Genomewide Association Studies of LRRK2 Modifiers of Parkinson’s Disease

Dongbing Lai, PhD†,1 Babak Alipanahi, PhD,2,3† Pierre Fontanillas, PhD,2†
Tae-Hwi Schwantes-An, PhD,‡ Jan Aasly, MD, PhD,§ Roy N. Alcalay, MD, MS,∥
Gary W. Beecham, PhD,¶ Daniela Berg, MD,†① Susan Bressman, MD,② Alexis Brice, MD,③
Kathrin Brockman, MD①,④ Lorraine Clark, PhD,⑤ Mark Cookson, PhD,⑥
Sayantan Das, PhD,⑦ Vivianna Van Deerlin, MD, PhD,⑧ Jordan Follett, PhD,⑨
Matthew J. Farrer, PhD,⑩ Joanne Trinh, PhD,⑪ Thomas Gasser, MD,⑫
Stefano Goldwurm, MD, PhD,⑬ Emil Gustavsson, PhD,⑭ Christine Klein, MD⑮,⑯
Anthony E. Lang, MD,⑰ J. William Langston, MD,⑱ Jeanne Latourelle, DSc,⑲
Timothy Lynch, MD,⑳ Karen Marder, MD, MPH,㉑ Connie Marras, MD, PhD,㉒
Eden R. Martin, PhD,㉓ Cory Y. McLean, PhD,㉔ Helen Mejia-Santana, MS,㉕
Eric Molho, MD,㉖ Richard H. Myers, PhD,㉗ Karen Nuytemans, PhD,㉘ Laurie Ozelius, PhD,㉙
Haydeh Payami, PhD,㉚ Deborah Raymond, MS,㉛ Ekaterina Rogaeva, PhD,㉜
Michael P. Rogers, MD,㉝ Owen A. Ross, PhD①,③,④ Ali Samii, MD,㉗
Rachel Saunders-Pullman, MD,⑨ Birgitt Schüle, MD,④ Claudia Schulte, MS,⑧,⑪
William K. Scott, PhD,⑥ Caroline Tanner, MD,⑤ Eduardo Tolosa, MD, PhD,⑥
James E. Tomkins, PhD,㉝ Dolores Vilas, MD,㉞ John Q. Trojanowski, MD, PhD,㉜
The 23andMe Research Team,② Ryan Uitti, MD,⑧ Jeffery M. Vance, MD, PhD,⑥
Naomi P. Visanji, PhD①,⑨ Zbigniew K. Wszolek, MD,⑧ Cyrus P. Zabetian, MD, MS①,③
Anat Mirelman, PhD①,⑨ Nir Giladi, MD,⑨ Avi Orr Urtreger, PhD,⑨ Paul Cannon, PhD,①
Brian Fiske, PhD,② and Tatiana Foroud, PhD①

Objective: The aim of this study was to search for genes/variants that modify the effect of LRRK2 mutations in terms of penetrance and age-at-onset of Parkinson’s disease.

Methods: We performed the first genomewide association study of penetrance and age-at-onset of Parkinson’s disease in LRRK2 mutation carriers (776 cases and 1,103 non-cases at their last evaluation). Cox proportional hazard models and linear mixed models were used to identify modifiers of penetrance and age-at-onset of LRRK2 mutations, respectively. We also investigated whether a polygenic risk score derived from a published genomewide association study of Parkinson’s disease was able to explain variability in penetrance and age-at-onset in LRRK2 mutation carriers.

Results: A variant located in the intronic region of CORO1C on chromosome 12 (rs77395454; p value = 2.5E-08, beta = 1.27, SE = 0.23, risk allele: C) met genomewide significance for the penetrance model. Co-immunoprecipitation
analyses of LRRK2 and CORO1C supported an interaction between these 2 proteins. A region on chromosome 3, within a previously reported linkage peak for Parkinson’s disease susceptibility, showed suggestive associations in both models (penetrance top variant: \(p\) value = 1.1E-07; age-at-onset top variant: \(p\) value = 9.3E-07). A polygenic risk score derived from publicly available Parkinson’s disease summary statistics was a significant predictor of penetrance, but not of age-at-onset.

**Interpretation:** This study suggests that variants within or near CORO1C may modify the penetrance of LRRK2 mutations. In addition, common Parkinson’s disease associated variants collectively increase the penetrance of LRRK2 mutations.

Parkinson’s disease (PD) is the second most common neurodegenerative disease in older adults. Several genes showing autosomal dominant (SNCA, LRRK2, and VPS35) or recessive (PRKN, PINK1, and DJ-1) inheritance patterns have been identified as the cause of familial PD. These genes harbor rare, high penetrance mutations that explain up to 10% of familial PD cases in different populations. Recently, large genomewide association studies (GWAS) have identified over 90 loci with small individual effects on disease risk in both familial and sporadic PD.

Mutations in LRRK2 are among the most common genetic causes of PD. The most frequent mutation is G2019S (rs34637584), which explains up to 10% of familial PD cases and 1% to 2% of all PD cases. Among patients with PD, the frequency of the G2019S mutation is approximately 3% in Europeans, 16% to 19% in Ashkenazi Jews and up to 42% in Arab-Berbers. Estimates of the risk for developing PD among LRRK2 G2019S mutation carriers range from 15% to 85%. To explain the incomplete penetrance of G2019S, it has long been hypothesized that there are other genes/variants outside of LRRK2 acting to modify its effect (LRRK2 modifiers). Identification of LRRK2 modifiers could aid the development of novel prevention and treatment strategies for PD.

Most studies of LRRK2 modifiers, to date, have focused on candidate genes. Because the protein product of LRRK2 may interact with \(\alpha\)-synuclein (encoded by SNCA), and tau (encoded by MAPT), variants in SNCA and MAPT were widely investigated. However, the results have been inconsistent, possibly due to small sample sizes and differences in variants and populations investigated.

Other PD associated genes, such as GBA, BST1, GAK, and PARK16 have also been investigated. However, the number of studies is limited and findings remain to be replicated. Genomewide searches for LRRK2 modifiers are sparse and limited to linkage studies.

Using 85 LRRK2 carriers from 38 families, a genomewide linkage study of LRRK2 modifiers found a suggestive linkage region at 1q32 (limit of detection [LOD] = 2.43); but that study did not identify any candidate genes/variants.
variants underlying the linkage peak.\textsuperscript{33} A genomewide linkage scan in Arab-Berber PD families found DNM3 as a \textit{LRRK2} modifier.\textsuperscript{34} This finding was not independently replicated, although it was still significant in a meta-analysis including the participants reported in the original finding.\textsuperscript{25,35} GWASs have successfully detected many disease genes/variants, including those associated with PD. However, to date, no GWAS for \textit{LRRK2} modifiers has been reported, probably due to limitations in sample size and corresponding statistical power.

In this study, we recruited \textit{LRRK2} mutation carriers from multiple centers and performed the first GWAS to identify genes/variants that modify the penetrance and age-at-onset of PD among \textit{LRRK2} mutation carriers. Using the largest cohort to date, which consisted of 1,879 \textit{LRRK2} mutation carriers (including 776 PD cases), one genomewide significant association signal was found in the intronic region of the \textit{CORO1C} gene. Using co-immunoprecipitation analyses, we demonstrated that the protein product of \textit{CORO1C} interacted with \textit{LRRK2}. In addition, we found that a polygenic risk score (PRS) derived from publicly available PD GWAS summary statistics, was associated with penetrance, but not age-at-onset, of PD in \textit{LRRK2} mutation carriers.

**Methods**

**Study Participants**

The studies and the \textit{LRRK2} mutation carriers were grouped into 3 cohorts. The first cohort was primarily identified from The Michael J. Fox Foundation’s \textit{LRRK2} Consortium and consisted of research sites worldwide (referred to as the MJFF consortium cohort). We searched PubMed and identified study groups that reported \textit{LRRK2} mutation carriers and then asked them to participate in this study (PubMed IDs: 16240353, 16333314, 18986508, and 16960813).\textsuperscript{36–39} We also made announcements at international conferences to recruit more study mutation carriers. Details can be found in their publications.\textsuperscript{36–39} To maximize participation and facilitate uniform data preparation across sites, a minimal dataset was submitted for all subjects that included \textit{LRRK2} mutation status, sex, age-at-onset (for PD cases), age at last evaluation (for non-PD participants), and pedigree information, along with the availability of a minimal amount of DNA (approximately 2 ug). The minimal phenotypic data were sent to Indiana University and the subjects were assigned a unique identifier. The second cohort was from Tel Aviv University, Israel (referred to as the Israel cohort). Participants were of Ashkenazi origin and recruited from the Movement Disorders Unit at Tel Aviv Medical Center. PD diagnosis was confirmed by a movement disorders specialist and clinical disease status (PD or not diagnosed as PD) was evaluated at the time of blood draw for genetic testing. The third cohort (referred to as the 23andMe cohort) consisted of research participants of the personal genetics company 23andMe, Inc. who were \textit{LRRK2} G2019S carriers and whose PD status was known. Individuals who reported via an online survey that they had been diagnosed with PD by a medical professional, were asked to provide their age at diagnosis. For individuals who affirmed at least once that they had not been diagnosed with PD, their age at the most recent completion of the survey was recorded. The institutional review board at each participating site approved this study.

**Genotyping, Quality Review, and Imputation**

All study participants were genotyped on the Illumina Omni 2.5 Exome Array version 1.1 (Illumina, San Diego, CA, USA), except 166 participants from the Israel cohort, who were genotyped on an earlier version of the same array (version 1.0). This array has common, rare, and exonic variants that were selected from diverse world population samples included in the 1000 Genomes Project. In total, there were > 2.58 M variants, including > 567 K exonic variants. Participants from the MJFF consortium and 23andMe were genotyped at the Center for Inherited Disease Research (CIDR) at Johns Hopkins University (Baltimore, MD, USA). The Israel cohort was genotyped at Tel Aviv University and 2 samples from the MJFF consortium were included for quality control. There were 134 duplicated and unexpected identical participants among all 3 cohorts. Pairwise concordance rates were all > 99.97%, showing high consistency among the 2 genotyping laboratories and the 2 versions of the Illumina array.

Variants with genotypic missing rates > 5% and nonpolymorphic variants were excluded. In addition, variants with A/T or C/G alleles were also excluded due to strand ambiguity. Hardy–Weinberg equilibrium (HWE) was not used to filter variants because these participants were ascertained to be \textit{LRRK2} mutation carriers and this participant selection scheme would directly violate HWE and remove potential \textit{LRRK2} modifiers from the analysis.

To confirm the reported pedigree structure and detect cryptic relatedness, we used a set of 56,184 high quality (missing rate < 2%, HWE \textit{p} values > 0.001), common (minor allele frequency [MAF] > 0.1), and independent (linkage disequilibrium as measured by \textit{r} \textsuperscript{2} < 0.5) variants to calculate the pairwise identity by descent using PLINK.\textsuperscript{40} Reported pedigree structures were revised accordingly, if necessary. Mendelian error checking was performed in the revised pedigree structure. Any inconsistent genotypes were set to missing. The same set of variants
was also used to estimate the principal components (PCs) of population stratification using Eigenstrat. All samples were imputed to the Haplotype Reference Consortium (http://www.haplotype-reference-consortium.org/) using Minimac3. A total of 725,802 high quality genotyped variants were selected for imputation (MAF > 3%, HWE p value > 0.0001, and missing rate < 5%). EAGLE version 2.4 was used to phase genotyped variants for each sample. After filtering out variants with poor imputation quality score (R² < 0.6) and checking for Mendelian inconsistencies using PLINK, a final dataset of 7,934,276 imputed and genotyped variants was used for association analyses.

**Genomewide Association Studies**

Our association analysis tested two models: (1) variants modifying the penetrance for PD among LRRK2 mutation carriers (penetrance model), and (2) variants modifying the age-at-onset for PD among LRRK2 mutation carriers (age-at-onset model). For the penetrance model (including PD cases and those not diagnosed as PD at last evaluation), the association analysis was designed to identify variants associated with the time to PD diagnosis or last evaluation for undiagnosed mutation carriers. For the age-at-onset model (PD cases only), the association analysis tested whether variants contributed to the age-at-onset for PD cases among LRRK2 mutation carriers.

For the penetrance model, a mixed effect Cox proportional hazard model (frailty model) was used with sex, 10 PCs, array, and cohort indicators as covariates. Family relationships were adjusted by using a kinship matrix calculated using R package COXME (https://cran.r-project.org/web/packages/coxme/index.html). For the age-at-onset model, a linear mixed model was fit with the same covariates as the penetrance model and a kinship matrix to adjust family relationships. Although adjusting 10 PCs and family relationships could minimize the effects of population stratifications and shared genetic and environment factors among family members, the use of this mixed samples was designed to search for common genetic variants that have the same effects in participants from all populations and could miss population-specific or family specific findings. Variants with MAF > 1% were tested for association in these 2 models. In addition to the single variant analyses, we performed gene-based association analyses for both the

### TABLE 1. Summary of Study Cohorts

| Cohorts            | MJFF consortium | Israel | 23andMe | Total |
|--------------------|-----------------|--------|---------|-------|
| Number of participants | 768             | 185    | 926     | 1,879 |
| % PD cases (N)      | 67% (512)       | 66% (122) | 15% (142) | 41% (776) |
| % Females (N)       | 49% (378)       | 53% (98)  | 52% (480) | 51% (956) |
| Mean age at last evaluation (SD) among non-PD | 56.2 (15.8) | 53.6 (14.3) | 45.9 (17.3) | 48.7 (17.4) |
| Mean age at PD diagnosis (SD) among cases | 56.9 (12.2) | 57.5 (11.4) | 59.4 (10) | 57.4 (11.7) |

**LRRK2 mutation (% of total)**

|                | G2019S | Non-G2019S |
|----------------|--------|------------|
| Number of participants | 699 (91%) | 69 (9%)    |
| % European ancestry | 91%   | 100%       |
| % Ashkenazi Jewish | 37%   | 100%       |
| % African-American or Latinos ancestry | 9% | 0% |

**Families**

|                | MJFF consortium | Israel | 23andMe | Total |
|----------------|-----------------|--------|---------|-------|
| Number of participants with families (% total) | 473 (61%) | 96 (52%) | 284 (31%) | 853 (45%) |
| Total number of families | 138 | 38 | 118 | 294 |
| Average (max) family size | 3.4 (17) | 2.5 (4) | 2.4 (10) | 2.9 |
| Average PD ratio in families | 0.56 | 0.43 | 0.12 | 0.37 |

**Ancestries (% of total)**

|                | MJFF consortium | Israel | 23andMe | Total |
|----------------|-----------------|--------|---------|-------|
| European ancestry | 91%   | 100%   | 89%     | 91%   |
| Ashkenazi Jewish | 37%   | 100%   | 48%     | 49%   |
| African-American or Latinos ancestry | 9% | 0% | 11% | 9% |

MJFF = Michael J. Fox Foundation; N/A = not applicable; PD = Parkinson’s disease.
penetrance and age-at-onset models. We focused on rare exonic and splicing variants, based on annotations from Variant Effect Predictor (https://useast.ensembl.org/info/docs/tools/vep/index.html), and restricting to variants with MAF < 3%. Only genotyped variants (N = 725,802) were used in the gene-based analyses due to the low quality of imputation for rare variants. The R package COXME was used to perform all analyses (https://cran.r-project.org/web/packages/coxme/index.html). Conditional analysis was conducted using the most significant variant in an associated region as a covariate, and additional signals within the associated region were determined based on p values < 0.01.

**Cell Culture, Co-Immunoprecipitation, and Antibodies**

To test whether the protein product of the identified gene interacts with LRRK2, we performed co-immunoprecipitation analyses. HEK293FT (R70007; Invitrogen) cells were maintained as previously described and transiently transfected with pEGFP-C1, pEGFP-CORO1C, or co-transfected with 3xFlag-LRRK2 using Lipofectamine 2000 (Thermofisher Scientific) for 16 hours. Cells were washed once with ice cold PBS and lysed in 1 ml of GFP-trap buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 50 mM NaF, 0.3% Triton X-100, 5% Glycerol, Halt phosphatase inhibitor cocktail [Thermofisher Scientific], and protease inhibitor cocktail [Roche]) or Flag/GFP buffer (20 mM Tris–HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 50 mM NaF, 0.3% Triton X-100, 5% Glycerol, Halt phosphatase inhibitor cocktail [Thermofisher Scientific], and protease inhibitor cocktail [Roche]) for 30 minutes on ice, followed by centrifugation at 4°C for 10 minutes at 13,900 g to obtain the supernatant. The supernatant was then incubated with either pre-equilibrated GFP-Trap agarose beads (ChromoTek) for 1 hour at 4°C on an end over end rotator to recover the GFP tag or pre-cleared with Pierce Protein A/G Agarose (Thermofisher Scientific; 20 minutes at 4°C). Pre-cleared lysates were incubated with 1 µg of monoclonal anti-FLAG antibody (F3165; Millipore-Sigma) or rabbit polyclonal anti-GFP (Ab290; Abcam) for 2 hours at 4°C on an end over end rotator. Antibody complexes were captured by Protein A/G Agarose beads for 2 hours. Bead complexes were washed 6 times with lysis buffer, followed by elution in 1 times loading dye containing 2% β-mercaptoethanol for 6 minutes at 95°C. All samples were resolved by Western blot as previously described. Each co-immunoprecipitation was repeated in 2 to 3 independent experiments. Primary antibodies used in this study are as follows: mouse monoclonal anti-FLAG M2 (F3165) was purchased from Thermofisher Scientific; Rabbit polyclonal to GFP was from Abcam (ab290), and anti-LRRK2 was from UC Davis/NIH NeuroMab Facility (clone N241A/34).

**In Silico Functional Studies**

To evaluate whether the genomewide significant findings had immediate biological consequences on gene expression.
(expression quantitative trait locus [eQTL]) of nearby genes, we searched Open Targets Genetics (https://genetics.opentargets.org/) and GTEx (https://www.gtexportal.org/). In addition, protein–protein interaction (PPI) data were assessed to identify whether the protein product of the nominated gene either interacts directly with LRRK2 or has common interactors that are shared with LRRK2, using Protein Interaction Network Online Tool (PINOT) version 1.046 queried on June 16, 2020 (http://www.reading.ac.uk/bioinf/PINOT/PINOT_form.html). We also performed chromatin interaction mapping to check whether our top findings interact with LRRK2 distantly. Functional Mapping and Annotation of GenomeWide Association Studies (FUMA: https://fuma.ctglab.nl/) was used to perform chromatin interaction mapping.47 Hi-C data for chromatin interaction mapping were from Schmitt et al 2016 and Giusti-Rodriguez et al 2019,48,49 and are available in FUMA.

Polygenic Risk Score Analyses

In the largest GWAS analysis of PD susceptibility to date, Nalls et al meta-analyzed 17 datasets with 56,306 PD cases or proxy-cases and 1.4 million controls.4 Based on their results, they developed a PRS using summary statistics of 1,805 variants that can explain 26% of PD heritability.4 In this study, we performed PRS analysis using these 1,805 variants. Detailed information about how to select these 1,805 variants was described in Nalls et al.4 Because we were searching for LRRK2 modifiers, variants in the LRRK2 region (chr12: 40,118,913-41,263,086) were excluded. The PRS was calculated as a weighted summation of effective alleles with the logarithm of odds ratios as the weights. This derived PRS was used to fit the same models with the same set of covariates as described for the genomewide association analyses using the R package COXME.

Results

Study participants from the 3 cohorts are summarized in Table 1. In total, 1,879 participants (853 individuals from 294 families and 1,026 singletons) were included in the analyses. Among them, 776 had, or self-reported, a PD diagnosis and 1,103 were not classified as affected with PD at the last evaluation. The majority of participants were G2019S carriers, only 4% carried other LRRK2 mutations as reported by the contributing sites, all from the MJFF consortium cohort. In the 23andMe cohort, 85% of participants were not diagnosed with PD and

| CHR | BP       | rsid      | Alleles | Gene         | Annotation | MAF | BETA | SE  | p value | G2019S only p value |
|-----|----------|-----------|---------|--------------|------------|-----|------|-----|---------|-------------------|
| 1   | 221,173,137 | rs141686162 | A/G     | HLX, DUSP10  | Intergenic | 0.01 | 0.42 | 0.28 | 0.13   | 0.86               |
| 3   | 124,083,400 | rs145611031 | C/G     | KALRN       | Intron     | 0.02 | 1.14 | 0.23 | 5.2E-07 | 2.3E-07            |
| 3   | 140,288,373 | rs150382576 | A/G     | CLSTN2      | 3’UTR      | 0.02 | 0.78 | 0.22 | 5.3E-04 | 2.1E-03            |
| 3   | 152,841,926 | rs59679443  | A/G     | RAP2B, ARHGEF26 | Intergenic | 0.04 | 0.62 | 0.17 | 3.1E-04 | 1.6E-04            |
| 3   | 152,932,435 | rs16846845  | G/C     | RAP2B, ARHGEF26 | Intergenic | 0.05 | 0.83 | 0.16 | 1.1E-07 | 4.5E-08            |
| 4   | 160,854,320 | rs12272007  | A/G     | LOC107986324 | Intergenic | 0.01 | 1.22 | 0.36 | 6.0E-04 | 1.9E-03            |
| 5   | 115,786,384 | rs73781088  | C/T     | SEMA6A      | Intron     | 0.03 | 0.47 | 0.19 | 0.01   | 0.05               |
| 8   | 9,520,115  | rs28398294  | G/A     | TNKS        | Intron     | 0.03 | 1.09 | 0.22 | 5.7E-07 | 1.1E-06            |
| 9   | 127,532,973 | rs148922482 | C/T     | NR6A1       | Intron     | 0.01 | 0.90 | 0.28 | 1.4E-03 | 3.7E-03            |
| 11  | 120,585,515 | rs28470321  | G/A     | GRIK4       | Intron     | 0.01 | 1.91 | 0.36 | 9.0E-08 | 6.1E-08            |
| 12  | 109,080,567 | rs77395454  | C/T     | CORO1C      | Intron     | 0.02 | 1.27 | 0.23 | 2.5E-08 | 1.0E-06            |
| 14  | 90,982,388  | rs76788674  | A/G     | CALM1, TTC7B | Intergenic | 0.03 | 0.78 | 0.16 | 7.1E-07 | 2.8E-06            |
| X   | 123,652,525 | rs185981774 | A/G     | TENM1       | Intron     | 0.02 | 0.86 | 0.17 | 5.4E-07 | 3.9E-07            |

Note: Genomewide significant variant is in bold.

CHR = chromosome; MAF = minor allele frequency.
most of them were less than 50 years of age at the time of their last evaluation. Based on PCs, the majority of participants were of European ancestry.

Manhattan plots for the single variant analyses of the penetrance and age-at-onset models are shown in Figure 1A and 1B. Q-Q plots for both models are shown in Figure 1C. No obvious bias was detected in either model; and genomic controls were 1.055 and 1.052 for the penetrance model and age-at-onset model, respectively. Twelve loci showed variants with \( p \) values < 1.0E-6 (ie, meet the threshold for suggestive significance) in either the penetrance model or the age-at-onset model (Supplementary Table S1). One variant on chromosome 12 reached genomewide significance (rs77395454, \( p \) value = 2.5E-08) in the penetrance model (Table 2, Supplementary Table S1). Conditional analysis suggested that there were no additional association signals in this locus. The top variant (rs77395454) on the chromosome 12 region is located in an intron of CORO1C (coronin 1C; Fig 2). The causal haplotype(s) spanned SELPLG, CORO1C, and SSH1, with most of the variants within CORO1C. Figure 3A–C shows the survival curves stratified by rs77395454 genotypes for all samples, familial samples, and unrelated samples, respectively. Heterozygous rs77395454 carriers (20 familial and 37 unrelated samples) had an increased risk of PD. Six other loci met suggestive significance (\( p \) value < 1.0E-6) for the penetrance model (see Table 2).

For the age-at-onset model, no chromosomal region reached genomewide significance, but 7 loci met the suggestive association threshold. Except for variants on chromosome 3 identified in both models, rs73781088 on chromosome 5 (intron of SEMA6A) for the age-at-onset model and rs28398284 on chromosome 8 (intron of TNKS) for the penetrance model, all other variants had no or marginal LD support. Variants on chromosome 3 from both models cover the same region but identified different haplotypes.

For comparison purposes, we also performed analyses limited to only the LRRK2 G2019S carriers; overall the results were comparable (see Table 2, Supplementary Table S1) but rs16846845 on chromosome 3 (\( p \) value = 4.5E-08) was genomewide significant for the penetrance model. In addition, we performed analyses using only individuals of predicted Ashkenazi Jewish ancestry. Results are less significant due to dramatically decreased sample sizes as shown in Supplementary Table S2. No genomewide

---

**FIGURE 2:** Regional association plot of the chromosome 12 region for the penetrance model. Y-axis is the –log (\( p \) value) for associations. X-axis denotes physical positions on the chromosome (Mb). The color scale shows the extent of linkage disequilibrium (LD; as measured by \( r^2 \)) between each variant and the top variant (indicated by the purple diamond) with larger \( r^2 \) indicating greater LD. Peaks indicate the recombination hot spots. SNP = single nucleotide polymorphism.
significantly results were detected in the gene-based analysis using exonic variants for either model.

We next sought evidence of a physical interaction between CORO1C and LRRK2 using co-immunoprecipitation analyses. GFP-tagged CORO1C or eGFP empty vector control were transfected into HEK293FT and lysates were incubated with anti-GFP agarose beads to recover the GFP tag; relative to empty vector control, eGFP-CORO1C co-precipitated endogenous LRRK2 from HEK293FT cells (Fig 4A). To further support this interaction, 3xFlag-tagged LRRK2 with either GFP-tagged CORO1C or eGFP empty vector control were cotransfected into HEK293FT and, after 16 hours, lysates were incubated with anti-GFP or anti-Flag antibodies to immunoprecipitate the GFP or Flag tag, respectively. Immunoprecipitation of LRRK2 via the Flag antibody co-precipitated eGFP-CORO1C, but not the eGFP from the empty vector control. (Fig 4B). Reciprocally, immunoprecipitation of the GFP tag via a GFP antibody co-precipitated Flag-LRRK2 only in the presence of GFP-CORO1C but not the empty vector GFP control (Fig 4C). Thus, co-immunoprecipitation analysis of Flag-LRRK2 and GFP-CORO1C support an interaction between these 2 proteins.

By searching Open Targets Genetics and GTEx, we found that the most significant variant, rs77395454, is an eQTL of CORO1C in blood and MYO1H in visceral adipose (omentum) but not in any brain tissues. The minor allele (C allele) is associated with higher expression of CORO1C and MYO1H. We did not find any previous report that LRRK2 interacts with CORO1C or MYO1H directly, however, there are several proteins that are the common interactors of both LRRK2 and CORO1C: ABCE1, ACTR2, CDC42, DAPK1, MYO1C, RAC1, and TP53, as identified by using PINOT, and it remains to be determined whether the interaction of LRRK2 with CORO1C is within a single complex or dependent upon these common interacting proteins. Chromatin interaction mapping did not find any variant that interacts with LRRK2 distantly.

Among those 1,805 variants that were obtained from the study of Nalls et al 2019, 20 variants were not present in our datasets. An additional 27 variants were located in the LRRK2 region and were excluded, and 1,758 variants were included in the PRS calculation. The PRS was a significant predictor in the penetrance model (p value = 7.8E-4) but not in the age-at-onset model (p value = 0.75). These results suggest that a high genetic risk of PD significantly increases the chance of developing PD among LRRK2 mutation carriers.
Discussion

Two major unresolved questions in PD research are why some, but not all, LRRK2 mutation carriers develop PD, and why the age-at-onset is so variable in those that do. This work represents the first GWAS study to report LRRK2 modifiers of PD penetrance and age-at-onset. One variant on chromosome 12 reached genomewide significance in the penetrance model (rs77395454 in an intronic region of CORO1C). Several loci reached suggestive significance in either the penetrance model or the age-at-onset model. One region on chromosome 3 showed suggestive associations in both models and reached genomewide significance in penetrance model when focused on G2019S carriers only. PRS derived from a publicly available PD GWAS was a significant predictor of penetrance of PD among LRRK2 mutation carriers.

The genomewide significant variant, rs77395454 on chromosome 12, is located in an intronic region of CORO1C. Our co-immunoprecipitation experiments from HEK293 cells showed significant associations in both models and reached genomewide significance in penetrance model when focused on G2019S carriers only. PRS derived from a publicly available PD GWAS was a significant predictor of penetrance of PD among LRRK2 mutation carriers.

The accumulation of CORO1C in knockout mice might represent compensation for diminished LRRK2 function. The CORO1C protein is a member of the WD repeat protein family that has been implicated in signal transduction, gene regulation (https://www.ncbi.nlm.nih.gov/gene/23603). In a zebrafish model of spinal muscular atrophy, overexpression of CORO1C rescued the phenotype caused by SMN deficiency. Using mass spectrometry, Malty et al showed that the product of CORO1C interacts with mitochondrial proteins associated with neurodegeneration. Collectively, these complementary results support that CORO1C is a more likely functional interactor of LRRK2 and all of these warrant more in-depth cell and in vitro studies, including mapping the domains of LRRK2 responsible for the interaction between LRRK2 and CORO1C. However, it is possible that other genes in this region may underpin the observed association. For example, the protein product of SSH1 regulates actin filament dynamics, which has been linked to LRRK2 mutations. SELPLG has been linked to neuropsychiatric disorders, such as conduct disorder. Further studies are needed to conclusively determine the gene(s) underlying the observed association.

Multiple variants on chromosome 3 were supported by both models, although they identified different associated haplotypes. The most significant variants were rs16846845 in the penetrance model and rs150382576 in the age-at-onset model. Furthermore, rs16846845 reached genomewide significance in G2019S only analysis for penetrance model. This region is under a known linkage peak

FIGURE 4: Evidence of a LRRK2-CORO1C complex. (A) Immunoprecipitation (IP) analysis of lysates from HEK293FT cells transiently expressing eGFP-C1 empty vector control or eGFP-CORO1C. IP with GFP-Trap was followed with immunoblot analysis with an anti-GFP (bottom panel) or an anti-LRRK2 antibody (top panel) as indicated. Input denotes whole cell lysate material that was used for immunoprecipitation analysis. (B, C) Co-immunoprecipitation analysis of lysates from HEK293FT cells transiently co-expressing 3xFLAG (3FL) epitope-tagged LRRK2 with either eGFP-C1 empty vector control or eGFP-CORO1C. Epitope-tagged proteins were recovered from lysates using anti-Flag (IP: Flag) or anti-GFP (IP: GFP) antibodies, followed by immunoblot detection using anti-Flag (top panels) and anti-GFP (bottom panels) antibodies. Input denotes whole cell lysate starting material from HEK293FT cells used for immunoprecipitation analysis; +/− indicates cDNA transfections.
The significant effect of the PRS in the penetrance model supports the polygenic nature of the \(LRRK2\) modifiers (ie, there are many genetic variants each with a small effect that collectively have a significant effect on the risk of PD in \(LRRK2\) mutation carriers). This result is in line with the recent analysis of Iwaki et al.\(^63\) In that study, a PRS was derived using 89 genomewide significant variants (some variants were also included in our PRS) identified in a PD GWAS of Nalls et al.\(^4\) Iwaki et al found that the PRS was significantly associated with \(LRRK2\) G2019S penetrance. Potential overlap between the participants in our study and that of Iwaki et al means that the results of these studies do not represent independent replication. We did not detect a significant association in the age-at-onset model. One reason for this may be the smaller sample size (less than half of that in the penetrance model, only 776 affected from 1,879 total participants analyzed), and the resulting lack of statistical power. Another possible explanation is that the PRS was derived from a GWAS comparing PD cases and controls, and these risk-associated genes/variants are not necessarily associated with age-at-onset. Note, some participants of our study were included in the study of Nalls et al.\(^4\) Although we were unable to directly check for overlapping samples and the results were potentially biased, the overlapping samples is at most 0.13% of the total sample in Nalls et al, therefore, our samples had a minimal influence on the weight estimation that was used to calculate the PRS.

There are several limitations of this study. First, despite the effort to enroll as many participants as possible, the sample size of this study still resulted in only modest statistical power. With this sample size, assuming a linear model, for a variant with MAF 3%, a change of at least 6 years of age-at-onset can be detected with 80% power at a genomewide significant level. Second, to maximize the number of eligible studies to join this collaboration, our inclusion criteria was quite minimal. Although this approach dramatically increased the sample size, many potentially important covariates were not collected and could not be adjusted for in subsequent analyses. Third, approximately 96% of our participants were G2019S carriers. However, there are carriers of other \(LRRK2\) mutations in the MJFF cohort. Although in a sensitivity analysis using only G2019S carriers, we observed similar effects for those top variants that we identified in both models, these mutations may have different effects that cannot be detected in the small number of carriers. Fourth, our study cohorts consisted of family participants and unrelated participants. Family history was not collected for every participant. Therefore, some unrelated individuals may be sporadic PD and have different penetrance from familial PD participants. Fifth, there was a lack of information on subjects with subtle signs of PD but who did not yet merit a diagnosis of PD. Sixth, although we included 10 PCs to adjust population stratifications, there may still exist fine-scale population stratifications that cannot be detected by those 10 PCs thus could potentially cause false positive findings. Nevertheless, we detected a genomewide significant variant. We provide experimental data to show CORO1C and \(LRRK2\) interact, and support that observation by proteomics literature. The PRS analysis suggested that there is unlikely to
be one or several single LRRK2 modifiers, but similar to overall PD risk, penetrance of LRRK2 mutations is affected by multiple genetic variants. Given the significant therapeutic efforts underway to develop targets for patients with PD carrying LRRK2 mutations, further replication of these results is essential. Furthermore, the genetic variants identified in this study and the PRS evaluated in the LRRK2 mutation carriers, may be used in the future to make personalized prevention and treatment possible.

Acknowledgments
The authors thank the research participants from all sites who made this study possible and Tara Candido for collecting the data.

Members of the 23andMe Research Team are: Michelle Agee, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Nicholas A. Furlotte, David A. Hinds, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Joyce Y. Tung, Vladimir Vacic, and Catherine H. Wilson.

Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268201200008I.

The authors acknowledge the Indiana University Pervasive Technology Institute for providing [HPC (Big Red II, Karst, Carbonate)] visualization, database, storage, or consulting resources that have contributed to the research results reported within this paper.

pEGFP-C1 and pEGFP-CORO1C were kind gifts from Prof. Dr. Brunhilde Wirth (University of Cologne; Germany). The 3xFlag-LRRK2 is as described.

This work is supported by The Michael J. Fox Foundation for Parkinson’s Research Grants 7984, 7984.01, 7984.02, and 8981; The Albertson Parkinson’s Research Foundation; The Brookdale Foundation; The Else Kroener Fresenius Foundation; The Haworth Family Professorship in Neurodegenerative Diseases fund; The Little Family Foundation; The Parkinson’s Foundation; The Sol Goldman Charitable Trust; NIH AG010124, AG062418, K02NS080915, NS036630, NS053488, NS071674, P50NS039764, P50NS062684, P50NS072187, R01NS065070, R01NS078086, R01NS096740, U54NS100693, U54NS110435, UL1TR000040, and UL1TR001873; The Department of Veterans Affairs 5101CX001702; DOD W81XWH-17-1-0249; The DFG FOR2488; The Canadian Institutes of Health Research; The 2019 Biomarkers Across Neurodegenerative Diseases Grant Program; BAND3 18063.

Author Contributions
T.F. and P.C. contributed to the conception and design of the study. D.L., B.A., P.F., P.C., and T.F. drafted the text and prepared the figures. D.L., B.A., P.F., T.S., J.A., R.N.A., G.W.B., D.B., S.B., A.B., K.B., L.C., M.C., S.D., V.V.D., M.F., J. Trinh, T.G., S.G., E.G., C.K., A.E.L., J.W.L., J.L., T.L., K.M., C.M., E.R.M., C.Y.M., H.M., E.M., R.H.M., K.N., L.O., H.P., D.R., E.R., M.P.R., O.A.R., A.S., R.S., B.S., C.S., W.K.S., C.T., E.T., J.E.T., D.V., J. Trojanowski, R.U., J.M.V., N.P.V., Z.K.W., C.P.Z., A.M., N.G., A.O.U., and B.F. contributed to acquisition and analysis of data. All authors reviewed and approved the submission.

Potential Conflicts of Interest
B.A., P.F., P.C., S.D., C.Y.M., and members of the 23andMe Research Team are current or former employees of 23andMe, Inc., and hold stock or stock options in 23andMe.

Data Availability
Aggregate-level data included in this study will be made available to qualified investigators upon request. Investigators interested in receiving 23andMe data, either alone or in combination with data from other cohorts, will need to sign a Data Transfer Agreement with 23andMe that protects 23andMe research participant privacy, and should visit https://research.23andme.com/dataset-access/ to submit a request.

References
1. Corti O, Lesage S, Brice A. What genetics tells us about the causes and mechanisms of Parkinson’s disease. Physiol Rev 2011;91:1161–1218.
2. Singleton AB, Farrer MJ, Bonifati V. The genetics of Parkinson’s disease: progress and therapeutic implications. Mov Disord 2013;28:14–23.
3. Chang D, Nalls MA, Hallgrimsdottir I.B, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson’s disease risk loci. Nat Genet 2017;49:1511–1516.
4. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson’s disease: a meta-analysis of genome-wide association studies. Lancet Neurol 2019;18:1091–1102.
5. Alessi DR, Sammler E. LRRK2 kinase in Parkinson’s disease. Science 2018;360:36–37.
6. Correia Guedes L, Ferreira JJ, Rosa MM, et al. Worldwide frequency of G2019S LRRK2 mutation in Parkinson’s disease: a systematic review. Parkinsonism Relat Disord 2010;16:237–242.
7. Di Fonzo A, Rohe CF, Ferreira J, et al. A frequent LRRK2 gene mutation associated with autosomal dominant Parkinson’s disease. Lancet 2005;365:412–415.
48. Schmitt AD, Hu M, Jung I, et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep 2016;17:2042–2059.

49. Giusti-Rodríguez P, Lu L, Yang Y, et al. Using three-dimensional regulatory chromatin interactions from adult and fetal cortex to interpret genetic results for psychiatric disorders and cognitive traits. bioRxiv 2019: 406330.

50. Boon JY, Dusonchet J, Trengrove C, Wolozin B. Interaction of LRRK2 with kinase and GTPase signaling cascades. Front Mol Neurosci 2014;7:64.

51. Bonet-Ponce L, Beilina A, Williamson CD, et al. LRRK2 mediates tubulation and vesicle sorting from membrane damaged lysosomes. bioRxiv 2020: 2020.01.23.917252.

52. Pellegrini L, Hauser DN, Li Y, et al. Proteomic analysis reveals coordinated alterations in protein synthesis and degradation pathways in LRRK2 knockout mice. Hum Mol Genet 2018;27:3257–3271.

53. Hosseinibarkooie S, Peters M, Torres-Benito L, et al. The power of human protective modifiers: PLS3 and CORO1C unravel impaired endocytosis in spinal muscular atrophy and rescue SMA phenotype. Am J Hum Genet 2016;99:647–665.

54. Malty RH, Aoki H, Kumar A, et al. A map of human mitochondrial protein interactions linked to neurodegeneration reveals new mechanisms of redox homeostasis and NF-kappaB signaling. Cell Syst 2017;5:564–577.e12.

55. Bardai FH, Ordonez DG, Bailey RM, et al. Lnk promotes tau neurotoxicity through dysregulation of actin and mitochondrial dynamics. PLoS Biol 2018;16:e2006265.

56. Caesar M, Felik S, Asaly JO, Gillardon F. Changes in actin dynamics and F-actin structure both in synaptoneurosomes of LRRK2(R1441G) mutant mice and in primary human fibroblasts of LRRK2(G2019S) mutation carriers. Neuroscience 2015;284:311–324.

57. Dick DM, Aliev F, Krueger RF, et al. Genome-wide association study of conduct disorder symptomatology. Mol Psychiatry 2011;16:800–808.

58. Gao X, Martin ER, Liu Y, et al. Genome-wide linkage screen in familial Parkinson disease identifies loci on chromosomes 3 and 18. Am J Hum Genet 2009;84:499–504.

59. Lee SL, Murdock DG, McCauley JL, et al. A genome-wide scan in an Amish pedigree with parkinsonism. Ann Hum Genet 2008;72:621–629.

60. Landers JE, Melki J, Meiningier V, et al. Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. Proc Natl Acad Sci U S A 2009;106:9004–9009.

61. Mez J, Chung J, Jun G, et al. Two novel loci, COBL and SLC10A2, for Alzheimer’s disease in African Americans. Alzheimers Dement 2017;13:119–129.

62. Sanchez-Contreras MY, Kouri N, Cook CN, et al. Replication of progressive supranuclear palsy genome-wide association study identifies SLC01A2 and DUSP10 as new susceptibility loci. Mol Neurodegener 2018;13:37.

63. Iwaki H, Blauwendraat C, Makarious MB, et al. Penetrance of Parkinson’s disease in LRRK2 p.G2019S carriers is modified by a polygenic risk score. Mov Disord 2020;35:774–780.

64. Beilina A, Bonet-Ponce L, Kumaran R, et al. The Parkinson’s disease protein LRRK2 interacts with the GARP complex to promote retrograde transport to the trans-Golgi network. Cell Rep 2020;31:107614.