A two-stage meta-analysis identifies several new loci for Parkinson's disease

Citation for published version:

Digital Object Identifier (DOI):
10.1371/journal.pgen.1002142

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
PLoS Genetics

Publisher Rights Statement:
This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
A Two-Stage Meta-Analysis Identifies Several New Loci for Parkinson’s Disease

International Parkinson’s Disease Genomics Consortium (IPDGC), Wellcome Trust Case Control Consortium 2 (WTCCC2)

Abstract

A previous genome-wide association (GWA) meta-analysis of 12,386 PD cases and 21,026 controls conducted by the International Parkinson’s Disease Genomics Consortium (IPDGC) discovered or confirmed 11 Parkinson’s disease (PD) loci. This first analysis of the two-stage IPDGC study focused on the set of loci that passed genome-wide significance in the first stage GWA scan. However, the second stage genotyping array, the ImmunoChip, included a larger set of 1,920 SNPs selected on the basis of the GWA analysis. Here, we have analyzed this set of 1,920 SNPs, and we identified five additional PD risk loci (combined \( p < 5 \times 10^{-10} \), PARK16/1q32, STX1B/16p11, FGF20/8p22, STBD1/4q21, and GPNMB/7p15). Two of these five loci have been suggested by previous association studies (PARK16/1q32, FGF20/8p22), and this study provides further support for these findings. Using a dataset of post-mortem brain samples assayed for gene expression (\( n = 292 \)) and methylation (\( n = 399 \)) and expression changes associated with PD risk variants in PARK16/1q32, GPNMB/7p15, and STX1B/16p11 loci, hence suggesting potential molecular mechanisms and candidate genes at these risk loci.

Introduction

Until the recent developments of high throughput genotyping and genome-wide association (GWA) studies, little was known of the genetics of typical Parkinson’s disease (PD). Studies of the genetic basis of familial forms of PD first identified rare highly penetrant mutations in LRRK2 [2,12], PARK7 [6] and PARK2 [5] and PARK7 [6]. Following these findings, GWA scans for idiopathic PD identified SNCA and MAPT as unequivocal risk loci [7,8,9,10,11] as well as implicated BST1 [8], GAK [12], and HLA-DR [13]. Using sequence based imputation methods [14], the meta-analysis of several GWA scans [7,9,10,11] conducted by the International Parkinson’s Disease Genomics Consortium (IPDGC) identified and replicated five new loci: AGMSD, STX9, MCCC1/ LAMPS, STT11, and CCDC62/HIP1R [15] and confirmed association at SNCA, LRRK2, MAPT, BST1, GAK and HLA-DR [15].

We conducted a two-stage association study. Combining stage 1 and stage 2, the data consist of 12,386 PD cases and 21,026 controls genotyped using a variety of platforms (Table 1). Stage 1 used genome-wide genotyping arrays and our initial analysis [15] focused on the subset of SNPs that passed genome-wide significance in stage 1. For stage 2 genotyping, we used a custom content Illumina iSelect array, the ImmunoChip and additional...
Author Summary

This paper describes the largest case-control analysis of Parkinson’s disease to date, with a combined sample set of over 12,000 cases and 21,000 controls. After combining our findings with an independent replication dataset of more than 3,000 cases and 29,000 controls, we found five additional PD risk loci in addition to the 11 loci previously identified in earlier consortium efforts. This successful study further demonstrates the power of the GWA scan experimental design to find new loci contributing to disease risk, even in the context of complex disorders like Parkinson’s disease. These new findings provide insights into the etiology of PD and will promote a better understanding of its pathogenesis.

GWAS typing as previously described [15]. The primary content of the ImmunoChip data focuses on autoimmune disorders but, as part of a collaborative agreement with the Wellcome Trust Case Control Consortium 2, we included 1,920 ImmunoChip SNPs on the basis of the stage 1 GWA PD results.

Here, we report the combined analysis for this full set of 1,920 SNPs. This step1+2 analysis identified seven new loci that passed genome-wide significance in the meta-analysis. During the process of analyzing these data and preparing for publication, we became aware that another group was also preparing a large independent GWA scan in PD for publication (Do et al, submitted). Following discussion with this group we agreed to cross validate the top hits from each study by exchanging summary statistics for this small number of loci.

To provide further insights into the molecular function of these associated variants, we tested risk alleles at these loci for correlation with the expression of physically close gene (expression quantitative trait locus, eQTL) and the methylation status (methQTL) of proximal DNA CpG sites in a dataset of 399 control frontal cortex and cerebellar tissue samples extracted post-mortem from individuals without a history of neurological disorders.

Results

In addition to eleven loci that passed genome-wide significance in stage 1 [15], we identified over 100 regions of interest defined as 10 kb windows containing at least one SNP associated at \( p < 10^{-3} \). We submitted the most associated SNP in each region for probe design and follow-up genotyping using the ImmunoChip platform. For each region of interest, we also added four SNPs in high level of linkage disequilibrium (LD) to provide redundancy where the most associated SNP would not pass the Illumina probe design step or the assay for that SNP would fail. To complete the array design we also added all non-synonymous dbSNPs located in known PD associated regions [1,2,3,4,5,6]. Of these 2,400 submitted SNPs, 1,920 passed QC and were included in the final array design. For these 1,920 SNPs we combined stage 1 and stage 2 associated data in a meta-analysis of 12,386 cases and 21,026 controls (Table 1) from the IPDGC. We exchanged summary statistics for these most significant hits with an additional large, case-control replication dataset (3,426 PD cases and 29,624 controls) in an attempt to demonstrate independent replication.

On the basis of stage 1+2 results, seven new SNPs passed our defined genome-wide significance threshold \( p < 5 \times 10^{-8} \), Table 2 and Figure 1. These loci are either novel or the previous evidence of association was not entirely convincing in individuals of European descent. We combined these results with the independent replication. Five of these seven loci replicated and showed strong combined evidence of PD association \( p < 10^{-4} \) overall. Taking either the nearest gene (or the strongest candidate when available) to designate these regions, these five loci are 4q21/STBD1, 7p15/GPNMB, 8p22/FGF20 [16] and 16p11/STXB.

rs700723/1q32 has been previously reported as PD associated \( \text{p } = 9.47 \times 10^{-4} \) in stage 2 only. To understand the potential biological consequences of risk variation at this locus we tested whether rs700723 was correlated with either gene expression or DNA methylation status of proximal transcripts or CpG sites respectively (Table 3). We found correlations with the expression of \( NU/CKS1 \) \( \text{p } = 1.8 \times 10^{-7} \) and \( RAB7L1 \) \( \text{p } = 7.3 \times 10^{-5} \). We also found correlations with the methylation state of CpG sites located in the \( FLJ3269 \) gene \( \text{p } = 3.9 \times 10^{-22} \).

In the case of 16p11/STXB, the proximal gene to the most associated SNP rs4889603 is \( SETD14 \). However, \( STXB \) is located 18 kb upstream of rs4889603 and is a more plausible PD candidate gene [17] owing to its synaptic receptor function. We therefore used this gene to designate this region. Our methQTL/eQTL dataset identified a correlation between the rs4889603 risk allele and increased methylation of a CpG dinucleotide in \( STXB \) (Table 3).

The SNP rs591323 in the 8p22 region is located \(~150\text{ kb downstream of the FGF20 gene} (\text{NCBI build 36.3})\), for which association with PD has been suggested previously in familial PD samples [16,18] but which remained controversial [19]. Our findings provide further support for a PD association at this locus, but again, whether the functionally affected transcript is \( FGF20 \) or not remains unclear.

The regions 4q21/STBD1 and 7p15/GPNMB have not been previously implicated in PD etiology. We found that the risk allele of rs156429, the most associated SNP in the 7p15 region, is associated in our eQTL dataset with decreased expression of the proximal transcript encoded by \( NUPL2 \) (Table 3). The same risk allele is also associated with increased methylation of multiple CpG sites proximal to \( GPNMB \) itself (Table 3). Neither of these regions contains an obvious candidate gene.

Two additional loci (3q26/NAIM3 and 8q21/AMPL1) showed strong evidence of association in stage 1 and 2 but were not disease associated in the Do et al dataset. Further replication is required to clarify the role of variation at these loci in risk for PD.

The strongly associated G2019S variant in the \( LRRKE2 \) gene [20] was included in the Immunochip design and we replicated the published association: control frequency: \( 0.045\% \) case frequency \( 0.61\% \), estimated odds ratio: 13.5 with 95% confidence interval: 5.5–43. However, the case collections have been partially screened for this variant therefore its frequency in cases and the odds ratio is likely to be underestimated.

The ImmunoChip array design provides some power to detect whether multiple distinct association signals exist at individual loci. Indeed, if a SNP showed an independent and sufficiently strong association in stage 1, it would have been included in stage 2 provided that it was not located in the same 10 kb window as the primary SNP in the region. There is precedent for this in PD, with the previous identification of independent risk signals at the \( SCN4A \) locus [11]. We therefore used the Immunochip data to test whether any of the seven loci in Table 2 showed some evidence of more than one independent signal. None of these seven loci showed any association \( \text{p } \geq 0.01 \) after conditioning on the main SNP in the region. In contrast, after conditioning on the most associated SNPs rs356182 in the \( SCN4A \) region, several SNPs...
remained convincingly associated ($p = 9.7 \times 10^{-8}$ for rs2245801 being the most significant).

Lastly, we performed a risk profile analysis to investigate the power to discriminate cases and controls on the basis of the 16 confirmed common associated variants (Table 4). For each locus, we estimated the odds ratio on the basis of stage 1 data and we applied these estimates to compute for each individual in the ImmunoChip cohort a combined risk score. Solely based on these 16 common variants, and therefore not considering rare highly penetrant variants such as G2019S in LRRK2 [20], we found that individuals in the top quintile of the risk score have an estimated three-fold increase in PD risk compared to individuals in the bottom quintile (Table 4). We note however that the effect size of several of these associated variants could be over-estimated (an effect known as winner’s curse, see [21]) but given the consistent estimates of odds ratio across studies (Table 4) we expect this bias to be minimal.

**Discussion**

The combination of GWA scans and imputation methods in large cohorts of PD cases and controls has enabled us to identify five PD associated loci in addition to the 11 previously reported by us. Two of these loci (1q32/PARK16, 8p22/FGF20) implicate regions that had been previously associated with PD risk [8,16]. The 1q32/PARK16 showed convincing evidence of association in the Japanese population [8] but until now the association P-value had not passed a stringent genome-wide significance threshold in samples of European descent [7]. The 8p22/FGF20 locus had been previously reported in a study of familial PD [16] and we provide the first evidence of association in a case-control study. The remaining three loci (STX1B/16p11, STBD1/4q21 and GPNMB/7p15) are new.

Adding the eleven previously reported common variants [15] to the five convincingly associated loci identified in this study, common variants at 16 loci have now been associated with PD. Controlling for the risk score based on the 11 SNPs previously identified [15] in the risk profile analysis (Table 4), the addition of these five new loci provides a modest but significant ($p = 2.2 \times 10^{-5}$) improvement of our ability to discriminate PD cases from controls.

Combining eQTL/methylation and case-control data implicates potential mechanisms which could explain the increased PD risk associated some of these variants. In particular, the strong eQTL in the 1q32/PARK16 region with the RAB7L1 and NUCKS1 genes (Table 3) suggests that either one of these genes could be the biological effector of this risk locus. However, existing data show that eQTLs are known to be involved in dopaminergic neurons survival [30]. Nevertheless, the newly reported secondary association (rs2245801) is in low LD ($r^2 = 0.21$) with rs2301134, the former is a GTP-binding protein that regulates the number and strength of excitatory synapses [27], it makes STX1B a very interesting candidate from a biologic perspective.

FGF20 at 8p22 has been suggested to be involved in PD [16], albeit negative results in smaller cohorts have followed the original finding [28]. FGF20 is a neurotrophic factor that exerts strong neurotrophic properties within brain tissue, and regulates central nervous development and function [29]. It is preferentially expressed in the substantia nigra [30], and it has been reported to be involved in dopaminergic neurons survival [30].

The ImmunoChip data provide limited resolution for the detection of multiple independent association signals in these regions. A previous study [31] reported some evidence of allelic heterogeneity at the 1q32/PARK16 locus but the ImmunoChip data do not support this result. A previous study [11] also reported two independent associations at the 4q22/SNCA locus and our data are consistent with this scenario. However, the newly reported secondary association (rs2245801) is in low LD ($r^2 = 0.21$) with rs2301134, the SNP reported in [11] as an independent association. Taken together, these findings suggest that at least three independent associations exist at SNCA/4q22. A more exhaustive fine-mapping analysis using either sequencing of large cohorts or targeted genotyping arrays will also be required to fully explore this locus.

As yet, we do not know which of the variants and which genes within each region are exerting the pathogenic effect. We cannot exclude that some of the currently reported variants are in fact tagging high penetrance, but rare, mutations [32]. Nevertheless, the successful identification of these 16 risk loci further demonstrates the power of the GWA study design, even in the context of disorders like Parkinson’s disease.

**Table 1.** Sample size and genotyping platform for the cohorts included in stage 1 (top set of rows), stage 2 (middle set of rows), and independent replication (bottom row).

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Controls</th>
<th>Cases</th>
<th>Genotyping platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>5,200</td>
<td>1,705</td>
<td>Illumina 660W-Quad</td>
</tr>
<tr>
<td>USA-NIA</td>
<td>3,034</td>
<td>971</td>
<td>Illumina HumanHap550</td>
</tr>
<tr>
<td>USA-dbGAP</td>
<td>857</td>
<td>876</td>
<td>Illumina 370 K</td>
</tr>
<tr>
<td>German</td>
<td>944</td>
<td>742</td>
<td>Illumina HumanHap550</td>
</tr>
<tr>
<td>French</td>
<td>1,984</td>
<td>1039</td>
<td>Illumina 610-Quad</td>
</tr>
<tr>
<td><strong>Total Stage 1</strong></td>
<td><strong>12,019</strong></td>
<td><strong>5,333</strong></td>
<td></td>
</tr>
<tr>
<td>Icelandic</td>
<td>1,427</td>
<td>479</td>
<td>Illumina HumanHap300</td>
</tr>
<tr>
<td>Dutch</td>
<td>2,024</td>
<td>772</td>
<td>Illumina 610-Quad</td>
</tr>
<tr>
<td>USA</td>
<td>2,215</td>
<td>2,807</td>
<td>ImmunoChip</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1,864</td>
<td>1,271</td>
<td>ImmunoChip</td>
</tr>
<tr>
<td>Dutch</td>
<td>402</td>
<td>304</td>
<td>ImmunoChip</td>
</tr>
<tr>
<td>French</td>
<td>363</td>
<td>267</td>
<td>ImmunoChip</td>
</tr>
<tr>
<td>German</td>
<td>712</td>
<td>1,153</td>
<td>ImmunoChip</td>
</tr>
<tr>
<td><strong>Total Stage 2</strong></td>
<td><strong>9,007</strong></td>
<td><strong>7,053</strong></td>
<td></td>
</tr>
<tr>
<td>Stage 1+Stage 2</td>
<td>21,026</td>
<td>12,386</td>
<td></td>
</tr>
<tr>
<td>Do et al- USA</td>
<td>29,624</td>
<td>3,426</td>
<td></td>
</tr>
</tbody>
</table>

 dof:10.1371/journal.pgen.1002142.t001
Two-Stage Meta-Analysis of Parkinson’s Disease

rs708723

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.71
0.89
1.12
1.26

rs34016896

dbGAP-GWAS
FR-GWAS
GER-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.79
0.89
1.00
1.12
1.26
1.41

rs6812193

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.63
0.79
1.00
1.26

rs156429

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.63
0.71
0.79
0.89
1.00
1.12

rs591323

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.71
0.89
1.12
1.26

chr8:89442157

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

0.79
1.26
2.00
3.16
5.01

rs4889603

dbGAP-GWAS
FR-GWAS
GER-GWAS
ICE-GWAS
NL-GWAS
UK-GWAS
US-GWAS
US-Immuno

Summary

Odds Ratio

1.00
1.26
1.58
2.00
PD that have a complex genetic component. We therefore expect that further and larger association analyses, perhaps using dedicated high-throughput genotyping arrays like the ImmunoChip, will continue to yield new insights into PD etiology.

Material and Methods

Genotyping and case control cohorts

Participating studies were either genotyped using the ImmunoChip as part of a collaborative agreement with the ImmunoChip Consortium, or as part of previous GWA studies provided by members of the IPDGC or freely available from dbGaP [7,9,10,11]. Genotyping of the UK cases using the Immunochip was undertaken by the WTCCC2 at the Wellcome Trust Sanger Institute which also genotyped the UK control samples. The constituent studies comprising the IPDGC have been described in detail elsewhere [15], although a summary of individual study quality control is available as part of Table S1. In brief all studies followed relatively uniform quality control procedures such as: minimum call rate per sample of 95%, mandatory concordance between self-reported and X-chromosome-heterogeneity estimated sex, exclusion of SNPs with greater than 5% missingness, HardyWeinberg equilibrium p-values at a minimum of 10^−7, minor allele frequencies at a minimum of 1%, exclusion of first degree relatives, and the exclusion of ancestry outliers based on either principal components or multidimensional scaling analyses using either PLINK [33] or EIGENSTRAT [34] to remove non-European ancestry samples. All GWAS studies utilized in this analysis (and in the QTL analyses) were imputed using MACHv1.0.16 [14] to conduct a two-stage imputation based on the August 2009 haplotypes from initial low coverage sequencing of 112 European ancestry samples in the 1000 Genomes Project [35], filtering the data for a minimum imputation quality of (RSQR>0.3) [14]. Logistic regression models were utilized to quantify associations with PD incorporating allele dosages as the primary predictor of disease. Imputed data was analyzed using MACH2DAT, and genotyped SNPs were analyzed using PLINK. All models were adjusted for covariates of components 1 and 2 from either principal components or multidimensional scaling analyses to account for population substructure and stochastic genotypic variation (except in the UK-GWAS data which were not adjusted for population substructure).

Association test statistics

Single SNP test statistics were combined across datasets using a score test methodology, essentially assuming equal odds ratio across cohorts. In addition, fixed and random effects meta-analyses were implemented in R (version 2.11) to confirm that the score test approximation does not affect the interpretation of the results. We also tested the relevant SNPs heterogeneity across cohorts and no significant heterogeneity was detected (Table S2).

Data exchange

We communicated to our colleagues in charge of the independent study (Do et al) the seven SNPs listed in Table 2. For this subset of SNPs they selected the marker with the highest r^2 value on their genotyping platform and provided us with the following summary statistics: odds ratio, direction of effect, standard error for the estimated odds ratio and one degree-of-freedom trend test P-value.

eQTL analysis and methylation analysis

Quantitative trait analyses were conducted to infer effects of risk SNPs on proximal CpG methylation and gene expression. For the five replicated SNP associations (Table 2), all available CpG probes and expression probes within +/- 1 MB of the target SNP were

**Table 2.** Summary statistics for the seven SNPs that pass genome-wide significance (p<5x10^-8) in the combined stage 1+2 analysis and that have either not been reported in published PD association studies.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chrom</th>
<th>Gene(s)</th>
<th>Alleles</th>
<th>MAF</th>
<th>Stage 1 OR (95%CI)</th>
<th>P</th>
<th>Stage 2 OR (95%CI)</th>
<th>P</th>
<th>Stage 1+2 OR (95%CI)</th>
<th>P</th>
<th>Do et al OR (95%CI)</th>
<th>P</th>
<th>Combined OR (95%CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs708723</td>
<td>1q32</td>
<td>RAB7L1/PARK16</td>
<td>T&gt;C</td>
<td>0.439</td>
<td>0.905 (0.862–0.95)</td>
<td>6.68×10^-5</td>
<td>0.863 (0.824–0.905)</td>
<td>9.47×10^-10</td>
<td>1.00×10^-12</td>
<td>0.758 (0.65–0.88)</td>
<td>2.12×10^-6</td>
<td>8.82×10^-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs34016896</td>
<td>3q26</td>
<td>NMD3</td>
<td>C&gt;T</td>
<td>0.305</td>
<td>1.14 (1.09–1.2)</td>
<td>3.00×10^-7</td>
<td>1.08 (1.02–1.14)</td>
<td>0.00399</td>
<td>1.81×10^-8</td>
<td>1.002 (0.95–1.06)</td>
<td>0.954</td>
<td>1.31×10^-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6812193</td>
<td>4q21</td>
<td>STBD1</td>
<td>C&gt;T</td>
<td>0.36</td>
<td>0.886 (0.843–0.932)</td>
<td>2.52×10^-10</td>
<td>0.906 (0.864–0.95)</td>
<td>5.29×10^-10</td>
<td>7.46×10^-10</td>
<td>0.839 (0.79–0.89)</td>
<td>7.55×10^-10</td>
<td>1.17×10^-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs156429</td>
<td>7p15</td>
<td>GPNMB</td>
<td>A&gt;G</td>
<td>0.403</td>
<td>0.894 (0.849–0.942)</td>
<td>2.15×10^-5</td>
<td>0.893 (0.852–0.917)</td>
<td>3.86×10^-10</td>
<td>3.27×10^-10</td>
<td>0.901 (0.85–0.95)</td>
<td>0.000193</td>
<td>3.05×10^-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs591323</td>
<td>8p22</td>
<td>FG20</td>
<td>G&gt;A</td>
<td>0.271</td>
<td>0.884 (0.836–0.935)</td>
<td>1.59×10^-5</td>
<td>0.875 (0.83–0.923)</td>
<td>8.49E×10^-7</td>
<td>7.45×10^-11</td>
<td>0.932 (0.88–0.99)</td>
<td>0.023</td>
<td>1.92×10^-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:894421578q21</td>
<td>MMP16</td>
<td>C&gt;T</td>
<td>0.0247</td>
<td>1.38 (1.21–1.57)</td>
<td>1.10×10^-6</td>
<td>1.29 (1.12–1.49)</td>
<td>0.000451</td>
<td>2.26×10^-9</td>
<td>0.969 (0.86–1.09)</td>
<td>0.589</td>
<td>2.36×10^-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4889003</td>
<td>16p11</td>
<td>STX1B</td>
<td>A&gt;G</td>
<td>0.413</td>
<td>1.12 (1.06–1.18)</td>
<td>4.13×10^-5</td>
<td>1.15 (1.11–1.21)</td>
<td>8.21×10^-9</td>
<td>2.66×10^-12</td>
<td>1.070 (1.01–1.13)</td>
<td>0.014</td>
<td>6.98×10^-13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1q32/PARK16 has been reported previously but is included because these data provide for the first time unequivocal evidence of association. P-values are computed using a one-degree-of-freedom trend test, including two principal components as covariates and combining the results across cohorts using a score test methodology. P-values are two-tailed and odds ratios are reported for the minor alleles. The notation X>Y indicates that X is the major allele and Y the minor allele. Allele frequencies were estimated using the UK control data. OR: odds ratio.
investigated as candidate QTL associations in frontal cortex and cerebellar tissue samples. 399 samples were assayed for genome-wide gene expression on Illumina HumanHT-12 v3 Expression Beadchips and 292 samples were assayed using Infinium HumanMethylation27 Beadchips, both per manufacturer’s protocols in each brain region. A more in depth description of the sample series comprising the QTL analyses, relevant laboratory procedures and quality requirements may be found in [15]. The QTL analysis utilized multivariate linear regression models to estimate effects of allele dosages per SNP on expression and methylation levels adjusted for covariates of age at death, gender, the first 2 component vectors from multi-dimensional scaling, post mortem interval (PMI), brain bank from where the

doi:10.1371/journal.pgen.1002142.003

Table 3. Significant eQTL associations (p<0.01) between the five SNPs with positive replication data (Table 2) and proximal (cis) changes in gene expression/methylation in frontal cortex and cerebellar tissue.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Region</th>
<th>SNP</th>
<th>Gene Tagged by</th>
<th>Illumina Probe</th>
<th>Alleles</th>
<th>Effect Estimate</th>
<th>Standard Error</th>
<th>Unadjusted P</th>
<th>Adjusted P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression</td>
<td>Frontal Cortex</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>NPL2</td>
<td>A&gt;G</td>
<td>0.083</td>
<td>0.018</td>
<td>3.6E-06</td>
<td>1.0E-04</td>
</tr>
<tr>
<td></td>
<td>Frontal Cortex</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>NPL2</td>
<td>A&gt;G</td>
<td>0.078</td>
<td>0.017</td>
<td>3.1E-06</td>
<td>1.0E-04</td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>NPL2</td>
<td>A&gt;G</td>
<td>0.133</td>
<td>0.025</td>
<td>1.0E-07</td>
<td>3.7E-06</td>
</tr>
<tr>
<td></td>
<td>Frontal Cortex</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>NPL2</td>
<td>A&gt;G</td>
<td>0.131</td>
<td>0.023</td>
<td>1.2E-08</td>
<td>1.0E-06</td>
</tr>
<tr>
<td>Methylation</td>
<td>Frontal Cortex</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>GPNMB</td>
<td>A&gt;G</td>
<td>0.027</td>
<td>0.005</td>
<td>5.1E-07</td>
<td>3.2E-05</td>
</tr>
<tr>
<td></td>
<td>Frontal Cortex</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>GPNMB</td>
<td>A&gt;G</td>
<td>0.009</td>
<td>0.002</td>
<td>1.6E-07</td>
<td>1.3E-05</td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>rs156429</td>
<td>7p15/GPNMB</td>
<td>GPNMB</td>
<td>A&gt;G</td>
<td>0.012</td>
<td>0.003</td>
<td>8.2E-05</td>
<td>3.4E-03</td>
</tr>
</tbody>
</table>

Table 4. Estimated PD risk profile for the five cohorts genotyped using the Immunochip.

<table>
<thead>
<tr>
<th>Study</th>
<th>Trend P-value</th>
<th>AUC</th>
<th>1st quintile</th>
<th>2nd quintile</th>
<th>3rd quintile</th>
<th>4th quintile</th>
<th>5th quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>OR</td>
<td>95% CI</td>
<td>OR</td>
</tr>
<tr>
<td>USA</td>
<td>~2E-16</td>
<td>0.614</td>
<td>1</td>
<td>–</td>
<td>1.54</td>
<td>1.29–1.84</td>
<td>1.92</td>
</tr>
<tr>
<td>UK</td>
<td>~2E-16</td>
<td>0.636</td>
<td>1</td>
<td>–</td>
<td>1.34</td>
<td>1.05–1.71</td>
<td>1.79</td>
</tr>
<tr>
<td>Germany</td>
<td>1.29E-11</td>
<td>0.692</td>
<td>1</td>
<td>–</td>
<td>1.32</td>
<td>0.98–1.79</td>
<td>1.88</td>
</tr>
<tr>
<td>France</td>
<td>5.19E-13</td>
<td>0.675</td>
<td>1</td>
<td>–</td>
<td>1.69</td>
<td>0.99–2.92</td>
<td>1.13</td>
</tr>
<tr>
<td>Netherlands</td>
<td>5.08E-05</td>
<td>0.601</td>
<td>1</td>
<td>–</td>
<td>1.06</td>
<td>0.65–1.74</td>
<td>1.35</td>
</tr>
<tr>
<td>Combined</td>
<td>~2E-16</td>
<td>0.645</td>
<td>1</td>
<td>–</td>
<td>1.43</td>
<td>1.26–1.61</td>
<td>1.79</td>
</tr>
<tr>
<td>% Cases per Quintile</td>
<td></td>
<td>37.90</td>
<td>46.06</td>
<td>51.15</td>
<td>56.56</td>
<td>63.75</td>
<td></td>
</tr>
</tbody>
</table>

Risk scores for the 16 confirmed loci were computed using the odds ratio estimated from the genome-wide case-control genotype data. Individuals were split into quintile on the basis of their risk scores. The odds ratios quantify the effect of the computed risk quintile on the probability of being a PD case (one-degree-of-freedom logistic trend test with the PD status as a binary outcome variable and the quintiles, coded as 1–5, as covariates). The first quintile group was taken as a reference group. OR: odds ratio, CI: confidence interval.
doi:10.1371/journal.pgen.1002142.004
samples were provided and in which preparation/hybridization batch the samples were processed. A total of 670 candidate QTL associations were tested: 87 expression QTLs in the cerebellum samples, 83 expression QTLs in the frontal cortex samples, 249 methylation QTLs in the cerebellum samples and 249 methylation QTLs in the frontal cortex samples. Multiple test correction was undertaken using false discovery rate adjusted p-values<0.05 to dictate significance, with the p-value adjustment undertaken in each series separately, stratified by brain region and assay. A complete list of all QTL associations tested is included in Table S3.

Supporting Information

Table S1 Summary of results for fixed and random effects meta-analysis, as estimates of effect heterogeneity across cohorts and SNP used at the Do et al replication stage. (XLSX)

Table S2 Summary of the quality control parameters applied to the GWA datasets included in this study. (XLSX)

Table S3 Complete list of tested QTL associations (expression and methylation). (XLSX)

Text S1 Membership of the Wellcome Trust Case Control Consortium 2. (DOC)

Acknowledgments

This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Maryland (http://biowulf.nih.gov). DNA panels and samples from the NINDS Human Genetics Resource Center DNA and Cell Line Repository (http://ccr.coriell.org/ninds) were used in this study, as well as clinical data. The submitter that contributed samples are acknowledged in detailed descriptions of each panel (http://ccr.coriell.org/sections/Collections/NINDS/?ssid=10).


1 UCL Genetics Institute, University College London, London, United Kingdom, 2 Laboratory of Neurogenetics. National Institute on Aging, National Institutes of Health, Bethesda, Maryland, United States of America, 3 Department of Molecular Neuroscience, Institute of Neurology, University College London, London, United Kingdom, 4 Institut National de la Sante et de la Recherche Medicale, UMR 1043, Centre de Biologie du Sommeil et de la Régulation du Sommeil, Toulouse, France, 5 UCL Genetics Institute, University College London, London, United Kingdom, 6 Department of Social Medicine, Department of Molecular Neuroscience, Institute of Neurology, City Hospital, Sandwell, United Kingdom, 7 Department of Neurology, University Nijmegen Medical Centre, Nijmegen, The Netherlands, 8 Department of Molecular Genetics, South Carolina, United States of America, 9 Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, United States of America, 10 Department of Molecular Genetics, South Carolina, United States of America, 11 Department of Social Medicine, Department of Molecular Neuroscience, Institute of Neurology, City Hospital, Sandwell, United Kingdom, 12 Department of Neurology, Free University, Toulouse, France, 13 Department of Neurology, Toulouse, France, 14 Department of Medical Genetics, Institute of Human Genetics, University of Tübingen, Tübingen, Germany, 15 Department of Neurology, Landspítali University Hospital, Reykjavik, Iceland, 16 Department of Neurology, Maastricht University Medical Centre, Maastricht, The Netherlands, 17 Department of Neurology, University of Massachusetts, Boston, Massachusetts, United States of America, 18 Department of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands, 19 Department of Motor Neuroscience, University College London Institute of Neurology, London, United Kingdom, 20 Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, 21 Department of Motor Neuroscience, University College London Institute of Neurology, London, United Kingdom, 22 Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, 23 Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, United States of America, 24 Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts, United States of America, 25 Program in Medical and Population Genetics, Broad Institute, Cambridge, United States of America, 26 Department of Neurology, Radboud University Nijmegen, Nijmegen, The Netherlands, 27 Department of Medical Genetics, Institute of Human Genetics, University of Tübingen, Tübingen, Germany, 28 Newcastle University Clinical Ageing Research Unit, Campus for Ageing and Vitality, Newcastle upon Tyne, United Kingdom, 29 Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, North Carolina, United States of America, 30 Neurology Department, The Medical School, Newcastle upon Tyne, Newcastle University, United Kingdom, 31 School of Clinical and Experimental Medicine, University of Birmingham, Edgbaston, Birmingham, United Kingdom, 32 Department of Neurology, City Hospital, Sandwell, United Kingdom, 33 West

PLoS Genetics | www.plosgenetics.org 7 June 2011 | Volume 7 | Issue 6 | e1002142
Birmingham Hospitals NHS Trust, Birmingham, United Kingdom, 34 Department of Clinical Neurosciences, University College London Institute of Neurology, London, United Kingdom, 35 Institut National de la Sante et de la Recherche Medicale, CIC-9503, Hôpital Pitié-Salpêtrière, Paris, France, 36 University of Aberdeen, Division of Applied Health Sciences, Population Health Section, Aberdeen, United Kingdom, 37 Centre Hospitalier Universitaire Nantes, CIC0004, Service de Neurologie, Nantes, France, 38 Institut National de la Sante et de la Recherche Medicale, U997, Université Victor Segalen, Bordeaux, France, 39 WellcomeTrust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom, 40 Klinik für Neurologie, Universitätssklinikum Schleswig-Holstein, Campus Kiel, Christian-Albrechts-Universität Kiel, Kiel, Germany, 41 Parkinson’s Disease Research Group, Faculty of Medicine, Imperial College London, London, United Kingdom, 42 Service de Neurologie, Hôpital Gabriel Montpied, Clermont-Ferrand, France, 43 AP-HP, Pitié-Salpêtrière Hospital, Department of Genetics and Cytogenetics, Paris, France, 44 Cambridge Centre for Brain Repair, University of Cambridge, Cambridge, United Kingdom, 45 Institute of Neurology, University College London, London, United Kingdom, 46 Department of Psychiatry, Department of Neurology, Washington University School of Medicine, St. Louis, Missouri, United States of America, 47 14 deCODE genetics, Reykjavik, Iceland, 48 Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands, 49 Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands, 50 American Association of Retired Persons, Washington DC, United States of America, 51 Queen Square Brain Bank for Neurological Disorders, Institute of Neurology, University College London, London, United Kingdom, 52 Department of Clinical Neurology, John Radcliffe Hospital, Oxford, United Kingdom, 53 Department of Neurology, Radboud University Nijmegen, Netherlands, 54 S. Hershey Medical Center, Hershey, Pennsylvania, United States of America, 55 German Research Centre for Environmental Health, Neuherberg, Germany, 56 Department of Geriatrics, Landsportslag University Hospital, Reykjavik, Iceland, 56 Institute of Human Genetics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany, 57 Sobell Department, Unit of Functional Neurosurgery, University College London Institute of Neurology, London, United Kingdom, 58 Section on Molecular Neurogenetics, Medical Genetics Branch, NHGRI, National Institutes of Health, Bethesda, Maryland, United States of America, 59 Medical Research Council Centre for Neuropsychiatry Genetics and Genomics, Cardiff University School of Medicine, Cardiff, United Kingdom, 60 Neurosciences Department, Questron, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom, 61 Neurogenetics Unit, University College London, Institute of Neurology/National Hospital for Neurology and Neurosurgery, London, United Kingdom, 62 Service de Neurologie, Centre Hospitalier Universitaire de Grenoble, Grenoble, France, 63 Translational Neurology, Biogen Idec, Cambridge, Massachusetts, United States of America, 64 Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands, 65 University of Cambridge, Department of Clinical Neurosciences, Addenbrooke’s Hospital, Cambridge, United Kingdom, 66 Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, 67 Department of Neurology, University of Rochester, Rochester, New York, United States of America, 68 Department of Pathology, University of Edinburgh, Edinburgh, United Kingdom, 69 University of Oxford, Department of Clinical Neurology, John Radcliffe Hospital, Oxford, United Kingdom, 70 Clinical Research Department, The Parkinson’s Institute and Clinical Center, Sunnyvale, California, United States of America, 71 Service de Neurologie, Hôpital Haut-Lévêque, Pessac, France, 72 Department of Medical and Molecular Genetics, King’s College London, London, United Kingdom, 73 Department of Neurology, Cardiff University, Cardiff, United Kingdom, 74 Department of Psychiatry and Psychotherapy, Wellcome Trust Behavioural and Clinical Neurosciences Institute, University of Cambridge, Cambridge, United Kingdom.

Author Contributions


References


PLoS Genetics | www.plosgenetics.org 8 June 2011 | Volume 7 | Issue 6 | e1002142

8