RESEARCH ARTICLE

Genetic analysis of indel markers in three loci associated with Parkinson's disease

Zhixin Huo[1], Xiaoguang Luo[2], Xiaoni Zhan[1], Qiaohong Chu[1], Qin Xu[1], Jun Yao[1], Hao Pang[1]*

1 Department of Forensic Genetics and Biology, China Medical University, Shenyang North New Area, Shenyang, P.R., China, 2 Department of Neurology, 1st Affiliated Hospital of China Medical University, Shenyang, P.R., China

* hpang@cmu.edu.cn
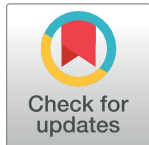
Abstract

The causal mutations and genetic polymorphisms associated with susceptibility to Parkinson's disease (PD) have been extensively described. To explore the potential contribution of insertion (I)/deletion (D) polymorphisms (indels) to the risk of PD in a Chinese population, we performed genetic analyses of indel loci in *ACE*, *DJ-1*, and *GIGYF2* genes. Genomic DNA was extracted from venous blood of 348 PD patients and 325 age- and sex-matched controls without neurodegenerative disease. Genotyping of the indel loci was performed by fragment length analysis after PCR and DNA sequencing. Our results showed a statistically significant association for both allele *X* (alleles without *5*) vs. *5* (odds ratio = 1.378, 95% confidence interval = 1.112–1.708, *P* = 0.003) and genotype *5/X+X/X* vs. *5/5* (odds ratio = 1.681, 95% confidence interval = 1.174–2.407, *P* = 0.004) in the *GIGYF2* locus; however, no significant differences were detected for the *ACE* and *DJ-1* indels. After stratification by gender, no significant differences were observed in any indels. These results indicate that the *GIGYF2* indel may be associated with increased risk of PD in northern China.

Introduction

Parkinson's disease (PD) is the most common neurodegenerative disorder besides Alzheimer disease (AD) [1]. Due to the presence of Lewy bodies (LBs) and loss of dopaminergic neurons in the substantia nigra of the midbrain, the clinical symptoms of PD are mainly characterized by tremors at rest, rigidity, bradykinesia, and posture disorders [2,3]. The etiology of slowly progressive parkinsonian syndromes remains elusive, although certain genetic and environmental factors have been shown to contribute to PD [3].

Over the last decade, research has shed light on the relationship between some candidate genetic markers and PD in different populations [4–8]. These genetic markers include single nucleotide polymorphisms (SNPs), which occupy the majority of PD genetic investigations; rearrangements; and insertion (I)/deletion (D) polymorphisms (indels). However, very few indels have been evaluated in PD studies, especially in Chinese individuals. Some of these indels are located in functionally important sites of human genes and can, therefore, potentially influence pathogenicity [9]. Additionally, some indel polymorphisms that are present in

non-functional regions are susceptible to disease and may also be linked to other pathogenic markers. Therefore, we analyzed the association between indels and PD risk in the current study.

Angiotensin converting enzyme (ACE), a membrane-bound zinc metallopeptidase, regulates blood pressure and blood volume by converting angiotensin I to angiotensin II and by degrading bradykinin. Previous studies have suggested that it may be involved in the pathogenesis of neurodegenerative disorders such as PD, AD, Huntington's disease, and autism [10,11]. Furthermore, the primary loss of substance P due to ACE hydrolysis has been shown to be a contributing factor in the pathogenesis of PD [12–14]. The polymorphism in the *ACE* gene is an indel of the 287-bp *Alu* repetitive sequence in intron 16 (rs1799752), which leads to a deletion allele and insertion allele [15].

The *DJ-1* (*PARK7*) gene is located on chromosome 1p36 [16], the product of which is a peptidase C56 family protein. It has been reported that oxidative stress is involved in the pathology of neurodegenerative diseases, such as PD and AD [17]. Interestingly, the DJ-1 protein protects dopaminergic neurons from oxidative stress [18,19]. In 2003, Vincenzo Bonifati *et al.* reported that *DJ-1* gene mutations are associated with patients in autosomal recessive early onset Parkinsonism [19]. Likewise, the *DJ-1* gene with an 18-bp indel polymorphism (g.168_185del; rs864309640; ss1966650459) in its 5' UTR has been shown to increase the risk of PD [20,21].

Grb10 interacting GYF Protein 2 (GIGYF2), which cooperates with the Grb10 adapter protein to stimulate IGF-1 receptors [22], was identified in dopaminergic neurons in the substantia nigra pars compacta (SNpc) of PD brain tissue and is involved in the regulation of tyrosine kinase receptor signaling at endosomes [23]. The *GIGYF2* gene was reported to be a susceptibility gene- that is accountable for PARK11-linked PD [24] and is located in the chromosomal region 2q36-37 [25]. Lautier *et al.* [24] described all *GIGYF2* coding exons through complete sequencing in Italian and French cohorts; their data strongly support that *GIGYF2* at the PARK 11 locus has a causal role in familial PD. Among them, one of the mutations showed a trinucleotide indel in exon 29 of the *GIGYF2* gene (NG_011847), which is described as being encoded in exon 25 of *GIGYF2* in most other published reports.

Three indel polymorphisms in *ACE*, *DJ-1* and *GIGYF2* have been reported to be associated with, as potential susceptibility factors, PD and AD pathophysiology [12,14,26,27]. However, the results from association studies evaluating the relationship between the above indel markers and distribution of PD risk are still unclear. In a northern Chinese population, we investigated many point mutations associated with Parkinson's disease. Nevertheless, these were insufficient to explain the association with PD. Therefore, we investigated the frequency distributions of indels markers in three loci to explain the pathogenesis of PD in this population.

Materials and methods

Study population

A total of 348 ethnic Chinese individuals with idiopathic PD [mean age (standard deviation) = 63.80 (13.74) years; age at onset (standard deviation) = 56.96 (10.55)] from north China represented the patient group, including 183 male and 165 female patients. All patients were diagnosed by movement disorder neurologists from the First Affiliated Hospital of China Medical University in Liaoning province in China according to the British Parkinson's Disease Society Brain Bank criteria for the clinical diagnosis of PD. The patients exhibited at least two of the three cardinal symptoms and signs for PD (tremor, rigidity, and bradykinesia) and also had a positive response to levodopa therapy. A total of 325 unrelated volunteers [mean age (standard deviation) = 66.87 (16.12) years] from the local community formed the control group,

including 220 males and 105 females. The number of healthy female volunteers was limited in this investigation. Control participants were in good health and without a history of neurodegenerative disease. The scientific experimental protocol was approved by the ethics committee on Human Research of the China Medical University. All methods were performed in accordance with the guidelines of the Declaration of Helsinki. An informed consent agreement was signed by all individuals.

Genomic DNA isolation and genotyping

Venous blood samples were obtained from the participants. Genomic DNA was extracted via the sodium dodecyl sulfate-proteinase K phenol-chloroform method.

The three indels were genotyped by fragment length analysis after PCR. The design of specific primers and related parameters is shown in Table 1. PCR was conducted in 20 μl volumes containing approximately 50 ng of genomic DNA, 2 × PowerTaq PCR MasterMix (BioTeke, Beijing, China) and 10 μM of each specific primer. PCR was carried out on a Thermo Cycler Block (Thermo Fisher Scientific, Vantaa, Finland) with initial denaturation at 94˚C for 3 min, followed by 35 cycles of denaturing for 20 s at 94˚C, annealing for 20 s (60˚C for *ACE* and *GIGYF2*; 65˚C for *DJ-1*), extension at 72˚C (30 s for *ACE*; 15 s for *DJ-1*; 8 s for *GIGYF2*); and a final extension at 72˚C for 1 min. PCR products were separated by electrophoresis on a 1% agarose gel for the *ACE* indel and on a 6% non-denaturing polyacrylamide gel for *DJ-1* and *GIGYF2* and were then visualized under a UV transilluminator after staining with ethidium bromide. All polymorphisms of *ACE*, *DJ-1*, and *GIGYF2* were analyzed in the 348 PD cases and 325 controls.

Capillary electrophoresis for *GIGYF2*

To precisely determine the fragment length of the detected alleles in the *GIGYF2* locus, the forward primer for PCR amplification was labeled with a fluorescent dye, 6-carboxylfluorescein (6-FAM). PCR products with fluorescence were detected using capillary electrophoresis on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems, USA) and analyzed using GeneMapper 2.1 analysis software. Additionally, the fluorescent ladder GeneScan™ 600 Liz (Applied Biosystems, USA) was used as an internal size marker.

DNA sequencing

To further confirm the sequences of all alleles containing the three indel polymorphisms, homozygous samples representing different alleles were first amplified and then directly sequenced. For heterozygous samples, mainly in the *GIGYF2* locus, we performed PCR amplification first. After PCR products were separated via a 6% polyacrylamide gel electrophoresis (PAGE) with 1 × TBE buffer and stained with ethidium bromide, the target fragments were excised from the gel and eluted with distilled water. The eluates were reamplified using the

Table 1. The information of indel loci and PCR parameters.

Genes	Location in gene	Indel polymorphism	Primer direction	Sequence (5'→ 3')	Allele size (base pairs)
ACE	Intron 16	287 bp	Forward	GGGGACTCTGTAAGCCACTG	532/245
			Reverse	TCCCATGCCCATAACAGGTC	
DJ-1	Intron 1	18 bp	Forward	GTGGGGTGAGTGGTACCCAA	248/230
			Reverse	GTCAGTCAAATCCAACGCCA	
GIGYF2	Exon 29	3 bp	Forward	TTCTGAGGCCAAGGAGTTTGC	159–192
			Reverse	ATCCTCAGAGGTACCGCATACA	

https://doi.org/10.1371/journal.pone.0184269.t001

same primers and purified in Suprec-02 tubes (Takara, Shiga, Japan). The DNA sequences were determined using a Bigdye Terminator Cycle Sequencing Reaction Kit on an ABI Prism 3130 Genetic Analyzer (Applied Biosystems, USA).

Statistical analysis

The Hardy-Weinberg equilibrium (HWE) test for frequency distribution was calculated using PowerMarker version 3.0 to detect whether the population was in Hardy-Weinberg equilibrium. For all genes, a mutant vs. wild-type model was used to test different distributions of genotype and alleles between patients and controls. To find the PD effect on gender we have divided our data into different groups (male PD cases vs. male controls, female PD cases vs. female controls, PD male vs. female cases and control male vs. female). We tested allele *5*, *6* and *7* against the rest of the alleles, but only allele *5* comparing with others showed the statistical significance. Therefore, in this study, *5* was treated as the major allele, and other allele classes (*X*) were merged to increase the statistics power for *GIGYF2*.

The $\chi 2$ test was used to test for association of the indel polymorphism and PD risk, using SPSS version 20.0 software. The statistical power to detect association of polymorphisms with PD exceeded 0.80 as estimated with PASS version 11.0 software. All analyses were carried out using a two-tailed $\chi 2$ test; $P < 0.05$ was considered as statistically significant. Bonferroni correction was utilized for multiple testing. Odds ratios (ORs), with their corresponding 95% confidence intervals (CIs), were used to evaluate the strength of the association between the polymorphisms and PD risk.

Results

Genotyping of three indel polymorphisms

The results showed that the frequency distribution for all indels did not deviate from Hardy-Weinberg equilibrium in each group (P > 0.05). The allele and genotype frequencies of the *ACE* and *DJ-1* loci and corresponding statistical power are shown in Table 2 and S1 Table. There were no statistically significant differences between PD patients and controls, as analyzed by the $\chi 2$ test (P > 0.05). For the indel polymorphism found presenting in exon 29 of the *GIGYF2* gene, more than three genotypes were observed, and all alleles were identified as a 3-bp repeat extension in the investigated samples, as indicated by both PAGE and capillary electrophoresis separation. To better understand the sequence architectures of these different alleles at the locus, we performed DNA sequencing. We found a total of nine sequence discrepancies. As shown in Fig 1, using allele 6 (A6) as a reference sequence [24], the indel polymorphism of the *GIGYF2* gene is more likely to be a short tandem repeat (STR) compared with

Table 2. Allele frequency distribution and association analysis on indel loci of *ACE* and *DJ-1* genes.

Variables	PD Case (%)	Control (%)	P	OR (95% CI)	Power
ACE					
I[a]	458 (66)	414 (64)	0.418	1.097 (0.877, 1.372)	0.128
D	238 (34)	236 (36)			
DJ-1					
D[b]	66 (9)	62 (10)	0.972	0.994 (0.690, 1.430)	0.050
I	630 (91)	588 (90)			

a: variant allele in *ACE*

b: variant allele in *DJ-1*.

```
A6   cagcagcagcagctgccacagcagcagcagcagcagccgccacagcagccgccacagcagccacaacagcag
     Q  Q  Q  Q  L  P  Q  Q  Q  Q  Q  Q  Q  P  P  Q  Q  P  P  Q  Q  P  Q  Q  Q

A5   *  *  *  *  *  *  ( Q  Q  Q  Q  Q )  *  *  *  *  *  *  *  *  *  *  *  *  *

A7   *  *  *  *  *  *  ( Q Q Q Q Q Q Q )  *  *  *  *  *  *  *  *  *  *  *  *  *

A4   *  *  *  *  *  *  ( Q   Q   Q   Q )  *  *  *  *  *  *  *  *  *  *  *  *  *

Av1  *  *  *  *  -  -  (     Q       )   *  *  *  *  *  *  *  *  *  *  *  *  *

Av2  *  *  *  *  -  -  -  -  -  -  -  -   *  *  *  *  *  *  *  *  *  *  *  *  *

Av3  *  *  *  *  *  *  *  *  *  *  *  *   -  -  -  -  *  *  *  *  *  *  *  *

Av4  *  *  *  *  *  *  ( Q Q Q Q Q Q Q )  -  -  -  -  *  *  *  *  *  *  *  *

Av5  *  *  *  *  *  *  (QQQLP + QQQQQQQ)  *  *  *  *  *  *  *  *  *  *  *  *  *
```
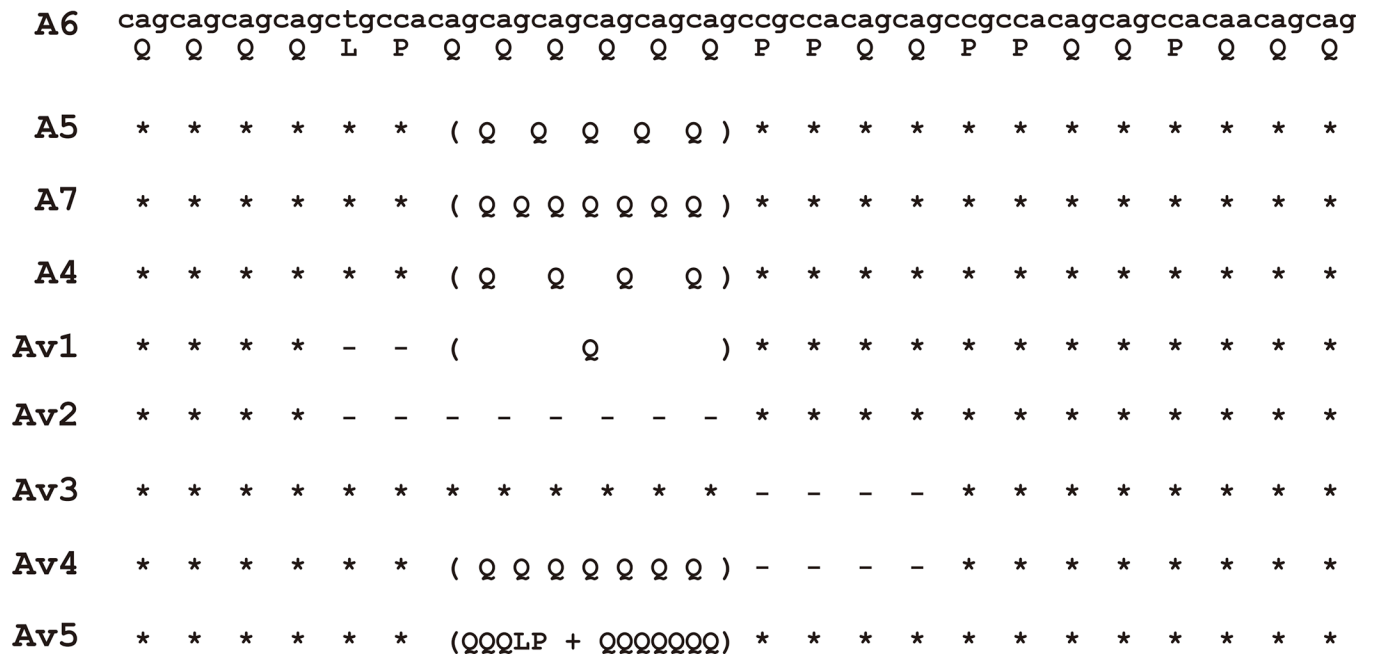
Fig 1. Nine allelic sequences were identified in the *GIGYF2* locus in this study. A6 (27.6%) is used as a reference sequence. Compared with A6, A5 (47.0%) and A7 (24.4%) were 3-bp repeated indels, and the others indels were different structural variations enclosing the poly-Q core unit (approximately 1.2%). The upper sequence represents base pairs of nucleic acids, and the lower sequence represents the corresponding amino acids.

alleles A5 and A7. Others indels exhibited a different structural variation enclosing the poly-Q core unit, which were identified as variant (*V*) alleles in our study.

The indel in the *GIGYF2* gene is associated with PD

For the complex indel located in exon 29 of *GIGYF2*, nine alleles and thirteen genotypes were identified, and the genotypic and allelic frequency distributions are shown in Table 3. The case-control analysis found a statistically significant difference after comparing the frequencies

Table 3. Frequency distribution and association analysis on indel locus of *GIGYF2* gene.

GIGYF2	PD Case (%)	Control (%)	χ^2	*P*	Power
5	300 (0.431)	332 (0.511)	9.845	0.020	0.754
6	202 (0.290)	170 (0.262)			
7	185 (0.266)	144 (0.222)			
V	9 (0.013)	4 (0.006)			
5/5	67 (0.192)	93 (0.286)	11.959	0.153	0.683
6/6	25 (0.072)	23 (0.071)			
7/7	19 (0.055)	17 (0.052)			
5/6	83 (0.239)	79 (0.243)			
5/7	77 (0.221)	64 (0.197)			
6/7	68 (0.195)	45 (0.138)			
5/V	6 (0.017)	3 (0.009)			
6/V	1 (0.003)	0 (0.000)			
7/V	2 (0.006)	1 (0.003)			

V: variants (alleles A4, Av1-Av5).

Table 4. Association study on indel locus of *GIGYF2* by genotypic and allelic stratification.

GIGYF2	PD Case (%)	Control (%)	$\chi 2$	*P*	OR (95% CI)	Power
X	396 (0.569)	318 (0.489)	8.579	0.003	1.378 (1.112, 1.708)	0.833
5	300 (0.431)	332 (0.511)				
X/X	115 (0.330)	86 (0.265)	8.916	0.012	-	0.767
5/X	166 (0.477)	146 (0.449)				
5/5	67 (0.193)	93 (0.287)				
5/X+X/X	281 (0.807)	232 (0.663)	8.129	0.004	1.681 (1.174, 2.407)	0.814
5/5	67 (0.193)	93 (0.287)				

X: alleles *6*, *7*, A4 and Av1-Av5. *P* < 0.05/2 represents statistics significant for multiple testing.

of all alleles (*P* < 0.05), but no difference was found after testing the frequencies of all genotypes. Considering that there are multiple alleles and genotypes present for this particular locus, we performed further tests of association for allelic and genotypic grouping analyses. Our results indicated that there was a significant difference between PD cases and controls when the *X* allele and *5/X+X/X* genotype of the indel in exon 29 of the *GIGYF2* gene was present compared with the *5* allele (OR = 1.378, 95% CI = 1.112–1.708, *P* = 0.003) and *5/5* genotype (OR = 1.681, 95% CI = 1.174–2.407, *P* = 0.004) (Table 4). Bonferroni correction was used in multiple testing, and the *P*-value was set at 0.025 (0.05/2) since two types of genetic models were used to analyze the association of the *GIGYF2* indel with PD. However, other allelic and genotypic grouping comparisons did not show any differences between PD cases and controls (S3 and S4 Tables).

Differences in the genetic distribution of PD-associated indels according to gender

We performed gender stratification to better understand the association between the three indels and PD. As expected, no association was found between the indels in the *ACE* and *DJ-1* genes and PD in the male or female datasets when studied separately (S2 Table). Although statistical significance was found in males, but not females, between patients and controls for the *GIGYF2* indel (Table 5), no difference between males and females was found in either PD cases or controls (Table 6).

Discussion

In the current study, we selected three different indels to evaluate their role as predisposing factors for PD. Our results showed no significant differences in the allelic frequencies of indels

Table 5. Association study on indel locus of *GIGYF2* by gender stratification.

GIGYF2	Male					Female				
	PD Case	Control	*P*	OR (95% CI)	Power	PD Case	Control	*P*	OR (95% CI)	Power
X	210	217	0.022	1.383 (1.047, 1.829)	0.626	186	101	0.061	1.394 (0.985, 1.973)	0.467
5	156	223				144	109			
X/X	60	59	0.051	-	0.580	55	27	0.185	-	0.357
5/X	90	99				76	47			
5/5	33	62				34	31			
5/X+X/X	150	158	0.017	1.784 (1.106, 2.876)	0.666	131	74	0.095	1.614 (0.918, 2.837)	0.386
5/5	33	62				34	31			

X: alleles *6*, *7*, A4 and Av1-Av5.

Table 6. Comparison of *GIGYF2* indel locus between male and female.

GIGYF2	PD Case					Control				
	Male	Female	χ2	*P*	OR (95% CI)	Male	Female	χ2	*P*	OR (95% CI)
X	210	186	0.073	0.787	1.042 (0.772, 1.407)	217	101	0.085	0.771	1.050 (0.756, 1.459)
5	156	144				223	109			
X/X	60	55	0.483	0.785	-	59	27	0.078	0.962	-
5/X	90	76				99	47			
5/5	33	34				62	31			
5/X+X/X	150	131	0.370	0.543	1.180 (0.682, 2.011)	158	74	0.062	0.803	1.068 (0.640, 1.781)
5/5	33	34				62	31			

X: alleles *6*, *7*, A4 and Av1-Av5.

in the *ACE* and *DJ-1* genes between PD patients and controls. However, an association between the *GIGYF2* indel polymorphism and PD was found in the population. With regard to the genetic factors associated with complex diseases or phenotypes, comprehensive understanding for different variations may be needed [28]. Single nonfunctional variants as susceptible locus, are often related to a disease or phenotype, which implied that these variants are actually in strong linkage with causal mutations. Therefore, results based on single variant in related loci can lead to incorrect explanation for complex phenotypes. Some studies have demonstrated loci with multiple variants in complete linkage disequilibrium with the indel polymorphisms [29,30]. It is noteworthy that the causal mutation may be better captured by haplotype-based methods [31,32]. Based on these issues, it is plausible that the indel polymorphisms in this study, whether they are susceptible for PD or not, must be considered for haplotype investigations to make clear relationship between indels and the true risk variants.

The indel polymorphism in the *ACE* gene is a transposon insertion in intron 16. These mobile genetic elements produce a number of small indel variations in humans [33], i.e., the *Alu* transposon in *ACE*. It was originally reported that the ACE activity levels in *D/D* carriers were approximately twice those found in *I/I* genotype individuals [34]. Therefore, this indel polymorphism was used as a marker for studying associations with pathophysiological conditions, including neurodegenerative diseases, such as PD [10]. Moreover, the *ACE* indel polymorphism in the Australian Caucasian population was not associated with PD in the first association study [35]. No statistically significant differences were found in alleles or genotypes for *ACE* indel polymorphism during investigations of Caucasian populations in Greece, Italy, and Columbia [36–38]. However, one report showed that the frequency of the homozygote *D/D* genotype in the *ACE* gene was considerably increased in patients with PD compared to controls in a Taiwanese population, despite there being no significant difference in allelic frequency [12]. In the current study, the *I* vs. *D* genetic model in the *ACE* indel locus did not result in a statistically significant association. Certain other reports from the literature have implied that different ages and ethnicities might be an explanation for the disparate results. On the other hand, power analysis was performed after the experiment was finished and was lower than 0.80, which suggest that more samples were needed to estimate the distributions. Although this study reported a negative result, it was the first to investigate the *ACE* indel polymorphism's influence on PD patients from the Chinese mainland. Further studies are needed to analyze the linkage disequilibrium between indels and multiple loci in this region of *ACE* and to evaluate the role of this indel in the incidence of Parkinson's disease in populations from other areas of the Chinese mainland.

Given that variants within the promoter region of both α-synuclein and parkin are associated with increased risk for sporadic PD [2,27], the *DJ-1* indel polymorphism was of particular

interest. The 18-bp indel polymorphism in the *DJ-1* gene can be classified as a random sequence type. Hague *et al.* was the first to identify several non-coding variants in *DJ-1*, including the indel in the *DJ-1* gene located in the promoter region and near the transcription regulatory SP1 site [39]. This polymorphism is hypothesized to affect gene expression due to its proximity to the transcriptional regulatory SP1 site. Furthermore, this indel may be involved in DJ-1 protein expression, subsequently influencing neuronal response to oxidative stress. Its association with genetic susceptibility to PD has been previously reported in several different populations although some association studies have indicated mixed results. The *DJ-1* indel polymorphism did not confer a risk for PD in studies conducted in Finland and Great Britain [40,41]. Nevertheless, the deletion allele of the indel locus was found to be a risk factor for PD in Italian and Indian case-control studies [20,21]. In addition, two investigations in southern Africa and west China, determined that the deletion allele occurred at very low frequencies in these populations and was, therefore, unlikely to play a significant role in PD [42,43]. In the present study, we also did not find that the *DJ-1* indel was associated with PD in our population from northern China. Because the statistical power was low, the efficiency may increase for this polymorphism with a larger sample size.

In this study, we identified the *X* allele the *5/X+X/X* genotype of the *GIGYF2* gene as risk factors that enhanced the risk of PD (P < 0.05). The power was approximately 0.80. Our results are the first to suggest that indels in the *GIGYF2* gene may be associated with PD in a northern Chinese population. The GIGYF2 protein is involved in insulin-like growth factor (IGF) and insulin signaling in the central nervous system [22,23]. Moreover, the *GIGYF2* gene has been reported to be a susceptibility gene responsible for PARK11-linked PD [24]. Given the current data and previously reported findings, we conclude that the architecture of the indel locus most likely represents a repeat extension of an irregular 3-bp multiple model. The *GIGYF2* indel was first identified when screening for *GIGYF2* gene mutations in Italian and French patients with familial PD [24]. Until now, only 3-bp insertion or deletion alleles in exon 29 of the *GIGYF2* gene have been detected. Although the indel locus might not play a pathogenic role in PD, the existence of multiple alleles and genotypes in the poly Q region is still meaningful because polyglutamine expansions can cause neurodegenerative disorders [44,45]. Lautier *et al.* identified seven different alleles in the indel locus, but did not confirm a causal role in familial PD [24]. Studies in Asian (Taiwanese and Singapore) [46] (Japanese) [47], Portuguese and North American [48], and Italian cohorts [49] suggest that the indel in exon 29 of *GIGYF2* is not strongly related to the development of PD based on mutational analysis, although different alleles were identified. For the association of indel polymorphism with PD, it appears to be comparatively lower because the 3-bp indel locus in exon 29 displays an irregular trait. Nichols *et al.* identified seven alleles in Caucasian population in North America [50]. Compared with the wild type (allele *6* in our study), an allele with one Q deleted (allele *5* in our study) was present at a higher frequency, while other alleles were present at a lower frequency. However, there was no evidence that any of these *GIGYF2* indels increased the risk for PD. Moreover, Cao *et al.* analyzed the DNA sequence surrounding exon 29 of *GIGYF2* and found that two alleles with one deleted (allele *5*) and inserted Q (allele *7*) were present at higher frequencies. The allele with an inserted Q was more prevalent in the control group, which suggested it was more likely to be a benign variant [51]. In the current study, we combined DNA sequencing with allele length analysis and found eight different alleles in northern China. First, we applied allele length polymorphism analysis to evaluate the association of the *GIGYF2* indel locus with PD. The results of this study indicated that: (1) all alleles are 3-bp fold different and alleles *5*, *6*, and *7* are present at higher frequencies, as reported by Cao *et al.* in Shanghai, China. No deviation from Hardy–Weinberg equilibrium was observed based on genotype distribution, which suggests that allele length analysis can be used in association studies. (2) Although the

distributions of genotypic frequency between both groups indicated that the indel is unlikely to influence PD development in northern Han Chinese, allele *X* conveys an increased risk in both the allelic *X* vs. *5* and *5/X +X/X* vs. *5/5* genetic models. These results suggest that other genetic groups and districts can be comparaed using these models. (3) The uncommon alleles found in this study seem to present preferentially in PD patients, which suggests that the complexity and diversity in exon 29 of the *GIGYF2* gene may influence its function and lead to cellular degeneration.

Subsequently, we conducted a gender stratification analysis of the relationship between the frequency distributions of the three indel polymorphisms in different models and susceptibility to PD. Many mutations and polymorphisms leading to gene deficiency or linkage association with diseases exhibit gender-linked differences [52,53]. Simunovic *et al.* found differential gene expression in both males and females with distinct PD-association patterns and described a bias of the male gender towards PD at the molecular level [54]. Furthermore, a common mutation associated with PD, G2019S in the *LRRK2* gene, was found to have a relatively greater genetic load in women with PD among individuals of Jewish descent [55]. Interestingly, rs12817488 in the *CCDC62* gene was associated with PD risk, but the results were unclear after stratification by gender. Yu *et al.* found that a significant association only existed in males [56], while Liu *et al.* only found this association in a female population [57]. In summary, several aspects may contribute to the different results between the two studies, including age of onset, EOPD/LOPD ratio, sex ratio, and MAF [58]. Our previous study also observed a genetic female distribution bias in the mitochondrial genome of PD patients in northern China [59]. Taken together, the above data indicate that the genetic effect might be specific to either males or females. In the present study, we continued to analyze gender-relevant traits associated with several genetic markers specific for PD in northern China. However, no sex dependent effects on PD were found for these three indel loci. Comparing previous data with present data, the genetic markers that we describe in this study are associated with PD in northern China and show complex characteristics in both overall and gender stratification, suggesting that it is vital to accurately assess the genetic architecture of PD patient populations as the number of genetic markers increases.

Conclusions

Overall, our result revealed that genotype *5/5* in the *GIGYF2* gene may be a protective factor for PD patients in northern China, but no differences were found in either *ACE* or *DJ-1*. Although there was the statistical significance comparing with frequencies of allele *5*, the allele X is a composite class of alleles. Therefore, these alleles may be associated markers and can be linked to the causative mutation.

Supporting information

S1 Table. The genotype frequency distribution and association analysis of the indel loci of the *ACE* and *DJ-1* genes.
(DOC)

S2 Table. Allele association analysis after gender stratification of the indel loci in the *ACE*, *DJ-1*, and *GIGYF2* genes. a: variant allele in *ACE*; b: variant allele in *DJ-1*; *V*: variants (alleles A4, Av1-Av5).
(DOC)

S3 Table. Association study of the indel in the *GIGYF2* gene stratified by allele. *X*: alleles *5*, *7*, A4, and Av1-Av5.
(DOC)

S4 Table. Association study of the indel in the *GIGYF2* gene stratified by allele. *X*: alleles *5*, *6*, A4, and Av1-Av5.
(DOC)

Author Contributions

Data curation: Xiaoguang Luo, Jun Yao.

Funding acquisition: Hao Pang.

Investigation: Zhixin Huo, Xiaoguang Luo, Qin Xu, Hao Pang.

Methodology: Xiaoni Zhan, Qiaohong Chu, Qin Xu, Hao Pang.

Software: Jun Yao.

Supervision: Hao Pang.

Validation: Xiaoni Zhan, Qiaohong Chu.

Writing – original draft: Zhixin Huo.

Writing – review & editing: Hao Pang.

References

1. Marsden CD. Parkinson's disease. J Neurol Neurosurg Psychiatry. 1994; 57: 672–681. https://doi.org/10.1371/journal.pone.0165266 PMID: 7755681

2. Cummings JL, Henchcliffe C, Schaier S, Simuni T, Waxman A, Kemp P. The role of dopaminergic imaging in patients with symptoms of dopaminergic system neurodegeneration. Brain. 2011; 134: 3146–3166. https://doi.org/10.1093/brain/awr177 PMID: 21810889

3. Fahn S. Description of Parkinson's disease as a clinical syndrome. Ann N Y Acad Sci. 2003; 991: 1–14. https://doi.org/10.1111/j.1749-6632.2003.tb07458.x PMID: 12846969

4. Rozenkrantz L, Gan-Or Z, Gana-Weisz M, Mirelman A, Giladi N, Bar-Shira A, et al. SEPT14 Is Associated with a Reduced Risk for Parkinson's Disease and Expressed in Human Brain. J Mol Neurosci. 2016; 59: 343–350. https://doi.org/10.1007/s12031-016-0738-3 PMID: 27115672

5. Yang X, Zhao Q, An R, Zhou H, Lin Z, Xu Y. SNP rs1805874 of the Calbindin1 Gene Is Associated with Parkinson's Disease in Han Chinese. Genet Test Mol Biomarkers. 2016; 20: 753–757. https://doi.org/10.1089/gtmb.2016.0149 PMID: 27611799

6. Wang L, Maldonado L, Beecham GW, Martin ER, Evatt ML, Ritchie JC, et al. DNA variants in CACNA1C modify Parkinson disease risk only when vitamin D level is deficient. Neurol Genet. 2016; 2: e72. https://doi.org/10.1212/NXG.0000000000000072 PMID: 27123490

7. Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, et al. Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. Nature. 2016; 533: 95–99. https://doi.org/10.1038/nature17939 PMID: 27096366

8. von Otter M, Bergstrom P, Quattrone A, De Marco EV, Annesi G, Soderkvist P, et al. Genetic associations of Nrf2-encoding NFE2L2 variants with Parkinson's disease—a multicenter study. BMC Med Genet. 2014; 15: 131. https://doi.org/10.1186/s12881-014-0131-4 PMID: 25496089

9. Mullaney JM, Mills RE, Pittard WS, Devine SE. Small insertions and deletions (INDELs) in human genomes. Hum Mol Genet. 2010; 19: R131–136. https://doi.org/10.1093/hmg/ddq400 PMID: 20858594

10. Kaur P, Muthuraman A, Kaur M. The implications of angiotensin-converting enzymes and their modulators in neurodegenerative disorders: current and future perspectives. ACS Chem Neurosci. 2015; 6: 508–521. https://doi.org/10.1021/cn500363g PMID: 25680080

11. Firouzabadi N, Ghazanfari N, Alavi Shoushtari A, Erfani N, Fathi F, Bazrafkan M, et al. Genetic Variants of Angiotensin-Converting Enzyme Are Linked to Autism: A Case-Control Study. PLoS One. 2016; 11: e0153667. https://doi.org/10.1371/journal.pone.0153667 PMID: 27082637

12. Lin JJ, Yueh KC, Chang DC, Lin SZ. Association between genetic polymorphism of angiotensin-converting enzyme gene and Parkinson's disease. J Neurol Sci. 2002; 199: 25–29. PMID: 12084438

13. Lieb K, Treffurth Y, Berger M, Fiebich BL. Substance P and affective disorders: new treatment opportunities by neurokinin 1 receptor antagonists? Neuropsychobiology. 2002; 45 Suppl 1: 2–6. 49254 PMID: 11893870

14. Song GG, Lee YH. The insertion/deletion polymorphism in the angiotensin-converting enzyme and susceptibility to schizophrenia or Parkinson's disease: A meta-analysis. J Renin Angiotensin Aldosterone Syst. 2015; 16: 434–442. https://doi.org/10.1177/1470320313495909 PMID: 25143327

15. Rigat B, Hubert C, Corvol P, Soubrier F. PCR detection of the insertion/deletion polymorphism of the human angiotensin converting enzyme gene (DCP1) (dipeptidyl carboxypeptidase 1). Nucleic Acids Res. 1992; 20: 1433. PMID: 1313972

16. van Duijn CM, Dekker MC, Bonifati V, Galjaard RJ, Houwing-Duistermaat JJ, Snijders PJ, et al. Park7, a novel locus for autosomal recessive early-onset parkinsonism, on chromosome 1p36. Am J Hum Genet. 2001; 69: 629–634. https://doi.org/10.1086/322996 PMID: 11462174

17. Giasson BI, Ischiropoulos H, Lee VM, Trojanowski JQ. The relationship between oxidative/nitrative stress and pathological inclusions in Alzheimer's and Parkinson's diseases. Free Radic Biol Med. 2002; 32: 1264–1275. PMID: 12057764

18. van der Merwe C, Jalali Sefid Dashti Z, Christoffels A, Loos B, Bardien S. Evidence for a common biological pathway linking three Parkinson's disease-causing genes: parkin, PINK1 and DJ-1. Eur J Neurosci. 2015; 41: 1113–1125. https://doi.org/10.1111/ejn.12872 PMID: 25761903

19. Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science. 2003; 299: 256–259. https://doi.org/10.1126/science.1077209 PMID: 12446870

20. De Marco EV, Annesi G, Tarantino P, Nicoletti G, Civitelli D, Messina D, et al. DJ-1 is a Parkinson's disease susceptibility gene in southern Italy. Clin Genet. 2010; 77: 183–188. https://doi.org/10.1111/j.1399-0004.2009.01310.x PMID: 19968671

21. Sadhukhan T, Biswas A, Das SK, Ray K, Ray J. DJ-1 variants in Indian Parkinson's disease patients. Dis Markers. 2012; 33: 127–135. https://doi.org/10.3233/DMA-2012-0913 PMID: 22960331

22. Giovannone B, Lee E, Laviola L, Giorgino F, Cleveland KA, Smith RJ. Two novel proteins that are linked to insulin-like growth factor (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling. J Biol Chem. 2003; 278: 31564–31573. https://doi.org/10.1074/jbc.M211572200 PMID: 12771153

23. Higashi S, Iseki E, Minegishi M, Togo T, Kabuta T, Wada K. GIGYF2 is present in endosomal compartments in the mammalian brains and enhances IGF-1-induced ERK1/2 activation. J Neurochem. 2010; 115: 423–437. https://doi.org/10.1111/j.1471-4159.2010.06930.x PMID: 20670374

24. Lautier C, Goldwurm S, Durr A, Giovannone B, Tsiaras WG, Pezzoli G, et al. Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease. Am J Hum Genet. 2008; 82: 822–833. https://doi.org/10.1016/j.ajhg.2008.01.015 PMID: 18358451

25. Pankratz N, Nichols WC, Uniacke SK, Halter C, Rudolph A, Shults C, et al. Significant linkage of Parkinson disease to chromosome 2q36-37. Am J Hum Genet. 2003; 72: 1053–1057. https://doi.org/10.1086/374383 PMID: 12638082

26. Kolsch H, Jessen F, Freymann N, Kreis M, Hentschel F, Maier W, et al. ACE I/D polymorphism is a risk factor of Alzheimer's disease but not of vascular dementia. Neurosci Lett. 2005; 377: 37–39. https://doi.org/10.1016/j.neulet.2004.11.062 PMID: 15722183

27. Kehoe PG, Russ C, McIlory S, Williams H, Holmans P, Holmes C, et al. Variation in DCP1, encoding ACE, is associated with susceptibility to Alzheimer disease. Nat Genet. 1999; 21: 71–72. https://doi.org/10.1038/5009 PMID: 9916793

28. Zhu X, Bouzekri N, Southam L, Cooper RS, Adeyemo A, McKenzie CA, et al. Linkage and association analysis of angiotensin I-converting enzyme (ACE)-gene polymorphisms with ACE concentration and blood pressure. Am J Hum Genet. 2001; 68: 1139–1148. https://doi.org/10.1086/320104 PMID: 11283791

29. Rieder MJ, Taylor SL, Clark AG, Nickerson DA. Sequence variation in the human angiotensin converting enzyme. Nat Genet. 1999; 22: 59–62. https://doi.org/10.1038/8760 PMID: 10319862

30. Martinez-Rodriguez N, Posadas-Romero C, Villarreal-Molina T, Vallejo M, Del-Valle-Mondragon L, Ramirez-Bello J, et al. Single nucleotide polymorphisms of the angiotensin-converting enzyme (ACE) gene are associated with essential hypertension and increased ACE enzyme levels in Mexican individuals. PLoS One. 2013; 8: e65700. https://doi.org/10.1371/journal.pone.0065700 PMID: 23741507

31. Weiss KM, Clark AG. Linkage disequilibrium and the mapping of complex human traits. Trends Genet. 2002; 18: 19–24. PMID: 11750696

32. Akey J, Jin L, Xiong M. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? Eur J Hum Genet. 2001; 9: 291–300. https://doi.org/10.1038/sj.ejhg.5200619 PMID: 11313774

33. Mills RE, Bennett EA, Iskow RC, Devine SE. Which transposable elements are active in the human genome? Trends Genet. 2007; 23: 183–191. https://doi.org/10.1016/j.tig.2007.02.006 PMID: 17331616

34. Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, Soubrier F. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. J Clin Invest. 1990; 86: 1343–1346. https://doi.org/10.1172/JCI114844 PMID: 1976655

35. Mellick GD, Buchanan DD, McCann SJ, Davis DR, Le Couteur DG, Chan D, et al. The ACE deletion polymorphism is not associated with Parkinson's disease. Eur Neurol. 1999; 41: 103–106. 8012 PMID: 10023114

36. Papapetropoulos S, Glynos K, Zhou Z, Orfanos SE, Mitsi G, Papapetropoulos A. The insertion/deletion polymorphism of the angiotensin converting enzyme (ACE) in Parkinson's disease. Open Neurol J. 2008; 2: 66–70. https://doi.org/10.2174/1874205X00802010066 PMID: 19452013

37. Pascale E, Purcaro C, Passarelli E, Guglielmi R, Vestri AR, Passarelli F, et al. Genetic polymorphism of Angiotensin-Converting Enzyme is not associated with the development of Parkinson's disease and of L-dopa-induced adverse effects. J Neurol Sci. 2009; 276: 18–21. https://doi.org/10.1016/j.jns.2008.08.017 PMID: 18805557

38. Benitez BA, Forero DA, Arboleda GH, Granados LA, Yunis JJ, Fernandez W, et al. Exploration of genetic susceptibility factors for Parkinson's disease in a South American sample. J Genet. 2010; 89: 229–232. PMID: 20861575

39. Hague S, Rogaeva E, Hernandez D, Gulick C, Singleton A, Hanson M, et al. Early-onset Parkinson's disease caused by a compound heterozygous DJ-1 mutation. Ann Neurol. 2003; 54: 271–274. https://doi.org/10.1002/ana.10663 PMID: 12891685

40. Eerola J, Hernandez D, Launes J, Hellstrom O, Hague S, Gulick C, et al. Assessment of a DJ-1 (PARK7) polymorphism in Finnish PD. Neurology. 2003; 61: 1000–1002. PMID: 14557580

41. Morris CM, O'Brien KK, Gibson AM, Hardy JA, Singleton AB. Polymorphism in the human DJ-1 gene is not associated with sporadic dementia with Lewy bodies or Parkinson's disease. Neurosci Lett. 2003; 352: 151–153. PMID: 14625045

42. Chen W, Peng R, Li T, Wu Y, Zhang J, Wang Y, et al. [Association of the DJ-1 gene polymorphism with sporadic Parkinson's disease in Sichuan province of China]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi. 2008; 25: 566–569. PMID: 18841573

43. Glanzmann B, Lombard D, Carr J, Bardien S. Screening of two indel polymorphisms in the 5'UTR of the DJ-1 gene in South African Parkinson's disease patients. J Neural Transm (Vienna). 2014; 121: 135–138. https://doi.org/10.1007/s00702-013-1094-x PMID: 24052110

44. Jones L, Houlden H, Tabrizi SJ. DNA repair in the trinucleotide repeat disorders. Lancet Neurol. 2017; 16: 88–96. https://doi.org/10.1016/S1474-4422(16)30350-7 PMID: 27979358

45. Kraus-Perrotta C, Lagalwar S. Expansion, mosaicism and interruption: mechanisms of the CAG repeat mutation in spinocerebellar ataxia type 1. Cerebellum Ataxias. 2016; 3: 20. https://doi.org/10.1186/s40673-016-0058-y PMID: 27895927

46. Tan EK, Lin CH, Tai CH, Tan LC, Chen ML, Li R, et al. Non-synonymous GIGYF2 variants in Parkinson's disease from two Asian populations. Hum Genet. 2009; 126: 425–430. https://doi.org/10.1007/s00439-009-0678-x PMID: 19449032

47. Li L, Funayama M, Tomiyama H, Li Y, Yoshino H, Sasaki R, et al. No evidence for pathogenic role of GIGYF2 mutation in Parkinson disease in Japanese patients. Neurosci Lett. 2010; https://doi.org/10.1016/j.neulet.2010.05.071 PMID: 20685231

48. Bras J, Simon-Sanchez J, Federoff M, Morgadinho A, Januario C, Ribeiro M, et al. Lack of replication of association between GIGYF2 variants and Parkinson disease. Hum Mol Genet. 2009; 18: 341–346. https://doi.org/10.1093/hmg/ddn340 PMID: 18923002

49. Guella I, Pistocchi A, Asselta R, Rimoldi V, Ghilardi A, Sironi F, et al. Mutational screening and zebrafish functional analysis of GIGYF2 as a Parkinson-disease gene. Neurobiol Aging. 2011; 32: 1994–2005. https://doi.org/10.1016/j.neurobiolaging.2009.12.016 PMID: 20060621

50. Nichols WC, Kissell DK, Pankratz N, Pauciulo MW, Elsaesser VE, Clark KA, et al. Variation in GIGYF2 is not associated with Parkinson disease. Neurology. 2009; 72: 1886–1892. https://doi.org/10.1212/01.wnl.0000346517.98982.1b PMID: 19279319

51. Cao L, Zhang T, Zheng L, Wang Y, Wang G, Zhang J, et al. The GIGYF2 variants are not associated with Parkinson's disease in the mainland Chinese population. Parkinsonism Relat Disord. 2010; 16: 294–297. https://doi.org/10.1016/j.parkreldis.2009.11.009 PMID: 20044296

52. Verma D, Chakraborti B, Karmakar A, Bandyopadhyay T, Singh AS, Sinha S, et al. Sexual dimorphic effect in the genetic association of monoamine oxidase A (MAOA) markers with autism spectrum disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2014; 50: 11–20. https://doi.org/10.1016/j.pnpbp. 2013.11.010 PMID: 24291416

53. Lin R, Zhang Y, Yan D, Liao X, Wang X, Fu Y, et al. Genetic Association Analysis of Common Variants in FOXO3 Related to Longevity in a Chinese Population. PLoS One. 2016; 11: e0167918. https://doi. org/10.1371/journal.pone.0167918 PMID: 27936216

54. Simunovic F, Yi M, Wang Y, Stephens R, Sonntag KC. Evidence for gender-specific transcriptional profiles of nigral dopamine neurons in Parkinson disease. PLoS One. 2010; 5: e8856. https://doi.org/10. 1371/journal.pone.0008856 PMID: 20111594

55. Saunders-Pullman R, Stanley K, San Luciano M, Barrett MJ, Shanker V, Raymond D, et al. Gender differences in the risk of familial parkinsonism: beyond LRRK2? Neurosci Lett. 2011; 496: 125–128. https://doi.org/10.1016/j.neulet.2011.03.098 PMID: 21511009

56. Liu RR, Zhou LL, Cheng X, Sun MX, Hu YB, Chen SF, et al. CCDC62 variant rs12817488 is associated with the risk of Parkinson's disease in a Han Chinese population. Eur Neurol. 2014; 71: 77–83. https:// doi.org/10.1159/000354333 PMID: 24335092

57. Yu RL, Guo JF, Wang YQ, Liu ZH, Sun ZF, Su L, et al. The single nucleotide polymorphism Rs12817488 is associated with Parkinson's disease in the Chinese population. J Clin Neurosci. 2015; 22: 1002–1004. https://doi.org/10.1016/j.jocn.2014.11.024 PMID: 25818163

58. Lu Y, Tan L, Shen N, Peng J, Wang C, Zhu Y, et al. Possible association of CCDC62 rs12817488 polymorphism and Parkinson's disease risk in Chinese population: a meta-analysis. Sci Rep. 2016; 6: 23991. https://doi.org/10.1038/srep23991 PMID: 27035708

59. Chu Q, Luo X, Zhan X, Ren Y, Pang H. Female genetic distribution bias in mitochondrial genome observed in Parkinson's Disease patients in northern China. Sci Rep. 2015; 5: 17170. https://doi.org/ 10.1038/srep17170 PMID: 26602989