Association between the LRP1B and APOE loci and the development of Parkinson’s disease dementia

Raquel Real, Alejandro Martinez-Carrasco, Regina H. Reynolds, Michael A. Lawton, Manuela M. X. Tan, Maryam Shoai, Jean-Christophe Corvol, Mina Ryten, Catherine Bresner, Leon Hubbard, Alexis Brice, Suzanne Lesage, Johann Faouzi, Alexis Elbaz, Fanny Artaud, Nigel Williams, Michele T. M. Hu, Yoav Ben-Shlomo, Donald G. Grosset, John Hardy and Huw R. Morris

Parkinson’s disease is one of the most common age-related neurodegenerative disorders. Although predominantly a motor disorder, cognitive impairment and dementia are important features of Parkinson’s disease, particularly in the later stages of the disease. However, the rate of cognitive decline varies among Parkinson’s disease patients, and the genetic basis for this heterogeneity is incompletely understood.

To explore the genetic factors associated with rate of progression to Parkinson’s disease dementia, we performed a genome-wide survival meta-analysis of 3923 clinically diagnosed Parkinson’s disease cases of European ancestry from four longitudinal cohorts. In total, 6.7% of individuals with Parkinson’s disease developed dementia during study follow-up, on average 4.4 ± 2.4 years from disease diagnosis.

We have identified the APOE ε4 allele as a major risk factor for the conversion to Parkinson’s disease dementia [hazard ratio = 2.41 (1.94–3.00), \( P = 2.32 \times 10^{-15} \)], as well as a new locus within the ApoE and APP receptor LRP1B gene [hazard ratio = 3.23 (2.17–4.81), \( P = 7.07 \times 10^{-09} \)]. In a candidate gene analysis, GBA variants were also identified to be associated with higher risk of progression to dementia [hazard ratio = 2.02 (1.21–3.32), \( P = 0.007 \)]. CSF biomarker analysis also implicated the amyloid pathway in Parkinson’s disease dementia, with significantly reduced levels of amyloid \( \beta_{42} \) (\( P = 0.0012 \)) in Parkinson’s disease dementia compared to Parkinson’s disease without dementia.

These results identify a new candidate gene associated with faster conversion to dementia in Parkinson’s disease and suggest that amyloid-targeting therapy may have a role in preventing Parkinson’s disease dementia.

1 Department of Clinical and Movement Neurosciences, UCL Queen Square Institute of Neurology, University College London, London WC1N 3BG, UK
2 UCL Movement Disorders Centre, University College London, London WC1N 3BG, UK
3 Aligning Science Across Parkinson’s (ASAP) Collaborative Research Network, Chevy Chase, MD, 20815, USA
4 Genetics and Genomic Medicine, UCL Great Ormond Street Institute of Child Health, University College London, London WC1N 1EH, UK
5 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2PS, UK
6 Department of Neurology, Oslo University Hospital, 0424 Oslo, Norway
7 Department of Neurodegenerative Diseases, UCL Queen Square Institute of Neurology, University College London, London WC1N 3BG, UK
8 UK Dementia Research Institute, University College London, London WC1E 6BT, UK

Received June 09, 2022. Revised October 04, 2022. Accepted October 16, 2022. Advance access publication November 9, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of the Guarantors of Brain. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Introduction

Parkinson’s disease (PD) is the second most common neurodegenerative disease, with an estimated worldwide prevalence of 100–200 cases per 100,000 individuals. Although PD is mainly viewed as a motor disorder, the development of dementia in PD is an important determinant of morbidity, mortality and need for social support. The clinico-pathological phenotype of Parkinson’s disease dementia (PDD) can be indistinguishable from dementia with Lewy bodies, although in PDD motor symptoms must by definition precede the development of dementia by at least 1 year. Neuropsychiatric manifestations of PDD include cognitive fluctuation with visual misperception, hallucinations and delusions together with deficits in attention, executive and visuo-spatial function. Cholinergic denervation seems to be important in PDD and cholinesterase inhibitors can improve PDD symptoms, but there is no treatment for the underlying disease pathology. Identifying the causal mechanisms will be an important step in defining new treatments.

Age is the single most important risk factor for PDD. It is estimated that by the age of 90, 80–90% of individuals with PD will have developed dementia. Other clinical predictors of progression to dementia include disease severity and longer disease duration. However, the rates of progression to PDD vary substantially among individuals, which has important implications for prognosis and quality of life. Several genetic factors have been reported to increase the risk or rate of progression to PDD. The most widely reported genetic risk factor associated with increased risk of conversion to PDD is the APOE ε4 allele. A meta-analysis of 17 studies found a significantly higher risk of developing dementia in PD carriers of the ε4 allele. Single rare variants in the GBA gene increase the risk of developing PDD, and the risk may relate to the pathogenicity of the variant. Several studies have also reported that the MAPT H1 haplotype is associated with dementia, although this has not been universally replicated. More recently, the RIMS2 locus has been described in association with progression to PDD, as well as suggestive association signals at the TMEM108 and WWOX loci. Genome-wide association studies in neurodegenerative disease have largely defined case-control risk factors for disease susceptibility, but the increasing availability of high-quality longitudinal clinical datasets enables a systematic search for disease modifying factors. Here, we use a genome-wide survival meta-analysis approach to identify new genetic factors that contribute to the progression to PDD.

Materials and methods

Patient cohorts

We have studied four independent longitudinal PD cohorts: Tracking Parkinson’s Disease (TPD, www.parkinsons.org.uk), Oxford Parkinson’s Disease Centre Discovery Cohort (OPDC, www.dpag.ox.ac.uk/opdc), Accelerating Medicines Partnership: Parkinson’s Disease (AMP-PD v2.5, www.amp-pd.org), which consists of harmonized data from multiple cohorts, and Drug Interaction With Genes.
in Parkinson’s Disease (DIGPD, clinicaltrials.gov/ct2/show/NCT01564992), comprising a total of 3923 participants after clinical and genetic data cleaning (Supplementary Fig. 1 and Table 1). Each subject provided written informed consent for participation according to the Declaration of Helsinki and all cohort studies were approved by the relevant ethics committee. Methods for clinical data collection, including setting, inclusion criteria and periods of recruitment, are available from the websites of the corresponding cohorts. All participants were diagnosed with PD according to the Queen Square Brain Bank criteria. Participants were excluded from the analysis if an alternative diagnosis was made during the follow-up period (including a diagnosis of dementia with Lewy bodies) and/or the probability of a PD diagnosis as assessed by a clinician at the last available visit was <90%. In AMP-PD, only individuals in the PD study arm were included to avoid selection bias of monogenic cases. Criteria for PDD were based on the Movement Disorders Society–Unified Parkinson’s disease Rating Scale part I score ≥4/5, visuospatial/executive ≤4/5, a cognitive deficit severe enough to affect activities of daily living (Movement Disorders Society–Unified Parkinson’s disease Rating Scale part I 1.1 score ≥2), and absence of severe depression (Movement Disorders Society–Unified Parkinson’s disease Rating Scale part I 1.3 score <4), except participants from the DIGPD cohort, for whom only Mini-Mental State Examination (MMSE) scores were available together with a clinician assigned diagnosis of dementia. Participants were excluded from the study (left censored) if they met criteria for PDD at study baseline (Supplementary Table 1). Time-to-event was calculated as the number of years from disease diagnosis until the midpoint between the date of the first visit where criteria for PDD were met or of study withdrawal due to dementia and the date of the previous available visit. The time interval between the last normal assessment and withdrawal due to dementia was on average 1.66±0.77 years for TPD and 2.57±1.18 years for OPDC. Individuals with missing data regarding time-to-event or event classification were also excluded from the study. Time intervals between visits varied across studies, with assessments being carried out every 18 months in the TPD and OPDC cohorts and every 12 months in the DIGPD and AMP-PD cohorts. Comparisons across cohorts were performed in R (R Project for Statistical Computing, RRID:SCR_001905; v.4.1.3; https://www.R-project.org/) using Pearson’s Chi-squared test (rstatix package, v.0.7.0; RRID:SCR_021240; https://CRAN.R-project.org/package=rstatix) for categorical variables, and Kruskal-Wallis test with Dunn’s test for post hoc multiple pairwise comparisons for continuous variables, with P-values adjusted by the Bonferroni method (stats package, v.4.1.3; https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html). Significance was set at α = 0.05.

### Data quality control

Whole-genome sequence data were available from participants in AMP-PD cohorts. The remainder of samples were genotyped with the Illumina HumanCoreExome array (TPD), Illumina HumanCoreExome-12 v.1.1 or Illumina Infinium HumanCoreExome-24 v.1.1 arrays (OPDC) and the Illumina Infinium Multi-Ethnic Global (MEGA) array (DIGPD). Sample quality control (QC) included the exclusion of samples with call rates <98%, samples with excess heterozygosity [defined as samples deviating more than two standard deviations (>2 SD) from the mean heterozygosity rate], samples with a mismatch between clinical sex and genetically determined sex from chromosome × heterogeneity, and samples from related individuals (pi-hat >0.125). Variants with missingness rate >5%, minor allele frequency <0.01 and Hardy–Weinberg equilibrium P < 1 × 10−5 were excluded. To identify the ancestry, variants in linkage disequilibrium were removed and samples clustered against the HapMap3 reference panel, using principal component analysis. Individuals who deviated >6 SD from the mean of the first 10 principal components of the HapMap3 CEU+TSI population were excluded from the analysis (Supplementary Fig. 12A). To avoid inclusion of individuals related to each other across the different cohorts, we merged the genetic data from all cohorts and performed a second relatedness check (pi-hat >0.125). For each pair of related individuals, the one with the highest missingness rate was excluded from the respective cohort. After extraction of European-ancestry samples and non-related individuals from each cohort, principal components were re-calculated to use as covariates. The genotyping array data were then imputed against the Haplotype Reference Consortium reference panel (v.r1.1 2016; http://www.haplotype-reference-consortium.org/) in the Michigan Imputation Server (RRID:SCR_017579; https://imputersoftware.sph.umich.edu/) using Minimac4 (v.1.0.0; https://genome.sph.umich.edu/wiki/Minimac4 v.1.0.0). Imputed variants were excluded if the imputation info R2 score was <0.3. Following imputation, variants with missingness >5% and minor allele frequencies <1% were also excluded. Data cleaning was performed using PLINK v.1.9 (RRID:SCR_001757; https://www.cog-genomics.org/plink/1.9/).

<table>
<thead>
<tr>
<th>Table 1 Demographic characteristics per cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohorts</strong></td>
</tr>
<tr>
<td><strong>TPD</strong></td>
</tr>
<tr>
<td><strong>OPDC</strong></td>
</tr>
<tr>
<td><strong>DIGPD</strong></td>
</tr>
<tr>
<td><strong>AMP-PD</strong></td>
</tr>
<tr>
<td><strong>n (% female)</strong></td>
</tr>
<tr>
<td>1424 (35.4)</td>
</tr>
<tr>
<td>772 (35.0)</td>
</tr>
<tr>
<td>370 (39.7)</td>
</tr>
<tr>
<td>1357 (38.2)</td>
</tr>
<tr>
<td><strong>Age at diagnosis, years</strong></td>
</tr>
<tr>
<td>65.8 ± 9.1</td>
</tr>
<tr>
<td>65.9 ± 9.5</td>
</tr>
<tr>
<td>59.6 ± 9.9</td>
</tr>
<tr>
<td>60.0 ± 9.7</td>
</tr>
<tr>
<td><strong>Age at baseline, years</strong></td>
</tr>
<tr>
<td>67.2 ± 9.1</td>
</tr>
<tr>
<td>67.1 ± 9.4</td>
</tr>
<tr>
<td>62.2 ± 9.9</td>
</tr>
<tr>
<td>63.7 ± 9.1</td>
</tr>
<tr>
<td><strong>MoCA at baseline</strong></td>
</tr>
<tr>
<td>25.5 ± 3.1</td>
</tr>
<tr>
<td>25.1 ± 3.2</td>
</tr>
<tr>
<td>n/a</td>
</tr>
<tr>
<td>26.5 ± 2.8</td>
</tr>
<tr>
<td><strong>MMSE at baseline</strong></td>
</tr>
<tr>
<td>n/a</td>
</tr>
<tr>
<td>n/a</td>
</tr>
<tr>
<td>28.3 ± 1.73</td>
</tr>
<tr>
<td>n/a</td>
</tr>
<tr>
<td><strong>Years from diagnosis to baseline</strong></td>
</tr>
<tr>
<td>1.32 ± 0.90</td>
</tr>
<tr>
<td>1.21 ± 0.93</td>
</tr>
<tr>
<td>2.55 ± 1.52</td>
</tr>
<tr>
<td>3.70 ± 4.68</td>
</tr>
<tr>
<td><strong>Education ≤12 years (%)</strong></td>
</tr>
<tr>
<td>31.5</td>
</tr>
<tr>
<td>38.3</td>
</tr>
<tr>
<td>37.8</td>
</tr>
<tr>
<td>12.3</td>
</tr>
<tr>
<td><strong>Event rate (%)</strong></td>
</tr>
<tr>
<td>7.2</td>
</tr>
<tr>
<td>12.4</td>
</tr>
<tr>
<td>5.13</td>
</tr>
<tr>
<td>3.54</td>
</tr>
<tr>
<td><strong>Years from PD diagnosis to dementia</strong></td>
</tr>
<tr>
<td>3.80 ± 2.00</td>
</tr>
<tr>
<td>4.38 ± 2.25</td>
</tr>
<tr>
<td>6.24 ± 2.49</td>
</tr>
<tr>
<td>4.91 ± 2.79</td>
</tr>
<tr>
<td><strong>Months in study (median)</strong></td>
</tr>
<tr>
<td>44.1</td>
</tr>
<tr>
<td>53.8</td>
</tr>
<tr>
<td>60.2</td>
</tr>
<tr>
<td>30</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation, except where indicated.

Investigation for New Discovery of Biomarkers (BioFIND) study (n = 88); Parkinson’s Disease Biomarker Program (PDBP) study (n = 670); Parkinson’s Progression Markers Initiative (PFMI) study (n = 368); SURE-PD3 study (n = 231).
Time-to-event genome-wide survival study and meta-analysis

A time-to-event genome-wide survival study (GWSS) was performed in R (v.4.1.2) in each cohort, using the Cox proportional hazards (CPH) function in the survival package (v.3.2.13; RRID:SCR_021137; https://CRAN.R-project.org/package=survival), in which time to PDD was regressed against each single nucleotide polymorphism (SNP), with age at diagnosis, sex and first five principal components as covariates. AMP-PD summary statistics were converted from hg38 to hg19 using the binary liftOver tool (RRID:SCR_018160; https://genome.sph.umich.edu/wiki/LiftOver). The summary results from each cohort were then meta-analysed using METAL software in a random-effects model, using genomic control correction (version released on the 25 March 2011; RRID:SCR_002013; http://csg.sph.umich.edu/abecasis/Metal/).25 The genomic inflation factor ($\lambda_G$) for each cohort varied between 0.863 and 0.9773. After the meta-analysis, the $\lambda_G$ was 1.035 (Supplementary Fig. 12B). On completion of the meta-analysis, variants that were not present in all samples were excluded, as well as variants with minor allele frequency variability >15% across studies. Variants were also excluded if the $P$-value for the Cochran’s Q-test for heterogeneity was $<0.05$ and the $I^2$ statistic was $\geq80$. Forest plots of variants of interest were prepared using the R package forestplot (v.2.0.1; https://CRAN.R-project.org/package=forestplot). Results of the meta-analysis were annotated using FUMA (Functional Mapping and Annotation of Genome-Wide Association Studies, RRID: SCR_017521; v.1.3.8; https://fuma.ctglab.nl/).26 Regional association plots were generated in LocusZoom (RRID:SCR_012174; http://locuszoom.org/).27 LDproxy (https://dl/dlink.nci.nih.gov/?tab=ldproxy)28 was used to identify variants in high linkage disequilibrium with variants of interest.

Tissue and cell-type specificity measures

Specificity represents the proportion of a gene’s total expression attributable to one cell type/tissue. To determine specificity of a gene to a tissue or cell type, specificity values from three independent gene expression datasets were generated. Briefly, these datasets included (i) bulk-tissue RNA-sequencing of 53 human tissues from the Genotype-Tissue Expression consortium (GTEx, v.8; RRID:SCR_013042);29 (ii) human single-nucleus RNA-sequencing of the middle temporal gyrus from the Allen Institute for Brain Science (AIBS, Allen Cell Types Database—Human MTG Smart-Seq 2018 dataset, available from celltypes.brain-map.org/rnaseq; RRID:SCR_014806);30 and (iii) human single-nucleus RNA-sequencing of the substantia nigra.31 Generation of specificity values for GTEx and AIBS were previously described in Chia et al.32 Briefly, specificity values for GTEx were generated using code modified from a previous publication (https://github.com/jbryois/sRNA_disease),33 to reduce redundancy among brain regions and to include protein- and non-protein-coding genes. Specificity values for the AIBS-derived dataset were generated using gene-level exonic reads and the ‘generate.celltype.data’ function of the EWCE R package (v.1.2.0).34 Likewise, specificity values from Agarwal et al.34 were generated using EWCE. Specificity values for all three datasets and the code used to generate these values are openly available at https://github.com/RHReynolds/MarkerGenes.35

Conditional analysis

To understand whether one or more genome-wide significant variants at the same locus were contributing to the signal, we performed conditional analysis on single SNPs using a conditional and joint association analysis approach. We used the GWSS meta-analysis summary statistics and the entire AMP-PD cohort ($n=10418$) as the reference sample for linkage disequilibrium estimation. The reference sample was subjected to the same QC steps as described before. We then used GTGA-COJO software (v.1.9.3.0 beta for Linux; https://yanglab.westlake.edu.cn/software/gtga/#Overview)36 to perform association analysis conditional on SNPs of interest.

Colocalization analysis

To investigate whether there is an overlap between PDD loci and expression quantitative trait loci (eQTLs), we used the coloc R package (v.5.1.0; https://cran.r-project.org/web/packages/coloc/index.html).37 We also used the R package colochelpR (v.0.99.0)38 to help prepare datasets for use with coloc. We took a Bayesian inference approach to test the H4 null hypothesis that there is a shared causal variant associated with both progression to PDD and gene expression regulation. The Bayesian inference approach additionally computes the posterior probability that there is no association with either trait (H0), there is association with the PDD trait but not the eQTL trait (H1), there is association with the eQTL trait but not the PDD trait (H2) and that there is association with both traits, but the causal variants are independent (H3). We extracted all the genes within 1 Mb of each significant locus in the PDD GWSS. Coloc was run using default $p_1=10^{-8}$, $p_2=10^{-4}$ and $p_{12}=10^{-12}$ priors ($p_1$ and $p_2$ are the prior probability that any random SNP in the region is associated with trait 1 and 2, respectively, while $p_{12}$ is the prior probability that any random SNP in the region is associated with both traits). A PP$H_4>0.9$ was considered evidence for the presence of a shared variant between traits, i.e. signal colocalization. Coloc calculates Bayes factors under the assumption that a single causal variant exists within a particular locus. This assumption may be relaxed by successively conditioning on the most significant variants for each trait, and testing for colocalization between each pair of conditioned signals.39 We therefore performed conditional analysis beforehand to confirm that there were no additional independent signals, thus meeting the assumption of a single causal variant at each locus. Cis-eQTL data were obtained from (i) eQTLGen, comprising bulk blood-derived gene expression from 31684 individuals (https://www.eqtigen.org/cis-eqtls.html, accessed on the 7 June 2021); and (ii) PsychENCODE, comprising gene expression from bulk RNA-sequencing from the prefrontal cortex of 1387 individuals (http://resource.psychencode.org/, accessed on the 7 June 2021).40,41 Next, to understand if LRP1B or BBS9 loci regulate alternative splicing, we used a similar approach using frontal cortex and substantia nigra splicing QTLs (eQTLs) data from the GTEx v.8 database containing all variant-gene associations from 183 and 100 individuals, respectively, based on LeafCutter (v.0.2.9; RRID:SCR_017639; https://davidknoewles.github.io/leafcutter/42) to identify any exon excision phenotypes. For LRPIB, we tested the alternative splicing from eight different introns. In addition, false discovery rate-filtered transcript-per-million transcript expression QTLs (tQTLs) (false discovery rate <0.001) were obtained from PsychENCODE and used to generate regional association plots overlapping with LRP1B signals. A full colocalization analysis for tQTLs was not possible due to the unavailability of unfiltered tQTL summary statistics from PsychENCODE.

Signal interaction between APOE and LRP1B

Given the affinity of LRP1B for ApoE-carrying lipoproteins, we conducted a survival analysis based on APOE $\alpha4$ allele and LRP1B...
rs80306347 carrier status to understand whether the effect of LRP1B rs80306347 signal was dependent on APOE. APOE genotypes were inferred from the imputed genotypes of rs7412 and rs429358 variants. Participants of the combined cohorts (n = 3923) were grouped according to the presence of the two markers either simultaneously or exclusively, and a CPH model adjusted for age at diagnosis, gender, the first five principal components and a cohort term was performed. We also conditioned the analysis on APOE ε4 carrier status by performing a survival analysis of LRP1B rs80306347 on APOE ε4 carriers and non-carriers separately. We have in addition performed the analysis including an interaction term between LRP1B rs80306347 and APOE ε4 carrier status.

Candidate loci analysis

We additionally performed a candidate loci analysis of specific loci or variants of interest in the combined cohorts to increase power (n = 3923), using CPH models adjusted for age at diagnosis, sex, the first five principal components and a cohort term. The regions of interest consisted of genetic variants or loci previously identified in association with cognitive impairment in PD and/or dementia with Lewy bodies: APOE ε4 allele (rs429358), GBA variants E365K (or E326K, rs2290288), T408M (or T369M, rs75548401) and N490S (or N370S, rs76763715), SNCA (rs356219, rs7680557, rs7681440, rs11931074, rs7684318), MHT1 haplotype (rs1800547), RIMS2 (rs182987047), TME108 (rs138073281) and WWOX (rs8050111). In addition, participants from DIGPD and a subset of individuals from the TPD study were Sanger sequenced for GBA ε4 allele status, for which we defined GBA mutation carriers as individuals with at least one Gaucher disease-causing mutation or PD-risk factor (Supplementary Table 2).

Genetic risk scores

To understand whether there is overlap in the risk of development of PDD and the risk of PD or Alzheimer’s disease, we performed a genetic risk score (GRS) analysis using PLINK v.1.9 software. Scores were calculated using the summary statistics from the largest PD genome-wide association study (GWAS) to date and the 2019 genome-wide association meta-analysis of Alzheimer’s disease, respectively. Only the independent genome-wide significant risk signals were used in the analysis. Scores were then z-transformed and added as a covariate in a logistic regression model, together with age at diagnosis, sex and the first five principal components. Each cohort was analysed independently, and results were meta-analysed using the meta R package (v.5.1-1; RRID: http://cran.r-project.org/web/packages/meta). We conducted the Alzheimer’s disease-GRS analysis also without the APOE signal to investigate if the effect of Alzheimer’s disease-GRS in the risk of developing PDD was mediated by factors independent of APOE. For the survival analysis based on Alzheimer’s disease-GRS, individuals were stratified into low-, middle- and high-risk tertiles of raw Alzheimer’s disease-GRS. We used CPH models adjusted for age at diagnosis, sex and the first five principal components in each cohort; results were then meta-analysed with the R package meta.

Association of clinical phenotype and APOE genotype with CSF biomarkers

A subset of AMP-PD participants [from the Investigation for New Discovery of Biomarkers (BioFIND) and Parkinson’s Progression Markers Initiative (PPMI) studies] included in the analysis have longitudinal CSF Alzheimer’s disease biomarker data available (n = 434). We investigated the association of phenotype (PDD versus PD) and APOE ε4 carrier status with average levels of amyloid beta (Aβ42, total tau and tau phosphorylated at threonine 181 (p-Tau181)) using unpaired two-sample Wilcoxon rank-sum tests (R stats package, v.4.1.2) at baseline, 12, 24 and 36 months of follow-up. Significance was set at α = 0.05.

Statistical power modelling

The R package survSNP (https://cran.r-project.org/web/packages/survSNP/index.html; v.0.25) was used to model statistical power across a range of minor allele frequencies and effect sizes. The time-to-event was fixed at 4.5 years. Modelling accounted for the event rates observed in the different cohorts.

Data availability

Meta-analysis summary statistics are available for download from https://pdgenetics.org/resources. TPD data are available on access request from https://www.trackingparkinsons.org.uk/about-1/data/. BioFIND, PPMI, Parkinson’s Disease Biomarker Program (PDBP) and SURE-PD3 cohorts were accessed from AMP-PD and data are available on registration at https://www.amp-pd.org/. OPDC data are available on request from the Dementias Platform UK (https://portal.dementiasplatform.uk/Apply). DIGPD data are available on request to the principal investigator (J.C. Corvol, Assistance Publique Hôpitaux de Paris). HapMap phase 3 data (HapMap3) are available for download at ftp://ftp.ncbi.nlm.nih.gov/hapmap/. The Ashkenazi Jewish population panel is accessible at https://www.ncbi.nlm.nih.gov/gds (accession ID: GSE23636). Cis-QTL data were obtained from eQTLGen (https://www.eqtlgen.org/cis-eqtlGen) and PsychENCODE (http://resource.psychencode.org). False discovery rate-filtered tQTL data were obtained from PsychENCODE (http://resource.psychencode.org/). Cortical sQTL data were accessed from the GTEx v.8 database (https://gtexportal.org/home/). GTEx bulk-tissue RNA-seq data are available at https://www.gtexportal.org/home/datasets. AIBS human single-nucleus RNA-seq data are available at https://portal.brain-map.org/atlases-and-data/maseq. Human single-nucleus RNA-seq of the substantia nigra data can be accessed from https://www.ncbi.nlm.nih.gov/geo/ (accession ID: GSE140231). Summary statistics from the PD GWAS (Nalls et al.47) used to perform the GRS analysis are available from https://pdgenetics.org/resources. Code used in the analysis is available from https://github.com/huw-morris-lab/PDD_GWSS (https://doi.org/10.5281/zenodo.6535455).

Results

CoHORT CHARACTERIZATION

Following data cleaning (Supplementary Fig. 1), a total of 3923 individuals diagnosed with PD were available for analysis, with an overall mean follow-up time of 43.4 ± 27.7 months. Demographic characteristics of each patient cohort are shown in Table 1. Participants in DIGPD and AMP-PD cohorts were significantly younger at PD diagnosis (Kruskal–Wallis chi-squared value = 348, d.f. = 3, P < 2 × 10^{-16}, post hoc Dunn’s multiple comparison test in Supplementary Table 3) and at study baseline (Kruskal–Wallis chi-squared value = 160, d.f. = 3, P < 2 × 10^{-16}, post hoc Dunn’s multiple comparison test in Supplementary Table 3), which is probably reflected in the significantly reduced event rates in these two
cohorts (Table 1). Adjusted MoCA or MMSE scores over time in cases of PD who did not develop dementia during the study follow-up remained constant over time, while they were consistently lower and showed greater decline in individuals who went on to develop PDD during the study follow-up (Supplementary Fig. 2).

Identification of genetic determinants of Parkinson’s disease dementia

In our genome-wide survival meta-analysis assessing the role of 610 7418 SNPs in the development of PDD, we identified three genome-wide significant genetic loci (Fig. 1 and Table 2; regional association plots in Supplementary Fig. 3). The most significant SNP was the e4 allele-tagging SNP rs429358 in APOE [hazard ratio (HR) = 2.41, 95% confidence interval (CI) = 1.94–3.00, \( P = 2.32 \times 10^{-15} \)]. APOE is the most important genetic risk factor for the development of Alzheimer’s disease and has also been shown in multiple studies to contribute to cognitive decline and dementia in PD.8–10,51 Conditional analysis on the lead SNP at the APOE locus did not reveal any other independent SNPs contributing to the signal at this location (Supplementary Fig. 4A and B).

The second genome-wide significant genetic locus was on chromosome 2. The lead SNP at this locus was rs80306347 (HR = 3.23, 95% CI = 2.17–4.81, \( P = 7.07 \times 10^{-09} \)). This is an intronic variant located in intron 5 of the LRP1B gene (ENSG00000168702). This gene encodes the low-density lipoprotein (LDL) receptor-related protein 1B, a member of the LDL receptor superfamily. LRP1B is a receptor for ApoE-carrying lipoproteins and is highly expressed in the adult human brain (Supplementary Fig. 5A and C).52 In addition, LRP1B was found to be significantly upregulated in excitatory neurons of the anterior cingulate cortex of PDD compared to PD and control brain samples (Supplementary Fig. 6).53 Similar to other LDL receptors, it is involved in the intracellular processing of the amyloid precursor protein (APP).54 Therefore, LRP1B constitutes a promising candidate for regulating the development of dementia in PD. Conditioning on the rs80306347 variant abolished genome-wide significance at the LRP1B locus, confirming that a single independent SNP is responsible for the signal at this location (Supplementary Fig. 4C and D).

rs78294974 is an intronic variant in the BBS9 gene (ENSG00000122507) in chromosome 7 and was associated with progression to dementia with genome-wide significance (HR = 3.90, 95% CI = 2.40–6.32, \( P = 3.59 \times 10^{-08} \)). This gene is ubiquitously expressed (Supplementary Fig. 5B and D) and encodes the Parathyroid Hormone-Responsive B1 (PTHB1) protein. This protein is part of a stable evolutionary conserved protein complex required for primary cilium biogenesis. The Bardet–Biedl syndrome complex is predominantly responsible for the vesicular trafficking of membrane proteins to the primary cilium, but there is some evidence that it might be involved in other vesicular transport pathways.55 The Bardet–Biedl

---

**Table 2** Top independent SNPs from GWSS meta-analysis

<table>
<thead>
<tr>
<th>CHR</th>
<th>BP</th>
<th>SNP ID</th>
<th>Effect allele</th>
<th>Nearest gene</th>
<th>Effect allele frequency</th>
<th>HR</th>
<th>95% CI</th>
<th>P-values*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PD</td>
<td>PDD</td>
<td>NFE</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>45 411 941</td>
<td>rs429358</td>
<td>C</td>
<td>APOE</td>
<td>0.1322</td>
<td>0.2245</td>
<td>0.1486</td>
<td>2.41</td>
</tr>
<tr>
<td>2</td>
<td>142 000 271</td>
<td>rs80306347</td>
<td>C</td>
<td>LRP1B</td>
<td>0.0212</td>
<td>0.0547</td>
<td>0.0277</td>
<td>3.23</td>
</tr>
<tr>
<td>7</td>
<td>33 184 022</td>
<td>rs78294974</td>
<td>A</td>
<td>BBS9</td>
<td>0.0174</td>
<td>0.0358</td>
<td>0.0224</td>
<td>3.90</td>
</tr>
</tbody>
</table>

BP, base pair position in hg19; CHR, chromosome; CI, confidence interval; HR, hazard ratio; NFE, non-Finnish European from gnomAD (https://gnomad.broadinstitute.org/); *Genome-wide significance level set at \( 5 \times 10^{-8} \).
syndrome complex has been shown to bind to Rabin8, which acts as a GTP/GDP exchange factor for the small GTPase Rab8, a substrate of LRRK2. Mutations in LRRK2 that increase its kinase activity lead to enhanced phosphorylation of RAB GTPases, thus causing RAB-mediated vesicular membrane trafficking and centrosomal defects. Because pathogenic LRRK2 mutations interfere with primary cilia formation, it has been suggested that defective ciliogenesis could contribute to the pathogenesis of LRRK2-related PD. Interestingly, the Bardet–Biedl syndrome complex has also been shown to be present at the postsynaptic density of hippocampal neurons and to be important for dendritic spine homeostasis, which could have important implications for cognition.

Forest plots of the GWSS meta-analysis (Supplementary Fig. 7A–C) show that the direction of the effect is consistent across cohorts in all the genome-wide significant loci, albeit with smaller contributions from AMP-PD to the LRP18 and BB59 signals (note that due to reduced number of events in individuals in the DIGPD cohort, infinite estimates were generated by the CPH analysis of this cohort). Several factors could be contributing to these differential observations between cohorts, namely the reduced event rate in the AMP-PD cohort compared to TPD and OPDC (7.2 and 12.4%, respectively). This could in turn be related to the younger age at baseline and shorter follow-up times in the AMP-PD cohorts, since increasing age is the most significant clinical risk factor for the development of dementia in PD (Table 1). To evaluate the effect of the different event rates on the power to detect a genome-wide significant effect on dementia-free survival, we modelled statistical power across a range of minor allele frequencies and effect sizes, assuming a median time to the event of 4.5 years, under an additive genetic risk model (Supplementary Fig. 8). At the current sample size, the detection of an association with genome-wide significance at 80% statistical power requires far more common alleles and/or higher effect sizes at the event rate of AMP-PD than at the event rates of TPD and OPDC cohorts. As an example, for a SNP with effect size of the magnitude observed with LRP1B rs80306347, only SNPs with minor allele frequency of 0.15 and higher can be detected at the event rate of AMP-PD, while SNPs with a minor allele frequency of 0.05 and 0.03 can be detected at the event rates observed in TPD and OPDC, respectively, thus demonstrating how a low event rate can hinder the ability of the survival analysis to detect significant effects of variants with rarer minor allele frequencies.

Finally, we searched for potential candidate genes with a P-value near the genome-wide significance threshold that could...
be relevant for neurodegeneration. Eighty-six variants in 33 independent loci had a suggestive $P < 1 \times 10^{-06}$ (Supplementary Table 4). The nearest genes of some of these variants are involved in pathways known to contribute to neurodegeneration, such as vesicle trafficking (VTI1A), ubiquitin signalling (DDB1), synaptic homeostasis (PTPRD) and endoplasmic reticulum protein QC and translocation (UGGT2, SSR1).59-62 Interestingly, SSR1 expression has recently been shown to be upregulated in an early PD mouse model and to be highly correlated with the loss of dopaminergic neurons.63 An intronic variant in SLC6A3, which encodes the dopamine transporter (DAT), was also observed to be nominally associated with faster progression to dementia in PD. This receptor is specifically expressed in nigro-striatal neurons and is essential in the regulation of dopamine metabolism and neurotransmission. Given its prominent role in the metabolism of dopamine, there has been long-standing interest in this gene in relation to the pathophysiology of PD.64,65 Future studies with larger samples sizes are needed to enable the identification of associations with suggestive variants of smaller effect sizes and allele frequencies.

Colocalization analysis

We did not identify proxy coding variants in high linkage disequilibrium with the lead variants in LRP1B or BBS9. To determine whether any of the GWSS genome-wide significant signals are involved in the regulation of gene expression, we performed colocalization analysis using eQTLs from eQTLGen61 and PsychENCODE,66 which represent large human blood and brain gene expression datasets, respectively. We found no colocalization between PDD GWSS loci and eQTLs from either dataset, indicating that there is currently no evidence of shared causal variants driving both gene expression and the three association signals for PDD progression.
The non-Gaucher

Supplementary Fig. 9

whiskers indicate
groups. Significance threshold: * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001.

Figure 5  CSF measurements of Alzheimer’s disease biomarkers. Box plots representing the measurements (in pg/ml) of the CSF biomarkers Aβ42, p-Tau181 and total Tau in a subset of individuals from the AMP-PD cohort (n = 352) across time (M0 = study baseline, M12 = 12 months, M24 = 24 months, M36 = 36 months). (A) CSF biomarker levels by phenotype (n = 28 PDD and n = 324 PD cases). (B) CSF biomarker levels by APOE ε4 allele carrier status (n = 86 APOE ε4 allele carriers and n = 266 APOE ε4 allele non-carriers). Box plots display a median line, the box limits indicate the first and third quartiles, the whiskers indicate ±1.5 × IQR, and the data points indicate the outliers. The Wilcoxon rank-sum test was used to compare medians across phenotypic groups. Significance threshold: * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001.

(Supplementary Table 5). Of note, LRP1B is not expressed in blood, thus no LRP1B eQTLs (significant or non-significant) were available from eQTLGen (Supplementary Fig. 5A). Next, we explored whether non-coding GWSS significant signals could have a role in alternative splicing by performing colocalization analysis using cortical and nigral sQTLs from the Genotype-Tissue Expression (GTEx) dataset.29 Again, we found no colocalization of PDD GWSS loci and sQTLs (Supplementary Table 6). We also generated regional associations plots for tQTLs from PsychENCODE and PDD GWSS signals in the region surrounding LRP1B, which on visual assessment did not suggest the presence of signal colocalization between LRP1B loci and tQTLs (Supplementary Fig. 9). Despite the power limitations of existing QTL datasets, the available data does not currently support LRP1B signals regulating the expression of transcript isoforms via alternative splicing.

APOE and LRP1B interaction

One of the ligands of LRP1B at the cell surface is APOE. To investigate whether the LRP1B signal was independent of APOE status, we defined four groups of PD patients in the combined cohorts (n = 3923): non-carriers of either APOE ε4 or LRP1B rs80306347 alleles, exclusive carriers of APOE ε4 allele, exclusive carriers of LRP1B rs80306347 allele and carriers of both alleles. We then used a CPH model in the combined cohorts to calculate the hazards of survival dementia-free in each of these groups, adjusting for sex, age at disease onset or diagnosis, the first five principal components and the cohort each individual originated from (Fig. 2A). Compared to non-carriers, participants exclusively carrying the LRP1B rs80306347 allele had an increased risk of progressing to PDD (HR = 2.33, 95% CI = 1.34–4.05, P = 0.00273). In addition, we also performed survival analysis controlling for APOE status (Fig. 2B and C). An increased hazard of progression to PDD was present in LRP1B rs80306347 carriers in both APOE ε4 allele carriers (HR = 3.47, 95% CI = 1.96–6.13, P = 1.82 × 10−55) and APOE ε4 allele non-carriers (HR = 2.25, 95% CI = 1.29–3.92, P = 0.00422), confirming that the effect of rs80306347 is independent of the effect of APOE. Finally, individuals carrying both APOE ε4 and LRP1B rs80306347 alleles had a much higher hazard of progression to PDD than carriers of each allele separately (HR = 8.08, 95% CI = 4.64–14.1, P = 1.55 × 10−13), indicating an increased risk of progression to PDD in carriers of both alleles (Fig. 2A). However, the addition of an interaction term in the regression model did not confirm an interaction between the two alleles.

Candidate gene analysis

Several other genes have been suggested to increase the risk of cognitive decline or dementia in PD. One of the most widely reported genes is GBA, which has also been described as a risk factor for PD and an earlier age of disease onset.57,67 The non-Gaucher disease-causing GBA PD-risk variants E365K (rs2230288, also known as E326K) has been described in association with cognitive progression in PD.13,14 We therefore performed a candidate loci survival analysis in the combined cohorts (n = 3923) based on E365K carrier status, which confirmed a significant HR for progression to dementia (HR = 2.24, 95% CI = 1.45–3.48, P = 3.12 × 10−54, Fig. 3A and Supplementary Table 7). Conversely, the PD-risk factor T408M (also known as T369M, rs75548401) showed a trend toward a faster rate of cognitive decline that did not reach statistical significance, in keeping with a previous study.15 The mild GD-causing variant N409S (also known as N370S, rs76763715) has shown inconsistent association with cognitive decline in PD. In our candidate loci analysis, PD patients carrying this variant had a HR of 4.96 (95% CI = 2.30–10.7, P = 4.42 × 10−56) of developing dementia. In addition, GBA Sanger sequencing data were available for 1793 individuals originating from the DIGPD and TPD cohorts. Mutations causing Gaucher’s disease and PD-risk variants were combined for survival analysis and were present in 9.3% of the cases
Alzheimer’s disease and Parkinson’s disease genetic risk scores in Parkinson’s disease dementia

Given the role of both APOE and LRP1B in APP metabolism, we next investigated the overlap between the Alzheimer’s disease-risk profile with that of PD cases with and without dementia. We calculated the normalized individual-level GRS in each of the cohorts, on the basis of the summary statistics from a recent large-scale GWAS meta-analysis of Alzheimer’s disease. A generalized linear model was used to test the association of Alzheimer’s disease GRSs with dementia status in each cohort, with results further meta-analysed using a random-effects model (Fig. 4A and Supplementary Fig. 11A). PDD was associated with a higher GRS for Alzheimer’s disease (odds ratio = 1.48, 95% CI = 1.32–1.66, $P = 4.47 \times 10^{-11}$). In contrast, the normalized GRS for PD, derived from the latest Parkinson’s disease GWAS study, was similar between PDD and non-demented PD cases ($OR = 0.99, 95\% CI = 0.82–1.19, P = 0.9078$; Fig. 4B and Supplementary Fig. 11C). This suggests that the genetic risk of developing PDD overlaps with the risk of developing Alzheimer’s disease. Interestingly, in a subset of PD samples from the AMP-PD cohort who have been tested for Alzheimer’s disease biomarkers in CSF, PDD cases had decreased $A\beta_{42}$ levels [median ± interquartile range (IQR): 581 ± 493 pg/ml versus 867 ± 478 pg/ml, $P = 0.001193$, Wilcoxon rank-sum test] and increased total tau (208 ± 129 pg/ml versus 158 ± 70 pg/ml, $P = 0.01617$, Wilcoxon rank-sum test) and t-Pau181 (18.3 ± 14.3 pg/ml versus 13.3 ± 5.8 pg/ml, $P = 0.002544$, Wilcoxon rank-sum test) levels at baseline (Fig. 5A), supporting the hypothesis that APP metabolism is important for the development of PDD. In addition, APOE $ε4$ carriers also had significantly decreased CSF $A\beta_{42}$ levels at baseline (median ± IQR: 689 ± 386 pg/ml versus 896 ± 543 pg/ml, $P = 1.7 \times 10^{-6}$, Wilcoxon rank-sum test) and subsequent time points, with no change in total tau or t-Pau181 levels (Fig. 5B).

The overlap is the most significant genetic determinant of the risk of developing Alzheimer’s disease, and was also confirmed to be significantly associated with the risk of progression to PDD in individuals previously diagnosed with PD. Therefore, to establish that the association between Alzheimer’s disease-GRS and progression to PDD is not exclusively due to the overlap of the APOE signal between these two conditions, we adjusted the generalized linear models for APOE $ε4$ carrier status. When adjusting for APOE $ε4$ carrier status, there was no significant association between PDD and the GRS for Alzheimer’s disease (OR = 1.06, 95% CI = 0.93–1.21, $P = 0.374$; Supplementary Fig. 11B), indicating that APOE $ε4$ carrier status alone is driving the risk of progression to dementia among Alzheimer’s disease GWAS top hits.

Finally, we assessed whether a higher Alzheimer’s disease-GRS could be contributing to decreased dementia-free survival, i.e. faster progression to PDD. We performed survival analysis using CPH models to calculate the hazards of survival dementia-free after stratification of PD individuals into low-, middle- and high-risk on the basis of Alzheimer’s disease-GRSs. Individuals in the higher tertile of Alzheimer’s disease-GRS had faster progression to dementia (HR = 2.38, 95% CI = 1.66–3.40, $P = 1.98 \times 10^{-09}$), but as with the overall risk of PDD, faster progression to dementia was abolished after exclusion of the APOE signal (HR = 1.16, 95% CI = 0.85–1.60, $P = 0.3438$, Fig. 4C and D).

Discussion

We have conducted a large GWAS of progression to dementia in PD patients. APOE has consistently been implicated as a risk factor for Alzheimer’s disease, PDD and dementia with Lewy bodies. 9.32,44,47,51,75 Our results confirm that APOE $ε4$ is also a significant contributing factor in the rate of progression to PDD, while a candidate gene approach confirmed the role of non-Gaucher disease-pathogenic GBA $E365K$ PD-risk variant and Gaucher disease-pathogenic N409S mutation in accelerating cognitive decline in PD. In addition, we identified a novel locus associated with PDD, and SNCA variants in the risk of cognitive decline or dementia in PD. Importantly, variants identified in dementia with Lewy bodies case-control GWAS studies do not appear to contribute to risk of progression to dementia in PD, suggesting that the mechanisms leading to dementia with Lewy bodies and PDD do not entirely overlap.

Alzheimer’s disease and Parkinson’s disease-mutation carriers had a HR for progression to PDD of 2.02 (95% CI = 1.21–3.32, $P = 0.007$), confirming the observation from several previous studies that GBA mutations increase the risk of dementia (Supplementary Fig. 10). A similar candidate locus approach in the combined cohorts confirmed the strong association of APOE $ε4$ carrier status (HR = 2.56, 95% CI = 2.00–3.28, $P = 6.36 \times 10^{-14}$) and LRP1B rs80306347 carrier status (HR = 2.71, 95% CI = 1.82–4.02, $P = 7.71 \times 10^{-05}$) with earlier progression to PDD (Fig. 3B and C).

Some studies have found that the APOE $ε4$ H1 haplotype is a risk factor for cognitive decline in PD and can increase the susceptibility to dementia with Lewy bodies. However, this finding has not been consistently replicated. 10,11 Similarly, we did not find any association between APOE haplotypes and time to dementia in PD (Supplementary Table 7).

Recently, common variants in RIMS2, TMEM108 and WWOX have been suggested to associate with faster progression to PDD. Using similar methodology and sample size, we did not replicate these findings (Supplementary Table 7), indicating that further studies are needed to confirm the role of these genes in the risk of cognitive decline in PD.
as beta-secretase is most active in the acidic pH of the endosome, which appears to be a key site for the production of Aβ.\textsuperscript{78} Therefore, modulation of intracellular APP trafficking by LDL receptors with opposing activities is postulated to be a crucial determinant of APP processing and subsequent neurodegeneration.\textsuperscript{79} For example, binding of LRP1 and LRAD3 to APP at the cell surface leads to its enhanced endocytic trafficking and increased processing to Aβ.\textsuperscript{80,81} In contrast, binding of LRP1B and LRP10 to APP leads to decreased trafficking of APP to the endosome, thus resulting in reduced amyloidogenic processing of APP.\textsuperscript{54,82} LRP10 mutants that disrupt the distribution of LRP10 from the trans-Golgi network to early endosomes lead to increased presence of APP in the endosomes and consequently to increased amyloidogenic processing of APP.\textsuperscript{82} Interestingly, loss of function mutations in LRP10 have recently been implicated in familial PD.\textsuperscript{83} Similarly, due to a slower rate of endocytosis that leads to APP accumulation at the cell surface, the binding of APP to LRP1B receptors reduces APP processing into Aβ and increases secretion of soluble APP instead, suggesting that enhanced LRP1B activity could protect against the pathogenesis of Alzheimer’s disease.\textsuperscript{84} Interestingly, a genome-wide study comparing elderly individuals without cognitive decline and those with late onset Alzheimer’s disease identified variants in LRP1B as protective against cognitive decline in old age.\textsuperscript{84}

It is likely that dementia in PD can be driven by distinct mechanisms. Research in dementia with Lewy bodies, a condition closely related to PDD, has shown that GBA is more strongly associated with risk of ‘pure’ dementia with Lewy bodies, while APOE ε4 is more strongly associated with dementia with Lewy bodies with Alzheimer’s disease co-pathology.\textsuperscript{85,86} This suggests that the genetic drivers of dementia in α-synucleinopathies are different in cases with and without Aβ co-pathology, with GBA predisposing to pure Lewy body pathology and APOE predisposing to concomitant Aβ deposition. While PD neuropathology is primarily characterized by deposition of α-synuclein aggregates, dementia in PD can also be associated with Aβ deposition.\textsuperscript{87–89} This leads to the question of whether Aβ metabolism could also play an important role in the development of PDD. In fact, increased cortical Aβ deposition has been shown to be associated with a faster progression to dementia in PD.\textsuperscript{89,90} and a low CSF Aβ$_{42}$-to-total tau ratio at baseline has been associated with cognitive decline in early PD.\textsuperscript{91} Our results on Alzheimer’s disease CSF biomarkers also suggest that PDD is associated with increased Aβ brain pathology, and it is likely that APOE ε4 is the main driver of this association. Furthermore, APOE is known to facilitate endocytosis of Aβ via LDL receptors at the cell surface,\textsuperscript{92} which could offer a mechanistic link between APOE and LRP1B and a possible explanation as to why PD carriers of both APOE ε4 and LRP1B rs80306347-C alleles appear to have a faster progression to dementia. Nonetheless, there is evidence that APOE ε4 can also contribute to neurodegeneration by non-amyloidogenic mechanisms: APOE ε4 allele carriers can present with ‘pure’ Lewy body dementia; α-synuclein pathology is increased in Lewy body dementia APOE ε4 carriers with minimal amyloid pathology, compared to age-matched non-carriers; APOE ε4 exacerbates α-synuclein pathology and leads to worse neurodegeneration and cognitive performances in mice.\textsuperscript{93,94}

Other genetic variants previously reported in association to dementia in PD were not confirmed. In particular, a large recent study using a similar genome-wide survival approach identified that a variant in RIMS2 was a stronger predictor of PDD than APOE and GBA.\textsuperscript{88} We were unable to replicate this finding, which could be the result of small variations in the post-imputation background allele frequencies in different cohorts. Given the relatively rare minor allele frequency of this SNP in the general population, it is possible that small changes in the allele frequency may significantly change the results of the analysis. The apparent discrepancies between studies will probably be resolved as larger longitudinal datasets become available.

Our study has some limitations. First, the analysis was conducted only in individuals of European ancestry, as data from this population was more readily available. It is therefore not possible to generalize our findings to other populations. Future studies including individuals from non-European ancestries are needed. Second, statistical power to detect a significant association is likely to be reduced by the fact that some individuals did not complete the study protocol because of early study withdrawal. It is possible that some individuals who were censored as non-dementia cases would have developed dementia if the follow-up duration had been longer. To mitigate this, individuals with normal longitudinal assessments who withdrew from the study due to the development of dementia were classified as PDD, where this information was available. This creates the potential for a skewed estimation of time to dementia in these cases. However, given the relatively short time interval between the last normal assessment and study withdrawal, the risk of disproportionate skewness is reduced. In addition, estimating time to dementia using the midpoint between the last normal assessment and withdrawal should further reduce that risk. Statistical power to detect a significant association is a function of sample size and event rates, which for dementia are likely to be influenced by mean age at baseline and duration of follow-up. Two of the cohorts (TPD and OPDC) recruited individuals of similar age to incident cohorts of Northern European ancestry, namely the cohorts included in the Parkinson’s Incidence Cohorts Collaboration.\textsuperscript{95} However, the remaining cohorts have a mean younger age than the observed average in incident population-based cohorts, which suggests these cohorts might not be representative of the wider PD population. Given these are non-incident cohorts, it is not possible to know when individuals who met criteria for dementia at baseline developed PDD, and so these were excluded from further analysis. It is therefore possible that individuals who develop dementia early in the disease course are not adequately represented in the dataset analysed. Despite being one of the largest genome-wide survival studies of progression to PDD, sample size and event rates are relatively small, and larger incident cohorts with longer follow-up times are needed to detect variants of small effect size. It is nevertheless reassuring that our study has identified some of the same genetic factors associated with higher risk of progression to dementia as large, incident population-based cohorts with long follow-up times such as APOE ε4 and GBA mutations, despite the potential limitations of large non-incident longitudinal cohorts.\textsuperscript{95}

In conclusion, this large genome-wide study identifies several interesting and plausible new gene candidates associated with faster progression to dementia in PD, while also corroborating the importance of the previously described APOE and GBA variants for cognitive outcomes in PD. In addition, our results provide further evidence that β-amyloid metabolism might play an important role in the pathophysiology of PDD, which has important therapeutic implications, as strategies aimed at Alzheimer’s disease could also prove effective in PD patients at risk of dementia.

**Acknowledgements**

Data used in the preparation of this article were obtained from the AMP-PD Knowledge Platform (https://www.amp-pd.org). AMP-PD is
a public-private partnership managed by the FNIH and funded by Celgene, GSK, Michael J. Fox Foundation for Parkinson’s Research, the National Institute of Neurological Disorders and Stroke (NINDS), Pfizer and Verily.

Clinical data and biosamples used in preparation of this article were obtained from the Fox Investigation for New Discovery of Biomarkers (BioFIND), the Parkinson’s Progression Markers Initiative (PPMI), the Parkinson’s Disease Biomarkers Program (PDBP) and the SURE-PD3 Study. BioFIND is sponsored by the Michael J. Fox Foundation for Parkinson’s Research with support from NINDS.

Data used in the preparation of this article were obtained from the Fox ‘BioFIND’ database (http://biofind.loni.usc.edu/). For up-to-date information on the study, visit michaeljfox.org/news/biofind.

The PPMI—a public–private partnership—is funded by the Michael J. Fox Foundation for Parkinson’s Research and funding partners (a full list of all the PPMI funding partners can be found at www.ppmi-info.org/fundingpartners). The PPMI Investigators have not participated in reviewing the data analysis or content of the paper. For up-to-date information on the study, visit www.ppmi-info.org.

The PDBP consortium is supported by NINDS at the National Institutes of Health. A full list of PDBP Investigators can be found at https://pdbp.ninds.nih.gov/policy. The PDBP Investigators have not participated in reviewing the data analysis or content of the paper.

The DIGPD cohort (ClinicalTrials.gov, NCT01564992) is a multi-center longitudinal cohort conducted in four Universities and four General Hospitals in France between 2009 and 2019, sponsored by Assistance Publique Hôpitaux de Paris, funded by a grant from the French Ministry of Health (PHRC 2008, AOR0810) and a grant from the Agence Nationale de Sécurité et des Médicaments (ANSM-2013). We thank the DIGPD Study Group that collected the data made available for this work.

Both the TPD and OPDC cohorts are primarily funded and supported by Parkinson’s UK (https://www.parkinsons.org.uk/) and supported by the National Institute for Health Research (NIHR) Dementias and Neurodegenerative Diseases Research Network (DeNDRoN). The TPD study is also supported by NHS Greater Glasgow and Clyde. The OPDC cohort is also supported by the NIHR Oxford Biomedical Research Centre, based at the Oxford University Hospitals NHS Trust, and the University of Oxford. TPD has multi-centre research ethics approval from the West of Scotland Research Ethics Committee: IRAS 70980, MREC 11/AL/0163 (ClinicalTrials.gov, NCT02881099). OPDC has multi-centre research ethics approval from the South Central Oxford A Research Ethics Committee 16/SC/0108.

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH and NINDS. The data used for the analyses described in this paper were obtained from https://console.cloud.google.com/storage/browser/gtex-resources on 26 January 2022.

**Funding**

This research was funded in whole or in part by Aligning Science Across Parkinson’s (grant number ASAP-000478) through the Michael J. Fox Foundation for Parkinson’s Research (MJFF). For the purpose of open access, the author has applied a CC BY public copyright licence to all Author Accepted Manuscripts arising from this submission.

This research was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. The UCL Movement Disorders Centre is supported by the Edmond J. Safra Philanthropic Foundation.

**Competing interests**

H.R.M. reports paid consultancy from Roche. Research Grants from Parkinson’s UK, Cure Parkinson’s Trust, PSP Association, CBD Solutions, Drake Foundation, Medical Research Council (MRC), Michael J. Fox Foundation. H.R.M. is a co-applicant on a patent application related to C9ORF72—Method for diagnosing a neurodegenerative disease (PCT/GB2012/052140). D.G.G. has received grants from Michael’s Movers, the Neurosciences Foundation and Parkinson’s UK, and honoraria from AbbVie, BIAL Pharma, Britannia Pharmaceuticals, GE Healthcare and consultancy fees from Acorda Therapeutics and the Glasgow Memory Clinic.

M.T.M.H. received funding/grant support from Parkinson’s UK, Oxford NIHR BRC, University of Oxford, CPT, Lab10X, NIHR, Michael J. Fox Foundation, H2020 European Union, GE Healthcare and the PSP Association. She also received payment for Advisory Board attendance/consultancy for Biogen, Roche, Sanofi, CuraSen Therapeutics, Evidera, Manus Neurodynamica, Lundbeck. Y.B.-S. has received grant funding from the MRC, NIHR, Parkinson’s UK, NIH and ESRC. J.C.C. has served on advisory boards for Biogen, Denali, Idorsia, Prevail Therapeutic, Servier, Theranexus, UCB and received grants from Sanofi and the Michael J. Fox Foundation outside of this work. A.E. received funding/grant support by Agence Nationale de la Recherche, France Parkinson and the Michael J. Fox Foundation. J.H. is supported by the UK Dementia Research Institute, which receives its funding from DRI Ltd, funded by the UK Medical Research Council, Alzheimer’s Society and Alzheimer’s Research UK. He is also supported by the MRC, Welcome Trust, Dolby Family Fund, National Institute for Health Research University College London Hospitals Biomedical Research Centre. All other authors report no competing interests.

**Supplementary material**

Supplementary material is available at Brain online.

**Appendix 1**

An appendix with the DIGPD Study Group members is available online.

**References**


