

## Supplementary Figure 1. Genetic ancestry estimation.

(A) A two-dimensional representation of the raw genetic data was learned on 1000G, which consists of 5 super populations. The PPMI (B) and DIGPD (C) subjects were projected on this space to estimate their genetic ancestry and excluded if their projection was too far from the European cluster. Excluded subjects are highlighted with a black circle.

1000G: 1000 Genomes Project; DIGPD: Drug Interaction With Genes in Parkinson's Disease; PPMI: Parkinson's Progression Markers Initiative.



# Supplementary Figure 2. Forest plot for the association analysis between

# the GRS of lack of premeditation and ICDs in PD.

CI: Confidence interval; DIGPD: Drug Interaction With Genes in Parkinson's Disease; OR: Odds ratio; PPMI: Parkinson's Progression Markers Initiative. seTE: standard error of treatment effect; TE: treatment effect.



Supplementary Figure 3. Power analysis.

Power calculation for different prevalences were performed. At the 0.80 power threshold, the sample sizes allow for discovering associations with OR = 2.2 in PPMI and OR = 2.0 in

DIGPD. Combining both cohorts allows for discovering associations with OR = 1.7. These are high odds ratios that are unlikely to exist for a genetic risk score for another phenotype. At the same power threshold, odds ratios of 1.3 and 1.1 (which are more reasonable) would require around 1k and 10k samples to be discovered respectively. This highlights the importance of the sample size for genetic analyses for which each effect size / odds ratio is usually very small.

DIGPD: Drug Interaction With Genes in Parkinson's Disease; OR: Odds ratio; PPMI: Parkinson's Progression Markers Initiative.

## Supplementary materials and methods

#### **Populations**

We used data from two research cohorts: the Parkinson's Progression Markers Initiative (PPMI) database and the Drug Interaction With Genes in Parkinson's Disease (DIGPD) study.

PPMI (<u>https://www.ppmi-info.org</u>) is a multicenter observational clinical study using advanced imaging, biologic sampling and clinical and behavioral assessments to identify biomarkers of PD progression. Data was gathered during face-to-face visits every 6-12 months. PD subjects were de-novo and drug-naïve at baseline. We downloaded the clinical and genetic data from the PPMI database (<u>https://www.ppmi-info.org/data</u>) on the 17th of October 2019.

DIGPD is a French multicenter longitudinal cohort with annual follow-up of PD patients. Eligible criteria consist in recent PD diagnosis (UK Parkinson's Disease Society Brain Bank criteria) with disease duration less than 5 years at recruitment. Data was gathered during face-to-face visits every 12 months following standard procedures.

Both studies were conducted according to good clinical practice, obtained approval from local ethic committees and regulatory authorities, and all patients provided informed consent prior to inclusion.

### **Participants**

Inclusion criteria in our analyses included having: (i) a PD diagnosis, (ii) at least two visits measuring ICDs, (iii) clinical and genetic data available, and (iv) a European genetic ancestry. We identified 378 subjects in PPMI and 382 subjects in DIGPD matching the first three criteria.

ICDs were assessed at each visit using the Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease - Rating Scale in PPMI, and through semi-structured interviews by a movement disorders specialist in DIGPD. The ICD phenotype was defined as the lifetime presence of ICDs.

#### **Genetic ancestry**

To date, most GWAS have been conducted in populations of European ancestry, which limits the use of GWAS-derived GRS in non-European ancestry populations, and their transferability to other populations depends on many factors such as linkage disequilibrium, allele frequencies, and genetic architecture. Directly computing GRS in another ancestry group that the one from the corresponding GWAS can lead to biased GRS.

To estimate the genetic ancestry of the PD subjects in PPMI and DIGPD, we used data from the 1000 genomes (1000G) project to learn a low-dimensional representation of the genetic data, which captures the main dimension of ancestry. Using the 50,842 common raw SNPs between 1000G, PPMI and DIGPD, we applied the Uniform Manifold Approximation Projection algorithm on the 1000G data to learn a low-dimensional space of the raw SNPs. Finally, we projected the PPMI and DIGPD subjects onto the main principal components to identify in which clusters they were the closest to. Subjects projected on another cluster than the European cluster were excluded.

#### Genotyping and quality control

Genotype data was acquired using NeuroX arrays in PPMI (267,607 variants measured), and Illumina Multi-Ethnic Genotyping Arrays in DIGPD (1,779,819 variants). We excluded variants with missing rates greater than 2% and variants deviating from Hardy-Weinberg equilibrium (p < 1E-8). We excluded related individuals (third-degree family relationships), individuals with mismatching between reported sex and genetically determined sex, and individuals with outlying heterozygosity ( $\pm$  3 SD). We then imputed missing SNPs using the Michigan Imputation Server for PPMI and the Sanger Imputation Server for DIGPD, using the reference panel of the Haplotype Reference Consortium (release 1.1).

For GRS calculation, we selected SNPs that were (i) biallelic, (ii) frequent enough (minor allele frequency > 1%), and (iii) imputed with sufficient accuracy (R2 > 0.8 for PPMI, INFO Score > 0.9 for DIGPD).

#### Phenotypes and genome-wide association studies

Phenotypes of interest included known or putative factors clinically associated with ICDs in PD, such as anxiety, depression, personality traits including impulsivity, eating and sleep disorders. We were also interested in more general phenotypes such as body height, body mass index (BMI), intelligence, and number of years of education, more because of the sample size of the corresponding GWAS rather than their prior association with ICDs in PD. In particular, body height and body mass index are phenotypes that are easy to collect with precision, and for which very large GWAS are available and the corresponding GRS explain a large part of the variance. These phenotypes are also usually collected in research cohorts, allowing for comparing the GRS with the true phenotypes, and thus validating our computation of the GRS.

We used the NHGRI-EBI GWAS Catalog to select the largest GWAS to date on samples of European ancestry. When summary statistics from several GWAS were available for a given phenotype, we only included the largest study.

#### **Computation of genetic risk scores**

When summary statistics were fully available, we estimated the coefficients of the GRS using the SBLUP algorithm implemented in the GCTA software. SBLUP directly estimates GRS coefficients from summary statistics, using a reference sample to estimate the linkage disequilibrium between SNPs. SBLUP is a random effects model that converts the coefficients, estimated from a GWAS using the BLUP (best linear unbiased predictor) algorithm and available in the corresponding summary statistics, into approximate best linear unbiased predictors. When summary statistics were not available in full, we computed small GRS by performing clumping to select the most significant, low correlated variants ( $r^2 < 0.1$  or distance > 10,000kb), and directly using the coefficients provided in the summary statistics. Clumping and GRS computation were performed using the PLINK software.

#### **Statistical analyses**

We estimated the association between the binary ICD phenotype and GRS using logistic regression, while correcting for age, sex, genetic ancestry (first four components), and the number of visits. We added the correction for the number of visits to reflect the fact that lifetime phenotype may be more likely as the number of visits increases. We performed the analyses in each cohort independently as the contributions of all the SNPs were estimated altogether, and the number of SNPs was much lower in PPMI than in DIGPD. We applied per-sample Bonferroni correction for multiple comparisons. We also investigated the association of the combination of the 40 GRS altogether with the likelihood-ratio test.

As the sample sizes were relatively small in both cohorts, we also performed a metaanalysis to estimate the combined effects of each GRS separately and combined altogether using fixed effects models with the inverse-variance weighting method.

Logistic regressions were performed using the statsmodels Python package. Metaanalyses were performed using the meta R package. Processing of the different text-like files was performed using the pandas and NumPy packages.