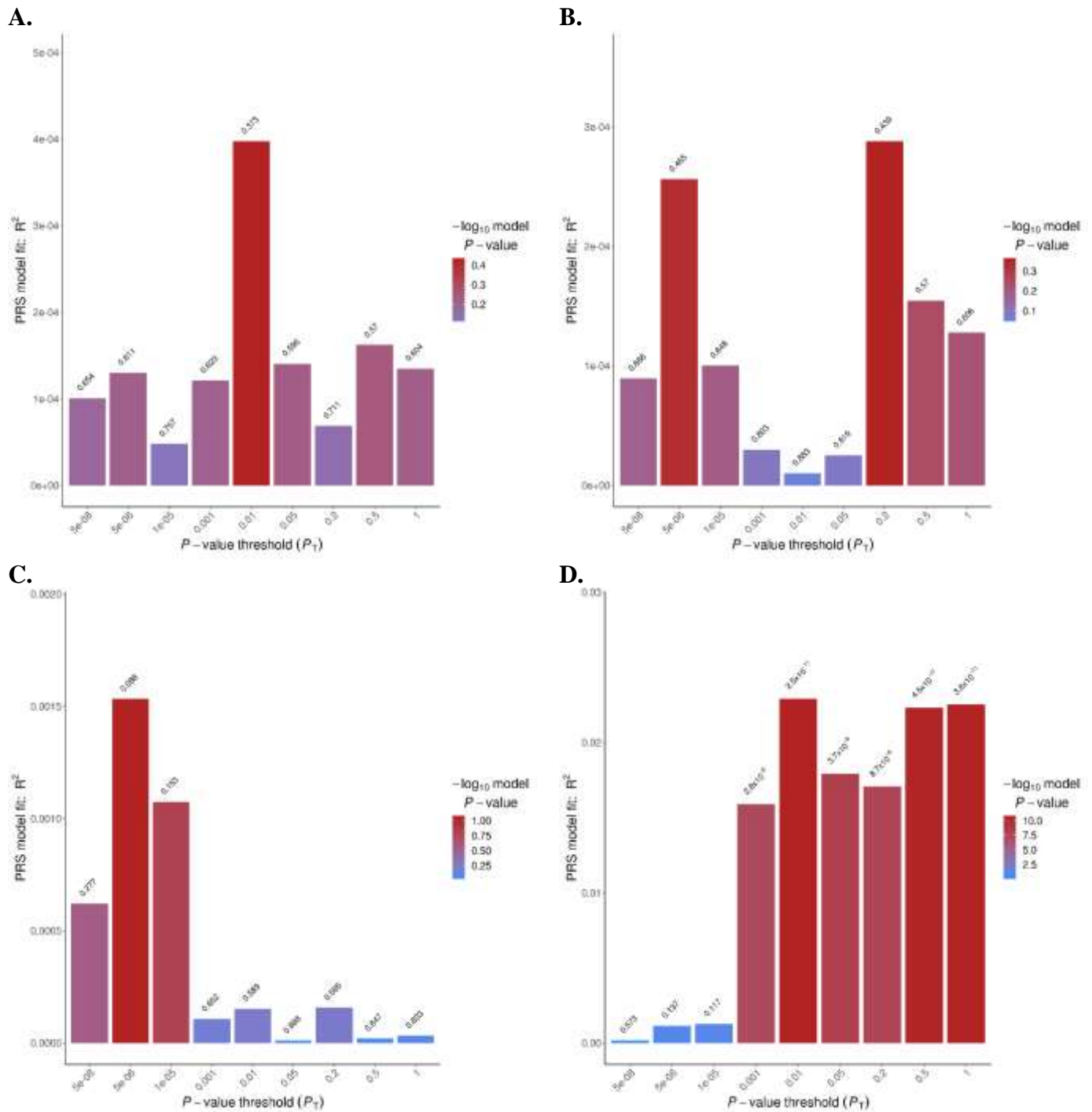
**Figure 1. CSF  $\alpha$ -Syn, A $\beta$ 42, t-tau, and p-tau<sub>181</sub> levels are lower in PD than in controls.** Box plot of the normalized CSF levels of **A.**  $\alpha$ -Syn. **B.** total tau. **C.** phosphorylated tau and **D.** A $\beta$ 42 in controls (gray) and PD cases (orange). PD cases (N=700) and controls (N=189) from two independent datasets (WUSTL and PPMI). The means for each group are represented by a horizontal line. A generalized linear model (CSF biomarker levels~Age+Sex+PD<sub>Status</sub>) was used to calculate the statistical differences between the CSF protein levels in PD cases and controls.



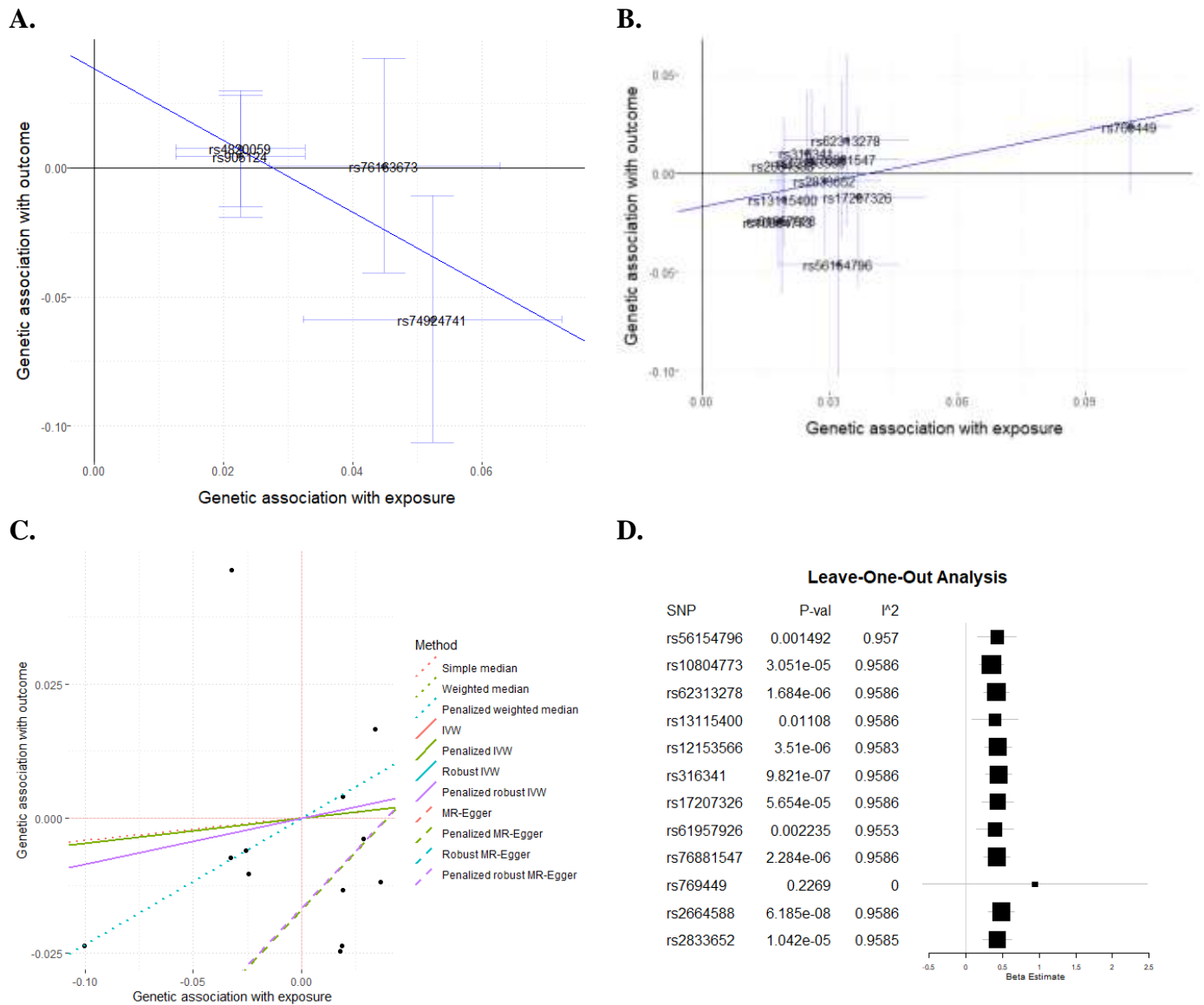
**Figure 2. Association plot of single variant analyses of CSF  $\alpha$ -Syn, t-tau, p-tau<sub>181</sub> and  $A\beta_{42}$  levels.** Manhattan plot shows negative  $\log_{10}$ -transformed p-values from the meta-analysis of **A.**  $\alpha$ -Syn. **B.** total tau. **C.** phosphorylated tau and **D.**  $A\beta_{42}$  CSF levels. The lowest p-value on chr19 (*APOE* locus) was  $p=4.5 \times 10^{-43}$ . The horizontal lines represent the genome-wide significance threshold,  $p=5 \times 10^{-8}$  (red) and suggestive threshold,

$p=1\times 10^{-5}$  (blue). **E, F** Regional association plots of loci are shown for SNPs associated with CSF A $\beta$ 42 levels near *HLA* (**E**) and near *APOE* locus (**F**). The SNPs labeled on each regional plot had the lowest p-value at each locus and are represented by a purple diamond. Each dot represents an SNP, and dot colors indicate linkage disequilibrium with the labeled SNP. Blue vertical lines show the recombination rate marked on the right-hand y-axis of each regional plot. Suggestive SNPs for  $\alpha$ -Syn, t-tau, p-tau<sub>181</sub> can be found in Tables S3 to S6.

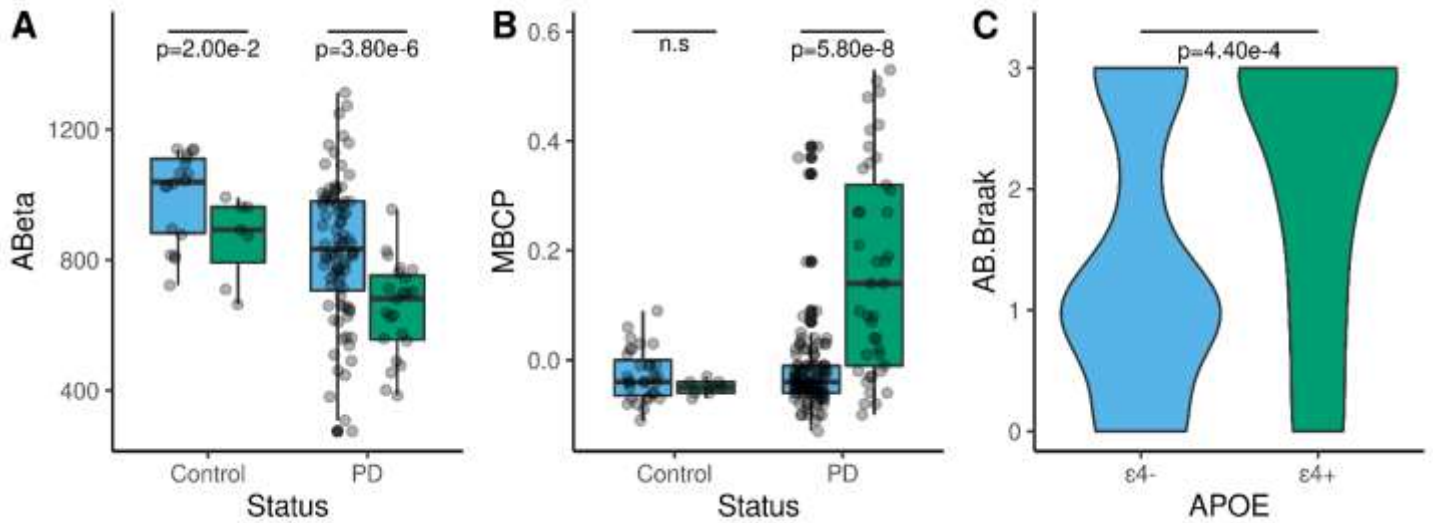




**Figure 3. Genetic architecture correlations of PD risk with CSF  $\alpha$ -Syn, t-tau, p-tau<sub>181</sub> and A $\beta$ 42 levels.** PRSice sum sum bar plots for PD risk and CSF biomarkers. Nagelkerke pseudo-R-squared fit for the model of **A.** CSF  $\alpha$ -Syn levels polygenic risk scores (PRS) and PD risk. **B.** CSF t-tau levels PRS—PD risk. **C.** CSF p-tau<sub>181</sub> PRS and PD risk. **D.** CSF A $\beta$ 42 levels PRS and PD risk. Total variance explained by the PRS for multiple p-value thresholds for the inclusion of SNPs, with the red bar indicating the optimal p-value threshold ( $P_T$ ), explaining the maximum amount of variance ( $R^2$ ) in PD risk in the target sample.



**Figure 4. MR regressions on PD risk genetic architecture and CSF  $\alpha$ -Syn and A $\beta$ 42 levels.** **A.** Association between META-PD risk and CSF  $\alpha$ -Syn levels (four variants). Robust regression MR-Egger method effect=-1.40 and  $p=0.06$ , which is not consistent with causality. **B.** Association between META-PD risk and CSF A $\beta$ 42 levels (twelve variants). Robust regression with MR-Egger method effect=0.43 and  $p=1.44 \times 10^{-05}$ , which is consistent with causality. Each dot corresponds to one genetic variant, with a 95% confidence interval (CI) of its genetic association with the exposure ( $\alpha$ -Syn and A $\beta$ 42 levels) and the outcome (META-PD risk). Regression lines correspond to the robust MR-Egger method regression; numerical results are given for all tested methods in Table S8. **C.** CSF A $\beta$ 42 regression using multiple MR methods. Each dot is one of the twelve variants included in this test; the effect of CSF A $\beta$ 42 levels on the x-axis and PD risk on the y-axis. Each line represents the regression of one MR-method of CSF A $\beta$ 42 levels on PD risk with one MR method. Additional details on the data sources and analysis methods to generate these figures are provided in Table S8. **D.** The forest plot illustrates the leave-one-out sensitivity analysis between CSF A $\beta$ 42 and META-PD risk. MR analysis without rs769449 decreased the I<sup>2</sup> statistic (I<sup>2</sup>=0.0%) and increased the p-value to non-significant levels, suggesting that the association is mainly driven by this variant.



**Figure 5. *APOE*  $\epsilon 4$  is associated with  $A\beta 42$  deposition in the brains of PD individuals.** **A.** Comparison of the levels of CSF  $A\beta 42$  in control (N=26) and PD (N=108) participants stratified by the presence ( $\epsilon 4+$ ; green) or absence ( $\epsilon 4-$ ; blue) of the *APOE*  $\epsilon 4$  allele. **B.** Effect of *APOE*  $\epsilon 4$  allele on the levels of mean cortical binding potentials (MCBP) in controls (N=44) and PD (N=156). **C.** PD patients carrying the *APOE*  $\epsilon 4$  allele exhibit a higher Braak  $A\beta$  score than non-carriers (N=92). Differences between *APOE*  $\epsilon 4$  carriers and non-carriers were statistically significant by the Mann-Whitney U test.

## References

- 1 Rizek, P., Kumar, N. & Jog, M. S. An update on the diagnosis and treatment of Parkinson disease. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* **188**, 1157-1165, doi:10.1503/cmaj.151179 (2016).
- 2 Dorsey, E. R., Sherer, T., Okun, M. S. & Bloem, B. R. The Emerging Evidence of the Parkinson Pandemic. *Journal of Parkinson's disease* **8**, S3-S8, doi:10.3233/JPD-181474 (2018).
- 3 Eusebi, P. *et al.* Diagnostic utility of cerebrospinal fluid alpha-synuclein in Parkinson's disease: A systematic review and meta-analysis. *Movement disorders : official journal of the Movement Disorder Society* **32**, 1389-1400, doi:10.1002/mds.27110 (2017).
- 4 Atik, A., Stewart, T. & Zhang, J. Alpha-Synuclein as a Biomarker for Parkinson's Disease. *Brain pathology* **26**, 410-418, doi:10.1111/bpa.12370 (2016).
- 5 Gao, L. *et al.* Cerebrospinal fluid alpha-synuclein as a biomarker for Parkinson's disease diagnosis: a systematic review and meta-analysis. *The International journal of neuroscience* **125**, 645-654, doi:10.3109/00207454.2014.961454 (2015).
- 6 Campbell, M. C. *et al.* Proteinopathy and longitudinal changes in functional connectivity networks in Parkinson disease. *Neurology* **94**, e718-e728, doi:10.1212/WNL.00000000000008677 (2020).
- 7 Polymeropoulos, M. H. *et al.* Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045-2047, doi:10.1126/science.276.5321.2045 (1997).
- 8 Eriksen, J. L., Przedborski, S. & Petrucelli, L. Gene dosage and pathogenesis of Parkinson's disease. *Trends in molecular medicine* **11**, 91-96, doi:10.1016/j.molmed.2005.01.001 (2005).
- 9 Nalls, M. A. *et al.* 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**, 1091-1102, doi:10.1016/S1474-4422(19)30320-5 (2019).
- 10 Mollenhauer, B. *et al.* Baseline predictors for progression 4 years after Parkinson's disease diagnosis in the De Novo Parkinson Cohort (DeNoPa). *Movement disorders : official journal of the Movement Disorder Society* **34**, 67-77, doi:10.1002/mds.27492 (2019).
- 11 Stewart, T. *et al.* Cerebrospinal fluid alpha-synuclein predicts cognitive decline in Parkinson disease progression in the DATATOP cohort. *Am J Pathol* **184**, 966-975, doi:10.1016/j.ajpath.2013.12.007 (2014).
- 12 Robinson, J. L. *et al.* Neurodegenerative disease concomitant proteinopathies are prevalent, age-related and APOE4-associated. *Brain : a journal of neurology* **141**, 2181-2193, doi:10.1093/brain/awy146 (2018).
- 13 Irwin, D. J. *et al.* Neuropathologic substrates of Parkinson disease dementia. *Ann Neurol* **72**, 587-598, doi:10.1002/ana.23659 (2012).
- 14 Brunnstrom, H., Hansson, O., Zetterberg, H., Londos, E. & Englund, E. Correlations of CSF tau and amyloid levels with Alzheimer pathology in neuropathologically verified dementia with Lewy bodies. *International journal of geriatric psychiatry* **28**, 738-744, doi:10.1002/gps.3881 (2013).
- 15 Campbell, M. C. *et al.* CSF proteins and resting-state functional connectivity in Parkinson disease. *Neurology* **84**, 2413-2421, doi:10.1212/WNL.0000000000001681 (2015).
- 16 Mollenhauer, B. *et al.* Longitudinal CSF biomarkers in patients with early Parkinson disease and healthy controls. *Neurology* **89**, 1959-1969, doi:10.1212/WNL.0000000000004609 (2017).
- 17 Siderowf, A. *et al.* CSF amyloid {beta} 1-42 predicts cognitive decline in Parkinson disease. *Neurology* **75**, 1055-1061, doi:10.1212/WNL.0b013e3181f39a78 (2010).
- 18 Kotzbauer, P. T. *et al.* Pathologic accumulation of alpha-synuclein and Abeta in Parkinson disease patients with dementia. *Archives of neurology* **69**, 1326-1331, doi:10.1001/archneurol.2012.1608 (2012).
- 19 Deming, Y. *et al.* Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta neuropathologica*, doi:10.1007/s00401-017-1685-y (2017).
- 20 Buddhala, C., Campbell, M. C., Perlmutter, J. S. & Kotzbauer, P. T. Correlation between decreased CSF alpha-synuclein and Abeta(1)(-)(4)(2) in Parkinson disease. *Neurobiology of aging* **36**, 476-484, doi:10.1016/j.neurobiolaging.2014.07.043 (2015).
- 21 Hughes, A. J., Daniel, S. E., Kilford, L. & Lees, A. J. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of neurology, neurosurgery, and psychiatry* **55**, 181-184 (1992).
- 22 Davis, A. A. *et al.* Variants in GBA, SNCA, and MAPT influence Parkinson disease risk, age at onset, and progression. *Neurobiology of aging* **37**, 209 e201-207, doi:10.1016/j.neurobiolaging.2015.09.014 (2016).

- 23 Benitez, B. A. *et al.* Resequencing analysis of five Mendelian genes and the top genes from genome-wide association studies in Parkinson's Disease. *Molecular neurodegeneration* **11**, 29, doi:10.1186/s13024-016-0097-0 (2016).
- 24 Diez-Fairen, M. *et al.* Pooled-DNA target sequencing of Parkinson genes reveals novel phenotypic associations in Spanish population. *Neurobiology of aging* **70**, 325 e321-325 e325, doi:10.1016/j.neurobiolaging.2018.05.008 (2018).
- 25 Kang, J. H. *et al.* Association of cerebrospinal fluid beta-amyloid 1-42, T-tau, P-tau181, and alpha-synuclein levels with clinical features of drug-naive patients with early Parkinson disease. *JAMA neurology* **70**, 1277-1287, doi:10.1001/jamaneurol.2013.3861 (2013).
- 26 Foster, E. R. *et al.* Amyloid imaging of Lewy body-associated disorders. *Movement disorders : official journal of the Movement Disorder Society* **25**, 2516-2523, doi:10.1002/mds.23393 (2010).
- 27 Rowland, D. J., Garbow, J. R., Laforest, R. & Snyder, A. Z. Registration of [18F]FDG microPET and small-animal MRI. *Nuclear medicine and biology* **32**, 567-572, doi:10.1016/j.nucmedbio.2005.05.002 (2005).
- 28 Mintun, M. A. *et al.* [11C]PIB in a nondemented population: potential antecedent marker of Alzheimer disease. *Neurology* **67**, 446-452, doi:10.1212/01.wnl.0000228230.26044.a4 (2006).
- 29 Logan, J. *et al.* Distribution volume ratios without blood sampling from graphical analysis of PET data. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* **16**, 834-840, doi:10.1097/00004647-199609000-00008 (1996).
- 30 Burack, M. A. *et al.* In vivo amyloid imaging in autopsy-confirmed Parkinson disease with dementia. *Neurology* **74**, 77-84, doi:10.1212/WNL.0b013e3181c7da8e (2010).
- 31 Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta neuropathologica* **82**, 239-259, doi:10.1007/BF00308809 (1991).
- 32 Khachaturian, Z. S. Diagnosis of Alzheimer's disease. *Archives of neurology* **42**, 1097-1105, doi:10.1001/archneur.1985.04060100083029 (1985).
- 33 Mirra, S. S. *et al.* The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* **41**, 479-486, doi:10.1212/wnl.41.4.479 (1991).
- 34 Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. *Neurobiology of aging* **18**, S1-2 (1997).
- 35 Delaneau, O., Coulonges, C. & Zagury, J. F. Shape-IT: new rapid and accurate algorithm for haplotype inference. *BMC bioinformatics* **9**, 540, doi:10.1186/1471-2105-9-540 (2008).
- 36 Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* **1**, 457-470, doi:10.1534/g3.111.001198 (2011).
- 37 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).
- 38 Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7, doi:10.1186/s13742-015-0047-8 (2015).
- 39 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191, doi:10.1093/bioinformatics/btq340 (2010).
- 40 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82, doi:10.1016/j.ajhg.2010.11.011 (2011).
- 41 Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* **50**, 229-237, doi:10.1038/s41588-017-0009-4 (2018).
- 42 Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-1468, doi:10.1093/bioinformatics/btu848 (2015).
- 43 Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience* **8**, doi:10.1093/gigascience/giz082 (2019).
- 44 Deming, Y. *et al.* Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. *Acta neuropathologica* **133**, 839-856, doi:10.1007/s00401-017-1685-y (2017).
- 45 Vassos, E. *et al.* An Examination of Polygenic Score Risk Prediction in Individuals With First-Episode Psychosis. *Biological psychiatry*, doi:10.1016/j.biopsych.2016.06.028 (2016).

- 46 Cruchaga, C. *et al.* Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, doi:10.1016/j.jalz.2017.08.013 (2017).
- 47 Ibanez, L. *et al.* Parkinson disease polygenic risk score is associated with Parkinson disease status and age at onset but not with alpha-synuclein cerebrospinal fluid levels. *BMC neurology* **17**, 198, doi:10.1186/s12883-017-0978-z (2017).
- 48 Ibanez, L. *et al.* Overlap in the Genetic Architecture of Stroke Risk, Early Neurological Changes, and Cardiovascular Risk Factors. *Stroke* **50**, 1339-1345, doi:10.1161/STROKEAHA.118.023097 (2019).
- 49 Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, doi:10.7554/eLife.34408 (2018).
- 50 Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *International journal of epidemiology* **46**, 1734-1739, doi:10.1093/ije/dyx034 (2017).
- 51 Burgess, S., Zuber, V., Gkatzionis, A. & Foley, C. N. Modal-based estimation via heterogeneity-penalized weighting: model averaging for consistent and efficient estimation in Mendelian randomization when a plurality of candidate instruments are valid. *International journal of epidemiology* **47**, 1242-1254, doi:10.1093/ije/dyy080 (2018).
- 52 Burgess, S., Foley, C. N., Allara, E., Staley, J. R. & Howson, J. M. M. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nature communications* **11**, 376, doi:10.1038/s41467-019-14156-4 (2020).
- 53 Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I<sup>2</sup> statistic. *International journal of epidemiology* **45**, 1961-1974, doi:10.1093/ije/dyw220 (2016).
- 54 Nalls, M. A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet* **46**, 989-993, doi:10.1038/ng.3043 (2014).
- 55 Chang, D. *et al.* A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat Genet* **49**, 1511-1516, doi:10.1038/ng.3955 (2017).
- 56 Blauwendraat, C. *et al.* Parkinson's disease age at onset genome-wide association study: Defining heritability, genetic loci, and alpha-synuclein mechanisms. *Movement disorders : official journal of the Movement Disorder Society* **34**, 866-875, doi:10.1002/mds.27659 (2019).
- 57 Zhong, X. L. *et al.* A Genome-Wide Association Study of alpha-Synuclein Levels in Cerebrospinal Fluid. *Neurotoxicity research* **35**, 41-48, doi:10.1007/s12640-018-9922-2 (2019).
- 58 Zhang, X. *et al.* Tau Pathology in Parkinson's Disease. *Frontiers in neurology* **9**, 809, doi:10.3389/fneur.2018.00809 (2018).
- 59 Kang, J. H. *et al.* CSF biomarkers associated with disease heterogeneity in early Parkinson's disease: the Parkinson's Progression Markers Initiative study. *Acta neuropathologica* **131**, 935-949, doi:10.1007/s00401-016-1552-2 (2016).
- 60 Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* **51**, 414-430, doi:10.1038/s41588-019-0358-2 (2019).
- 61 Ibanez, L. *et al.* Pleiotropic Effects of Variants in Dementia Genes in Parkinson Disease. *Frontiers in neuroscience* **12**, 230, doi:10.3389/fnins.2018.00230 (2018).
- 62 Rongve, A. *et al.* GBA and APOE epsilon4 associate with sporadic dementia with Lewy bodies in European genome wide association study. *Sci Rep* **9**, 7013, doi:10.1038/s41598-019-43458-2 (2019).
- 63 Rongve, A. *et al.* Author Correction: GBA and APOE epsilon4 associate with sporadic dementia with Lewy bodies in European genome wide association study. *Sci Rep* **9**, 15168, doi:10.1038/s41598-019-51827-0 (2019).
- 64 Aarsland, D. *et al.* Cognitive decline in Parkinson disease. *Nature reviews. Neurology* **13**, 217-231, doi:10.1038/nrneurol.2017.27 (2017).
- 65 Irwin, D. J. *et al.* Neuropathological and genetic correlates of survival and dementia onset in synucleinopathies: a retrospective analysis. *The Lancet. Neurology* **16**, 55-65, doi:10.1016/S1474-4422(16)30291-5 (2017).
- 66 Bassil, F. *et al.* Amyloid-Beta (Abeta) Plaques Promote Seeding and Spreading of Alpha-Synuclein and Tau in a Mouse Model of Lewy Body Disorders with Abeta Pathology. *Neuron* **105**, 260-275 e266, doi:10.1016/j.neuron.2019.10.010 (2020).

- 67 Hinrichs, A. L. *et al.* Cortical binding of pittsburgh compound B, an endophenotype for genetic studies of Alzheimer's disease. *Biological psychiatry* **67**, 581-583, doi:10.1016/j.biopsych.2009.09.012 (2010).
- 68 Shi, Y. *et al.* ApoE4 markedly exacerbates tau-mediated neurodegeneration in a mouse model of tauopathy. *Nature* **549**, 523-527, doi:10.1038/nature24016 (2017).
- 69 Dickson, D. W. *et al.* APOE epsilon4 is associated with severity of Lewy body pathology independent of Alzheimer pathology. *Neurology* **91**, e1182-e1195, doi:10.1212/WNL.0000000000006212 (2018).
- 70 Tsuang, D. *et al.* APOE epsilon4 increases risk for dementia in pure synucleinopathies. *JAMA neurology* **70**, 223-228, doi:10.1001/jamaneurol.2013.600 (2013).
- 71 Davis, A. A. *et al.* APOE genotype regulates pathology and disease progression in synucleinopathy. *Science translational medicine* **12**, doi:10.1126/scitranslmed.aay3069 (2020).
- 72 Espay, A. J. *et al.* Disease modification and biomarker development in Parkinson disease: Revision or reconstruction? *Neurology* **94**, 481-494, doi:10.1212/WNL.0000000000009107 (2020).
- 73 Espay, A. J. *et al.* Revisiting protein aggregation as pathogenic in sporadic Parkinson and Alzheimer diseases. *Neurology* **92**, 329-337, doi:10.1212/WNL.0000000000006926 (2019).
- 74 Dhavale, D. D. *et al.* A sensitive assay reveals structural requirements for alpha-synuclein fibril growth. *J Biol Chem* **292**, 9034-9050, doi:10.1074/jbc.M116.767053 (2017).
- 75 Kang, U. J. *et al.* Comparative study of cerebrospinal fluid alpha-synuclein seeding aggregation assays for diagnosis of Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society* **34**, 536-544, doi:10.1002/mds.27646 (2019).