European Journal of Human Genetics (2010) 18, 1356–1359
© 2010 Macmillan Publishers Limited All rights reserved 1018-4813/10
www.nature.com/ejhg

SHORT REPORT

Genetic variability at the PARK16 locus

Arianna Tucci[1], Mike A Nalls[2], Henry Houlden[1], Tamas Revesz[1], Andrew B Singleton[2], Nicholas W Wood[1], John Hardy[1] and Coro Paisán-Ruiz*[,1]

Parkinson's disease (PD) is a complex neurodegenerative disease which is clinically heterogeneous and pathologically consists of loss of dopaminergic neurons in the substantia nigra and intracytoplasmic neuronal inclusions containing alpha-synuclein aggregations known as Lewy bodies. Although the majority of PD is idiopathic, pathogenic mutations in several mendelian genes have been successfully identified through linkage analyses. To identify susceptibility loci for idiopathic PD, several genome-wide association studies (GWAS) within different populations have recently been conducted in both idiopathic and familial forms of PD. These analyses have confirmed *SNCA* and *MAPT* as loci harboring PD susceptibility. In addition, the GWAS identified several other genetic loci suggestively associated with the risk of PD; among these, only one was replicated by two different studies of European and Asian ancestries. Hence, we investigated this novel locus known as PARK16 for coding mutations in a large series of idiopathic pathologically proven PD cases, and also conducted an association study in a case–control cohort from the United Kingdom. An association between a novel *RAB7L1* mutation, c.379-12insT, and disease (*P*-value=0.0325) was identified. Two novel coding variants present only in the PD cohort were also identified within the *RAB7L1* (p.K157R) and *SLC41A1* (p.A350V) genes. No copy number variation analyses have yet been performed within this recently identified locus. We concluded that, although both coding variants and risk alleles within the PARK16 locus seem to be rare, further molecular analyses within the PARK16 locus and within different populations are required in order to examine its biochemical role in the disease process.
European Journal of Human Genetics (2010) **18,** 1356–1359; doi:10.1038/ejhg.2010.125; published online 4 August 2010

Keywords: PARK16 locus; genetics; association studies

INTRODUCTION

Parkinson's disease (PD; MIM #68600) is a common complex disease clinically characterized by resting tremor, bradykinesia, postural instability and rigidity, and pathologically by the presence of severe pars-compacta nigral-cell loss and an accumulation of aggregated α-synuclein in specific brain stem, spinal cord and cortical regions.[1] Although the majority of PD is idiopathic, pathogenic mutations have successfully been identified in some mendelian forms.[2] Many of these mendelian genes have also been investigated in the idiopathic disease, but only SNPs at the *SNCA* and *LRRK2* loci have shown susceptibility for idiopathic PD (IPD): several SNPs at the *SNCA* locus have been characterized as risk factors for IPD in different populations,[3] a *LRRK2*-associated haplotype showed an increased disease risk in the Chinese population,[4] two *LRRK2* mutations absent in European ancestry populations are overrepresented in PD in some Asian populations[5,6] and common *LRRK2* variation may also contribute to the risk for IPD in the North American population.[7] Similarly, the frequency and distribution of *GBA* mutations in PD vary within populations, being more prevalent among the Ashkenazi Jewish population and rare among Asians.[8] Taken together, PD is a complex genetic disorder in which the prevalence of some pathogenic mutations may vary widely within ethnicities.[8,9]

Genome-wide (GW) SNP genotyping assays have been proven to be a powerful technique to identify genetic risk factors in many complex disorders.[10] Consequently, three large PD-associated genome-wide association studies (GWAS) from two European ancestry and one Asian population have recently identified genetic risks underlying PD, of which *SNCA* (all three studies) and *MAPT* (only European ancestry studies) loci showed the strongest evidences of association with PD;[11–13] these associations have recently been corroborated by a meta-analysis carried out in European ancestry PD-associated GWAS.[14] In addition, one out of three additional genetic risk loci for PD was independently identified by two studies.[11–13] Therefore, in this study, we investigate whether novel genetic variants within this locus designated as PARK16 may predispose to the risk for PD in a British cohort of pathologically proven PD cases and neurologically normal individuals. PARK16 located on chromosome 1q32 comprises 169.6 kb and contains five different genes (Table 1).

MATERIALS AND METHODS

Subjects

The PD cohort was collected from brain tissues at The Queen Square Brain Bank for Neurological Disorders in the United Kingdom. Cases (n=453) were clinically and pathologically diagnosed according to the PD Brain Bank criteria.[15,16] The mean age at onset was 59 years (ranged from 35 to 86 years) and the average age of death was 78 years (ranged from 51 to 94 years). The male-to-female ratio was 3.5:1. Family history was reported in <1% of individuals. DNA samples from 82 PD cases reporting a positive family history were also employed. Positive family history was compatible with the diagnosis of PD in at least one first- or second-degree relative. The mean age of disease onset in these familial cases was 57 years (ranged from 29 to 71 years). Patients and all relatives of patients gave informed consent for scientific research. The control cohort (n=483) analyzed here was the '1958 British birth cohort', which

[1]Department of Molecular Neuroscience and Reta Lila Weston Institute, UCL Institute of Neurology, London, UK; [2]Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA
*Correspondence: Dr C Paisán-Ruiz, Department of Neurology, Mount Sinai School of Medicine, 1 Gustave L. Levy Place, New York, NY 10029, USA; Tel: (212) 241 7317; Fax:(212) 348 1310; E-mail: cpaisan@gmail.com
Received 22 February 2010; revised 11 May 2010; accepted 24 June 2010; published online 4 August 2010

Table 1 Previously reported PD-associated SNPs within the PARK16 locus

PARK16 locus (SNPs)	*Chr*	*Position (bp)*	*Alleles (minor/major)*	*P-values (combined stages I and II)*	*Chr Localization*
rs16856139	1	203 905 087	T/C	1.02×10^{-07}	SLC45A3
rs823128	**1**	**203 980 001**	**G/A**	$\mathbf{7.29\times10^{-08}}$	**NUCKS1**
rs823122	1	203 991 651	C/T	4.88×10^{-09}	Genomic region
rs947211	**1**	**204 019 288**	**A/G**	$\mathbf{1.52\times10^{-12}}$	**Genomic region**
rs823156	1	204 031 263	G/A	7.60×10^{-04}	SLC41A1
rs823156	1	204 031 263	G/A	3.60×10^{-09}	SLC41A1
rs708730	1	204 044 403	G/A	2.43×10^{-08}	SLC41A1
rs11240572	1	204 074 636	A/C	6.11×10^{-07}	PM20D1
rs11240572	1	204 074 636	A/C	1.08×10^{-07}	PM20D1

SNPs showing association with the disease in the Japanese case–control cohort are represented in black; SNPs showing association with PD in both European ancestry case–control cohorts are shown in gray.[12,13] The SNPs showing the highest *P*-value for each independent study are highlighted in bold. *RAB7L1* chromosomal localization: 204 003 738–204 011 233 bp. *RAB7L1* is located in the 27.6 kb interval between rs823122 (at ~12 kb away) and rs947211 (at ~8 kb away) SNPs.

comprises individuals who were all born in March 1958 in England, Scotland or Wales and which is used in all disease-related studies carried out by the Wellcome Trust Case Control Consortium (WTCCC; http://www.b58cgene.sgul.ac.uk).

PCR and sequencing analyses

In the first instance, PCR and sequencing analyses of all open-reading frames (ORFs) of *NUCKS1* (RefSeq NM_022731, seven exons), *RAB7L1* (RefSeq NM_003929, five exons) and *SLC41A1* (RefSeq NM_173854.4, 11 exons) genes were performed in 182 PD cases. Later, each variant identified in the PD cohort (n=9) was also analyzed in 351 neurologically normal individuals. Thereafter, every SNP showing association with the disease (n=1; c.379-12insT) and each coding variant absent in controls (n=2; p.K157R and p.A350V) were further analyzed in a larger sample size, resulting in a total of 454 PD cases and 483 controls being analyzed. In addition, the two coding variants absent in the control population were also tested in 82 familial PD cases. All PCR analyses were performed using both forward and reverse genomic primers (all primer sequences are available on request) previously designed by ExonPrimer (http://www.ihg.gsf.de/ihg/ExonPrimer.html) and FastStart Taq DNA polymerase (http://www.roche-applied-science.com). Each purified product was sequenced using both forward and reverse primers with Applied Biosystems BigDye terminator v3.1 sequencing chemistry as per the manufacturer's instructions; the resulting reactions were then resolved on an ABI3730XL genetic analyzer (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA) and analyzed with Sequencher software 4.9 (Gene Codes Corporation, Ann Arbor, MI, USA).

Alamut mutation interpretation software was used for determining amino-acid properties and for predicting the functional and structural effects of novel coding mutations (http://www.interactive-biosoftware.com/alamut.html). Multiple alignments for *RAB7L1* and *SLC41A1* encoding proteins were determined through the NCBI-associated homoloGene database using the MUSCLE program.[17] The Human Protein Reference Database (http://www.hprd.org/) was employed to search for predicted protein motifs and domains. The NCBI-BLAST database was also used to search for sequence similarities between Rab proteins; the RAB7L1 protein sequence (RefSeq: NP_001129134.1) was aligned with RAB1 (RefSeq: NP_004152.1), RAB3A (RefSeq: NP_002857.1), RAB7 (RefSeq: NP_004628.4) and RAB8A (RefSeq: CAG38820.1) proteins (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Statistical analyses

All statistical analyses (χ^2-tests of association and permutation analyses) were performed using Haploview 4.1 software (http://www.broad.mit.edu/haploview/). To compare PARK16-associated allelic frequencies between diverse populations, HapMap data corresponding to the PARK16 locus from Yoruba (YRI), Japanese (JPT), Han Chinese (CHB) and northern and western European (CEU-Utah residents) populations were also analyzed through Haploview software (http://www.hapmap.org).

RESULTS

To try and identify novel genetic variants underlying risk for PD in a British case–control cohort, the genomic area harboring the PARK16 locus was investigated in depth through sequencing analyses. In the first instance, it was decided to perform sequencing analyses of all coding regions and exon-intron boundaries of genes located within the genomic area shared by the PARK16 loci identified in both European ancestry and Asian populations;[12,13] this area, which is flanked by rs823128 (203 980 001 bp) and rs11240572 (204 074 636 bp) SNPs, contained four genes (*NUCKS1*, *RAB7L1*, *SLC41A1* and *PM20D1*) (Table 1). However, *NUCKS1*, *RAB7L1* and *SLC41A1* genes were located in the same LD block and were suggestively reported as the best candidates for the etiology of PD according to their functional roles.[12] In addition, the minor allele frequency of rs11240572 located in intron 10 of PM20D1 is <0.03 in the European ancestry population (Table 3); hence, only the coding regions of *NUCKS1*, *RAB7L1* and *SLC41A1* were analyzed in our 182 PD cases. PCR analyses of all ORFs revealed the presence of nine different genetic variants within *RAB7L1* (n=5) and *SLC41A1* (n=4) genes, whereas no genetic variation was identified across the *NUCKS1* gene. There were two coding variants (p.Gln104Glu (s41302139) and p.Lys157Arg (novel)), two novel intronic variants (c.197-49insG and c.379-12insT) and one UTR-5′ variant (rs708755) among the mutations identified within the *RAB7L1* gene; whereas three coding variants (p.Thr113Thr (rs11240569), p.Asn252Asn (rs708727) and p.Ala350Val (novel)) and a known intronic variant (rs41264905) were identified within the *SLC41A1* gene. All genetic variants, with the exception of both novel coding mutations that were identified in one PD patient each, were found in both cases and controls (Table 2). The coding mutations were a heterozygous c.470A>G transition causing p.Lys157Arg and a heterozygous c.1049C>T transition causing p.Ala350Val, which were located within *RAB7L1* (exon 4) and *SLC41A1* (exon 8) genes, respectively (Supplementary Figure 1). To inspect whether these novel coding mutations may or may not be the disease-causing mutations, both were tested in a larger sample size of additional pathologically proven IPD cases (n=272, n (total)=454), 82 familial cases clinically diagnosed with PD and 483 neurologically normal individuals; the investigation failed to detect any other mutation carrier in both PD and control populations. Both variants are also conserved among species (Supplementary Figure 1). Contradictory results were obtained with respect to the functional consequences for both novel mutations: the K157 amino acid of *RAB7L1* was predicted to be highly conserved whereas the A350 amino-acid of *SLC41A1* was shown to be moderately conserved (Alamut software, Interactive Biosoftware, Rouen, France). Both clinical and pathological features of K157R and A350V mutation carriers are described in Supplementary Material 1.

To test whether the remaining seven genetic variants identified may predispose to the risk for PD, they were additionally tested in 351

Table 2 χ^2-association tests for common variants identified within PARK16 locus performed by Haploview software

Gene	*Chr*	*Mutation*	*Position (bp)*	*Associated allele*	χ^2	*P-value*	*Case/control frequencies*
RAB7L1	1	rs708725	204 010 761	A	0.609	**0.4351**	0.443, 0.418
RAB7L1	1	c.197-49insG	204 007 453	Ins G	0.949	**0.3301**	0.072, 0.057
RAB7L1	1	rs41302139	204 007 291	C	0.075	**0.7844**	0.019, 0.017
RAB7L1	**1**	**c.379-12insT**	**204 006 575**	**Ins T**	**4.573**	**0.0325 (0.0399)**	**0.013, 0.004**
SLC41A1	1	rs708727	204 034 508	T	0.776	**0.3783**	0.438, 0.409
SLC41A1	1	rs41264905	204 034 656	T	1.402	**0.2364**	0.009, 0.003
SLC41A1	1	rs11240569	204 045 854	T	0.971	**0.3243**	0.301, 0.270

The only variant (c.379-12insT) that showed association with the disease is highlighted in bold; the 1×10^6 permutation value associated to this variant is also shown within parentheses.
The case/control frequencies for each variant are also shown.

Table 3 PARK16 core SNPs frequencies in diverse populations

	Position (bp)	*ObsHET*	*PredHET*	*HWpval*	*MAF*	*Alleles (m/M)*
CEU (SNPs)						
rs16856139	203 905 087	0.094	0.09	1	0.047	C/T
rs823128	203 980 001	0.034	0.034	1	0.017	A/G
rs823122	203 991 651	0.077	0.09	0.4364	0.047	T/C
rs947211	204 019 288	0.376	0.364	0.9746	0.239	G/A
rs823156	204 031 263	0.308	0.295	0.938	0.179	A/G
rs708730	204 044 403	0.222	0.248	0.3983	0.145	A/G
rs11240572	204 074 636	0.043	0.042	1	0.021	C/A
CHB+JPT (SNPs)						
rs16856139	203 905 087	0.224	0.217	1	0.124	C/T
rs823128	203 980 001	0.253	0.247	1	0.144	A/G
rs823122	203 991 651	0.253	0.247	1	0.144	T/C
rs947211	204 019 288	0.494	0.494	1	0.447	A/G
rs823156	204 031 263	0.365	0.347	0.7025	0.224	A/G
rs708730	204 044 403	0.371	0.35	0.6318	0.226	A/G
rs11240572	204 074 636	0.324	0.294	0.3199	0.179	C/A
YRI (SNPs)						
rs16856139	203 905 087	0.087	0.159	5×10^{-4}	0.087	C/T
rs823128	203 980 001	0.522	0.472	0.381	0.383	A/G
rs823122	203 991 651	0.478	0.499	0.7532	0.483	C/T
rs947211	204 019 288	0.487	0.476	1	0.391	A/G
rs823156	204 031 263	0.426	0.427	1	0.309	G/A
rs708730	204 044 403	0.217	0.258	0.1759	0.152	G/A
rs11240572	204 074 636	0	0	1	0	C/C

Abbreviations: CEU, CEPH (Utah residents with ancestry from northern and western Europe); CHB, Han Chinese in Beijing, China; JPT, Japanese in Tokyo, Japan; YRI, Yoruba in Ibadan, Nigeria, MAF, minor allele frequencies; m, minor allele; M, major allele.

neurologically normal individuals. A single-marker χ^2-test of association was then performed. This analysis revealed a slightly significant association between the c.379-12insT mutation within the *RAB7L1* gene (intron 3) and PD (frequentist *P*-value=0.0325), which remained significant after one million iterations of permutation testing to adjust for multiple comparisons (permuted *P*-value=0.0399; Table 2).

DISCUSSION

Sequencing of the genes within the PARK16 locus in a British cohort of 182 pathologically proven PD cases revealed the presence of two novel mutations; in one patient in the *RAB7L1* (K157R) and in another patient in the *SLC41A1* (A350V) gene. Both mutations carriers showed typical IPD and Lewy body pathology; however, even the lack of occurrence of both mutations in a large sample of ethnicity-matched control individuals (n=483) does not fully disclose their pathogenecity. In addition, seven intronic and exonic variants were identified during the sequencing process, and therefore, an association study was performed revealing a weak association between the c.379-12insT mutation and IPD. Curiously, no intronic variation was previously reported within the *RAB7L1* locus, suggesting that genetic variability within this locus is rare. In this study, large copy number variations were not examined; however, given the presence of a rare novel mutation and slightly associated risk allele within the *RAB7L1* gene, further molecular analyses are warranted to determine the precise biochemical role of *RAB7L1* in the etiology of PD. The *RAB7L1* encoding protein is a member of the Rab GTPases subfamily, which includes several small GTPases involved in intracellular cell signaling processes and vesicle trafficking. The K157 amino acid of *RAB7L1* lies in the Rab domain (8–176 amino acids) of the protein, which is predicted to be highly conserved among species and is also conserved among other Rab proteins such as RAB1A, RAB3A, RAB7A and RAB8A proteins (data not shown). Molecular links between PD and Rab proteins were already suggested: mutations in the Ras-like GTPase domain of dardarin cause PD[9,18] and elevated expression of RAB1, RAB3A and RAB8A proteins protects against alpha-syn-induced dopaminergic neuronal loss in animal models of PD.[19,20] *SLC41A1* is a Mg (2+) transporter that may have a role in magnesium homeostasis. Brain metal dyshomeostasis has often been speculated as a cause of neurodegeneration; nevertheless, the precise nature of its biochemical mechanisms underlying neurodegeneration is still vague.[21,22]

Although no association between the PARK16 locus and PD was identified in a GWAS meta-analysis,[14] analyses of PARK16-associated SNPs within the HapMap data revealed marked differences in the minor allelic frequencies between populations, thus affecting analytic power (Table 3). Similarly, population differences at the *BST1* and *MAPT* loci were recently reported,[12,13] and the haplotype H2 of *MAPT* reported to be almost exclusively of Caucasian origin is low in all populations.[23] By and large, different genetic markers should be used when investigating different populations, as some may not be relevant to all populations. Thereafter, we conclude that, although pathogenic mutations and risk alleles within the PARK16 locus seem to be rare in European ancestry populations, further molecular analyses within different populations are required to examine its biochemical role in PD and before undertaking any functional work on the encoded proteins associated with this locus.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all patients who participated in this study and all families who supported the donation of tissue for research. We also thank The Medical Research Council (MRC; HH: MRC fellowships G108/638 and G0802760 and JH: start-up funds) and The Michael J Fox Foundation (HH and CPR) for support. This study was also supported by the NIHR UCLH/UCL Comprehensive Biomedical Research Centre. MAN and ABS were partly supported in this research by the Intramural Research Program of the NIH, National Institute on Aging (AG000957-07 (2009) Assessment of Candidate Loci in Neurological diseases).

1 Lees AJ, Hardy J, Revesz T: Parkinson's disease. *Lancet* 2009; **373**: 2055–2066.
2 Hardy J, Lewis P, Revesz T, Lees A, Paisan-Ruiz C: The genetics of Parkinson's syndromes: a critical review. *Curr Opin Genet Dev* 2009; **19**: 254–265.
3 Winkler S, Hagenah J, Lincoln S *et al*: alpha-Synuclein and Parkinson disease susceptibility. *Neurology* 2007; **69**: 1745–1750.
4 Skipper L, Li Y, Bonnard C *et al*: Comprehensive evaluation of common genetic variation within LRRK2 reveals evidence for association with sporadic Parkinson's disease. *Hum Mol Genet* 2005; **14**: 3549–3556.
5 Di Fonzo A, Wu-Chou YH, Lu CS *et al*: A common missense variant in the LRRK2 gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan. *Neurogenetics* 2006; **7**: 133–138.
6 Ross OA, Wu YR, Lee MC *et al*: Analysis of Lrrk2 R1628P as a risk factor for Parkinson's disease. *Ann Neurol* 2008; **64**: 88–92.
7 Paisan-Ruiz C, Washecka N, Nath P, Singleton AB, Corder EH: Parkinson's disease and low frequency alleles found together throughout LRRK2. *Ann Hum Genet* 2009; **73**: 391–403.
8 Sidransky E, Nalls MA, Aasly JO *et al*: Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med* 2009; **361**: 1651–1661.
9 Paisan-Ruiz C: LRRK2 gene variation and its contribution to Parkinson disease. *Hum Mutat* 2009; **30**: 1153–1160.
10 Hardy J, Singleton A: Genomewide association studies and human disease. *N Engl J Med* 2009; **360**: 1759–1768.
11 Pankratz N, Wilk JB, Latourelle JC *et al*: Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *Hum Genet* 2009; **124**: 593–605.
12 Satake W, Nakabayashi Y, Mizuta I *et al*: Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* 2009; **41**: 1303–1307.
13 Simon-Sanchez J, Schulte C, Bras JM *et al*: Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet* 2009; **41**: 1308–1312.
14 Edwards TL, Scott WK, Almonte C *et al*: Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann Hum Genet* 2010; **74**: 97–109.
15 Gibb WR, Lees AJ: The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1988; **51**: 745–752.
16 Hughes AJ, Daniel SE, Kilford L, Lees AJ: Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; **55**: 181–184.
17 Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004; **32**: 1792–1797.
18 Greggio E, Cookson MR: Leucine-rich repeat kinase 2 mutations and Parkinson's disease: three questions. *ASN Neuro* 2009; **1**: e00002.
19 Cooper AA, Gitler AD, Cashikar A *et al*: Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* 2006; **313**: 324–328.
20 Gitler AD, Bevis BJ, Shorter J *et al*: The Parkinson's disease protein alpha-synuclein disrupts cellular Rab homeostasis. *Proc Natl Acad Sci USA* 2008; **105**: 145–150.
21 Bolognin S, Messori L, Zatta P: Metal ion physiopathology in neurodegenerative disorders. *Neuromolecular Med* 2009; **11**: 223–238.
22 Kolisek M, Launay P, Beck A *et al*: SLC41A1 is a novel mammalian Mg2+ carrier. *J Biol Chem* 2008; **283**: 16235–16247.
23 Evans W, Fung HC, Steele J *et al*: The tau H2 haplotype is almost exclusively Caucasian in origin. *Neurosci Lett* 2004; **369**: 183–185.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)