



**HAL**  
open science

## Genetic variability at the PARK16 locus

Coro Paisan-Ruiz, Arianna Tucci, Michael A Nalls, Henry Houlden, Tamas Revesz, Andrew Singleton, Nicholas W. Wood, John Hardy

► **To cite this version:**

Coro Paisan-Ruiz, Arianna Tucci, Michael A Nalls, Henry Houlden, Tamas Revesz, et al.. Genetic variability at the PARK16 locus. *European Journal of Human Genetics*, 2010, 10.1038/ejhg.2010.125 . hal-00563105

**HAL Id: hal-00563105**

**<https://hal.science/hal-00563105>**

Submitted on 4 Feb 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Genetic variability at the PARK16 locus**

2 Arianna Tucci<sup>1</sup>, Mike A. Nalls<sup>2</sup>, Henry Houlden<sup>1</sup>, Tamas Revesz<sup>1</sup>, Andrew B. Singleton<sup>2</sup>, Nicholas W.  
3 Wood<sup>1</sup>, John Hardy<sup>1</sup> and Coro Paisán-Ruiz<sup>1CA</sup>

4 <sup>1</sup>Department of Molecular Neuroscience and Reta Lila Weston Institute, UCL Institute of Neurology, London, Queen Square,  
5 London, United Kingdom

6 <sup>2</sup>Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda,  
7 MD, USA

8  
9

10 <sup>CA</sup> Correspondence should be addressed to:

11 Coro Paisán-Ruiz, PhD

12 Department of Molecular Neuroscience and Reta Lila Weston Institute, UCL Institute of Neurology, 9th Floor, Queen Square

13 House, Queen Square, London WC1N 3BG, England. Tel: 44-(0)-207-837-3611 (Ext 4015); Fax: 44-(0)-207-833-1016; Email:

14 C.Paisan-Ruiz@ion.ucl.ac.uk

15

16

17 Running Title: PARK16 Locus variability

18 **Abstract**

19 Parkinson's disease is a complex neurodegenerative disease whose hallmark pathological  
20 features are loss of dopaminergic neurons in the substantia nigra and intracytoplasmic  
21 neuronal inclusions containing alpha-synuclein aggregations known as Lewy bodies. Although  
22 the majority of PD is idiopathic, pathogenic mutations in several mendelian genes have  
23 successfully been identified through linkage analyses. To identify susceptibility loci for  
24 idiopathic Parkinson's disease, several genome-wide association studies (GWAS) within  
25 different populations have recently been conducted in both idiopathic and familial forms of PD.  
26 These analyses have confirmed *SNCA* and *MAPT* as loci harbouring PD susceptibility. In  
27 addition, the GWAS identified several other genetic loci suggestively associated with the risk of  
28 PD; among these, only one was replicated by two different studies of European and Asian  
29 ancestries. Hence, we investigated this novel locus known as PARK16 for coding mutations in a  
30 large series of idiopathic pathologically proven PD cases in addition to performing an  
31 association study in a case-control cohort from the UK. An association between a novel *RAB7L1*  
32 mutation, c.379-12insT, and disease (Pvalue = 0.0325) was identified. Two novel coding  
33 variants present only in the PD cohort were also identified within the *RAB7L1* (p.K157R) and  
34 *SLC41A1* (p.A350V) genes. No copy number variation (CNV) analyses have yet been performed  
35 within this recently identified locus. We concluded that although both coding variants and risk  
36 alleles within the PARK16 locus seem to be rare, further molecular analyses in these genes  
37 within different populations are required in order to examine its biochemical role in the disease  
38 process.

## 40 Introduction

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

56 Genome-wide (GW) SNP genotyping assays have been proven to be a powerful technique to  
57 identify genetic risk factors in many complex disorders<sup>10</sup>. Consequently, three large PD-  
58 associated genome-wide association studies (GWAS) from two European ancestry and one  
59 Asian populations have recently identified genetic risk underlying PD, of which *SNCA* (all three  
60 studies) and *MAPT* (only European ancestry studies) loci showed the strongest evidences of  
61 association with PD<sup>11-13</sup>; these associations have recently been corroborated by a meta-analysis

62 carried out in European ancestry PD-associated GWAS<sup>14</sup>. In addition, one out of three  
63 additional genetic risk loci for PD was independently identified by two studies<sup>11-13</sup>. Therefore,  
64 we here investigate whether novel genetic variants within this locus designated as PARK16 may  
65 predispose to the risk for PD in a British cohort of pathologically proven PD cases and  
66 neurologically normal individuals. PARK16 located on chromosome 1q32 comprises 169.6kb  
67 and contains 5 different genes (Table 1).

68

## 69 **Materials and Methods**

70 Subjects: The PD cohort was collected from brain tissues at The Queen Square Brain Bank for  
71 Neurological Disorders in the UK. Cases (n=453) were clinically and pathologically diagnosed  
72 according to the PD Brain Bank criteria<sup>15, 16</sup>. The mean age at onset was 59 years (ranged from  
73 35 to 86 years) and the average of death was 78 years (ranged from 51 to 94 years). The male-  
74 to-female ratio was 3.5: 1. Family history was reported in < 1% individuals. DNA samples from  
75 82 PD cases reporting positive family history were also employed. Positive family history was  
76 compatible with the diagnosis of PD in at least one first or second degree relative. The mean  
77 age of disease onset in these familial cases was 57 years (ranged from 29 to 71 years). Patients  
78 and all relatives of patients gave informed consent for scientific research. The control cohort  
79 (n=483) analyzed here was the “1958 British birth cohort” whose individuals were all born in  
80 March 1958 in England, Scotland or Wales and which is used in all disease-related studies  
81 carried out by the Wellcome Trust Case Control Consortium (WTCCC;  
82 <http://www.b58cgene.sgul.ac.uk>).

83

84 PCR and sequencing analyses: In first instance, PCR and sequencing analyses of all open-  
85 reading frames (ORFs) of *NUCKS1* (RefSeq NM\_022731, 7 exons), *RAB7L1* (RefSeq NM\_003929,  
86 5 exons) and *SLC41A1* (RefSeq NM\_173854.4, 11 exons) genes were performed in 182 PD cases.  
87 Later, each variant identified in the PD cohort (n=9) was also analyzed in 351 neurologically  
88 normal individuals. Hereafter, every SNP showing association with the disease (n=1; c.379-  
89 12insT) and each coding variant absent in controls (n=2; p.K157R and p.A350V) were further  
90 analyzed in larger sample size compiling a total of 454 PD cases and 483 controls analyzed. In  
91 addition, the two coding variants absent in control population were also tested in 82 familial PD  
92 cases. All PCR analyses were performed using both forward and reverse genomic primers (all  
93 primer-sequences are available upon request) previously designed by ExonPrimer  
94 (<http://ihg.gsf.de/ihg/ExonPrimer.html>) and FastStart Taq DNA polymerase ([www.roche-](http://www.roche-applied-science.com)  
95 [applied-science.com](http://www.roche-applied-science.com)). Each purified product was sequenced using both forward and reverse  
96 primers with Applied Biosystems BigDye terminator v3.1 sequencing chemistry as per the  
97 manufacturer's instructions; the resulting reactions were then resolved on an ABI3730XL  
98 genetic analyzer (Applied Biosystems) and analyzed with Sequencher software 4.9 (Gene Codes  
99 Corporation).

100 The Alamut mutation interpretation software was used to look for amino acid properties and  
101 for predictions of the functional and structural effects of novel coding mutations  
102 (<http://www.interactive-biosoftware.com/alamut.html>). Multiple alignments for *RAB7L1* and  
103 *SLC41A1* encoding proteins were performed through NCBI-associated homoloGene database by  
104 using the MUSCLE program<sup>17</sup>. The Human Protein Reference Database (HPRD;

105 <http://www.hprd.org/>) was employed to search for predicted protein motifs and domains. And  
106 NCBI-BLAST database was also used to search for sequence similarities between Rab proteins;  
107 the RAB7L1 protein sequence (RefSeq: NP\_001129134.1) was aligned with, RAB1 (RefSeq:  
108 NP\_004152.1), RAB3A (RefSeq: NP\_002857.1), RAB7 (RefSeq: NP\_004628.4) and RAB8A  
109 (RefSeq: CAG38820.1) proteins (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

110

111 *Statistical analyses:* All statistical analyses (chi square tests of association and permutation  
112 analyses) were performed using the Haploview 4.1 software  
113 (<http://www.broad.mit.edu/haploview/>). In order to compare PARK16-associated allelic  
114 frequencies between diverse populations, HapMap data corresponding to the PARK16 locus  
115 from Yoruba (YRI), Japan (JPT), Han Chinese (CHB) and Northern and Western European (CEU-  
116 Utah residents) populations was also analyzed through haploview software ([www.hapmap.org](http://www.hapmap.org)).

117

## 118 **Results**

119 To try and identify novel genetic variants underlying risk for PD in a British case-control cohort,  
120 the genomic area harboring PARK16 locus was deeply investigated through sequencing  
121 analyses. In first instance, it was decided to perform sequencing analyses of all coding regions  
122 and exon-intron boundaries of genes located within the genomic area shared by both PARK16  
123 loci identified in European ancestry and Asian populations, respectively<sup>12,13</sup>; this area flanked  
124 by rs823128 (203,980,001 bp) and rs11240572 (204,074,636 bp) SNPs contained four genes  
125 (*NUCKS1*, *RAB7L1*, *SLC41A1* and *PM20D1*) (Table 1). However, *NUCKS1*, *RAB7L1* and *SLC41A1*

126 genes were located in the same LD block and were suggestively reported as the best candidates  
127 for the etiology of PD according to their functional roles<sup>12</sup>. In addition, the minor allele  
128 frequency of rs11240572 located in intron 10 of PM20D1 is < 0.03 in European ancestry  
129 population (Table 3). Hence, only the coding regions of *NUCKS1*, *RAB7L1* and *SLC41A1* were  
130 analyzed in our 182 PD cases. PCR analyses of all ORFs revealed the presence of 9 different  
131 genetic variants within *RAB7L1* (n=5) and *SLC41A1* (n=4) genes, while no genetic variation was  
132 identified across the *NUCKS1* gene. There were two coding variants (p.Gln104Glu (s41302139)  
133 and p.Lys157Arg (novel)), two novel intronic variants (c.197-49insG and c.379-12insT) and one  
134 UTR-5' variant (rs708755) among the variants identified within the *RAB7L1* gene; whereas three  
135 coding mutations (p.Thr113Thr (rs11240569), p.Asn252Asn (rs708727) and p.Ala350Val (novel))  
136 and a known intronic variant (rs41264905) were identified within the *SLC41A1* gene. All genetic  
137 variants, with the exception of both novel coding mutations which were identified in one PD  
138 patient each, were found present in both cases and controls (Table 2). The coding mutations  
139 were a heterozygous c.470A>G transition causing p.Lys157Arg and a heterozygous c.1049C>T  
140 transition causing p.Ala350Val which were respectively located within *RAB7L1* (exon 4) and  
141 *SLC41A1* (exon 8) genes (supplemental figure 1). In order to inspect whether these novel  
142 coding mutations may or may not be the disease-causing mutations, both were tested in larger  
143 sample size of additional pathologically proven idiopathic PD cases (n =272, n (total) =454), 82  
144 familial cases clinically diagnosed of PD, and 483 neurologically normal individuals; failing to  
145 detect any other mutation carrier in neither PD nor control populations. Both variants are also  
146 conserved among species (supplemental figure 1). Contradictory results were obtained with  
147 respect to the functional consequences for both novel mutations; however, the K157 amino

148 acid of *RAB7L1* was predicted to be highly conserved whereas the A350 amino-acid of *SLC41A1*  
149 was shown to be moderately conserved (Alamut software). Both clinical and pathological  
150 features of K157R and A350V mutation carriers are described in the supplemental material 1.

151 To test the hypothesis whether the remaining seven genetic variants identified may predispose  
152 to the risk for PD, they were additionally tested in 351 neurologically normal individuals. A  
153 single-marker chi square test of association was then performed. This analysis revealed a  
154 slightly significant association between the c.379-12insT mutation within the *RAB7L1* gene  
155 (intron 3) and PD (frequentist P value = 0.0325), which remained significant after one million of  
156 iterations of permutation testing to adjust for multiple comparisons (permuted P value =  
157 0.0399; Table 2).

158

## 159 **Discussion**

160 Sequencing analyses of the coding region of *NUCKS1*, *RAB7L1* and *SLC41A1* genes in a British  
161 cohort of 182 pathologically proven PD cases revealed the presence of two novel mutations, in  
162 one patient each, within *RAB7L1* (K157R) and *SLC41A1* (A350V) genes. *PARK16* locus was not  
163 examined for the presence of large CNVs. Both mutations carriers showed typical IPD and Lewy  
164 body pathology. However, even the non-occurrence of both mutations in a large sample of  
165 ethnicity-matched control individuals (n =483) does not fully disclose their pathogenicity.  
166 Seven additional intronic and exonic variants were also identified during the sequencing  
167 process. Therefore, an association study which revealed a weak association between the c.379-  
168 12insT mutation and IPD was carried out. Curiously, no intronic variation was previously

169 reported within the *RAB7L1* locus, suggesting that genetic variability within this locus is rare.  
170 Given the presence of a rare novel mutation and slightly associated risk allele within *RAB7L1*,  
171 further investigations within this gene are warranted in order to determine its precise  
172 biochemical role in the pathogenesis of PD. The *RAB7L1* encoding protein is a member of the  
173 Rab GTPases subfamily which includes a large number of small GTPases involved in intracellular  
174 cell signaling processes and vesicle trafficking. The K157 amino acid of *RAB7L1* lies in the Rab  
175 domain (8 - 176 amino acids) of the protein which is predicted to be highly conserved among  
176 species and is also conserved among other Rab proteins, such as RAB1A, RAB3A, RAB7A and  
177 RAB8A proteins (data not shown). Molecular links between PD and Rab proteins were already  
178 suggested: mutations in the Ras-like GTPase domain of dardarin cause PD<sup>9,18</sup> and elevated  
179 expression of RAB1, RAB3A and RAB8A proteins protect against alpha-syn-induced  
180 dopaminergic neuron loss in animal models of PD<sup>19,20</sup>. *SLC41A1* is a Mg (2+) transporter that  
181 may play role in magnesium homeostasis. Brain metal dyshomeostasis has often been  
182 speculated as cause for neurodegeneration; nevertheless, the precise nature of its biochemical  
183 mechanisms underlying neurodegeneration is still vague<sup>21,22</sup>.  
184 Although no association between the PARK16 locus and PD was identified in a GWAS meta-  
185 analysis<sup>14</sup>, analyses of the PARK16-associated SNPs within the HapMap data revealed marked  
186 differences in the minor allelic frequencies between populations; thus, affecting analytic power  
187 (Table 3). Likewise, population differences at the *BST1* and *MAPT* loci were recently reported  
188<sup>12,13</sup> and the haplotype H2 of *MAPT* reported to be almost exclusively of Caucasian origin is low  
189 in all populations<sup>23</sup>. By and large, different genetic markers should be used when investigating  
190 different populations as some may be not relevant to all populations. Thereafter, we conclude

191 that although pathogenic mutations and risk alleles within the PARK16 locus seem to be rare in  
192 European ancestry populations, further molecular analyses within different populations are  
193 required in order to examine its biochemical role in the PD and prior to undertake any  
194 functional work on the encoded proteins associated with this locus.

195

## 196 **Acknowledgements**

197 We thank the patients for taking part of this study and to all families who support the donation of  
198 tissue for research. We would like also to thank The Medical Research Council (MRC; HH: MRC  
199 fellowships G108/638 and G0802760 and JH: Start up funds), and The Michael J Fox Foundation (HH  
200 and CPR) for support. This study was also supported by the NIHR UCLH/UCL Comprehensive  
201 Biomedical Research Centre. Participation by MAN and ABS in this research was supported in part by  
202 the Intramural Research Program of the NIH, National Institute on Aging (AG000957-07 (2009)  
203 Assessment of Candidate Loci in Neurological diseases).

204

## 205 **Conflict of Interest Statement**

206 The authors declare they have no conflict of interest.

207

208 **References**

209 1. Lees AJ, Hardy J, Revesz T: Parkinson's disease. *Lancet* 2009; **373**: 2055-2066.  
210

211 2. Hardy J, Lewis P, Revesz T, Lees A, Paisan-Ruiz C: The genetics of Parkinson's syndromes: a  
212 critical review. *Curr Opin Genet Dev* 2009; **19**: 254-265.  
213

214 3. Winkler S, Hagenah J, Lincoln S *et al*: alpha-Synuclein and Parkinson disease susceptibility.  
215 *Neurology* 2007; **69**: 1745-1750.  
216

217 4. Skipper L, Li Y, Bonnard C *et al*: Comprehensive evaluation of common genetic variation within  
218 LRRK2 reveals evidence for association with sporadic Parkinson's disease. *Hum Mol Genet* 2005;  
219 **14**: 3549-3556.  
220

221 5. Di Fonzo A, Wu-Chou YH, Lu CS *et al*: A common missense variant in the LRRK2 gene,  
222 Gly2385Arg, associated with Parkinson's disease risk in Taiwan. *Neurogenetics* 2006; **7**: 133-138.  
223

224 6. Ross OA, Wu YR, Lee MC *et al*: Analysis of Lrrk2 R1628P as a risk factor for Parkinson's disease.  
225 *Ann Neurol* 2008; **64**: 88-92.  
226

227 7. Paisan-Ruiz C, Washecka N, Nath P, Singleton AB, Corder EH: Parkinson's disease and low  
228 frequency alleles found together throughout LRRK2. *Ann Hum Genet* 2009; **73**: 391-403.  
229

230 8. Sidransky E, Nalls MA, Aasly JO *et al*: Multicenter analysis of glucocerebrosidase mutations in  
231 Parkinson's disease. *N Engl J Med* 2009; **361**: 1651-1661.  
232

233 9. Paisan-Ruiz C: LRRK2 gene variation and its contribution to Parkinson disease. *Hum Mutat* 2009;  
234 **30**: 1153-1160.  
235

236 10. Hardy J, Singleton A: Genomewide association studies and human disease. *N Engl J Med* 2009;  
237 **360**: 1759-1768.  
238

239 11. Pankratz N, Wilk JB, Latourelle JC *et al*: Genomewide association study for susceptibility genes  
240 contributing to familial Parkinson disease. *Hum Genet* 2009; **124**: 593-605.  
241

242 12. Satake W, Nakabayashi Y, Mizuta I *et al*: Genome-wide association study identifies common  
243 variants at four loci as genetic risk factors for Parkinson's disease. *Nat Genet* 2009; **41**: 1303-  
244 1307.  
245

246 13. Simon-Sanchez J, Schulte C, Bras JM *et al*: Genome-wide association study reveals genetic risk  
247 underlying Parkinson's disease. *Nat Genet* 2009; **41**: 1308-1312.  
248

249 14. Edwards TL, Scott WK, Almonte C *et al*: Genome-Wide Association Study Confirms SNPs in SNCA  
250 and the MAPT Region as Common Risk Factors for Parkinson Disease. *Ann Hum Genet* 2010; **74**:  
251 97-109.  
252

- 253 15. Gibb WR, Lees AJ: The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's  
254 disease. *J Neurol Neurosurg Psychiatry* 1988; **51**: 745-752.  
255
- 256 16. Hughes AJ, Daniel SE, Kilford L, Lees AJ: Accuracy of clinical diagnosis of idiopathic Parkinson's  
257 disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; **55**: 181-  
258 184.  
259
- 260 17. Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput.  
261 *Nucleic Acids Res* 2004; **32**: 1792-1797.  
262
- 263 18. Greggio E, Cookson MR: Leucine-rich repeat kinase 2 mutations and Parkinson's disease: three  
264 questions. *ASN Neuro* 2009; **1**: e00002.  
265
- 266 19. Cooper AA, Gitler AD, Cashikar A *et al*: Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues  
267 neuron loss in Parkinson's models. *Science* 2006; **313**: 324-328.  
268
- 269 20. Gitler AD, Bevis BJ, Shorter J *et al*: The Parkinson's disease protein alpha-synuclein disrupts  
270 cellular Rab homeostasis. *Proc Natl Acad Sci U S A* 2008; **105**: 145-150.  
271
- 272 21. Bolognin S, Messori L, Zatta P: Metal ion physiopathology in neurodegenerative disorders.  
273 *Neuromolecular Med* 2009; **11**: 223-238.  
274
- 275 22. Kolisek M, Launay P, Beck A *et al*: SLC41A1 is a novel mammalian Mg<sup>2+</sup> carrier. *J Biol Chem*  
276 2008; **283**: 16235-16247.  
277
- 278 23. Evans W, Fung HC, Steele J *et al*: The tau H2 haplotype is almost exclusively Caucasian in origin.  
279 *Neurosci Lett* 2004; **369**: 183-185.  
280  
281  
282

284 **Legend to supplemental figure 1**

285

286 **Supplemental figure 1: Upper panel:** Chromatograms of the sequences showing both novel  
287 mutations identified in *RAB7L1* and *SCL41A1* genes. **Lower panel:** Tables showing conservation  
288 of both lysine and alanine amino-acids among different species. *RAB7L1* (Hs: NP\_001129134.1,  
289 Pt: XP\_001162428.1, Clf: XP\_536104.2, Bt: NP\_001092564.1, Mm: NP\_659124.1, Gg:  
290 XP\_417967.2); *SLC41A1* (Hs: NP\_776253.3, Pt: XP\_525038.2, Clf: XP\_536105, Bt: XP\_613469.2,  
291 Mm: NP\_776290.1, Gg: XP\_417968.2,).

292

293 **Legends to tables**

294 **Table 1:** Previously reported PD-associated SNPs within the *PARK16* locus. SNPs showing  
295 association with the disease in the Japanese case-control cohort are represented in black; SNPs  
296 showing association with PD in both European ancestry case-control cohorts are shown in grey  
297 <sup>12,13</sup>. The SNPs showing the highest P value for each independent study are highlighted in bold.

298 *RAB7L1* chromosomal localization: 204,003,738 – 204,011,233bp. *RAB7L1* is located in the  
299 27.6kb interval between rs823122 (at ~12kb away) and rs947211 (at ~8kb away) SNPs.

300 **Table 2:** Chi square associations tests for common variants identified within *PARK16* locus  
301 performed by haploview software. The only variant (c.379-12insT) which showed association  
302 with the disease is highlighted in bold; the  $1 \times 10^6$  permutation value associated to this variant  
303 is also shown in brackets. The case/control frequencies for each variant are also shown.

304

305 **Table 3:** PARK16 core SNPs frequencies in diverse populations. CEU: CEPH (Utah residents with  
306 ancestry from northern and western Europe); CHB: Han Chinese in Beijing, China; JPT: Japanese  
307 in Tokyo, Japan; YRI: Yoruba in Ibadan, Nigeria. MAF = minor allele frequencies. m = minor  
308 allele, M = major allele.

309

310

**Tables : Genetic variability at the PARK16 locus (143-10-EJHG)**

<b>PARK16 locus (SNPs)</b>	<b>Chr</b>	<b>Position (bp)</b>	<b>Alleles (minor/major)</b>	<b>P values (combined Stage I and Stage II)</b>	<b>Chr Localization</b>
rs16856139	1	203905087	T/C	1.02 x10 <sup>-07</sup>	SLC45A3
<b>rs823128</b>	1	<b>203980001</b>	<b>G/A</b>	<b>7.29 x10<sup>-08</sup></b>	<b>NUCKS1</b>
rs823122	1	203991651	C/T	4.88 x10 <sup>-09</sup>	Genomic region
<b>rs947211</b>	<b>1</b>	<b>204019288</b>	<b>A/G</b>	<b>1.52 x10<sup>-12</sup></b>	<b>Genomic region</b>
rs823156	1	204031263	G/A	7.60 x10 <sup>-04</sup>	SLC41A1
rs823156	1	204031263	G/A	3.60 x10 <sup>-09</sup>	SLC41A1
rs708730	1	204044403	G/A	2.43 x10 <sup>-08</sup>	SLC41A1
rs11240572	1	204074636	A/C	6.11 x10 <sup>-07</sup>	PM20D1
rs11240572	1	204074636	A/C	1.08 x10 <sup>-07</sup>	PM20D1

Table 1: Previously reported PD-associated SNPs within the PARK16 locus. 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 grey (Satake, et al., 2009; Simon-Sanchez, et al., 2009). The SNPs showing the highest P value for each independent study are highlighted in bold. *RAB7L1* chromosomal localization: 204,003,738 – 204,011,233bp. *RAB7L1* is located in the 27.6kb interval between rs823122 (at ~12kb away) and rs947211 (at ~8kb away) SNPs.

<b>Gene</b>	<b>Chr</b>	<b>Mutation</b>	<b>Position (bp)</b>	<b>Associated</b>		<b>P value</b>	<b>Case/Control frequencies</b>
				<b>Allele</b>	<b>Chi square</b>		
RAB7L1	1	rs708725	204010761	A	0.609	<b>0.4351</b>	0.443, 0.418
RAB7L1	1	c.197-49insG	204007453	Ins G	0.949	<b>0.3301</b>	0.072, 0.057
RAB7L1	1	rs41302139	204007291	C	0.075	<b>0.7844</b>	0.019, 0.017
<b>RAB7L1</b>	<b>1</b>	<b>c.379-12insT</b>	<b>204006575</b>	<b>Ins T</b>	<b>4.573</b>	<b>0.0325 (0.0399)</b>	<b>0.013, 0.004</b>
SLC41A1	1	rs708727	204034508	T	0.776	<b>0.3783</b>	0.438, 0.409
SLC41A1	1	rs41264905	204034656	T	1.402	<b>0.2364</b>	0.009, 0.003
SLC41A1	1	rs11240569	204045854	T	0.971	<b>0.3243</b>	0.301, 0.270

Table 2: Chi square associations tests for common variants identified within PARK16 locus performed by haploview software. The only variant (c.379-12insT) which showed association with the disease is highlighted in bold; the 1 x 10<sup>6</sup> permutation value associated to this variant is also shown in brackets. The case/control frequencies for each variant are also shown.

<b>CEU (SNPs)</b>	<b>Position (bp)</b>	<b>ObsHET</b>	<b>PredHET</b>	<b>HWpval</b>	<b>MAF</b>	<b>Alleles (m:M)</b>
rs16856139	203905087	0.094	0.09	1	0.047	C:T
rs823128	203980001	0.034	0.034	1	0.017	A:G
rs823122	203991651	0.077	0.09	0.4364	0.047	T:C
rs947211	204019288	0.376	0.364	0.9746	0.239	G:A
rs823156	204031263	0.308	0.295	0.938	0.179	A:G
rs708730	204044403	0.222	0.248	0.3983	0.145	A:G
rs11240572	204074636	0.043	0.042	1	0.021	C:A
<b>CHB + JPT (SNPs)</b>	<b>Position (bp)</b>	<b>ObsHET</b>	<b>PredHET</b>	<b>HWpval</b>	<b>MAF</b>	<b>Alleles (m:M)</b>
rs16856139	203905087	0.224	0.217	1	0.124	C:T
rs823128	203980001	0.253	0.247	1	0.144	A:G
rs823122	203991651	0.253	0.247	1	0.144	T:C
rs947211	204019288	0.494	0.494	1	0.447	A:G
rs823156	204031263	0.365	0.347	0.7025	0.224	A:G
rs708730	204044403	0.371	0.35	0.6318	0.226	A:G
rs11240572	204074636	0.324	0.294	0.3199	0.179	C:A
<b>YRI (SNPs)</b>	<b>Position (bp)</b>	<b>ObsHET</b>	<b>PredHET</b>	<b>HWpval</b>	<b>MAF</b>	<b>Alleles (m:M)</b>
rs16856139	203905087	0.087	0.159	5 x10 <sup>-4</sup>	0.087	C:T
rs823128	203980001	0.522	0.472	0.381	0.383	A:G
rs823122	203991651	0.478	0.499	0.7532	0.483	C:T
rs947211	204019288	0.487	0.476	1	0.391	A:G
rs823156	204031263	0.426	0.427	1	0.309	G:A
rs708730	204044403	0.217	0.258	0.1759	0.152	G:A
rs11240572	204074636	0	0	1	0	C:C

Table 3: PARK16 core SNPs frequencies in diverse populations. 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.