



Editor's choice  
Scan to access more  
free content

## ORIGINAL ARTICLE

# Discovery of a frameshift mutation in podocalyxin-like (PODXL) gene, coding for a neural adhesion molecule, as causal for autosomal-recessive juvenile Parkinsonism

Sumedha Sudhaman,<sup>1</sup> Kameshwar Prasad,<sup>2</sup> Madhuri Behari,<sup>2</sup> Uday B Muthane,<sup>3</sup> Ramesh C Juyal,<sup>4</sup> BK Thelma<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2015-103459>).

<sup>1</sup>Department of Genetics, University of Delhi South Campus, New Delhi, India

<sup>2</sup>Department of Neurology, All India Institute of Medical Sciences, New Delhi, India

<sup>3</sup>Parkinson's and Ageing Research Foundation, Bengaluru, Karnataka, India

<sup>4</sup>Regional Center for Biotechnology, Faridabad, Haryana, India

## Correspondence to

Professor B K Thelma, Department of Genetics, University of Delhi South Campus, Benito Juarez Marg, New Delhi 110021, India; [thelma@delhi.ac.in](mailto:thelma@delhi.ac.in)

Received 15 August 2015

Revised 1 January 2016

Accepted 18 January 2016

Published Online First

10 February 2016

## ABSTRACT

**Background** Mutations in known genes for inherited forms of Parkinson's disease (PD) account for <30% of familial PD (FPD) implying that more causal gene(s) remain to be identified. We attempted to discover the putative causal variant in an Indian family with autosomal-recessive juvenile Parkinsonism (ARJP), tested negative for mutations in *PARK2*, *PINK1* and *DJ1*.

**Methods** Whole exomes of two affected siblings were sequenced. Variants prioritised were screened for segregation with disease in the family by targeted sequencing. Gene thus identified was screened for index/additional exonic mutations, if any, in an independent PD cohort by PCR sequencing. Variants observed were functionally validated in differentiated PC12 cells.

**Results** A novel homozygous frameshift mutation, c.89\_90insGTCGCCCC in exon 1 of podocalyxin-like gene (*PODXL*, 7q32–33), resulting in loss of protein, segregated with disease in the family. Mutant allele was absent in 186 healthy controls screened by PCR sequencing and in control exomes available in the laboratory and public databases. Screening of additional 212 sporadic and 68 FPD cases identified three novel heterozygous missense variants namely c.1285C>A, c.1118G>A and c.881G>A in three unrelated cases. Significant differences in neurite branching and length ( $p<0.0001$ ) were observed in PC12 cells with wild-type and mutant constructs.

**Conclusions** Based on the genetic and functional evidence in this study and literature support on the role of *PODXL* in neural development, a novel frameshift mutation in *PODXL* seems to be the likely cause of ARJP in this family. This is the first report suggesting the possible role of a neurodevelopmental pathway in PD aetiology.

## INTRODUCTION

Selective degeneration of dopaminergic neurons in the *substantia nigra pars compacta* of the midbrain is the primary pathology underlying Parkinson's disease (PD), the second most common neurodegenerative disorder. Despite intense research efforts, the exact mechanism of this degeneration, which occurs mostly in a sporadic manner in the elderly population, remains unknown. Xenobiotics were believed to be the primary contributors to this most common form of movement disorder,<sup>1</sup> but a genetic basis with several candidate genes has since

been shown to be associated with PD.<sup>2–3</sup> These genes come from a number of pathways including the dopaminergic pathway, ubiquitin proteasome system, synaptic homeostasis, autophagy, lysosomal pathway, mitochondrial maintenance, axonal transport, programmed cell death and inflammatory pathways, all of which have been implicated to lead to neuronal cell death.<sup>4–5</sup> Unlike a polygenic contribution from these pathways to the common sporadic forms of PD, mutation(s) in single genes are known to be causal in Mendelian forms of the disease, which constitute ~10% of all PD cases. Genome-wide linkage scans in well-characterised PD families were successful in discovering six genes namely *SNCA*, *PARK2*, *PINK1*, *DJ1*, *LRRK2* and *EIF4G1*.<sup>6–7</sup> More recently, next-generation sequencing (NGS) technology has helped to identify three more genes namely *VPS35*,<sup>8</sup> *DNAJC13*<sup>9</sup> and *CHCHD2*.<sup>10</sup> Though the contribution of *DNAJC13* to familial and sporadic PD across different populations remains to be investigated, mutations in the other eight genes account for <30% of all familial forms of the disease examined to date,<sup>7–11</sup> though with varied frequency across ethnic groups. For example, mutations in *LRRK2* were seen in 30%–42% and 2.6% of autosomal-dominant PD cases from Tunisia and USA, respectively,<sup>12</sup> but 0.3% in Japanese<sup>13</sup> and <0.1% in North Indian PD cases.<sup>14</sup> On the other hand, mutations in *PARK2* explain up to 50% familial PD (FPD) cases and ~15% sporadic PD cases globally,<sup>15</sup> but are seen in ~14.3% of FPD and 6.9% in young-onset and 5.9% in late-onset sporadic cases from India.<sup>16</sup> It is also noteworthy that almost all these FPD genes known to date are involved only in the ubiquitin proteasomal degradation pathway, mitochondrial maintenance or autophagy-lysosome pathway.<sup>4</sup> It is very likely that novel mutations in more genes from these and other known or yet unknown pathways remain to be identified to fully understand PD biology. Analysis of FPD cases continues to be one of the major approaches for discovery of additional causal variants, and the power of the recent NGS tools, irrespective of family size, makes this an attractive strategy.

With this background, we carried out exome sequencing of a 10-member Indian family with autosomal-recessive juvenile Parkinsonism (ARJP)



CrossMark

**To cite:** Sudhaman S, Prasad K, Behari M, et al. *J Med Genet* 2016;**53**:450–456.

and which was negative for mutations in three known ARJP genes namely *PARK2*, *PINK1* and *DJ1*. We identified a novel, putative disease-causing variant in podocalyxin-like gene (*PODXL*), coding for a neural adhesion molecule, and not implicated in PD aetiology to date.

METHODS

Recruitment of study subjects

A 10-member ARJP family (figure 1) was recruited from Aligarh, Uttar Pradesh, India by clinicians KP and MB. PD was diagnosed using the UK PD Brain Bank clinical diagnostic criteria.<sup>17</sup> There was a family history of consanguinity, though the parents (I1, I2, figure 1) and five remaining siblings (II1, II2, II6, II7 and II8, figure 1) were unaffected. A detailed description of the clinical features of all three affected siblings is presented in table 1. Following clearance from ethical committees of participating institutions, informed consent and 10 mL peripheral venous blood was obtained from each of the study subjects.

Genetic analysis

Genomic DNA was isolated using the conventional phenol chloroform extraction method used in the laboratory. Mutation screening in all exons of *PARK2*, *PINK1* and *DJ1* was carried out by PCR sequencing using primers described elsewhere<sup>16</sup> and listed in the online supplementary tables S1 and S2.

NGS and data analysis

Exome sequencing of two (II4, II5, figure 1) of the three affected siblings aged 16 and 13 years at disease onset, respectively, from the family under study was carried out. Library preparation and enrichment using NimbleGen V3 kit and sequencing on Illumina HiSeq2000 were done using a commercial facility (Axeq Technologies, USA). Raw data obtained were subjected to stringent pre-alignment quality control (QC) for quality-based filtering using FASTX-Toolkit,<sup>18</sup> alignment by BWA<sup>19</sup> and post-alignment QC with SAMtools.<sup>20</sup> Target coverage and depth (see online supplementary table S3) were assessed by BEDTools,<sup>21</sup> followed by variant-calling using SAMtools,<sup>20</sup> and GATK,<sup>22</sup> and annotation using wANNOVAR.<sup>23</sup>

All SNPs with minor allele frequency (MAF) >0.01 reported in databases such as dbSNP135, 137 and 138,<sup>24</sup> 1000 genomes<sup>25</sup> and NHBLI 6500 exomes<sup>26</sup> were excluded to start with. Keeping in mind the autosomal-recessive inheritance pattern, homozygous exonic variants shared between the two affected individuals were prioritised. In parallel, KGGseq<sup>27</sup> was also used to filter variants (see online supplementary table S4).

Prioritised variants (see online supplementary table S5) were then checked for segregation with the disease in the remaining family members by targeted sequencing using HaloPlex enrichment kit (Agilent, USA) on Illumina HiSeq2500 using a commercial facility (Strand Genomics, India). Index mutation(s)

were screened in unrelated control individuals (n=186) by PCR-Sanger sequencing; and checked in 144 healthy control exomes and 74 non-PD exomes (of subjects with other disorders), available in the laboratory.

Further, index mutation(s) and additional exonic variants, if any, in the gene(s) segregating with the disease phenotype were screened in a PD cohort (n=212 sporadic and 68 familial cases) available in the laboratory (see online supplementary table S6 for primer details) by PCR-Sanger sequencing. All the variant(s) identified as above were then taken forward for suitable functional studies *in vitro*.

In vitro functional analysis

Transfection and generation of stable PC12 cells

Stably transfected PC12<sup>28</sup> cells were generated with the wild-type (WT) and mutant open reading frame (ORF) constructs of the gene of interest using a standard protocol. Briefly, TrueORFGold expression-validated ORF clone of gene of interest with a C-terminal myc-DDK-tag was obtained in a pCMV6 vector (RC210816, OriGene Technologies, Rockville, Maryland, USA). Site-directed mutagenesis to generate the mutant(s) (see online supplementary table S7 for primer details)

Table 1 Clinical characteristics of affected individuals in the study family with ARJP

Parameter	II3	II4*	II5*
Relationship	Sister	Proband	Brother
Affection status	Affected	Affected	Affected
AAO	17 years	16 years	13 years
AAE	22 years	20 years	17 years
Gait disturbance	+	+	+
Tremors	+	+	+
Slowness	+	+	+
Stiffness	+	+	+
Loss of balance	+	+, 3 years after onset	+, 2 years after onset
Falls	+	+, 3 years after onset	+, 2 years after onset
LD/CD response	40%–50% improvement	50% improvement	40% improvement
Dystonia	+, in 'off' period	+	+
Dyskinesia	+, 4 years after onset	+	+
Rigidity	+, 4/4 on UPDRS III	+, 3/4 on UPDRS III	+, 3/4 on UPDRS III
Bradykinesia	nt	nt	2/4 during 'on' state
Postural instability	nt	nt	nt
a.d.I	Dependant	Dependant	Dependant
MMSE	nt	nt	28/28

II3, II4, II5, birth and milestones normal; no psychosis/hallucinations/urinary symptoms/postural hypotension/episodes of drenching/sweating; no history of exposure to drugs/toxins; no history of repeated head trauma/febrile illnesses associated with loss of consciousness/alertness; other systems in CNS and systemic examination normal. All the siblings investigated for secondary causes—complete blood count, liver, renal function tests, calcium, phosphorus, alkaline phosphatase, peripheral blood smear examination, copper studies and MRI profiles normal. II3 died 5 years after disease onset.

\* Exome sequenced.  
+, symptom present; a.d.I, activities for daily living; AAE, age at time of examination; AAO, age at disease onset; ARJP, autosomal-recessive juvenile Parkinsonism; II3, II4, II5, three affected siblings from family with ARJP; LD/CD, L-dopa/carbidopa combination dose; MMSE, Mini Mental State Examination; nt, not tested (as subject was physically unable to perform required tests); UPDRS III, Unified Parkinson's Disease Rating Scale III.

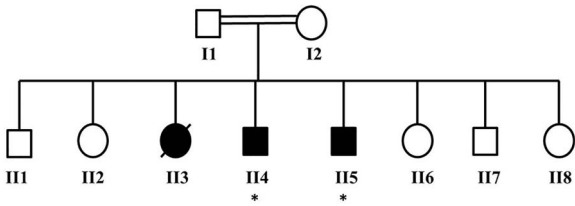


Figure 1 Pedigree of an Indian family with autosomal-recessive juvenile Parkinsonism. I1, I2 unaffected parents; II3, II4, II5 affected siblings; II1, II2, II6, II7, II8 unaffected siblings; \* exome sequenced.

was performed as per Stratagene Quikchange II XL kit-protocol from this WT ORF along with subsequent transformation of XL10Gold *Escherichia coli* cells followed by confirmation of the mutation(s) by Sanger sequencing. Stable PC12 cells for each of the constructs were generated by transfection-mediation using Lipofectamine LTX and Plus reagent (Thermo Fisher Scientific) and G418 selection. Transfections and subsequent stable-line generations were done twice per construct to obtain biological replicates. Untransfected (UT) PC12 cells were used as controls.

#### Protein expression confirmation

Expression of WT and mutant proteins from the respective stable PC12 cells was confirmed by western blot analysis using standard protocols. Briefly, equal quantities of total cell lysates prepared from the respective stable lines, after protein estimation, were run on 12% sodium dodecyl sulfate-polyacrylamide gels, blotted and probed with appropriate antibodies.  $\beta$ -actin was used as an internal control. Antibodies used were mouse monoclonal anti-DDK-tag (TA50011-100, OriGene Technologies) diluted to 1:2000; rabbit polyclonal anti- $\beta$ -actin (ab8227, Abcam) diluted to 1:2000; goat anti-mouse IgG-horseradish peroxidase (HRP) conjugate (32430, Thermo Fisher Scientific) diluted to 1:5000; and goat anti-rabbit IgG-HRP conjugate (ab97200, Abcam) diluted to 1:10 000.

#### Neurite branching assay

Eight thousand to 10 000 cells from each of the stable PC12 lines were seeded per chamber on collagen-coated four-well chamber slides (SPL Lifesciences, Korea) and grown for 24 h in 1× Dulbecco's Modified Eagle medium (DMEM) supplemented with 10% horse serum, 5% fetal bovine serum and additional supplements. UT PC12 cells were used as controls. After 24 h, medium was changed to reduced-serum differentiation medium (1× DMEM+1% horse serum+100 mg/mL nerve growth factor). Differentiation medium was changed every 48 h. Cells were allowed to differentiate for 4 days, and then permeabilised and fixed on day 5 with methanol. Slides were then prepared for immunofluorescence by treating with primary (mouse monoclonal anti-DDK-tag, TA50011-100 OriGene Technologies, diluted to 1:1000) and secondary (donkey anti-mouse IgG-AlexaFluor594, 715-585-151, Jackson ImmunoResearch, diluted to 1:500) antibodies and visualised on Olympus BX51 fluorescence microscope. Images captured on 20× were then analysed using IMAGEJ2 software<sup>29</sup> for two parameters, namely total neurite length per cell and total number of branch points per cell. At least 80 cells per category were analysed for both these parameters. Mann-Whitney test was performed to generate p values between WT and each of the mutants and p values <0.05 were considered to be significantly different. All statistical analyses were done using GraphPad Prism trial V6.01 for Windows, GraphPad software, La Jolla, California, USA.

## RESULTS

### Novel frameshift mutation in *PODXL* identified as a potential causal variant in the family with ARJP

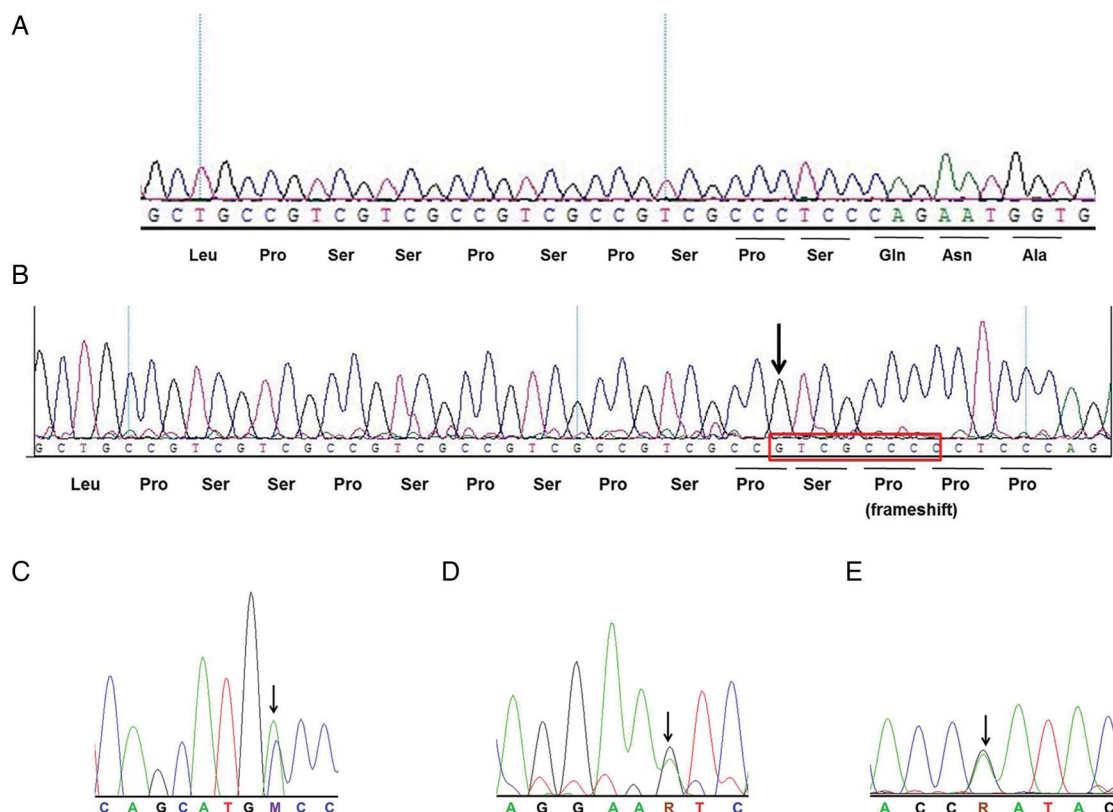
The 10-member family with two unaffected parents, three affected and five unaffected children included in this study were screened and found to be negative for mutations in *PARK2*, *PINK1* and *DJ1*. The oldest affected sibling (II3, figure 1) died after recruitment but prior to exome sequencing in this family (at the age of 22 years). Exome sequencing of the two younger affected brothers (aged 16 and 13 years at disease onset) yielded total target coverage (>10×) of 93.9% and 72.9×, and 64.3× mean target depth, respectively (see online supplementary

table S3). Downstream analysis with optimal stringency and removal of shared variants with frequencies >0.01 from different databases (as detailed in 'Methods' section) was done, which yielded 17 homozygous shared exonic variants in 17 different genes (see online supplementary table S5). On segregation analysis of these variants in the remaining family members by targeted sequencing using a customised HaloPlex enrichment kit (Agilent), a single novel, homozygous frameshift mutation in *PODXL* was found to be shared among all three affected siblings and thus segregating with the disease in this family (figure 2A, B). Both parents (I1, I2, figure 1) were found to be heterozygous for this change, and of the remaining five unaffected children, three (II1, II2 and II7) were heterozygous for this index mutation, and two (II6, II8) had homozygous WT alleles (figure 2A). This mutation, a tandem repeat expansion of the six bp motif namely GTCGCC followed by insertion CC (c.89\_90insGTCGCCCC, nomenclature as per Human Genome Variation Society (HGVS) convention), causes a frameshift immediately after a proline-serine stretch at Ser31, in the WT protein. While the proline-serine expansion (n=4) observed in affected individuals is a reported polymorphism (rs759639123, MAF=0.3), the frameshift insertion in exon 1 of the gene leads to a stop codon gain at codon 168 of the 558 (NM\_001018111.2, isoform 1) or 526 (NM\_005397.3, isoform 2) codon transcripts from the gene (see online supplementary figure S1). This change in the mRNA would trigger nuclear nonsense-mediated decay mechanism leading to absence of protein in the affected members with the homozygous variant. *PODXL*, or podocalyxin-like, located on chromosome 7q32-33 (with nine exons and two isoforms), is a known sialomucin, expressed largely in the brain, heart and kidneys.<sup>30</sup> There are 136 organisms having orthologs with this gene. Loss of function of the mouse ortholog of *PODXL* has been previously reported to distort neuronal branching,<sup>31</sup> which may have debilitating implications in neuronal development and neurodegenerative diseases. Thus, the observed novel mutation in *PODXL* may be potentially causing ARJP in the family under study.

### Additional heterozygous mutations in *PODXL* in an unrelated PD cohort

On screening all eight coding exons of *PODXL* by PCR sequencing in 280 PD cases from the PD DNA bank available in the laboratory, we identified three novel heterozygous missense mutations. Mutation NM\_005397.3:c.1285C>A:p.P429T (P429T) in exon 7 (figure 2C) was found in a young-onset PD (YOPD) male with age at disease onset (AAO) of 22 years and age at time of examination (AAE) of 25 years. Mutation NM\_005397.3:c.1118G>A:p.S373N (S373N) in exon 5 (figure 2D) was found in a late-onset PD (LOPD) male with AAO of 71 years and AAE of 73 years. Mutation NM\_005397.3:c.881G>A:p.R294Q (R294Q) in exon 3 (figure 2E) was observed in an FPD male with AAO of 42 years and AAE of 62 years. This FPD subject has an affected maternal uncle (deceased) and an affected maternal first cousin. All three unrelated heterozygous carriers of these novel mutations are from North India and have classical PD symptoms. They have also been previously screened and found negative for mutations in *PARK2*,<sup>16</sup> *PINK1* and *DJ1* (unpublished data). The three heterozygous mutations and the index mutation were absent in 186 age-matched unrelated control individuals screened by PCR-Sanger sequencing, 144 unrelated healthy control exomes and 74 non-PD exomes (of subjects with other disorders) available in the laboratory. These variants have also not been reported in over 60 000 unrelated non-PD exomes from Exome





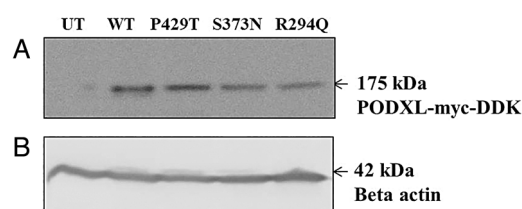
**Figure 2** Electropherograms depicting wild-type (WT)/mutant sequences in podocalyxin-like (*PODXL*). (A) Homozygous WT sequence found in two unaffected siblings and unrelated controls. (B) Homozygous frameshift mutation (c.89\_90insGTCGCCCC) found in three affected individuals; arrow indicates site of frameshift; red box indicates insertion sequence; (C) c.1285C>A, p.P429T denoted by 'S' in a young-onset Parkinson's disease case; (D) c.1118G>A, p.S373N denoted by 'R' in a late-onset PD case; and (E) c.881G>A, p.R294Q denoted by 'R' in a familial Parkinson's disease case; arrows indicate sites of mutations.

Aggregation Consortium (ExAC) database.<sup>32</sup> All the four mutations (SCV000258390, SCV000258391, SCV000258392 and SCV000258393) are in exons present in both isoforms of *PODXL* (see online supplementary figure S1). *In silico* prediction using SIFT,<sup>33</sup> PolyPhen2<sup>34</sup> and MutationTaster<sup>35</sup> suggested the likely pathogenic nature of the three non-synonymous mutations identified.

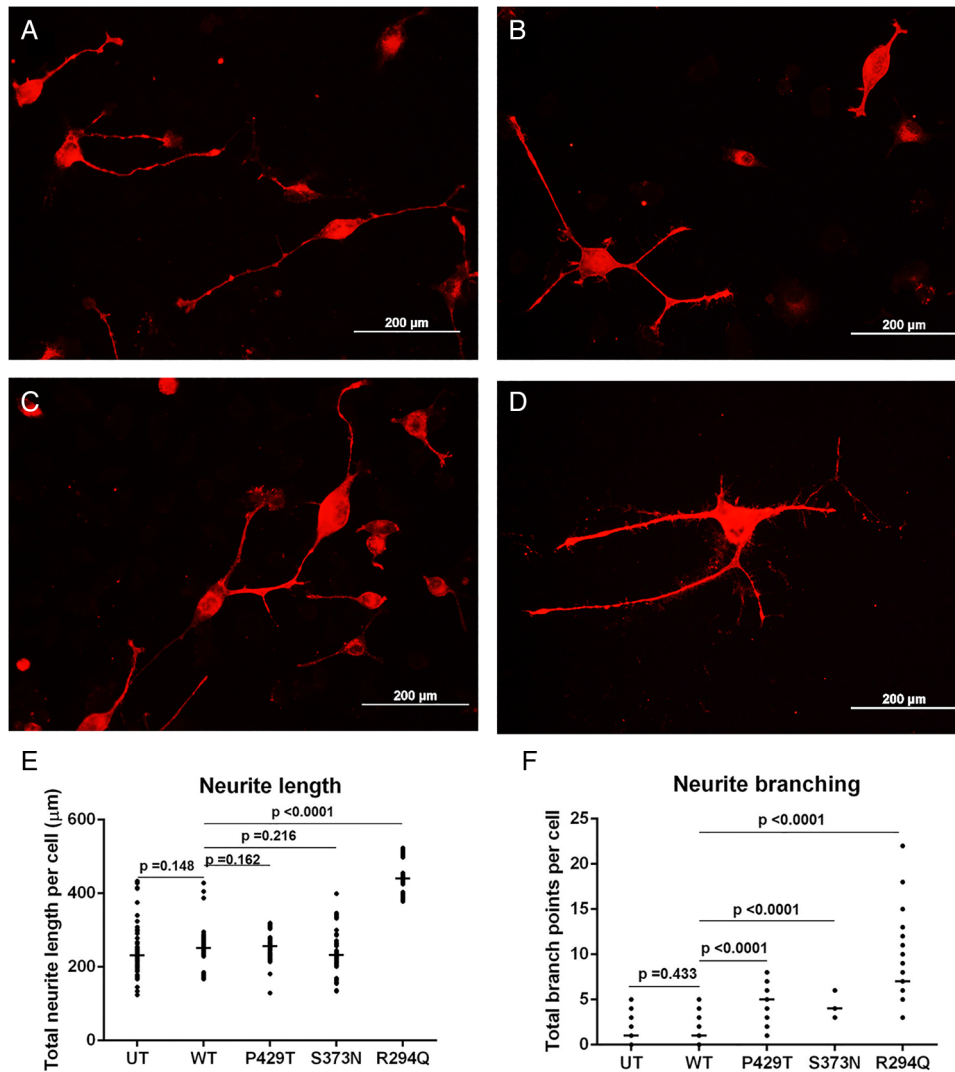
### Novel mutations in *PODXL* alter neurite branching pattern in PC12 cells differentiated into neurons

The homozygous frameshift mutation in the family under study, which would lead to a total loss of protein, as described above, was not analysed further. *PODXL* is mostly expressed in the brain, heart and kidney. Due to non-availability of either the tissue samples from the affected individuals or the autopsy samples from the deceased affected sibling, absence of protein

in these affected individuals could not be demonstrated. Further, it may be noted that loss of protein cannot be investigated in an overexpression model due to endogenous *PODXL* expression but a knock-out strategy would be useful. However, such a knock-out mouse model showing increased neurite outgrowth and branching has already been reported by another group.<sup>31</sup> This evidence lends substantial support for the functional significance of this protein. Based on this knowledge, the functional significance of the three additional novel mutations identified in our PD cohort was investigated *in vitro* using the neurite branching assay. This was done by first generating stable lines of the WT *PODXL* ORF and three mutant ORFs, namely P429T, S373N and R294Q, in PC12 cells after protein expression in all the lines was confirmed by western blotting (figure 3). After 5 days of differentiation into neurons, the neurite branching profile, scored for total neurite length and total number of branch points per cell, was compared between the WT and each of the three mutants (figure 4). Mutant R294Q (observed in FPD) showed a significant increase (Mann-Whitney  $p < 0.0001$ ) in both total length of neurites per cell and in the total number of branch points per cell. Mutants P429T and S373N showed a significant increase (Mann-Whitney  $p < 0.0001$ ) in the total number of branch points per cell, but no significant change in total neurite length per cell (figure 4). There was no significant difference between WT and UT control cells scored for the same parameters, which can be explained by the endogenous expression of *PODXL* in the rat cells. These findings are suggestive of the functional significance of the mutations in *PODXL*.



**Figure 3** Western blot profile confirming protein expression. (A) Podocalyxin-like (*PODXL*)-myc-DDK expression in wild-type (WT) and mutant lines P429T, S373N and R294Q; (B)  $\beta$ -actin expression across untransfected (UT), WT, P429T, S373N and R294Q lines.



**Figure 4** Neurite branching profile in wild-type (WT) and mutant stable PC12 cells differentiated into neurons. (A–D) Representative images captured under 20× objective, on Olympus BX51 fluorescence microscope. Stable PC12 lines expressing podocalyxin-like (PODXL)-myc-DDK: (A) WT. (B–D) Mutants P429T, S373N and R294Q, respectively. (E and F) Median (denoted by short horizontal lines for each group) with aligned dot plots. (E) Total neurite length per cell. (F) Total number of branch points per cell.

## DISCUSSION

Mendelian forms of PD have undoubtedly a major genetic component compared with the predominant late-onset sporadic PD cases.<sup>36</sup> Notable early efforts in PD gene discovery using genome-wide linkage scans identified three genes, namely *PARK2*, *PINK1* and *DJ1*, for ARJP. These have been extensively investigated across ethnic groups. However, mutations in these genes explain <15% of ARJP cases in India,<sup>16 37 38</sup> and ~1%–50% of ARJP cases and ~1%–20% of YOPD cases in Caucasians<sup>15</sup> implying that more disease-causing genes remain to be discovered. This is not unexpected, considering that PD is a genetically and clinically heterogeneous condition and genes from a range of pathways could contribute to the etiopathology of this neurodegenerative condition. It is at this juncture that the hypothesis-free exome sequencing approach has become a powerful tool and PD families (as well as families with other monogenic disorders), irrespective of their size, are an enviable resource to facilitate new disease-causing gene variant discovery.

Using this approach of exome sequencing in an ARJP family from North India, we identified a novel putative disease-causing variant (c.89\_90insGTCGCC) in *PODXL*, a gene whose

contribution to PD is hitherto unreported. As already mentioned in the ‘Results’ section, the index mutation led to loss of protein and, therefore, could not be validated in this overexpression study in PC12 cells differentiated into neurons. However, as mentioned earlier, loss of function of the mouse ortholog, in a knock-out study reported previously, has shown increased neurite outgrowth and branching.<sup>31</sup> This may be considered as evidence for the role of *PODXL* in the pathology observed in the study family.

*PODXL* is a highly sialylated adhesion glycoprotein, abundantly expressed in the brain, second only to the neural cell adhesion molecule (NCAM).<sup>31</sup> Several studies have demonstrated the role NCAM plays in modulating neural plasticity and development.<sup>39–41</sup> NCAM bound to polysialic acid (PSA) is known to regulate axonal fasciculation<sup>42</sup> and synaptic stabilisation.<sup>43</sup> To some extent, an early *in vivo* study on *PODXL*-deficient mice<sup>31</sup> has been insightful in understanding whether sialylated and non-sialylated forms of neural *PODXL* play a similar role to PSA-NCAM in neural development. The authors were able to show that though PSA-bound forms of *PODXL* were an essential negative regulator of neurite

outgrowth, similar to PSA-NCAM, non-sialylated PODXL was required for regulation of neurite branching. They reported increased neurite outgrowth and branching, and a reduction in the number of synapses formed in hippocampal and cortical neurons isolated from brains of these PODXL-deficient mice, suggesting this protein to be a negative regulator of neurite outgrowth and branching, with a direct involvement in synaptogenesis. The same study demonstrated interactions between PODXL from brains of mice, and sodium–hydrogen exchanger regulatory factors 1,2 (NHERF1,2), Ezrin and G-coupled protein and GTPase Rho A,G, thereby proposing that negative regulation of neurite outgrowth and branching was due to PODXL-mediated activation of NHERF1,2/Ezrin/RhoA,G complex. Independently, in another study, MRI of brains in PODXL-deficient mice has indicated a potential role of this protein in neuronal development.<sup>44</sup> In the same study, the absence of PODXL in the aetiology of schizophrenia and other neurodegenerative diseases has also been discussed. A significant ventricular enlargement was reported to correlate with the ablation of PODXL in mouse brains. In other studies, such ventricular enlargement was seen to correspond with cognitive decline,<sup>45</sup> and motor asymmetry<sup>46</sup> in PD cases, which they believed may serve as a marker for neurodegenerative disease progression. Interestingly, a reduction in sialic acid and sialylated membrane-bound molecules has been reported to trigger inflammatory responses in the brain by microglial activation.<sup>47</sup> The downstream neurotoxic effects of such microglia-mediated inflammatory pathways have been implicated as one of the mechanisms of PD pathogenesis.<sup>47</sup> Therefore, it may not be unexpected, that a loss of PODXL, a polysialylated cell adhesion molecule, might lead to PD through a microglia-activated immune response. These speculations are amply corroborated by our own findings of a frameshift mutation leading to loss of protein in three affected siblings, all of whom had a debilitating form of juvenile-onset PD, including the reported death of the oldest sibling during the study period.

On the other hand, on screening for the index mutation and/or additional variants in *PODXL*, we found three novel, although heterozygous, mutations one each in YOPD, LOPD and FPD cases whose ages at onset were considerably later than in the index family. These mutations were predicted to be functionally relevant using *in silico* tools. The experimental observations of alteration in neurite branching pattern, either in the total neurite length or total number of branch points or both (figure 4), support the *in silico* predictions. Though these findings are promising, the mechanism underlying *PODXL* variants influencing neurite branching is not yet well understood. In this study, we have been able to compare the neurite branching pattern of WT and mutants *in vitro*, but a direct correlation of the role of *PODXL* in PD pathogenesis remains to be demonstrated in an animal model. Therefore, we would refrain from concluding on the causality of these mutations in these three unrelated cases but may speculate that they probably contribute to the disease as risk variants in addition to other genetic and non-genetic determinants that the carriers may harbour. This may be similar to the observations of several sporadic PD cases with heterozygous mutations in *PARK2*.<sup>48</sup>

Finally, based on the demographic history of genetically distinct Indian populations including the recent report of Indians showing a higher average inbreeding coefficient compared with European populations,<sup>49</sup> an increased incidence of autosomal-recessive diseases therein is highly probable. Therefore, new disease causal variant discovery efforts with contemporary deep sequencing tools in Indian populations may be immediately

rewarding. In summary, with compelling genetic evidence and *in vitro* functional characterisation, *PODXL* may be considered as the fourth putative causal gene discovered for autosomal-recessive PD warranting investigation in other ethnic groups.

**Acknowledgements** We are thankful to all the participants in this study. We acknowledge contributions of Ms Anjali Dhyani for genomic DNA isolation, Ms Alina Genis for mutation screening of *PARK2*, *PINK1* and *DJ1* genes, Mr Surojit Bose, LeadInvent Ltd., India for assistance in establishing a bioinformatics pipeline for data analysis, Link Biotech Ltd., India and Central instrumentation facility, UDSC for Sanger sequencing services. We gratefully acknowledge the infrastructure support provided to the Department of Genetics, UDSC, by the University Grants Commission, New Delhi under Special Assistance Programme and Department of Science and Technology, New Delhi under FIST and DU–DST PURSE programmes. We are thankful to Prof Dr Peter Heutink, German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany for critical reading of the manuscript.

**Contributors** SS, MB, UBM, RCJ and BKT were responsible for the concept and design of the study. KP, MB and UBM were responsible for diagnosis, recruitment, clinical characterisation and blood collection for all study subjects. SS and BKT acquired, analysed and interpreted the data. SS and BKT wrote the first draft of the manuscript. All authors read the draft, provided their inputs and agreed on the final version of the manuscript before submission.

**Funding** This work was supported by the Department of Science and Technology, Govt. of India, New Delhi (grant number SP/50/B-55/2000), and Department of Biotechnology, Govt. of India, New Delhi (grants number BT/PR2425/Med/13/089/2001 and number BT/PR14500/MED/12/478/2010) to BKT, MB, UBM and RCJ; and Council of Scientific and Industrial Research, New Delhi, India for JRF and SRF to SS.

**Competing interests** None declared.

**Ethics approval** Institutional Ethical Committees of AIIMS and University of Delhi South Campus.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- 1 Stevenston GB, Heafield MT, Waring RH, Williams AC. Xenobiotic metabolism in Parkinson's disease. *Neurology* 1989;39:883–7.
- 2 Hauser MA, Li YJ, Takeuchi S, Walters R, Noureddine M, Maready M, Darden T, Hulette C, Martin E, Hauser E, Xu H, Schmechel D, Stenger JE, Dietrich F, Vance J. Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. *Hum Mol Genet* 2003;12:671–7.
- 3 Schulte C, Gasser T. Genetic basis of Parkinson's disease: inheritance, penetrance, and expression. *Appl Clin Genet* 2011;4:67–80.
- 4 Saiki S, Sato S, Hattori N. Molecular pathogenesis of Parkinson's disease: update. *J Neurol Neurosurg Psychiatr* 2012;83:430–6.
- 5 Jain S, Wood NW, Healy DG. Molecular genetic pathways in Parkinson's disease: a review. *Clin Sci (Lond)* 2005;109:355–64.
- 6 Nuytemans K, Theuns J, Cruts M, Van Broeckhoven C. Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update. *Hum Mutat* 2010;31:763–80.
- 7 Klein C, Westenberger A. Genetics of Parkinson's disease. *Cold Spring Harb Perspect Med* 2012;2:a008888.
- 8 Vilariño-Guell C, Wider C, Ross OA, Dachselt JC, Kachergus JM, Lincoln SJ, Soto-Ortolaza AI, Cobb SA, Wilhoite GJ, Bacon JA, Behrouz B, Melrose HL, Hentati E, Puschmann A, Evans DM, Conibear E, Wasserman WW, Aasly JO, Burkhard PR, Djaldetti R, Ghika J, Hentati F, Krygowska-Wajs A, Lynch T, Melamed E, Rajput A, Rajput AH, Solida A, Wu RM, Uitti RJ, Wszolek ZK, Vingerhoets F, Farrer MJ. VPS35 mutations in Parkinson disease. *Am J Hum Genet* 2011;89:162–7.
- 9 Vilariño-Guell C, Rajput A, Milnerwood AJ, Shah B, Szu-Tu C, Trinh J, Yu I, Encarnacion M, Munsie LN, Tapia L, Gustavsson EK, Chou P, Tatarnikov I, Evans DM, Pishotta FT, Volta M, Beccano-Kelly D, Thompson C, Lin MK, Sherman HE, Han HJ, Guenther BL, Wasserman WW, Bernard V, Ross CJ, Appel-Cresswell S, Stoessl AJ, Robinson CA, Dickson DW, Ross OA, Wszolek ZK, Aasly JO, Wu RM, Hentati F, Gibson RA, McPherson PS, Girard M, Rajput M, Rajput AH, Farrer MJ. DNAJC13 mutations in Parkinson disease. *Hum Mol Genet* 2014;23:1794–801.
- 10 Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, Li Y, Ogaki K, Ando M, Yoshino H, Tomiyama H, Nishioka K, Hasegawa K, Saiki H, Satake W, Mogushi K, Sasaki R, Kokubo Y, Kuzuhara S, Toda T, Mizuno Y, Uchiyama Y, Ohno K, Hattori N. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. *Lancet Neurol* 2015;14:274–82.
- 11 Verstraeten A, Theuns J, Van Broeckhoven C. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet* 2015;31:140–9.



- 12 Ishihara L, Gibson RA, Warren L, Amouri R, Lyons K, Wielinski C, Hunter C, Swartz JE, Elango R, Akkari PA, Leppert D, Surh L, Reeves KH, Thomas S, Ragone L, Hattori N, Pahwa R, Jankovic J, Nance M, Freeman A, Gouider-Khouja N, Kefi M, Zouari M, Ben Sassi S, Ben Yahmed S, El Euch-Fayeche G, Middleton L, Burn DJ, Watts RL, Hentati F. Screening for Lrrk2 G2019S and clinical comparison of Tunisian and North American Caucasian Parkinson's disease families. *Mov Disord* 2007;22:55–61.
- 13 Zabetian CP, Morino H, Ujike H, Yamamoto M, Oda M, Maruyama H, Izumi Y, Kaji R, Griffith A, Leis BC, Roberts JW, Yearout D, Samii A, Kawakami H. Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. *Neurology* 2006;67:697–9.
- 14 Punia S, Behari M, Govindappa ST, Swaminath PV, Jayaram S, Goyal V, Muthane UB, Juyal RC, Thelma BK. Absence/rarity of commonly reported LRRK2 mutations in Indian Parkinson's disease patients. *Neurosci Lett* 2006;409:83–8.
- 15 Lesage S, Brice A. Role of Mendelian genes in "sporadic" Parkinson's disease. *Parkinsonism Relat Disord* 2012;18(Suppl 1):S66–70.
- 16 Chaudhary S, Behari M, Dihana M, Swaminath PV, Govindappa ST, Jayaram S, Goyal V, Maitra A, Muthane UB, Juyal RC, Thelma BK. Parkin mutations in familial and sporadic Parkinson's disease among Indians. *Parkinsonism Relat Disord* 2006;12:239–45.
- 17 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 Psychiatr* 1992;55:181–4.
- 18 Hannonlab.cshl.edu. FASTX-Toolkit.2015. [http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html) (accessed 1 Aug 2015).
- 19 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–60.
- 20 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–9.
- 21 Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2.
- 22 DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- 23 Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet* 2012;49:433–6.
- 24 Database resources of The National Center for Biotechnology Information. *Nucleic Acids Res* 2015;43(Database issue):D6–17.
- 25 Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012;491:56–65.
- 26 Evs.gs.washington.edu. Exome Variant Server.2015. <http://evs.gs.washington.edu/EVS/> (accessed 1 Aug 2015).
- 27 Li MX, Gui HS, Kwan JS, Bao SY, Sham PC. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. *Nucleic Acids Res* 2012;40:e53.
- 28 Greene LA, Tischler AS. Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc Natl Acad Sci USA* 1976;73:2424–8.
- 29 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;9:676–82.
- 30 Kerjaschki D, Sharkey DJ, Farquhar MG. Identification and characterization of podocalyxin—the major sialoprotein of the renal glomerular epithelial cell. *J Cell Biol* 1984;98:1591–6.
- 31 Viturina N, Andrés R, Pérez-Martínez E, Martínez A, Bribián A, Blasi J, Chelliah S, López-Doménech G, De Castro F, Burgaya F, McNagny K, Soriano E. Podocalyxin is a novel polysialylated neural adhesion protein with multiple roles in neural development and synapse formation. *PLoS ONE* 2010;5:e12003.
- 32 Exac.broadinstitute.org. ExAC Browser. 2015. <http://exac.broadinstitute.org> (accessed 1 Aug 2015).
- 33 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
- 34 Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
- 35 Schwarz JM, Rödelberger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* 2010;7:575–6.
- 36 Gasser T. Mendelian forms of Parkinson's disease. *Biochim Biophys Acta* 2009;1792:587–96.
- 37 Biswas A, Sadhukhan T, Majumder S, Misra AK, Das SK, Variation Consortium IG, Ray K, Ray J. Evaluation of PINK1 variants in Indian Parkinson's disease patients. *Parkinsonism Relat Disord* 2010;16:167–71.
- 38 Sadhukhan T, Biswas A, Das SK, Ray K, Ray J. DJ-1 variants in Indian Parkinson's disease patients. *Dis Markers* 2012;33:127–35.
- 39 Bukalo O, Fentrop N, Lee AY, Salmen B, Law JW, Wotjak CT, Schweizer M, Dityatev A, Schachner M. Conditional ablation of the neural cell adhesion molecule reduces precision of spatial learning, long-term potentiation, and depression in the CA1 subfield of mouse hippocampus. *J Neurosci* 2004;24:1565–77.
- 40 Muller D, Wang C, Skibo G, Toni N, Cremer H, Calaora V, Rougon G, Kiss JZ. PSA-NCAM is required for activity-induced synaptic plasticity. *Neuron* 1996;17:413–22.
- 41 Senkov O, Sun M, Weinhold B, Gerardy-Schahn R, Schachner M, Dityatev A. Polysialylated neural cell adhesion molecule is involved in induction of long-term potentiation and memory acquisition and consolidation in a fear-conditioning paradigm. *J Neurosci* 2006;26:10888–10898.
- 42 Cremer H, Chazal G, Goridis C, Represa A. NCAM is essential for axonal growth and fasciculation in the hippocampus. *Mol Cell Neurosci* 1997;8:323–35.
- 43 Dityatev A, Dityateva G, Schachner M. Synaptic strength as a function of post-versus presynaptic expression of the neural cell adhesion molecule NCAM. *Neuron* 2000;26:207–17.
- 44 Nowakowski A, Alonso-Martín S, González-Manchón C, Larrucea S, Fernández D, Vilar M, Cerdán S, Ayuso MS, Parrilla R. Ventricular enlargement associated with the panneuronal ablation of the podocalyxin gene. *Mol Cell Neurosci* 2010;43:90–7.
- 45 Dalaker TO, Zivadinov R, Ramasamy DP, Beyer MK, Alves G, Bronnick KS, Tysnes OB, Aarsland D, Larsen JP. Ventricular enlargement and mild cognitive impairment in early Parkinson's disease. *Mov Disord* 2011;26:297–301.
- 46 Lewis MM, Smith AB, Styner M, Gu H, Poole R, Zhu H, Li Y, Barbero X, Gouttard S, McKeown MJ, Mailman RB, Huang X. Asymmetrical lateral ventricular enlargement in Parkinson's disease. *Eur J Neurol* 2009;16:475–81.
- 47 Perry VH. Innate inflammation in Parkinson's disease. *Cold Spring Harb Perspect Med* 2012;2:a009373.
- 48 Pankratz N, Kissell DK, Pauculo MW, Halter CA, Rudolph A, Pfeiffer RF, Marder KS, Foroud T, Nichols WC. Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations. *Neurology* 2009;73:279–86.
- 49 Juyal G, Mondal M, Luisi P, Laayouni H, Sood A, Midha V, Heutink P, Bertranpetit J, Thelma BK, Casals F. Population and genomic lessons from genetic analysis of two Indian populations. *Hum Genet* 2014;133:1273–87.

## Online Supplementary Information

**Supplementary Table S1. Primers for *PINK1* exons (5'-3')**

PINK1(1A)FP	GCCCAGGACCAGTGATGTT
PINK1(1A)RP	AAGAAGCGGAGACGGTTAGG
PINK1(1B)FP	CTGGGTTCGAGCGCTGCTG
PINK1(1B)RP	CCGCGCTTAGCTCCGTCCT
PINK1(2)FP	TCCCTTTTCTTGGGCCTTCC
PINK1(2)RP	GATGGGCATTTTGAGAACATCTCC
PINK1(3)FP	CTCAGCCTGCCAGTTAAGAC
PINK1(3)RP	AAATCCATCACCAACATTAAGCC
PINK1(4)FP	GTCAGTGCCAGTGTTGGTGTG
PINK1(4)RP	CTTCTCCGGAAGGGTGATG
PINK1(5)FP	CGTATTGGGAGTCGTCGAT
PINK1(5)RP	TTGTGTGCGGAGCTAAATAAA
PINK1(6)FP	GGCCAACACTGAGCCATTAG
PINK1(6)RP	AAGCAAAGTGGCAGGGAAG
PINK1(7)FP	GATTAGCCCATGGATCAGGT
PINK1(7)RP	AACCTGACCTTCACTCTGGAA
PINK1(8)FP	TCTAGCTACAGCTTCCCTTCCTGT
PINK1(8)RP	TGAACTCTCACTCAAGTTCTTCCATT

**Supplementary Table S2. Primers for *DJ1* exons (5'-3')**

DJ1(1)FP	CCTCTGACAACCCCAGTCC
DJ1(1)RP	GTCCAGCACAGGGACACC
DJ1(2)FP	TTTGGGGTATCTCAGGGTTG
DJ1(2)RP	GCGTTAAATGTGAGCAGTG
DJ1(3)FP	AGGGTGAGACCCCATCTCTC
DJ1(3)RP	GGGGAAGACATTCAAGCAAA
DJ1(4)FP	TTCCGTCATGTGGATACACC
DJ1(4)RP	CAGCCTCCTCCCGAAATATT
DJ1(5)FP	GCCTTGCTTGGGTTTAAGAA
DJ1(5)RP	ATCAAACCATCGAATGAAAGG
DJ1(6)FP	AAACATGGGCTTTTCTATATCTGCACT
DJ1(6)RP	CAGTAAGCCAAGATCACGCCACT
DJ1(7)FP	CACATAGCCCATTAGGATGTCA
DJ1(7)RP	AGCTGCAAATGAAGGTGATA



**Supplementary Table S3. Exome sequencing statistics**

Subject ID	II4	II5
Total reads	113,219,058	107,521,894
Total yield (bp)	11,435,124,858	10,859,711,294
Read length (bp)	101	101
Target regions (bp)	64,190,747	64,190,747
Average throughput depth of target regions	178.1	169.2
Mappable reads (=reads mapped to human genome)	77,742,112	69,579,491
Mappable yield (bp)	7,652,369,056	6,849,627,383
% Mappable reads (out of total reads)	68.70%	64.70%
On-target reads (=reads mapped to target regions)	58,803,203	52,004,764
On-target yield (bp)	4,678,327,999	4,124,980,529
% On-target reads (out of mappable reads)	75.60%	74.70%
% On-target reads (out of total reads)	51.90%	48.40%
% Coverage of target regions (more than 1X)	96.50%	96.50%
Number of on-target genotypes (more than 1X)	61,934,147	61,960,221
% Coverage of target regions (more than 10X)	93.90%	93.90%
Number of on-target genotypes (more than 10X)	60,257,749	60,304,932
Median read depth of target regions	65	59
Mean read depth of target regions	72.9	64.3

**Supplementary Table S4. Variant prioritisation**

Total variants (SNVs and indels) in both II4 & II5	2,701,114
Total unreported variants (after excluding variants reported in dbSNP135) in both II4 & II5	14,512
Total unreported homozygous in both II4 & II5	7,202
Total unreported homozygous shared between II4 & II5	1,142
Total unreported homozygous exonic shared between II4 & II5	347
After exclusion of variants with frequency >0.01 in 74 non-PD exomes, 144 control exomes, NHBLI6500 exomes, dbSNP138, 1000 genomes	17

SNVs- single nucleotide variants; indels- insertion deletions; II4 & II5- exome sequenced individuals

**Supplementary Table S5. List of 17 prioritised variants screened by HaloPlex-targeted sequencing**

Chromosome	Position	Reference/Alternative Allele	Gene	Region, change	Effect on amino acid	OMIM ID, name (associated diseases)
Chr1	22446108	CTTTTTT/++++++T	WNT4	3UTR	none	158330, Mullerian Aplasia and Hyperandrogenism
Chr1	101441384	GTTTTTTTTTTTTTTT TTTTTTTTTTTTT/GTTT TTTTTTTTTTTTTTTTT TT----- /GTTTTTTTTTTTTTTTTT TTTTTTTTTT----	SLC30A7	3UTR	none	none
Chr1	234565176	C/T	TARBP1	Exonic, missense	NM_005646 (30Exons):exon16:c.G2857A:p.V953M	none
Chr6	90374283	A/G	MDN1	Exonic, missense	NM_014611(102Exons):exon85:c.T14159C:p.I4720T	none
Chr7	21598500	A/G	DNAH11	Exonic, missense	NM_001277115(82Exons):exon3:c.A576G:p.I192M	611884, Ciliary Dyskinesia, Primary, 7; CILD7
Chr7	100637547	G/A	MUC12	Exonic, missense	NM_001164462(12Exons):exon2:c.G3703A:p.A1235T	none
Chr7	131241029	GGCGAC/GGGGCGAC GGCGAC	PODXL	Exonic, frameshift	NM_001018111(9Exons):exon1:c.89_90insGTCGCCCC:p.S31insfs; NM_005397.3(8Exons):exon1:c.89_90insGTCGCCCC:p.S31insfs	none
Chr8	7787850	C/T	ZNF705B	5UTR	none	none
Chr10	17841713	T/A	TMEM236	3UTR	none	none
Chr10	75434095	T/A	AGAP5	3UTR	none	none
Chr10	5682135	C/G	ASB13	3UTR	none	none
Chr10	86274785	C/T	CCSER2	3UTR	none	none
Chr16	129266	GTTT/++++T	MPG	5UTR	none	none
Chr16	10624465	GCA/+++CA	EMP2	3UTR	none	none
Chr16	30505559	C/T	ITGAL	Exonic, missense	NM_001114380(29Exons):exon10:c.C991T:p.R331W; NM_002209(31Exons):exon12:c.C1240T:p.R414W	none
Chr21	9569365	TACACA/TACA--	unknown	unknown	none	none
Chr21	9418848	A/G	unknown	unknown	none	none

**Supplementary Table S6. Primers for *PODXL* exons (5'-3')**

PODXL(1)FP	GGACGGGCCAGGAGTAG
PODXL(1)RP	CATGCAAACCCACTTAGCAC
PODXL(2A)FP	CCCATTCTCCACAAAAAGGA
PODXL(2A)RP	GTGAGGGGTCGTCAGATGTT
PODXL(2B)FP	AGGCAACCCTACTACCACCA
PODXL(2B)RP	AAGGCATGAGCCTTTTCAGA
PODXL(3)FP	CCGTCTCTGTGGGAGGTAAG
PODXL(3)RP	GGGTAAGTGCTGCTCAAAGC
PODXL(4)FP	GGGCTTTGAGCAGCACTTAC
PODXL(4)RP	ATGGGAAAGGACCACTTCCT
PODXL(5)FP	GGAGGCAGGTTCTAGCACAG
PODXL(5)RP	GGGTGTGGCTTGACAGTTCT
PODXL(6)FP	CCCGAGAGACACTCAGGAGA
PODXL(6)RP	ATGGGAAGTGGCAGAGAACA
PODXL(7)FP	CCTTGTTGGGAGGTCTGGTTA
PODXL(7)RP	CCTGCTCCCTTTCCTCTTCT
PODXL(8C)FP	AGAATGACCACCAGGAGCAG
PODXL(8C)RP	CTTCAGGTCTCGGCAATCTC

**Supplementary Table S7. Primers for site directed mutagenesis (5'-3')**

P429T PODXL FP	GACCGCTTCAGCATGACCCTCATCATCACCA
P429T PODXL RP	TGGTGATGATGAGGGTCATGCTGAAGCGGTC
S373N PODXL FP	GCATCTGTTCCAGGAAATCAGACCGTGGTCG
S373N PODXL RP	CGACCACGGTCTGATTTTCCTGGAACAGATGC
R294Q PODXL FP	AGCATCAACTACCCACCAATACCCCAAACACCTT
R294Q PODXL RP	AAGGTGTTTTGGGGTATTGGTGGGTAGTTGATGCT

# Supplementary Figure S1. Comparison of transcripts from two isoforms of wildtype and mutant alleles of *PODXL*

A	B	C
<p><b>558 codon transcript of WT isoform1</b></p> <p>MRCALALSALLLLSTPPLPSSPSPSPSQNAT  QTTDSSNKTAPTASSVTIMATDTAQQSTVPTS  KANEILASVKATTLGVSSDSPGTTTAAQQVSGPV  NTTVARGGGSGNPPTTIESPKSTKSADTTTVATS  TATAKPNTTSSQNGAEDTTNSGGKSSHVTTDL  TSTKAEHLTTPHPTSPLSRQPTSTHPVATPTSSG  HDHLMKISSSSSTVAIPGYTFTSPGMTTTLLETVF  HHVSQLAGLELLTSGDLPTLASQSAGITASSVISQR  TQQTSSQMPASSTAPSSQETVQPTSPATALRTP  LPETMSSSPTAASTTHRYPKTPSPTVAHESNWA  KCEDLETQTQSEKQLVNLNGNTLCAGGASDEKL  ISLICRAVKATFNPAQDKCGIRLASVPGSQTVVV  KEITIHTKLPKADVYERLKDWDDELKEAGVSDM  KLGDQGPPEEAEDRFSMPLIITIVCMASFLLVA  ALYGCCHQRLSQRKDQQLTEELQTVENGYHD  NPTLEVMTSSEMQEKKVVSINGELGDSWIVP  LDNLTKDDLDEEEDTHL <b>Stop</b></p>	<p><b>526 codon transcript of WT isoform 2</b></p> <p>MRCALALSALLLLSTPPLPSSPSPSPSQNAT  QTTDSSNKTAPTASSVTIMATDTAQQSTVPTS  KANEILASVKATTLGVSSDSPGTTTAAQQVSGPV  NTTVARGGGSGNPPTTIESPKSTKSADTTTVATS  TATAKPNTTSSQNGAEDTTNSGGKSSHVTTDL  TSTKAEHLTTPHPTSPLSRQPTSTHPVATPTSSG  HDHLMKISSSSSTVAIPGYTFTSPGMTTTLPSVI  SQRTQQTSSQMPASSTAPSSQETVQPTSPATAL  RTPPTLPETMSSSPTAASTTHRYPKTPSPTVAHES  NWAKCEDLETQTQSEKQLVNLNGNTLCAGGAS  DEKLISLICRAVKATFNPAQDKCGIRLASVPGSQ  VVVKEITIHTKLPKADVYERLKDWDDELKEAGVS  DMKLGQDQGPPEEAEDRFSMPLIITIVCMASFLL  VAALYGCCHQRLSQRKDQQLTEELQTVENGY  HDNPTLEVMTSSEMQEKKVVSINGELGDSWI  VPLDNLTKDDLDEEEDTHL <b>Stop</b></p>	<p><b>168 codon transcript of frameshift mutant</b></p> <p>MRCALALSALLLLSTPPLPSSPSPSPSPSPPPR  MQPRLRLRTHLTQHRLQHPVSPSWLQIQPSRA  QSPLPRPTKSWPRRRPPLVYPVTHRGLQPWLS  KSQAQSTLPWLEEAQATLLPPSRAPRAQKVQT  PLQLQPPQPQLNLTPQAARMEQKIQTGGKA  ATV <b>Stop</b></p>

Amino acids in bold, in line 1 are common across all three transcripts