

A pathway-based analysis provides additional support for an immune-related genetic susceptibility to Parkinson's disease

Peter Holmans¹, Valentina Moskvina¹, Lesley Jones¹, Manu Sharma^{4,10}, The International Parkinson's Disease Genomics Consortium (IPDGC)[†], Alexey Vedernikov¹, Finja Buchel³, Mohamad Sadd⁵, Jose M. Bras⁶, Francesco Bettella⁸, Nayia Nicolaou⁹, Javier Simón-Sánchez⁹, Florian Mittag³, J. Raphael Gibbs^{2,6}, Claudia Schulte^{4,10}, Alexandra Durr^{11,12}, Rita Guerreiro⁶, Dena Hernandez^{2,6}, Alexis Brice^{11,12,13,14}, Hreinn Stefánsson⁸, Kari Majamaa¹⁵, Thomas Gasser^{4,10}, Peter Heutink⁹, Nicholas W. Wood^{6,7}, Maria Martinez⁵, Andrew B. Singleton², Michael A. Nalls², John Hardy⁶, Huw R. Morris¹ and Nigel M. Williams^{1,*}

¹Department of Psychological Medicine and Neurology, Institute of Psychological Medicine and Clinical Neurosciences, MRC Centre in Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff CF14 4XN, UK, ²Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, ³Center for Bioinformatics Tuebingen (ZBIT) and ⁴Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany, ⁵Paul Sabatier University, Toulouse, France, ⁶Department of Molecular Neuroscience, Institute of Neurology and ⁷UCL Genetics Institute, University College London, London, UK, ⁸deCODE genetics, Scientific Services, Sturlugata, 8 IS-101 Reykjavik, Iceland, ⁹Section of Medical Genomics, Department of Clinical Genetics, VU University Medical Centre, Amsterdam, The Netherlands, ¹⁰DZNE – Deutsches Zentrum für Neurodegenerative Erkrankungen (German Center for Neurodegenerative Diseases), Tuebingen, Germany, ¹¹Inserm, UMRS975, CNRS UMR 7225, CRICM, F-75013 Paris, France, ¹²UPMC Univ Paris 06, UMRS975, F-75013, Paris, France, ¹³AP-HP, Hôpital de la Salpêtrière, Département de Génétique, Paris, France, ¹⁴Institut du Cerveau et de la Moelle Epinière, F-75013, Paris, France and ¹⁵Department of Medical Biochemistry and Molecular Biology, FIN-90014 University of Oulu, Finland

Received August 10, 2012; Revised October 31, 2012; Accepted November 16, 2012

Parkinson's disease (PD) is the second most common neurodegenerative disease affecting 1–2% in people >60 and 3–4% in people >80. Genome-wide association (GWA) studies have now implicated significant evidence for association in at least 18 genomic regions. We have studied a large PD-meta analysis and identified a significant excess of SNPs ($P < 1 \times 10^{-16}$) that are associated with PD but fall short of the genome-wide significance threshold. This result was independent of variants at the 18 previously implicated regions and implies the presence of additional polygenic risk alleles. To understand how these loci increase risk of PD, we applied a pathway-based analysis, testing for biological functions that were significantly enriched for genes containing variants associated with PD. Analysing two independent GWA studies, we identified that both had a significant excess in the number of functional categories enriched for PD-associated genes (minimum $P = 0.014$ and $P = 0.006$, respectively). Moreover, 58 categories were significantly enriched for associated genes in both GWA studies ($P < 0.001$), implicating genes involved in the 'regulation of leucocyte/lymphocyte activity' and also 'cytokine-mediated signalling' as conferring an increased susceptibility to PD. These results

*To whom correspondence should be addressed at: Department of Psychological Medicine and Neurology, Institute of Psychological Medicine and Clinical Neurosciences, MRC Centre in Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, UK. Tel: +44 2920687070; Email: williamsnm@cf.ac.uk

[†]A full list of the IPDGC members and affiliations appears in the appendix.

were unaltered by the exclusion of all 178 genes that were present at the 18 genomic regions previously reported to be strongly associated with PD (including the *HLA* locus). Our findings, therefore, provide independent support to the strong association signal at the *HLA* locus and imply that the immune-related genetic susceptibility to PD is likely to be more widespread in the genome than previously appreciated.

INTRODUCTION

Parkinson's disease (PD) is the second most common neurodegenerative disease and is characterized by bradykinesia with resting tremor, stiffness and gait disturbance. Pathologically, the disease involves the deposition of alpha-synuclein as Lewy bodies and Lewy neurites in multiple motor and non-motor brain areas. The population prevalence is estimated at 0.3%, but this increases with age—rising to 1–2% in people >60 and 3–4% in people >80 years of age (1). Although the average age at onset is 68 years old, the onset is extremely variable, ranging from adolescence to old age (2,3).

The identification of rare highly penetrant mutations in genes causing familial PD (4–8) has had a considerable impact on our understanding of the pathogenesis of this complex and common disorder. Studies of rare Mendelian forms of PD show that increased alpha-synuclein fibril formation and abnormalities of mitophagy are of central importance in at least some forms of the disease. More recently, our understanding of the idiopathic form of the illness has been greatly enhanced by a number of large genome-wide association (GWA) studies (9–19). As well as providing unequivocal evidence that *SNCA* and *MAPT* are also risk factors for idiopathic PD (9,12,17–19), these studies have collectively identified variants at over 18 loci that significantly increase risk for PD (9,10). The challenge now is to understand how these genetic loci influence the risk of PD and how their cognate biological functions are influenced by variation at these loci. It is, however, inevitable that the findings of the primary GWA analyses will be limited to a subset of the variants of the strongest genetic effect and, consequently, a large number of true polygenic risk alleles that generate weaker evidence for association may be overlooked. Thus, despite their high levels of significance, the current crop of loci that are strongly associated with PD are thought to account for only a very small amount (1–2%) of the expected heritability of PD (20). As in other complex disorders, the missing heritability in PD is expected to relate to a combination of some variants that yield only relatively weak association signals and others that are not tagged by the genotyping microarray. Given that there is a substantial increase in the estimated heritability detected in PD GWA studies (24%) when weak effect loci are also considered (20), this strongly implies that a large proportion of genetic signal must lie below the genome-wide significance thresholds set in the primary analyses.

PD risk alleles are unlikely to be randomly distributed among genes, but instead are more likely to be distributed among genes whose functions are related. Indeed, the relevance of mitochondrial dysfunction (21–24) and protein degradation pathways (25–28) to PD pathogenesis has been highlighted by converging functional studies of Mendelian genes implicated in

familial PD. In applying such a model to GWA data, one would expect to see an overall excess of association signal in a series of SNPs selected from genes which code for functionally related proteins and it might be anticipated that at least some of these shared biological functions relate to the current known Mendelian and non-Mendelian loci. Such effects have been identified in other complex disorders both when the analysis was limited to genes containing robustly associated SNPs (29–32) and also when the analysis was expanded to include genes containing SNPs whose evidence for association fell short of stringent genome-wide significant thresholds (30,32,33). To date, biological pathway-based analysis in PD has been limited to relatively small GWA data sets, and has implicated the axon guidance pathway as being relevant to PD (34), a finding that was recently replicated in a small independent data set (35).

In this study, we have used the ALIGATOR algorithm (32) to analyse SNPs associated with PD in two large independent GWA studies for enrichment in pre-defined sets of functionally related genes. Importantly, we have also confirmed our results using gene set enrichment analysis and established that our findings are not dependent on the most strongly associated loci. Our study implicates specific biological categories of genes that show evidence for association with PD and is an important next step in our translation of genetic susceptibility to an understanding of disease pathogenesis.

RESULTS

Assessment of SNPs passing sub-genome-wide significance thresholds

Analyses of the large International Parkinson's Disease Genetics Consortium (IPDGC) meta analysis consisting of 5333 PD cases and 12 219 controls have previously identified 18 regions showing evidence for genome-wide significant or suggestive evidence for association (9–12). In this study, we first set out to establish whether this data set harboured further evidence for additional PD genes. Our analysis revealed that a significant excess of SNPs surpassing different thresholds of significance remained even after all 18 previously implicated regions were excluded (Table 1). At our most stringent threshold ($P < 0.0001$), we identified a 4.2-fold excess of independent SNPs [following correction for linkage disequilibrium (LD)] significantly associated with PD than were expected (679.2 compared with 161.7, $P < 1 \times 10^{-16}$). This suggested the existence of genuine additional PD susceptibility loci that failed to meet the stringent significance thresholds that were set in the previous studies.

Table 1. Assessment of SNPs passing sub-genome-wide significance thresholds

	Significance threshold			
	0.05	0.01	0.001	0.0001
Expected <i>N</i> of independent significant tests ^a	69 362.2	14 119	1443.9	161.7
Estimated <i>N</i> of independent associated tests (observed) ^a	82 692	19 643.4	2907.9	679.2
<i>P</i> -value	$P < 1 \times 10^{-16}$	$P < 1 \times 10^{-16}$	$P < 1 \times 10^{-16}$	$P < 1 \times 10^{-16}$
Ratio observed/expected	1.2	1.4	2	4.2

Only the imputed genotypes for SNPs present in 1000 Genomes data (9) were assessed following the exclusion of 38 932 SNPs at 18 regions strongly associated with PD ($n = 4\,921\,315$ SNPs).

^aData have been corrected for LD.

ALIGATOR analysis

As it is not possible to implicate any individual locus from the observed enrichment of sub-GWA signal, we next chose to assess whether any biological functional categories were significantly enriched for genes carrying SNPs that were nominally associated with PD. To achieve this, we a priori considered the IPDGC meta-analysis (9) as two large independent PD-GWA data sets: UK-GWA (1705 PD patients and 5200 controls) and Meta-GWA (3628 PD patients and 7019 controls) and applied ALIGATOR (32) analysis to the top 5% of associated genes ($n = 1050$), which captured genes containing at least one SNP surpassing $P = 0.00205$ and $P = 0.0024$ in the Meta-GWA and UK-GWA studies, respectively. In each GWA study, ALIGATOR was then used to assess the functional categories that were enriched for these 1050 genes at varying thresholds of nominal significance ($P = 0.05$, $P = 0.01$, $P = 0.001$). This revealed a significant excess in the number of enriched functional categories compared with the simulated data, in both the Meta-GWA (minimum $P = 0.014$) and UK-GWA studies (minimum $P = 0.006$) (Table 2). In the largest data set (Meta-GWA), our most significant observation was of 498 functional categories that were enriched at $P < 0.05$ for associated genes ($P = 0.014$).

In order to further investigate the nature of this enrichment, we next assessed the level of overlap between the two studies by running ALIGATOR in the UK-GWA study, but this time restricting the analysis to only the 498 functional categories enriched at $P < 0.05$ in the Meta-GWA study. This revealed a significant excess of categories enriched in both GWA study, regardless of the threshold used (Table 3). The most significant evidence implicated 58 categories enriched for associated genes at $P < 0.05$ in both Meta-GWA and UK-GWA ($P < 0.001$) (Table 3).

These 58 functional categories contained a total of 269 different genes that were associated with PD (Supplementary Material, Table S1). However, given that a single gene can be present in multiple categories, grouping together categories based on their gene membership enabled us to further condense the 58 categories into 24 functionally related groups (Supplementary Material, Table S2), such that each category within a group shared at least 50% of its genes in common with at least one other category in the group. This revealed that 26 categories (45%) were present in just two functionally related groups which, importantly, also contained 9 of the 10 categories that showed the strongest evidence for enrichment in PD in both GWA studies ($P < 0.01$ in both). The categories

included in these two biological groups implicate genes involved in the ‘regulation of leucocyte/lymphocyte activity’ and also ‘cytokine-mediated signalling’ as conferring an increased susceptibility to PD (Table 4). In total, these two functional groups contained 132 (49%) of the 269 genes associated with PD (Supplementary Material, Table S1); however, as 38 genes are shared by both, it reflects the biological relationship between these groups.

It is plausible that our results are being artificially biased by genes whose evidence for association is merely a consequence of LD with the very strong association signal of known GWA study hits. To investigate this possibility, we repeated our analysis using identical gene-wide significance thresholds, but this time excluding all 178 genes that were present at 18 genomic regions previously reported to be strongly associated with PD, including the *HLA* locus. Given that each of these regions is likely to span at least one true PD susceptibility gene which would now be excluded from our ALIGATOR analysis, this approach is highly conservative. Nevertheless, this analysis again revealed a significant excess in the number of enriched functional categories compared with the simulated data, in both the Meta-GWA (minimum $P = 0.016$) and UK-GWA studies (minimum $P = 0.022$) (Table 2), as well as significant evidence in favour of a common set of functional categories that were enriched ($P < 0.05$) in both GWA studies ($P < 0.001$) (Table 3). This suggests that our findings are not dependent on either the previously identified susceptibility loci or the genes that are falsely associated with PD merely as a consequence of LD with the very strong association signals. Enrichment *P*-values in the absence of the 178 genes from the 18 previously associated regions are shown in Supplementary Material, Table S3 for each of the 58 categories significantly enriched in both GWA studies in the original analysis. It can be seen that, in general, the category-specific enrichment *P*-values are essentially unaltered by removal of the PD regions. To further confirm that the excess in significantly enriched functional categories is not a result of LD with strong signals, we repeated the analysis excluding not only the 18 genomic regions but also the entire MHC region (25–35 Mb on chromosome 6) and any gene within 1 Mb of an SNP that showed genome-wide significance for association ($P < 5 \times 10^{-8}$) in the IPDGC meta-analysis (a total of 1058 genes). This made little difference to the significant excess of enriched functional categories in the Meta-GWA and UK-GWA samples analysed separately (Table 2) or the significance of the overlap in categories enriched in both

Table 2. ALIGATOR analysis of two independent GWA studies

Sample	Functional categories enriched at $P < 0.05$		$P < 0.01$		$P < 0.001$	
	<i>N</i> categories	<i>P</i> -value	<i>N</i> categories	<i>P</i> -value	<i>N</i> categories	<i>P</i> -value
Full data set						
Meta-GWA	498	0.014	116	0.016	16	0.058
UK-GWA	416	0.091	108	0.037	25	0.006
After excluding known GWA study hits						
Meta-GWA	495	0.017	122	0.016	12	0.069
UK-GWA	430	0.064	100	0.049	19	0.022
After excluding the MHC region and all genes within 1 Mb of a genome-wide significant SNP						
Meta-GWA	513	0.011	129	0.01	13	0.105
UK-GWA	430	0.054	93	0.065	16	0.057

The top 5% of associated genes ($n = 1050$) were considered to be significant. Categories require two or more significant genes to be considered for enrichment.

studies (Table 3). Enrichment P -values under this analysis are given in Supplementary Material, Table S3 for each of the 58 categories significantly enriched in both GWA studies. These are very similar to those obtained in the original analysis.

Gene-set enrichment analysis

Category-specific GSEA (gene-set enrichment analysis) P -values are given in Table 4 for each of the 26 categories implicating genes involved in the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling, whereas the results for all 58 categories significantly enriched for associated genes are given in Supplementary Material, Table S2. As some of the less-significant categories identified by ALIGATOR do not have significant GSEA P -values, it suggests that some of the enrichment may be the result of the threshold used to define significant genes. Despite this, a closer inspection of the GSEA analysis reveals that 34 (59%) and 16 (28%) of the 58 categories also have significant GSEA P -values in the Meta-GWA and UK-GWA samples, respectively, with 14 (24%) being significant in both GWA studies. Moreover, of the 14 categories that are significant in both GWA studies, 10 (70%) are from the two biological groups that we have termed the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling, increasing confidence that the enrichment of PD-associated genes seen in these categories is genuine.

DISCUSSION

In this study, we have tested whether SNPs associated with PD in two large independent GWA studies are enriched in pre-defined sets of functionally related genes. By restricting our analysis to the functional categories defined in five publicly available ontology databases, it is inevitable that any biological functions that are not well curated by these sources will not have been analysed by our study. Despite this limitation, our analysis of two independent GWA studies has identified significant evidence that a series of functional categories were enriched for genes associated with PD, and that a subset of 58 categories was significantly enriched in both GWA studies. This significant overlap was further condensed to just 24 biological groups, of which the two with the strongest

evidence for association highlight processes involving the immune system. Based on their proposed functions, we have termed these two biological groups the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling. These findings imply that genetic variation in genes related to these biological processes has a role in increasing susceptibility to PD and is, therefore, a potential mechanism that should be the subject of further detailed genetic and functional analyses. Although in this study we were unable to replicate the previous reports of an enrichment of association signal in genes involved in axon guidance (34,35), it is worth noting that the T-cell receptor signalling pathway [Kyoto Encyclopedia of Genes and Genomes (KEGG) 04660] has previously been implicated in PD pathology by pathway analysis of GWA data (35). Clearly, the validity of our analysis is dependent on the quality with which biological categories are annotated and that this is likely to be variable. For this reason, we simultaneously analysed annotations from multiple ontology databases and as a result it is encouraging that our two strongest biological groups which contain genes involved in the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling are implicated by multiple functional categories independently curated in different ontology databases. We also obtained largely analogous results, using both ALIGATOR and GSEA analyses, implying that the biological groups implicated in this study are unlikely to be biased by our method of analysis. Moreover, as the application of ALIGATOR to GWA studies of different diseases has previously implicated biological groups that were different to those identified in this study (32,33), it suggests that the functional categories that we find enriched for genes associated with PD are not biased in favour of those that have been most comprehensively annotated.

It is important to note that this finding does not imply that all nominally significant genes in an enriched functional category are true susceptibility genes. Biological pathway-based analyses assume that variation in functionally related genes is more likely to be important in disease, so any functional category implicated in our data is more likely to be relevant to PD aetiology since it contains a significant excess of genes carrying SNPs associated with PD. With this in mind, it is worth noting that although we have demonstrated that there was a significant overlap of 58 categories enriched for genes

Table 3. ALIGATOR analysis of categories enriched in both GWA studies

Discovery sample	Replication sample	Functional categories enriched at $P < 0.05$		$P < 0.01$		$P < 0.001$	
		N categories	P -value	N categories	P -value	N categories	P -value
Full data set							
Meta-GWA	UK-GWA ^a	58	<0.001	18	0.004	6	0.004
After excluding known GWA study hits							
Meta-GWA	UK-GWA ^a	54	<0.001	13	0.007	4	0.005
After excluding the MHC region and all genes within 1 Mb of a genome-wide significant SNP							
Meta-GWA	UK-GWA ^a	62	<0.001	16	<0.001	9	<0.001

^aUK-GWA was analysed using only the 498 functional categories enriched at $P < 0.05$ in Meta-GWA.

associated with PD from two independent GWA studies, the signal from the two studies in each functional category does not necessarily reflect the same set of genes. This observation is, however, plausible when one considers that this study focused on susceptibility alleles of small effect which are likely to be carried by affected individuals in different genes in the same functional category.

Our findings are not dependent on the genes present at loci that have been previously reported as being strongly associated with PD and their exclusion has minimal effect on the pathways identified in our analysis. Nevertheless, it might be expected that the functional pathways implicated by our analysis would be supported by a framework of genes identified through Mendelian forms of the illness or as genome-wide significant by GWA. Our data show no strong evidence for such a pattern, with only 5 of the 18 loci implicated by GWA studies harbouring genes (a total of 10 genes) that are included in the biological categories related to the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling. Moreover, SNCA is the only gene implicated in familial PD that is also included in the two biological groups implicated in our study. This observation is similar to the findings in other complex diseases (30,32,33) and demonstrates that the aggregating power of testing whole pathways can generate useful biological insights from association signals that are below the thresholds set in GWA studies. Together with our observation that a significant excess of SNPs below the genome-wide significance threshold are present in the IPDGC GWA study (9), this implies that at least some 'missing' disease heritability is present at loci that the current GWA studies do not have the power to detect.

The strongest findings from our study show a striking enrichment of genes associated with PD being involved in biological processes related to the immune system. Specifically, these were identified as two distinct but overlapping biological groups related to the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling. Despite PD not typically being considered an immune disease, there are emerging data that are starting to demonstrate a relationship between PD progression and the immune response. First, in the degenerating PD brain, the activation of microglia and monocytes contributes to the blood-brain barrier becoming compromised, and this allows toxins and infections to reach the CNS (36). There is also evidence that altered immune components are coupled with PD progression. For example, CD4 (37)- and CD3-positive (38) T-cells have been found in the substantia

nigra and Lewy body lesions of PD patients, respectively, and several inflammatory cytokines such as IL1-beta, IL8, IL6, IL4, TNF-alpha, IFN-gamma are up-regulated in the CSF and sera of PD patients (39,40). LRRK2 has also been shown to regulate B2-lymphocyte function (41), potentially providing a direct link between cell-mediated immunity and PD pathology. Although these studies are typically interpreted as detecting the secondary effects of neurodegeneration in PD, they have recently been complemented by strong evidence for genetic association between SNPs at the *HLA* locus and PD (9,11). These genetic findings indicate that the immune system is likely to be aetiologically important in PD and not simply activated as a secondary response to neurodegeneration. It is also possible that future studies of larger GWA data sets will reveal additional biological mechanisms that are associated with PD.

Our findings, therefore, provide independent support to the strong association signal at the *HLA* locus which spans at least 10 immune-related genes. It is worth noting that five genes from this locus (*HLA-DQA1*, *HLA-DQB1*, *HLA-DRA*, *HLA-DRB1*, *HLA-DRB5*) were included in the biological groups, implicating the regulation of leucocyte/lymphocyte activity and cytokine-mediated signalling. However, the widespread LD at this locus is unlikely to have adversely influenced our results as we classed the *HLA* locus as a single signal in our analysis and it only counted once towards any pathway. Moreover, as previously mentioned, we obtained analogous results when previously implicated regions were excluded, demonstrating that our findings are not dependent on markers at the *HLA* locus. Our findings imply that the immune-related genetic susceptibility to PD is likely to be more widespread in the genome than previously appreciated. By implicating genes involved in the regulation of leucocyte/lymphocyte activity, our study indicates that components of the adaptive immune system have a role in PD pathology. However, given the large overlap in gene content with the cytokine-mediated signalling pathway which is itself related to both the innate and the adaptive immune responses as well as to inflammatory processes, the exact nature of any involvement remains to be elucidated. Future work will, therefore, be required to dissect out the exact involvement of the innate and adaptive immune system in PD. Functional and epidemiological analyses will be required to investigate whether genetic variants in these immune-related pathways influence disease progression, although the quantitative analysis of cytokine levels in peripheral blood could potentially be an avenue

Table 4. Functional categories significantly enriched for genes associated with PD in both GWA studies

Biological group	Functional category Source:ID	Meta-GWA				ALIGATOR P-value	GSEA P-value	UK-GWA			Proposed biological function of functional category
		N genes	N Significant genes observed	N significant genes expected	ALIGATOR P-value			N Significant genes observed	N significant genes expected	ALIGATOR P-value	
1	GO:51251	190	21	9.16	0.0002	0.0138	18	9.29	0.0028	0.0538	Positive regulation of lymphocyte activation
1	GO:2696	208	22	9.89	0.0004	0.0032	19	10.03	0.003	0.0088	Positive regulation of leucocyte activation
1	GO:50867	212	22	10.09	0.0004	0.0032	19	10.23	0.0048	0.0102	Positive regulation of cell activation
1	GO:51249	251	23	11.57	0.0006	0.0162	23	11.6	0.0004	0.0738	Regulation of lymphocyte activation
1	GO:2694	288	26	14.36	0.0014	0.0004	28	14.26	<0.0001	0.002	Regulation of leucocyte activation
1	GO:50865	300	27	15.02	0.0016	0.0004	28	14.89	0.0002	0.005	Regulation of cell activation
1	MGI:8661	16	3	0.33	0.0028	0.0312	2	0.33	0.0368	0.0166	Decreased interleukin-10 secretion
1	GO:50870	153	15	7.37	0.0042	0.0128	13	7.49	0.0272	0.072	Positive regulation of T-cell activation
1	MGI:5027	60	6	1.77	0.006	0.0432	6	1.81	0.0068	0.0114	Increased susceptibility to parasitic infection
1	MGI:8659	25	4	0.88	0.009	0.01	3	0.84	0.0446	0.034	Abnormal interleukin-10 secretion
1	GO:50863	197	16	9.1	0.0154	0.0102	17	9.18	0.007	0.0336	Regulation of T-cell activation
1	GO:50671	84	9	4.08	0.0166	0.2736	8	4.17	0.043	0.2526	Positive regulation of lymphocyte proliferation
1	GO:32946	85	9	4.1	0.0172	0.2726	8	4.19	0.045	0.192	Positive regulation of mononuclear cell proliferation
1	MGI:2406	256	17	10.3	0.02	0.4188	25	10.3	<0.0001	0.3384	Increased susceptibility to infection
1	KEGG:4672	45	4	1.31	0.021	0.006	4	1.31	0.0276	0.0002	Intestinal immune network for IgA production
1	GO:50864	72	7	3.11	0.03	0.5206	8	3.1	0.0084	0.23	Regulation of B-cell activation
1	GO:50670	118	10	5.28	0.0332	0.1768	11	5.38	0.0122	0.2292	Regulation of lymphocyte proliferation
1	GO:32944	119	10	5.29	0.034	0.1774	11	5.4	0.0122	0.1922	Regulation of mononuclear cell proliferation
1	GO:70663	122	10	5.4	0.0388	0.1742	11	5.52	0.0146	0.219	Regulation of leucocyte proliferation
1	GO:30888	44	5	2.09	0.0444	0.5298	5	2.11	0.0454	0.214	Regulation of B-cell proliferation
2	GO:19221	236	19	9.69	0.0016	0.0012	19	9.69	0.0018	0.03	Cytokine-mediated signalling pathway
2	GO:71345	285	20	11.08	0.005	0.0056	20	11.08	0.004	0.1142	Cellular response to cytokine stimulus
2	GO:1637	24	2	0.31	0.0342	0.385	2	0.31	0.0344	0.2646	G-protein-coupled chemoattractant receptor activity
2	GO:4950	24	2	0.31	0.0342	0.385	2	0.31	0.0344	0.2646	Chemokine receptor activity
2	GO:4896	72	7	3.42	0.0388	0.619	7	3.36	0.0442	0.3376	Cytokine receptor activity
2	KEGG:4630	146	10	5.46	0.039	0.0568	13	5.4	0.0022	0.049	Jak-STAT signalling pathway

Proposed biological functions were provided by the respective source of each ontology database. Biological groups indicate categories which share >50% of genes with at least one other category in the group.

for the investigation of biomarkers that increase risk to PD. Our findings do, however, contrast with those of a previous study which applied the same analytical strategy to GWA studies in Alzheimer's disease and specifically highlighted the involvement of the complement and innate immune system in the disease (33). Given the importance of the immune system, it is perhaps unsurprising that it has an impact upon the presentation of both of these common late-onset diseases. However, despite implicating the involvement of immune-related genes in both Alzheimer's disease and PD, different functional categories are highlighted by each study, suggesting that the increased risk to disease conferred by the immune system is likely to play a specific role in different neurodegenerations, and is unlikely to be a common degenerating effector end-pathway.

MATERIALS AND METHODS

Data summary

The GWA studies used in this study were performed as described (13,15,17–19,42). We have previously studied a large meta-analysis composed of these studies for allelic association (9). However, in order to provide independent discovery and replication samples, in this study we divided the meta-analysis of the IPDGC (9) into two large independent PD-GWA data sets: (i) the UK-based GWA study of Spencer *et al.* (17) (referred from here on as UK-GWA), which was composed of 1705 PD patients (56.7% male) and 5200 controls (49.5% male) and (ii) a trimmed meta-analysis composed of 3628 PD patients (59.5% male) and 7019 controls (52.7% male) which excluded the UK-GWA study and was, therefore, composed of the remaining four GWA studies from USA National Institute on Aging, Germany, France and the USA database of genotypes and phenotypes (13,15,18,19,42) (referred from here on as Meta-GWA). Study-specific quality control and genotype imputation using MACH (version 1.0.16) has been previously described for these data sets (9), which resulted in a total of 5 108 844 and 5 096 499 successfully imputed SNPs being included in our analyses of the Meta-GWA and UK-GWA studies, respectively. Notably, each of the GWA studies included in the meta-analysis had previously been corrected for population stratification, as detailed in Nalls *et al.* (9).

Statistical analysis

Excess of SNPs passing sub-genome-wide significance thresholds

Genomic regions harbouring known PD GWA signals were defined according to the LD structure of the region surrounding the SNP showing strongest evidence for association. This was performed for each of the 18 loci that had been previously reported to be strongly associated with PD (9–12). The genomic coordinates of these regions (genome build 36.1) were chr1: 153 274 689–154 282 163; chr2: 135 307 258–135 502 264; chr2: 168 509 679–168 847 014; chr3: 184 135 009–184 371 343; chr4: 812 847–991 302; chr4: 15 073 141–15 467 459; chr4: 90 755 160–91 514 165; chr6: 32 313 388–32 996 301; chr12: 38 433 501–39 059 770; chr12: 121

576 087–121 958 509; chr17: 40 640 223–41 828 247; chr1: 203 921 882–204 058 936; chr3: 161 856 058–162 852 739; chr4: 77 299 726–77 935 917; chr7: 23 067 304–23 484 157; chr8: 16 687 436–17 032 678; chr8: 89 060 020–89 833 080; chr16: 30 445 872–31 242 830. In total, there were 38 932 SNPs in these regions. The number of independent SNPs (N) in the whole genome excluding all spanned by genomic regions harbouring known PD GWA signals was estimated by the method of Moskvina and Schmidt (43), as were the observed number of independent SNPs significant at each P -value criterion. In the absence of excess association, the expected number of independent SNPs significant at level ' α ' is distributed as a binomial (N, α).

Pathway analyses

Assignment of SNPs to genes. SNPs that mapped to within the longest transcript of a gene (genome build 36.1) were assigned to that gene: if SNPs mapped within more than one gene, all such genes were included. Using this approach, we were able to map 1 945 730 SNPs to 21 013 genes in the Meta-GWA study and 1 941 147 SNPs to 21 004 genes in the UK-GWA study. The P -value of the SNP showing the strongest evidence for allelic association was then taken as the gene-wide significance measure.

Assignment of genes to functional categories. We next assigned genes to a series of functional categories defined by five independent sources: (i) Gene Ontology (GO) (44) (<http://www.geneontology.org/>, downloaded 8 November 2011), (ii) KEGG (45) (<http://www.genome.jp/kegg/>, downloaded 27 June 2011), (iii) PANTHER (<http://www.pantherdb.org/pathway/>, accessed on 20 August 2010), (iv) the 'canonical pathways' collection from the Molecular Signatures Database v3.0 (MsigDB) (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>, accessed on 2 February 2011), (v) the Mouse Genome Informatics (MGI) database (46) (<http://www.informatics.jax.org/>, accessed on 22 February 2010). The MGI contains a comprehensive catalogue of behavioural, physiological and anatomical phenotypes observed in mutant mice, and as previously described (33), we extracted phenotype data for single-gene studies (excluding all transgenes) and converted mouse genes to their human orthologues, using the MGI's mouse/human orthology assignment. This allowed us to map 6297 different phenotypic annotation terms to 5671 human genes. We restricted our analysis to a total of 15 381 functional categories containing between 3 and 300 genes: 9164 in GO, 231 in KEGG, 542 in PANTHER, 687 in MsigDB and 4757 in MGI categories. Moreover, to remove the possibility of a small category being deemed significantly enriched based on just one signal, we only classed categories as being enriched if they carried at least two signals.

ALIGATOR analysis. ALIGATOR was then used to test the list of gene-wide significance measures for enrichment within functional categories as previously described in Holmans *et al.* (32). Briefly, ALIGATOR defines a list of significant genes as those genes containing at least one SNP surpassing a pre-defined significance criterion, and comparing the number of such genes in each category with that observed in 5000 randomly generated gene lists of the same length. This procedure gives a measure of the enrichment significance for

each category. Multiple testing correction for multiple non-independent categories is carried out by a bootstrap procedure, wherein one random gene list is selected as the 'observed' data and is compared with a set of 5000 gene lists selected randomly (with replacement) from the remainder, with the process being repeated 1000 times. This procedure also gives a test of whether more categories achieve a given level of enrichment significance in the real data than would be expected by chance. Further details are given in Holmans *et al.* (32). Unlike methods designed for gene-expression data (where there is typically only one measurement per gene), ALIGATOR corrects for variable numbers of SNPs per gene. This is done by generating the random gene lists by sampling SNPs (not genes) at random, adding the gene(s) the SNP lies in to the list. This ensures that the random gene lists account for large genes containing many SNPs being more likely to contain significant SNPs by chance. Each gene was counted once regardless of how many significant SNPs it contains, thus eliminating the influence of LD between SNPs within genes. To prevent the analysis being biased by SNPs located in multiple functionally related genes which physically overlapped, we set a restriction that limited each SNP from contributing to more than one gene in any single category. Moreover, to address the potential of multiple significant genes that are close together reflecting the same association signal due to LD, we conservatively grouped significant genes that were <1 Mb apart and located in the same functional category into one signal. Replicate gene lists of the same length as the original were generated by randomly sampling SNPs (thus correcting for variable gene size). The lists were used to obtain *P*-values for enrichment for each category and to correct these for testing multiple non-independent categories, and also to test whether the number of significantly enriched categories is higher than expected. To assess the potential of any bias caused by LD with strong association signals that had been previously identified in these samples, we also performed ALIGATOR analysis after excluding all genes ($n = 178$) present within the list of genomic regions harbouring known PD GWA signals.

Gene-set enrichment analysis

As a further validation of the ALIGATOR results, and to show that the results of our analyses are not driven by the choice of *P*-value cut-off for defining significant genes, GSEA was performed using the method described in Wang *et al.* (31). Rather than defining a list of significant genes, GSEA ranks all genes in order of a gene-wide association statistic and tests whether the genes in a particular gene set have higher rank overall than would be expected by chance. Following Wang *et al.* (31) in order to allow for varying numbers of SNPs per gene, the gene-wide statistic used was the Simes-corrected single-SNP *P*-value (47).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We kindly acknowledge the patients who participated and the physicians who helped in recruitment. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by Parkinson's UK (formerly The PD society) (ref no K0906). Additionally, part of the study was undertaken at UCLH/UCL, using funding through a Department of Health NIHR Biomedical Research Centre. This work was also supported by Parkinson's UK (grants 8047 and J-0804) and the Medical Research Council (G0700943). The German work was also supported by the German National Genome Network (NGFNplus #01GS08134; German Ministry for Education and Research). This work was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project numbers Z01 AG000949-06 and Z01 AG000950-10. The French GWA scan work was supported by the French National Agency of Research (<http://www.agence-nationale-recherche.fr>, ANR-08-MNP-012) and by the National Research Funding Agency (ANR-08-NEUR-004-01) in the ERA-NET NEURON framework (<http://www.neuron-eranet.eu>). We also want to thank the Hersenstichting Nederland (<http://www.hersenstichting.nl>), the Neuroscience Campus Amsterdam and the section of Medical genomics, the Prinses Beatrix Fonds (<http://www.prinsesbeatrixfonds.nl>) for sponsoring this work.

REFERENCES

- Nussbaum, R.L. and Ellis, C.E. (2003) Alzheimer's disease and Parkinson's disease. *N. Engl. J. Med.*, **348**, 1356–1364.
- Wickremaratchi, M.M., Ben-Shlomo, Y. and Morris, H.R. (2009) The effect of onset age on the clinical features of Parkinson's disease. *Eur. J. Neurol.*, **16**, 450–456.
- Van Den Eeden, S.K., Tanner, C.M., Bernstein, A.L., Fross, R.D., Leimpeper, A., Bloch, D.A. and Nelson, L.M. (2003) Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. *Am. J. Epidemiol.*, **157**, 1015–1022.
- Gilks, W.P., Abou-Sleiman, P.M., Gandhi, S., Jain, S., Singleton, A., Lees, A.J., Shaw, K., Bhatia, K.P., Bonifati, V., Quinlan, N.P. *et al.* (2005) A common LRRK2 mutation in idiopathic Parkinson's disease. *Lancet*, **365**, 415–416.
- Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R. *et al.* (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*, **276**, 2045–2047.
- Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y. and Shimizu, N. (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*, **392**, 605–608.
- Bonifati, V., Rizzu, P., van Baren, M.J., Schaap, O., Breedveld, G.J., Krieger, E., Dekker, M.C.J., Squitieri, F., Ibanez, P., Joosse, M. *et al.* (2003) Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science*, **299**, 256–259.
- Valente, E.M., Abou-Sleiman, P.M., Caputo, V., Muqit, M.M.K., Harvey, K., Gispert, S., Ali, Z., Del Turco, D., Bentivoglio, A.R., Healy, D.G.

- et al.* (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, **304**, 1158–1160.
9. Nalls, M.A., Plagnol, V., Hernandez, D.G., Sharma, M., Sheerin, U.-M., Saad, M., Simón-Sánchez, J., Schulte, C., Lesage, S., Sveinbjörnsdóttir, S. *et al.* (2011) Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet*, **377**, 641–649.
 10. International Parkinson's Disease Genomics Consortium and Wellcome Trust Case Control Consortium (2011) A two-stage meta-analysis identifies several new loci for Parkinson's disease. *PLoS Genet.*, **7**, e1002142.
 11. Hamza, T.H., Zabetian, C.P., Tenesa, A., Laederach, A., Montimurro, J., Yearout, D., Kay, D.M., Doheny, K.F., Paschall, J., Pugh, E. *et al.* (2010) Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat. Genet.*, **42**, 781–785.
 12. Satake, W., Nakabayashi, Y., Mizuta, I., Hirota, Y., Ito, C., Kubo, M., Kawaguchi, T., Tsunoda, T., Watanabe, M., Takeda, A. *et al.* (2009) Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.*, **41**, 1303–1307.
 13. Pankratz, N., Wilk, J.B., Latourelle, J.C., DeStefano, A.L., Halter, C., Pugh, E.W., Doheny, K.F., Gusella, J.F., Nichols, W.C., Foroud, T. *et al.* (2009) Genome-wide association study for susceptibility genes contributing to familial Parkinson disease. *Hum. Genet.*, **124**, 593–605.
 14. Edwards, T.L., Scott, W.K., Almonte, C., Burt, A., Powell, E.H., Beecham, G.W., Wang, L., Züchner, S., Konidari, I., Wang, G. *et al.* (2010) Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann. Hum. Genet.*, **74**, 97–109.
 15. Fung, H.C., Scholz, S., Matarin, M., Simón-Sánchez, J., Hernandez, D., Britton, A., Gibbs, J.R., Langefeld, C., Stiegert, M.L., Schymick, J. *et al.* (2006) Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol.*, **5**, 911–916.
 16. Maraganore, D.M., de Andrade, M., Lesnick, T.G., Strain, K.J., Farrer, M.J., Rocca, W.A., Pant, P.V.K., Frazier, K.A., Cox, D.R. and Ballinger, D.G. (2005) High-resolution whole-genome association study of Parkinson disease. *Am. J. Hum. Genet.*, **77**, 685–693.
 17. Spencer, C.C., Plagnol, V., Strange, A., Gardner, M., Paisan-Ruiz, C., Band, G., Barker, R., Bellenguez, C., Bhatia, K., Blackburn, H. *et al.* (2011) Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum. Mol. Genet.*, **20**, 345–353.
 18. Simón-Sánchez, J., Schulte, C., Bras, J.M., Sharma, M., Gibbs, J.R., Berg, D., Paisan-Ruiz, C., Lichtner, P., Scholz, S.W., Hernandez, D.G. *et al.* (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.*, **41**, 1308–1312.
 19. Simón-Sánchez, J., van Hilten, J.J., van de Warrenburg, B., Post, B., Berendse, H.W., Arepalli, S., Hernandez, D.G., de Bie, R.M.A., Velseboer, D., Scheffer, H. *et al.* (2011) Genome-wide association study confirms extant PD risk loci among the Dutch. *Eur. J. Hum. Genet.*, **19**, 655–661.
 20. Keller, M., Saad, M., Bras, J.M., Bettella, F., Nicolaou, N., Simon-Sanchez, J., Mittag, F., Buechel, F., Sharma, M., Gibbs, J.R. *et al.* (2012) Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Hum. Mol. Genet.*, **21**, 4996–5009.
 21. Park, J., Lee, S.B., Lee, S., Kim, Y., Song, S., Kim, S., Bae, E., Kim, J., Shong, M., Kim, J.-M. *et al.* (2006) Mitochondrial dysfunction in *Drosophila* PINK1 mutants is complemented by parkin. *Nature*, **441**, 1157–1161.
 22. Kamp, F., Exner, N., Lutz, A.K., Wender, N., Hegermann, J., Brunner, B., Nuscher, B., Bartels, T., Giese, A., Beyer, K. *et al.* (2010) Inhibition of mitochondrial fusion by α -synuclein is rescued by PINK1, Parkin and DJ-1. *EMBO J.*, **29**, 3571–3589.
 23. Thomas, K.J., McCoy, M.K., Blackinton, J., Beilina, A., van der Brug, M., Sandebring, A., Miller, D., Maric, D., Cedazo-Minguez, A. and Cookson, M.R. (2011) DJ-1 acts in parallel to the PINK1/parkin pathway to control mitochondrial function and autophagy. *Hum. Mol. Genet.*, **20**, 40–50.
 24. Cremades, N., Cohen, S.I.A., Deas, E., Abramov, A.Y., Chen, A.Y., Orte, A., Sandal, M., Clarke, R.W., Dunne, P., Aprile, F.A. *et al.* (2012) Direct observation of the interconversion of normal and toxic forms of α -synuclein. *Cell*, **149**, 1048–1059.
 25. Sha, D., Chin, L.S. and Li, L. (2010) Phosphorylation of parkin by Parkinson disease-linked kinase PINK1 activates parkin E3 ligase function and NF-kappaB signaling. *Hum. Mol. Genet.*, **19**, 352–363.
 26. Xiong, H., Wang, D., Chen, L., Choo, Y.S., Ma, H., Tang, C., Xia, K., Jiang, W., Ronai, Z., Zhuang, X. *et al.* (2009) Parkin, PINK1, and DJ-1 form a ubiquitin E3 ligase complex promoting unfolded protein degradation. *J. Clin. Invest.*, **119**, 650–660.
 27. Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K. *et al.* (2000) Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat. Genet.*, **25**, 302–305.
 28. Zhang, Y., Gao, J., Chung, K.K., Huang, H., Dawson, V.L. and Dawson, T.M. (2000) Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. *Proc. Natl Acad. Sci. USA*, **97**, 13354–13359.
 29. Baranzini, S.E., Galwey, N.W., Wang, J., Khankhanian, P., Lindberg, R., Pelletier, D., Wu, W., Uitdehaag, B.M.J., Kappos, L., Polman, C.H. *et al.* (2009) Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum. Mol. Genet.*, **18**, 2078–2090.
 30. Ritchie, M.D. (2009) Using prior knowledge and genome-wide association to identify pathways involved in multiple sclerosis. *Genome Med.*, **1**, 65.
 31. Wang, K., Li, M. and Bucan, M. (2007) Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.*, **81**, 1278–1283.
 32. Holmans, P., Green, E.K., Pahwa, J.S., Ferreira, M.A.R., Purcell, S.M., Sklar, P., Owen, M.J., O'Donovan, M.C. and Craddock, N. (2009) Gene Ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am. J. Hum. Genet.*, **85**, 13–24.
 33. Jones, L., Holmans, P., Hamshe, M.L., Harold, D., Moskva, V., Ivanov, D., Pocklington, A., Abraham, R., Hollingworth, P., Sims, R. *et al.* (2010) Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease. *PLoS One*, **5**, e13950.
 34. Lesnick, T.G., Papapetropoulos, S., Mash, D.C., Ffrench-Mullen, J., Shehadeh, L., de Andrade, M., Henley, J.R., Rocca, W.A., Ahlskog, J.E. and Maraganore, D.M. (2007) A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet.*, **3**, e98.
 35. Srinivasan, B.S., Doostzadeh, J., Absalan, F., Mohandessi, S., Jalili, R., Bigdeli, S., Wang, J., Mahadevan, J., Lee, C.L.G., Davis, R.W. *et al.* (2009) Whole genome survey of coding SNPs reveals a reproducible pathway determinant of Parkinson disease. *Hum. Mut.*, **30**, 228–238.
 36. Stone, D.K., Reynolds, A.D., Mosley, R.L. and Gendelman, H.E. (2009) Innate and adaptive immunity for the pathobiology of Parkinson's disease. *Antioxid. Redox Signal.*, **11**, 2151–2166.
 37. Brochard, V., Combadière, B., Prigent, A., Laouar, Y., Perrin, A., Beray-Berthet, V., Bonduelle, O., Alvarez-Fischer, D., Callebert, J., Launay, J.M. *et al.* (2009) Infiltration of CD4+ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J. Clin. Invest.*, **119**, 182–192.
 38. Castellani, R.J., Nugent, S.L., Morrison, A.L., Zhu, X., Lee, H.G., Harris, P.L.R., Bajic, V., Sharma, H.S., Chen, S.G., Oetgen, P. *et al.* (2011) CD3 in Lewy pathology: does the abnormal recall of neurodevelopmental processes underlie Parkinson's disease. *J. Neural Transm.*, **118**, 23–26.
 39. Mogi, M., Kondo, T., Mizuno, Y. and Nagatsu, T. (2007) p53 protein, interferon-gamma, and NF-kappaB levels are elevated in the parkinsonian brain. *Neurosci. Lett.*, **414**, 94–97.
 40. Reale, M., Iarlori, C., Thomas, A., Gambi, D., Perfetti, B., Di Nicola, M. and Onofri, M. (2009) Peripheral cytokines profile in Parkinson's disease. *Brain Behav. Immun.*, **23**, 55–63.
 41. Kubo, M., Kamiya, Y., Nagashima, R., Maekawa, T., Eshima, K., Azuma, S., Ohta, E. and Obata, F. (2010) LRRK2 is expressed in B-2 but not in B-1 B cells, and downregulated by cellular activation. *J. Neuroimmunol.*, **229**, 123–128.
 42. Saad, M., Lesage, S., Saint-Pierre, A., Corvol, J.C., Zelenika, D., Lambert, J.C., Vidailhet, M., Mellick, G.D., Lohmann, E., Durif, F. *et al.* (2011) Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population. *Hum. Mol. Genet.*, **20**, 615–627.
 43. Moskva, V. and Schmidt, K.M. (2008) On multiple-testing correction in genome-wide association studies. *Genet. Epidemiol.*, **32**, 567–573.
 44. Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C. *et al.* (2004) The Gene

- Ontology (GO) database and informatics resource. *Nucleic Acids Res.*, **32**, D258–D261.
45. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. and Tanabe, M. (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.*, **40**, D109–D114.
46. Bult, C.J., Eppig, J.T., Kadin, J.A., Richardson, J.E. and Blake, J.A. (2008) The Mouse Genome Database (MGD): mouse biology and model systems. *Nucleic Acids Res.*, **36**, D724–D728.
47. Simes, R. (1986) An improved Bonferroni-type procedure for multiple tests of significance. *Biometrika*, **73**, 751–754.

APPENDIX

The International Parkinson's Disease Genomics Consortium

Sampath Arepalli¹, Roger Barker¹¹, Jeffrey Barrett¹², Yoav Ben-Shlomo¹³, Henk W. Berendse¹⁴, Daniela Berg⁴, Kailash Bhatia⁵, Rob M.A. de Bie¹⁵, Alessandro Biffi^{16,17,18}, Bas Bloem¹⁹, Alexis Brice^{9,10,33}, Zoltan Bochdanovits⁸, Michael Bonin²⁰, Jose M. Bras³, Kathrin Brockmann⁴, Janet Brooks¹, David J. Burn²¹, Gavin Charlesworth², Honglei Chen²², Patrick F. Chinnery²³, Sean Chong¹, Carl E. Clarke²⁴, Mark R. Cookson¹, Jonathan M. Cooper²⁵, Jen-Christophe Corvol^{9,10,26}, Carl Counsell²⁷, Philippe Damier²⁸, Jean Francois Dartigues²⁹, Panagiotis Deloukas¹², Günther Deuschl³⁰, David T. Dexter³¹, Karin D. van Dijk¹⁴, Allissa Dillman¹, Frank Durif³², Alexandra Durr³³, Sarah Edkins¹², Jonathan R. Evans³⁴, Thomas Foltynie³⁵, Jianjun Gao²², Michelle Gardner³⁵, Thomas Gasser⁴, J. Raphael Gibbs^{1,3}, Alison Goate³⁶, Emma Gray¹², Rita Guerreiro³, Ómar Guðstafsson³⁷, John Hardy³, Clare Harris²⁷, Dena G. Hernandez^{1,3}, Peter Heutink⁸, Jacobus J. van Hilten³⁸, Albert Hofman³⁹, Albert Hollenbeck⁴⁰, Peter Holmans⁴⁷, Janice Holton⁴¹, Michele Hu⁴², Heiko Huber⁴, Gavin Hudson²³, Sarah E. Hunt¹², Johanna Huttenlocher³⁷, Thomas Illig⁴³, Cordelia Langford³⁴, Andrew Lees⁴¹, Suzanne Lesage^{9,10}, Peter Lichtner⁴⁴, Patricia Limousin⁴⁵, Grisel Lopez⁴⁶, Delia Lorenz³⁰, Maria Martinez^{6,7}, Alisdair McNeill²⁵, Catriona Moorby²⁴, Matthew Moore¹, Huw Morris⁴⁷, Karen E. Morrison²⁴, Valentina Moskvina⁴⁷, Ese Mudanohwo⁴⁸, Michael A. Nalls¹, Justin Pearson⁴⁷, Joel S. Perlmutter³⁶, Hjörvar Pétursson³⁷, Vincent Plagnol², Pierre Pollak⁴⁹, Bart Post¹⁹, Simon Potter¹², Bernard Ravina⁵⁰, Tamas Revesz⁴¹, Olaf Riess²⁰, Fernando Rivadeneira^{39,51}, Patrizia Rizzu⁸, Mina Ryten³, Mohamad Saad^{6,7}, Stephen Sawcer⁵², Anthony Schapira²⁵, Hans Scheffer⁵³, Manu Sharma⁴, Karen Shaw⁴¹, Una-Marie Sheerin⁵, Ira Shoulson⁵⁴, Claudia Schulte⁴, Ellen Sidransky⁴⁶, Javier Simón-Sánchez⁸, Andrew B. Singleton¹, Colin Smith⁵⁵, Hreinn Stefánsson³⁷, Kári Stefánsson³⁷, Stacy Steinberg³⁷, Joanna D. Stockton²⁴, Sigurlaug Sveinbjörnsdóttir⁵⁶, Kevin Talbot⁵⁷, Carlie M. Tanner⁵⁸, Avazeh Tashakkori-Ghanbaria¹², François Tison⁵⁹, Daniah Trabzuni³, Bryan J. Traynor¹, André G. Uitterlinden³⁹, Daan Velseboer¹⁵, Marie Vidailhet^{9,10}, Robert Walker⁵⁵, Bart van de Warrenburg¹⁹, Mirdhu Wickremaratchi⁶⁰, Nigel Williams⁴⁷, Caroline H. Williams-Gray³⁴, Sophie Winder-Rhodes⁶¹ and Nicholas Wood³⁵

¹Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA; ²UCL Genetics Institute, Gower Place, London WC1E 6BT, UK; ³Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, London, UK; ⁴Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tuebingen and DZNE, German Center for Neurodegenerative Diseases, Tuebingen, Germany; ⁵UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK; ⁶INSERM U563, CPTP, Toulouse, France; ⁷Paul Sabatier University, Toulouse,

France; ⁸Section of Medical Genomics, Department of Clinical Genetics, VU University Medical Centre, Amsterdam, The Netherlands; ⁹INSERM, UMRS975 (formerly UMRS679), Paris, France; ¹⁰Université Pierre et Marie Curie, Paris, France; ¹¹Department of Neurology, Addenbrooke's Hospital, University of Cambridge, Hills Road, Cambridge CB2 2QQ, UK; ¹²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK; ¹³Department of Social Medicine, Bristol University, Bristol, UK; ¹⁴Department of Neurology and Alzheimer Center, VU University Medical Center, Amsterdam, The Netherlands; ¹⁵Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; ¹⁶Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA; ¹⁷Department of Neurology, Massachusetts General Hospital, Boston, MA 02114, USA; ¹⁸Program in Medical and Population Genetics, Broad Institute, 7 Cambridge Center, Cambridge, MA 02139, USA; ¹⁹Department of Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; ²⁰Department of Medical Genetics, Institute of Human Genetics, University of Tuebingen, Tuebingen, Germany; ²¹Newcastle University Clinical Ageing Research Unit, Campus for Ageing and Vitality, Newcastle upon Tyne NE4 5PL, UK; ²²Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, North Carolina, USA; ²³Neurology M4104, The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK; ²⁴School of Clinical and Experimental Medicine, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; ²⁵Department of Clinical Neurosciences, UCL Institute of Neurology, Royal Free Campus, Rowland Hill Street, London NW3 2PF, UK; ²⁶INSERM CIC-9503, Hôpital Pitié-Salpêtrière, Paris, France; ²⁷Population Health Section, Division of Applied Health Sciences, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen AB25 2ZD, UK; ²⁸CHU Nantes, CIC0004, Service de Neurologie, Nantes, France; ²⁹INSERM U897, Victor Segalen University, Bordeaux, France; ³⁰Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel, Christian-Albrechts-Universität Kiel, Kiel, Germany; ³¹Parkinson's Disease Research Group, Faculty of Medicine, Imperial College London, Fourth Floor, Burlington Danes Building, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK; ³²Service de Neurologie, Hôpital Gabriel Montpied, Clermont-Ferrand, France; ³³Department of Genetics and Cytogenetics, AP-HP, Pitié-Salpêtrière Hospital, Paris, France; ³⁴Cambridge Centre for Brain Repair, Forvie Site, Robinson Way, Cambridge CB22PY, UK; ³⁵Institute of Neurology, University College London, London, UK; ³⁶Department of Psychiatry and Department of Neurology, Washington University School of Medicine, St Louis, MO, USA; ³⁷deCODE genetics, Sturlugata 8, IS-101 Reykjavik, Iceland; ³⁸Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands; ³⁹Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; ⁴⁰AARP, Washington, DC, USA; ⁴¹Queen Square Brain Bank for Neurological Disorders, Institute of Neurology, University College London, London, UK; ⁴²Department of Clinical Neurology, West Wing, Level 3, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK; ⁴³Institute of Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany; ⁴⁴Institute of Human Genetics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany; ⁴⁵Unit of Functional Neurosurgery, Sobell Department, Institute of Neurology, Queen Square, London, UK; ⁴⁶Section on Molecular Neurogenetics, Medical

Genetics Branch, NHGRI, National Institutes of Health, Bethesda, MD, USA; ⁴⁷MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff CF14 4XN, UK; ⁴⁸Neurogenetics Unit, UCL Institute of Neurology/National Hospital for Neurology and Neurosurgery, Queen Square, London, UK; ⁴⁹Service de Neurologie, CHU de Grenoble, Grenoble, France; ⁵⁰Department of Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; ⁵¹Translational Neurology, Biogen Idec, 14 Cambridge Center, Bio 6, Cambridge, MA, USA; ⁵²Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; ⁵³Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's, Hospital Hills Road, Cambridge CB2 0QQ, UK; ⁵⁴Department of Human Genetics,

Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; ⁵⁵Department of Neurology, University of Rochester, Rochester, New York, NY 14620, USA; ⁵⁶Department of Pathology, Wilkie Building, Teviot Place, Edinburgh EH8 9AG, UK; ⁵⁷Department of Neurology, Landspítali, University Hospital, Reykjavík, Iceland; ⁵⁸Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK; ⁵⁹Clinical Research Department, The Parkinson's Institute and Clinical Center, Sunnyvale, CA, USA; ⁶⁰Service de Neurologie, Hôpital Haut-Lévêque, Pessac, France; ⁶¹Department of Neurology, Cardiff University, Cardiff, UK; ⁶²Department of Psychiatry and Medical Research Council/Wellcome Trust Behavioural and Clinical Neurosciences Institute, University of Cambridge, Cambridge, UK.