

<https://doi.org/10.1038/s41531-024-00826-8>

The p.Gly2019Ser is a common *LRRK2* pathogenic variant among Egyptians with familial and sporadic Parkinson's disease



Martina B. William^{1,2}, Sharifa Hamed³, Ali Shalash⁴, Eman M. Khedr³, Mohamed H. Yousef¹, Shaimaa El-Jaafary⁵, Gharib Fawi⁶, Asmaa Helmy⁴, Eman Hamid⁴, Mohamed Essam⁴, Hamin Lee⁷, Alina Jama⁷, Mohamed Koraym³, Doaa M. Mahmoud⁸, Sara Elfarrash⁸, Yasmin Elsaid⁹, Asmaa S. Gabr¹, Nourhan Shebl¹⁰, Nesreen Abdelwahhab⁹, Tamer M. Belal⁹, Nehal A. B. Elsayed¹⁰, Mohamed El-Gamal¹¹, Shimaa Elgamal¹², Salma Ragab¹², Jaidaa Mekky¹³, Lobna Aly¹³, Samir Nabhan¹, Gaafar Ragab^{14,15}, Mohamed A. Hussein¹⁴, Mohamed Tharwat Hegazy^{14,15}, Henry Houlden⁷, Mohamed Salama^{10,11} ✉ & Mie Rizig⁷ ✉

The impact of *LRRK2* variants on the risk of Parkinson's disease in Egyptians remains unknown. We examined 1210 Egyptians (611 PD patients and 599 controls) from 16 governorates across Egypt for 12 *LRRK2* pathogenic variants. The p.Gly2019Ser was the only variant detected, with a prevalence of 4.1% in sporadic cases, 6.5% in familial cases, and 0.68% in controls. Among p.Gly2019Ser carriers, all were heterozygous bar one homozygous patient, and all shared the common haplotype 1. Demographics and UPDRS scores did not differ between carriers and non-carriers, with most patients being males and developed PD in their fifties. Young and Early-onset PD prevalence was 37.5% in carriers and 33% in non-carriers. Familial cases were 16.6% in carriers and 11.5% in non-carriers. This study affirms that like other Mediterranean populations, Egyptians with PD have a higher prevalence of the p.Gly2019Ser variant compared to the global average. *LRRK2* inhibitors could be promising therapeutic options for further exploration in this population.

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms such as tremors, stiffness, and bradykinesia, in addition to other non-motor symptoms including sleep disturbances, autonomic dysfunction, gastrointestinal manifestations, as well as psychiatric and cognitive symptoms^{1,2}. While aging is the primary risk factor for PD, research suggests a significant influence of genetic and environmental factors in the development of the disease^{3,4}.

Situated at the crossroads of Africa, the Middle East, and Europe, Egypt possesses distinctive characteristics in terms of geography, geopolitics, and

history. Egypt's unique location has played a role in shaping the genetic composition of its people through historical invasion and migratory events. The modern-day Egyptians represent a homogeneous population with exclusive genetic variants and unique genetic admixture from Europeans, Arabs, Asians, and Africans ancestries⁵. Egypt, with a population of over 110 million, is the fourteenth most populous country in the world and the third most populous country in Africa⁶. As the number of elderly individuals in Egypt increases and life expectancy rises, it is expected that there will be a significant increase in the number of people developing neurodegenerative

¹Institute of Global Health and Human Ecology, The American University in Cairo, Cairo, Egypt. ²Department of Clinical Pharmacy, Faculty of Pharmacy, Assiut University, Assiut, Egypt. ³Department of Neurology, Faculty of Medicine, Assiut University, Assiut, Egypt. ⁴Department of Neurology, Faculty of Medicine, Ain Shams University, Cairo, Egypt. ⁵Department of Neurology, Faculty of Medicine, Cairo University, Cairo, Egypt. ⁶Department of Neurology, Faculty of Medicine, Sohag University, Sohag, Egypt. ⁷Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, UCL, London, UK. ⁸Department of Medical Physiology, Faculty of Medicine, Mansoura University, Dakahlia, Egypt. ⁹Department of Neurology, Faculty of Medicine, Mansoura University, Dakahlia, Egypt. ¹⁰Department of Neurology, Mansoura International Hospital, Dakahlia, Egypt. ¹¹Department of Forensic Medicine and Clinical Toxicology, Faculty of Medicine, Mansoura University, Dakahlia, Egypt. ¹²Department of Neuropsychiatry, Faculty of Medicine, Kafrelsheikh University, Kafr El-Sheikh, Egypt. ¹³Department of Neurology, Faculty of Medicine, Alexandria University, Alexandria, Egypt. ¹⁴Department of Internal Medicine, Rheumatology and Clinical Immunology Unit, Faculty of Medicine, Cairo University, Cairo, Egypt. ¹⁵Faculty of Medicine, Newgiza University (NGU), Giza, Egypt.

✉ e-mail: mohamed-salama@aucegypt.edu; m.rizig@ucl.ac.uk



disorders such as PD⁷. Although exact statistics are not readily available, the limited information suggests that the prevalence of PD in Egypt could be around 500 cases per 100,000 individuals, which is among the highest rates globally^{8,9}. To ensure that this expanding and ancestrally distinct patient population can take advantage of the progress being made in the diagnosis and treatment of PD, it is crucial to comprehensively examine their genetic makeup and risk to the disease.

In the past two decades, advances in genetics have greatly enhanced our understanding of the molecular mechanisms leading to the development of PD^{10,11}. Situated on chromosome 12, *LRRK2* is among the most researched genes in PD. It consists of 51 exons spanning 144 kb and encodes a large 286 kDa multidomain protein of the same name. More than a thousand missense genetic variants have been described in the *LRRK2* gene, with at least 20 confirmed as pathogenic or likely pathogenic and approximately 50 having conflicting classifications^{12–14}. These pathogenic variants are inherited in an autosomal dominant pattern and are primarily located in *LRRK2*'s catalytic core, which includes the ROC, COR, and Kinase domains. They result in enhanced *LRRK2* autophosphorylation activity and increased phosphorylation of surrogate substrates, leading to a harmful kinase gain of function^{15–17}. Inhibitors that control the activity of the *LRRK2* kinase are among the most promising new drug targets for PD currently undergoing clinical trials¹⁷.

It is widely recognized that *LRRK2* pathogenic variants linked to PD exhibit diverse geographic and ancestral distributions^{18,19}. For instance, the well-known p.Gly2019Ser pathogenic variant has a higher prevalence in PD patients from North African and Ashkenazi Jewish backgrounds compared to individuals of European ancestry²⁰. The p.Arg1441Gly pathogenic variant, on the other hand, is notably common in the Basque population, contributing to almost half of familial PD cases in that region²¹. The p.Gly2385Arg and p.Arg1628Pro pathogenic variants are more frequently observed in PD patients of East Asian descent^{19,22–24}. The pathogenic variant p.Asn1437Ser is mainly found in Germans, while p.Asn1437His is more prevalent in Scandinavians¹⁶. Other few rare *LRRK2* pathogenic variants

were exclusively found in specific families such as the p.Tyr1699Cys variant was observed in a German Canadian family²⁵, and the p.Ile2020Thr variant was detected in a Japanese family with a single-founder effect²⁶. Notably, to date no disease-causing *LRRK2* PD pathogenic variant was reported in individuals of sub-Saharan African descent.

Knowledge about the prevalence of the *LRRK2* pathogenic variants in Egypt is lacking. Only two small single-center studies examined the prevalence of the *LRRK2* p.Gly2019Ser in Egyptians. The first study²⁷, conducted in lower Egypt, included 113 PD patients and 87 controls, while the second study involved 69 patients and 96 controls from upper Egypt²⁸. The prevalence of the p.Gly2019Ser pathogenic variant was significantly different between the two studies, with 9.7% in Lower Egypt and 1.45% in Upper Egypt. There is only one case report of a family with the *LRRK2* Arg1441Cys pathogenic variant²⁹. A summary of the findings from these studies is included in Supplementary Materials 1. At present, there are no multicenter studies examining the *LRRK2* pathogenic variants in Egyptians, and the prevalence of these variants across the country remains uncertain.

This study aimed to investigate the prevalence of 12 known *LRRK2* pathogenic variants, including the previously reported p.Gly2019Ser, in a large cohort of PD patients and controls from multiple centers across Egypt. Additionally, it sought to assess the contribution of these variants to the genetic risk of developing PD in Egyptians.

Results

Recruitment and cohort characteristics

A total of 1210 participants were included in this study: 611 patients with PD (64.8% male) and 599 unrelated healthy controls (43.4% male). The recruitment process took place in 16 governorates, covering 90% of the inhabited regions, across Egypt: 12 in lower Egypt and 4 in upper Egypt. All participants self-identified as native Egyptians for at least two generations. Figure 1 illustrates the distribution of recruitment sites throughout Egypt and the number of participants (cases and controls) recruited at each site. Notably, approximately 70% of the cohort participants (PD 422, Controls

Fig. 1 | Map of Egypt illustrating the recruitment sites across the 16 governorates participated in this study and the numbers of PD patients (N_{PD}) and healthy controls (N_{CT}) recruited from each site. Upper Egypt is delineated by a bold black border, while Lower Egypt is delineated by a bold blue border.

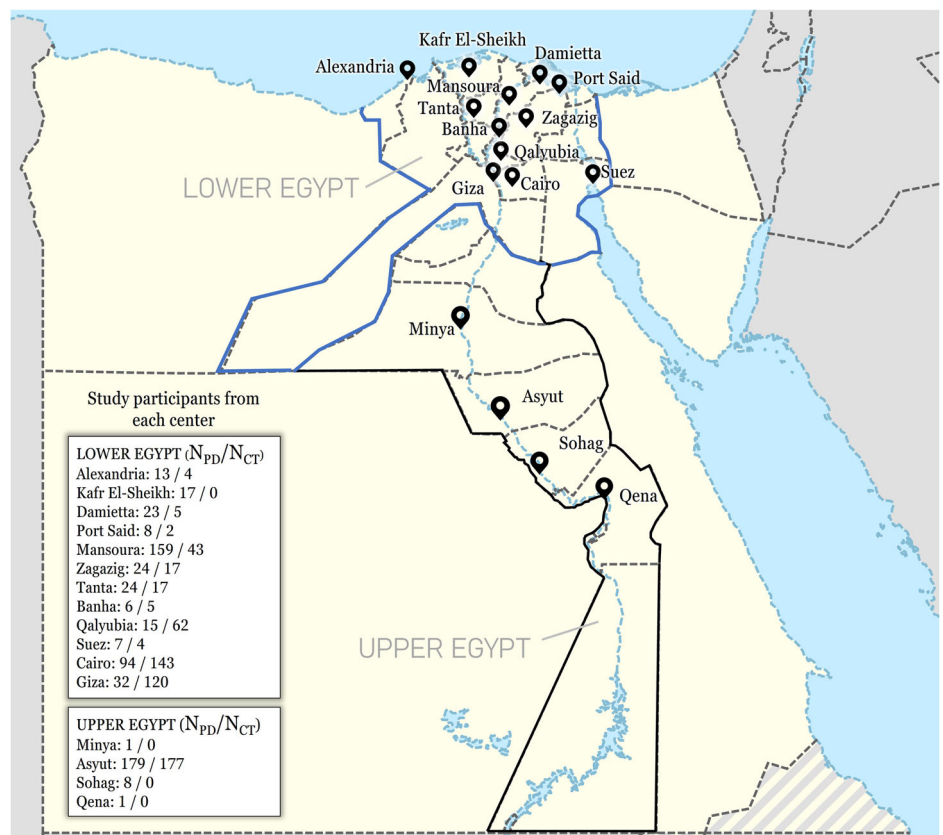


Table 1 | Genotypic and allelic frequencies of the *LRRK2* p.Gly2019Ser in carriers vs noncarriers across Egypt

Participants	Genotype frequencies		Allele frequencies		
	GG	GA	AA	G	A
Across Egypt					
All Cases (547)	523 (0.956)	23 (0.042)	1 (0.002)	1069 (0.977)	25 (0.023)
Familial (62)	58 (0.935)	3 (0.048)	1 (0.016)	119 (0.960)	5 (0.040)
Sporadic (485)	465 (0.959)	20 (0.041)	0 (0.000)	950 (0.979)	20 (0.021)
All Controls (587)	583 (0.993)	4 (0.007)	0 (0.000)	1170 (0.997)	4 (0.003)
Upper Egypt					
Cases (186)	177 (0.952)	9 (0.048)	0 (0.000)	363 (0.976)	9 (0.024)
Familial (28)	25 (0.893)	3 (0.107)	0 (0.000)	53 (0.946)	3 (0.054)
Sporadic (158)	152 (0.962)	6 (0.038)	0 (0.000)	310 (0.981)	6 (0.019)
Controls (177)	176 (0.994)	1 (0.006)	0 (0.000)	353 (0.997)	1 (0.003)
Lower Egypt					
Cases (361)	346 (0.958)	14 (0.039)	1 (0.003)	706 (0.978)	16 (0.022)
Familial (34)	33 (0.971)	0 (0.000)	1 (0.029)	66 (0.971)	2 (0.029)
Sporadic (327)	313 (0.957)	14 (0.043)	0 (0.000)	640 (0.978)	14 (0.021)
Controls (410)	407 (0.993)	3 (0.007)	0 (0.000)	817 (0.996)	3 (0.004)

422) were recruited from lower Egypt, while 30% of the participants (PD 189, Controls 177) were from upper Egypt.

The mean $\pm$ SD age of PD patients at recruitment was 61.01 ± 9.67 years (range: 31–94), while the mean $\pm$ SD age at recruitment of controls was 43 ± 14 years (range: 17–90). Among PD patients, 5.07% (31/611) had their first symptoms before reaching 40 years of age; classified as Young-onset PD (YOPD). Additionally, 28.64% (175/611) developed PD between the ages of 40 and 50, classified as Early-onset PD (EOPD). Family history of PD was identified in 11.46% (70/611) of patients. Consanguinity was uncommon, present in only 3.6% of the patients' families. The mean age at diagnosis was 54.33 ± 8.68 years (range 21–83), which aligns with the mean age at onset of initial motor symptoms at 53.81 ± 8.52 years (range 21–83). The mean $\pm$ SD MDS Unified Parkinson's Disease Rating Scale (UPDRS) score for patients was 100.35 ± 49.96 (range: 7–231). Supplementary Table 1 lists the demographic and UPDRS scores of the study participants and compare these characteristics between individuals from lower and upper Egypt. Except for the age of the patients during the study, duration of illness and the MDS UPDRS scores at the time of study, there were no significant differences observed between PD patients from upper and lower Egypt. At the time of the study, patients from Upper Egypt were, on average, approximately 2 years younger than those from Lower Egypt. Additionally, the duration of their illness was, on average, 3 years shorter, and their UPDRS scores were about 10 points higher compared to patients from Lower Egypt (P value < 0.001 , details in Supplementary Table 1). Notably, there were no significant differences in the age of onset between the two groups (Upper Egypt 53.66 ± 8.13 years (range 27–83), Lower Egypt 53.88 ± 8.71 years (21–79): U 38971.50, p value 0.652).

***LRRK2* pathogenic SNPs in patients and controls**

A validated allele Specific fluorescence-based PCR assays (KASP assays) were used to genotype 12 pathogenic *LRRK2* SNPs. The assay performed effectively and clearly distinguished genotyping clusters for all 12 SNPs. The allelic frequencies for all SNPs were in Hardy-Weinberg equilibrium ($p = 0.88$ –1).

In both the PD patients and controls, only the p.Gly2019Ser pathogenic variant was detected, while the other 11 *LRRK2* pathogenic variants were not found. A summary of the genotyping results for the 12 *LRRK2* SNPs is provided in Supplementary Table 2.

Among the participants who were successfully genotyped for p.Gly2019Ser, 4.38% (24/547) of PD patients and 0.68% (4/587) of healthy controls were found to carry the pathogenic p.Gly2019Ser allele. As

expected, the p.Gly2019Ser is significantly associated with PD status (OR 8.988, 95% CI 2.276–35.498, p value 0.002). The p.Gly2019Ser Minor Allelic Frequencies distribution in patients (MAF ~ 0.023) and in controls (MAF ~ 0.003) were uniform across Egypt with no significant differences observed between upper and lower regions (X^2 0.1367, p value: 0.71). The carrier rate (i.e. prevalence) was 6.45% (4/62) in familial PD cases and 4.12% (20/485) in non-familial/sporadic cases. Table 1 details the Genotypic and Allelic frequencies of the p.Gly2019Ser in this cohort, including a breakdown of its distribution in upper and lower Egypt.

Screening an external dataset of Whole Genome Sequence (WGS) from 110 healthy Egyptians aged 40–50, available through the Egyptian reference genome browser^{5,30}, yielded similar findings. Except for the p.Gly2019Ser variant, all the other 11 *LRRK2* variants were absent. Among the 110 healthy Egyptians available in this dataset, there was one heterozygous carrier of the p.Gly2019Ser (Prevalence 0.9%, MAF 0.004).

Demographic, clinical characteristics and haplotype of the p.Gly2019Ser carriers

Among the 24 PD patients with the p.Gly2019Ser, 15 were recruited from Lower Egypt and 9 were from Upper Egypt, with a male predominance of 70.83%. The average age at recruitment of these patients was 59.67 ± 8.32 years (range: 42–79), while the mean age of disease onset was 52 ± 8.07 years (range: 38–67). Most patients (79.17%) developed their first motor symptoms before the age of 60, with 4.17% experiencing symptoms before 40, 33.33% between 40 and 49, 41.67% between 50 and 59, and 20.83% between 60 and 69 years. Only 16% had a family history of PD, while 84% were sporadic cases. Consanguinity within the family was reported only in 4.17% (1/24). The Mean $\pm$ SD of MDS-UPDRS score for p.Gly2019Ser carriers with PD was 120.48 ± 59.61 . There were no significant differences in demographic characteristics or UPDRS scores between p.Gly2019Ser PD carriers and non-carriers (see Table 2).

All p.Gly2019Ser carriers were heterozygous, except for one patient who was homozygous. The homozygous patient was a male from Upper Egypt; he had a family history of PD on both parents. He experienced his first motor symptoms at 45 and initially responded well to levodopa therapy. However, his symptoms progressively worsened, and by the age of 54, when he joined the study, he began experiencing dyskinesia and rapid wearing-off effects.

The Four non-manifesting p.Gly2019Ser carriers in the control group were two females aged 26 and 29, and one male aged 55 from lower Egypt and one female aged 39, from upper Egypt. None of the control carriers self-

Table 2 | Demographic and clinical characteristics of p.Gly2019Ser carriers vs non-carriers

Parameter	All cohort		P value*	P ¹ value*
	Non-G2019S carriers (n = 523)	G2019S carriers (n = 24)		
Gender distribution, n (%)	183 (34.99)	7 (29.17)	0.664	—
Female	340 (65.01)	17 (70.83)		
Male				
Age at Study (y), mean ± SD (range)	60.69 ± 9.79 (31–94)	59.67 ± 8.32 (42–79)	0.523	—
Age at onset (y), mean ± SD (range)	53.82 ± 8.67 (21–83)	52 ± 8.07 (38–67)	0.377	—
Young onset PD (< 40 y), n (%)	28 (5.35)	1 (4.17)	> 0.9999	—
Early onset PD (40–50 y), n (%)	147 (28.11)	8 (33.33)	0.644	—
Late onset PD (> 50 y), n (%)	348 (66.54)	15 (62.50)	0.665	—
Age at Diagnosis (y), mean ± SD (range)	54.28 ± 8.77 (21–83)	52 ± 8.07 (38–67)	0.246	—
Family History n (%)	58 (11.09)	4 (16.67)	0.336	—
Duration of illness, mean ± SD (range)	6.88 ± 5.47 (0–33)	7.67 ± 5.90 (0–25)	0.393	—
MDS UPDRS mean ± SD (range)	118.83 ± 48.94 (7–223)	120.48 ± 59.61 (38–231)	0.151	0.099

* P values for categorical variables were calculated using Fisher's exact test and for continuous variables using the Mann–Whitney U test. P¹ is P values following adjustment for sex, age, and duration of disease using Univariate Analysis of Variance (ANCOVA).

reported a known family history of PD or experiencing PD prodromal symptoms particularly olfactory dysfunction and (or) any sleep disturbances. Table 3 details the characteristics and UPDRS scores of each of the 28 individuals (24 patients and 4 controls) who are positive for the p.Gly2019Ser.

All p.Gly2019Ser carriers (except for one control sample that failed during genotyping) had at least one copy of the minor allele (A) of the rs28903073 SNP, suggesting they shared haplotype 1. This haplotype is widely recognized as the predominant haplotype for the p.Gly2019Ser, commonly encountered in individuals carrying the p.Gly2019Ser pathogenic allele across Europe, North and South America, as well as North Africa³¹. Supplementary fig. 1 illustrates the Sanger sequencing chromatograms of the rs28903073 SNP in individuals carrying the p.Gly2019Ser variant.

Discussion

The existing understanding of the genetic risk of PD in the Middle East and North Africa (MENA), particularly in Egypt, the most populous country in this region, is limited and relies mainly on preliminary data from case reports and small cohorts^{27–29,32} (for details, refer to Supplementary Materials 1). In this study, the International Parkinson's Disease Genomic Consortium Africa (IPDGC Africa)³³, a consortium dedicated to advancing the understanding of PD in Africans, report on the largest multicenter genetic investigation in Egyptians with PD to date. The study examined the prevalence of 12 *LRRK2* variants with confirmed pathogenicity including the p.Gly2019Ser, among 611 PD patients and 599 controls from 16 out of Egypt's 27 governorates, encompassing around 90% of the nation's populated regions.

The p.Gly2019Ser was the only *LRRK2* pathogenic variant found in this cohort. None of the other 11 variants was detected. Carriers of the p.Gly2019Ser pathogenic allele were found across Upper and Lower Egypt, with a combined average prevalence of approximately 4.4% among PD patients. This prevalence was higher in familial cases (6.4%) compared to sporadic patients (4.1%) and did not show significant variability between upper and lower regions. The overall p.Gly2019Ser MAF in cases was 2.3%. These results affirm that the p.Gly2019Ser is a common *LRRK2* risk variant in Egyptians with PD. In addition, they suggest a widespread distribution of this pathogenic variant across the entire country, including Upper Egypt, contrary to previous reports that had localized its presence to Lower Egypt^{27,28}. Most PD patients carrying the p.Gly2019Ser developed symptoms in their forties and fifties with nearly 40% of carriers classify as YOPD or EOPD. There were no significant variations in the age of onset, male predominance,

or total UPDRS scores between those with and without the p.Gly2019Ser.

Certain regional trends were observed among PD patients, notably the shorter duration of illness in Upper Egypt compared to Lower Egypt. The reasons behind these differences remain unclear. However, patients from Upper Egypt were, on average, about 2 years younger than those from Lower Egypt and exhibited a more severe form of the disease, as indicated by UPDRS scores that were approximately 10 points higher (adjusted P value < 0.001). This severity and rapid progression might prompt patients in upper Egypt to seek treatment more swiftly. Additionally, environmental contributors such as the more prevalent use of pesticides and pollutants in Upper Egypt could influence disease severity and contribute to these regional differences. Another possible explanation is that many residents of Upper Egypt come from rural communities with lower educational levels and higher illiteracy rates. This could lead to early motor symptoms being overlooked or misinterpreted as normal aging, causing patients to seek treatment only at more advanced stages and creating a false impression of a shorter disease duration.

Around one third of PD patients who did not have the p.Gly2019Ser variant experienced symptoms before turning 50 years old. This percentage exceeds the global EOPD rates of 3–14% and aligns closely with data from previous clinical and epidemiological studies on EOPD in Egypt^{9,34,35}. Interestingly, the incidence of a family history of PD was relatively low, around 16% in p.Gly2019Ser carriers and 11% in non-carriers, consistent as well with previously reported rates in Egyptians (11–15%)^{36,37}. Furthermore, the consanguinity rates among families of PD patients were below 5% in all regions. The coexistence of high EOPD rates and lack of family history of PD and (or) consanguinity in Egyptian patients present an intriguing scenario. EOPD is typically associated with higher genetic risks and tends to be hereditary^{10,38}. Combining these observations with the fact that only 5% of EOPD cases in the cohort could be attributed to the *LRRK2* p.Gly2019Ser suggests that other variants likely in recessive genes such as *GBA1*, *DJ-1*, *PINK1*, *PRKN* may play a role in the development of PD among Egyptians. We anticipate finding novel variants exclusive to the Egyptians and (or) variants similar to those previously described in these genes within the Mediterranean region and North Africa. We speculate that the recently identified African *GBA1* c.1225-34 T>G intronic variant³⁹ found in Nigerians with PD will not significantly contribute to the genetic risk of PD in Egyptians. This is primarily due to differences in haplotypic structures and genetic admixture between the two populations, as well as the fact that the *GBA1* c.1225-34 T>G variant is exclusive to individuals of West African ancestry. Nonetheless, the absence of substantial familial rates hints at potential contributions of environmental factors as well^{40–43}.

Table 3 | Individual demographics and UPDRS scores of p.Gly2019Ser carriers

No	Phenotype	Region % from total	Governorate	Gender	Age (y)	Age at onset (y)	Family* History	Consanguinity	UPDRS score
1	PD	Lower Egypt 4.15% (15/361)	Cairo	Male	54	45	Yes	No	139
2	PD		Cairo	Male	69	67	No	No	230
3	PD		Giza	Male	53	45	No	No	163
4	PD		Giza	Male	63	52	No	No	72
5	PD		Giza	Female	54	40	No	Yes	231
6	PD		Qalyubia	Male	55	53	No	No	73
7	PD		Cairo	Male	58	58	No	No	43
8	PD		Cairo	Male	55	55	No	No	72
9	PD		Mansoura	Male	79	54	No	No	223
10	PD		Mansoura	Female	60	40	No	No	194
11	PD		Banha	Male	66	57	No	No	100
12	PD		Cairo	Female	46	40	No	No	69
13	PD		Alexandria	Female	42	38	No	No	Unavailable
14	PD		Alexandria	Male	49	42	No	No	38
15	PD		Alexandria	Male	67	57	No	No	87
16	PD	Upper Egypt 4.84% (9/186)	Asyut	Male	63	61	No	No	53
17	PD		Asyut	Male	59	49	Yes	No	85
18	PD		Asyut	Female	54	52	Yes	No	72
19	PD		Asyut	Male	59	50	Yes	No	127
20	PD		Asyut	Male	65	60	No	No	171
21	PD		Asyut	Male	68	62	No	No	128
22	PD		Asyut	Female	62	56	No	No	153
23	PD		Asyut	Male	62	55	No	No	119
24	PD		Asyut	Female	70	60	No	No	129
Summary				Males (70.83%) Females (29.17%)	59.67 ± 8.32 (42–79)	52.00 ± 8.07 (38–67)	Yes (16.67%) No (83.33%)	1/24 (4.2%)	120.48 ± 59.61 (38–231)
1	Control	Lower Egypt	Cairo	Female	29	NA	No/unknown	No	NA
2	Control		Giza	Male	55	NA	No/unknown	No	NA
3	Control		Cairo	Female	26	NA	No/unknown	No	NA
4	Control	Upper Egypt	Asyut	Female	39	NA	No/unknown	No	NA

It is worth noting that, in this cohort, the p. Gly2019Ser pathogenic variant was also identified among the control group at an average prevalence of 0.68% and MAF of 0.3%. Searching the Egyptian genome reference browser^{5,30}, which includes an independent WGS dataset from 110 healthy individuals, showed similar findings for p.Gly2019Ser, with a prevalence of 0.9% and a MAF of 0.4%. This prevalence figures ranks among the highest reported rates in controls globally and is like those found in non-manifesting carriers within the Maghrib (0.7–2%) and Ashkenazi Jewish (0.4–2%) populations. While it is possible that all asymptomatic carriers might eventually develop PD, the presence of the p.Gly2019Ser in healthy individuals; with no PD or prodromal signs of PD; at different ages (26,29,39 and 55 years) suggests that this pathogenic variant may exhibit reduced or incomplete penetrance in Egyptians.

Numerous studies have investigated the date and origin of the p. Gly2019Serin different populations, linking it to three distinct haplotypes identified as haplotype 1, 2, and 3⁴⁴. Haplotype 1 is prevalent among carriers of the pathogenic variant worldwide, identified in Europeans, Ashkenazi Jews, North Africans, and Latin Americans^{45,46}. Haplotype 2 is present in individuals of European descent, while Haplotype 3 was first observed in Turkish and Japanese families. Haplotypes 2 and 3 are both extremely rare, less studied and seem to have emerged more recently compared to Haplotype 1⁴⁴. Within our cohort, all individuals with the p.Gly2019Ser shared Haplotype 1, suggesting a common ancestral founder with most carriers of this pathogenic variant in other populations. It is

hypothesized that the founding event of the p.Gly2019Ser likely occurred in the Near East about 4000–5000 years ago before spreading to Europe and other regions due to migratory movements^{44,47,48}. Recently, Kmiecik and colleagues⁴⁹ analyzed a large global cohort of 1286 p.Gly2019Ser carriers. The analysis identified shared Identity by Descent (IBD) segments among p.Gly2019Ser carriers from genetically distinct ancestral groups in seven geographical locations: Ashkenazim from Eastern Europe, North Africa, Italy, Northwestern Europe, Puerto Rico, Cuba, and Mexico. Interestingly, most p.Gly2019Ser carriers were of North African and Ashkenazi descent, and these ancestries were present in all p.Gly2019Ser carriers, confirming an ancient founding event in these populations. The Ashkenazi p.Gly2019Ser carriers from Eastern Europe were found to be genetically closer to individuals from North Africa. Latin Caribbean groups (Puerto Rico, Mexico, and Cuba) showed the greatest genetic similarity to Eastern Europe, suggesting founder effects from European migrations to the Americas possibly during the Spanish colonial era. Northwestern Europeans and Italians were more distantly related, indicating that the p.Gly2019Ser might have reached these regions through different migration routes.

The route of introduction of the p.Gly2019Ser variant into Egypt is uncertain; however, the country is located at a critical juncture for historical migratory paths through which the pathogenic variant could have been introduced to its population. To the west, Egypt shares borders with Maghreb countries where the pathogenic variant is prevalent, particularly

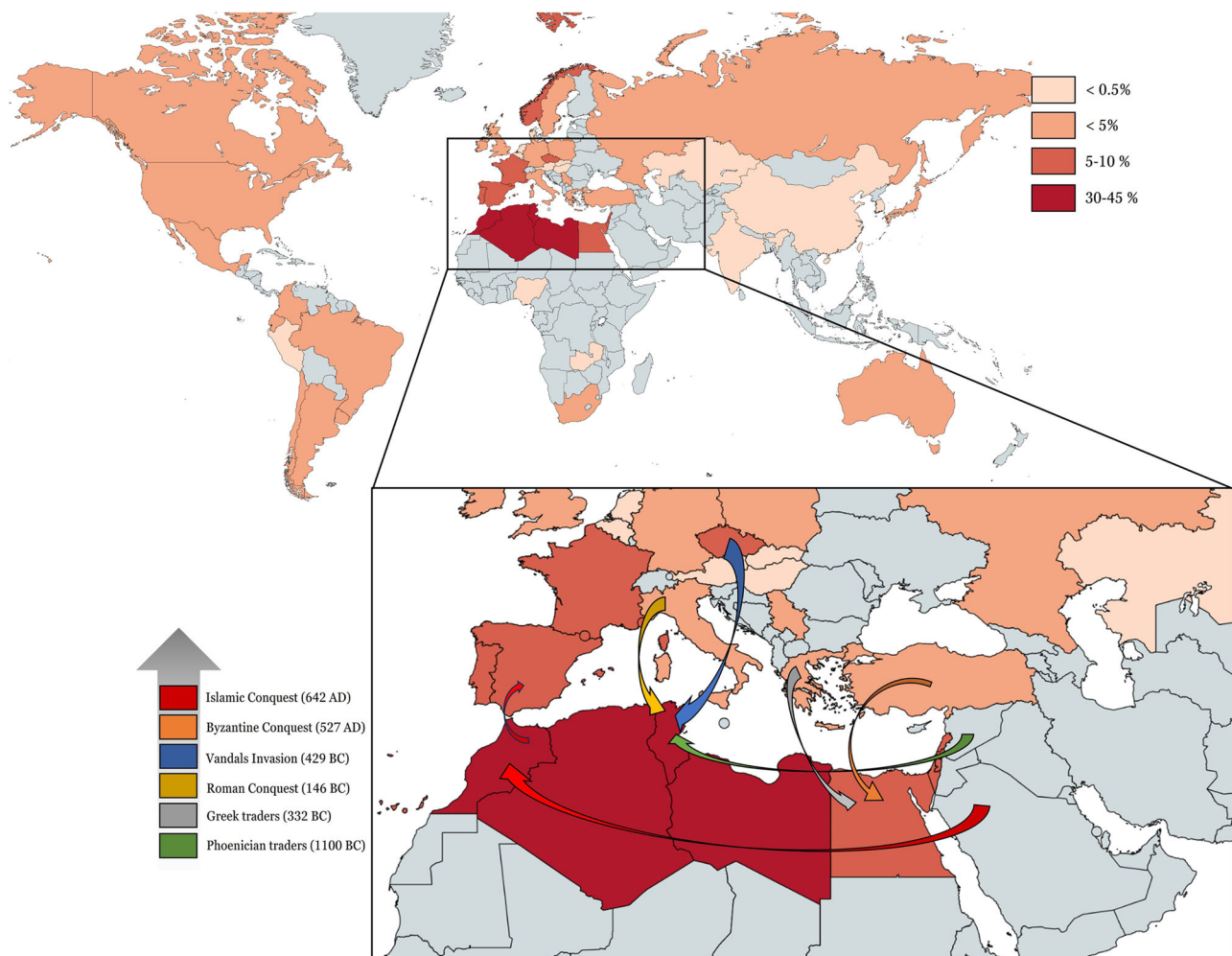


Fig. 2 | A map illustrates that Egypt is centrally located in a region with the highest prevalence of the p.Gly2019Ser pathogenic variant globally. The map's arrows depict potential migration routes through which this pathogenic variant could have been introduced to Egypt. We speculate that the p.Gly2019Ser pathogenic variant might have been brought to Egypt via east-west migration during the

movements of Islamic Arab and North African Berbers, known as the Maghrebi Effect (red arrow), or through maritime invasion and trades activities across the Mediterranean involving the Byzantines, Vandals, Romans, Greeks, and Phoenicians (orange, blue, yellow, gray, and green arrows), known as the “back to Africa” gene flow.

among Arab barbers, with rates of 39% in sporadic PD and 36% in familial PD cases¹⁸. In the east, the prevalence of the pathogenic variant is approximately 10.6% in sporadic PD cases and 26% in familial cases among Ashkenazi Jews, making them the second most affected ethnic group. Southern Mediterranean European countries exhibit higher frequencies of the pathogenic variant (2–4% in sporadic PD and 3–14% in familial PD) compared to Northern European nations (<1–2% in sporadic PD and 0–3% in familial PD)¹⁸. Notably, the p.Gly2019Ser is rare in Asian countries and absent in Sub-Saharan Africa^{50,51}. It is plausible that the p.Gly2019Ser entered Egypt via the east-west axis during the migratory movements of Arab and (or) North African Berbers, referred to as “the Maghrebi Effect” or through north-south axis by the maritime migrations across the Mediterranean, known as the “back to Africa” gene flow (Fig. 2)^{52,53}.

In conclusion, this study provides insight from the most extensive genetic investigation carried out on PD susceptibility in Egyptians to date. It verifies that the p.Gly2019Ser is common in Egypt with prevalence rates similar to those seen in Southern Mediterranean areas. Significantly, it confirms the presence of the p.Gly2019Ser in individuals indigenous to upper Egypt. These findings place Egyptians at the forefront of populations that may be able to host and benefit from *LRRK2* therapeutic programs that are presently in progress. However, our study raises several unresolved questions particularly regarding the penetrance rates and the route of entry of the p.Gly2019Ser into Egypt. Moreover, the lack of clinical scales within

this cohort has impeded the comprehension of clinical characteristics of patients including the understanding of both motor and non-motor symptoms, disease progression, and the efficacy of antiparkinsonian medications. It's crucial to highlight as well that the controls in this cohort are notably younger than the patients. While this does not influence the prevalence or minor allele frequencies of *LRRK2* pathogenic variants, some controls could potentially develop PD later in life, especially since a considerable number are currently under the age of 40. Efforts are underway within our team to increase the number of controls above the age of 50 and to gather additional clinical data to address these limitations. Additionally, we acknowledge that this study focused on assessing only 12 *LRRK2* variants using a targeted approach. A more comprehensive analysis utilizing untargeted next generation whole exomes and whole genome sequencing methods are necessary in this population to investigate both known and novel PD variants, especially in cases of YOPD and EOPD.

Finally, we hope that our study will encourage further investigations on *LRRK2* in the adjacent territories including the Levant, the Arabian Peninsula, and the Nile basin. It will be of interest to explore if the p.Gly2019Ser presence extends beyond upper Egypt into countries that have had historical migratory links with Egypt, such as Sudan, Ethiopia, Oman, and Yemen. The exploration of the *LRRK2* pathogenic variants distributions in these regions poses an interesting question that warrants deeper exploration.

Methods

Participants recruitment

The study was granted ethical approval by the Egyptian participating institutions review boards (Ethics Approval # 2021-2022-058 and 2021-2022-203), as well as University College London (UCL) ethical committee (REC # 22/NE/0080). Written informed consent was obtained from all participants at each study site in accordance with the Declaration of Helsinki and the Common Rule.

Participants, including both patients and controls, were recruited from clinical centers by the IPDGC Africa members in Egypt³³, from two separate cohorts. Cohort one recruited by the Egyptian Network of Neurodegenerative Diseases (ENND), which is a pan Egypt network formed in December 2013 and is based in multiple universities including Mansoura, Cairo, Ain-Shams, Al-Azhar, Alexandria, Asyut, Tanta, Zagazig, Menoufia, Sohag, Al-Minya, Banha, Suez Canal Universities, and The American University in Cairo. Cohort two recruited by independent team from the University of Assiut. The recruitment process took place between 2011 and 2022. Only subjects who self-declared themselves and their both parents as native Egyptians were included in the study.

PD patients were diagnosed by movement disorders specialists using the UK Brain Bank Criteria⁵⁴ and (or) the criteria established by the Movement Disorder Society Task Force⁵⁵. Controls were unrelated-healthy individuals who do not have PD, a known family history of PD or any other neurological condition. They also did not self-report any prodromal symptoms of PD, particularly olfactory dysfunction or sleep disturbances. Basic demographic information including age, gender, self-declared ethnicity, age at disease onset (defined as age at experiencing first motor symptoms), and age at diagnosis (for cases only), were collected from all participants. The Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) were also collected from patients. Familial PD was defined by having one or more first- or second-degree relatives diagnosed with PD, whereas Sporadic PD was identified by the lack of a known family history of the disease. We sought to gather data on antiparkinsonian drug treatments and the equivalent dosage of levodopa, but unfortunately, this information was unavailable or deemed unreliable for a significant number of patients. This was partly due to limited access to medications and patients' inability to consistently obtain PD drugs.

LRK2 SNPs genotyping

DNA was extracted from either 10 mL of venous whole blood samples or saliva samples using established protocols. Genotyping was carried out using the Kompetitive Allele Specific Polymerase Chain Reaction assay (KASP; LGC Genomics, Herts, UK) as described previously⁵⁰. Twelve *LRK2* single-nucleotide polymorphisms (SNPs) were selected for genotyping. These SNPs include rs34637584 (p.Gly2019Ser), rs34995376 (p.Arg1441His), rs74163686 (p.Asn1437His), rs35870237 (p.Ile2020Thr), rs35801418 (p.Tyr1699Cys), rs34410987 (p.Pro755Leu), rs34778348 (p.Gly2385Arg), rs281865052 (p.Met1869Val), rs34594498 (p.Ala419Val), rs33949390 (p.Arg1628Pro), rs35808389 (p.Leu1114Leu), and rs34805604 (p.Ile1122Val). These SNPs were carefully chosen based on strict criteria, including replicated association studies, robust disease segregation data, functional evidence, and the availability of a validated KASP assay for each SNP. To ensure the pathogenicity of each SNP, we checked the ClinVar and MDS Gene databases^{56,57}. Initially, we also included the SNP rs33939927, which codes for p.Arg1441Gly/Cys/Ser. However, we later determined that it was unsuitable for KASP genotyping due to its proximity to rs34995376 (p.Arg1441His) and its tri-allelic nature.

The haplotype background for p.Gly2019Ser carriers was determined by genotyping the rs28903073. The "A" allele for this SNP has a frequency of <0.1% and if observed indicates the presence of haplotype 1⁴⁸. Variant specific primer was designed using primer 3 software (<http://bioinfo.ut.ee/primer3-0.4.0>). PCR was performed using Touchdown with the following cycling conditions: 94 °C for 10 min; 8 cycles of [30 s at 94 °C, 30 s at 65 °C, 45 s at 72 °C]; 18 cycles of [30 s at 94 °C, 30 s at (65–55 °C), 45 s at 72 °C]; 14 cycles of 30 s at 94 °C, 30 s at 55 °C, 45 s at 72 °C; and a final extension step of

10 min at 72 °C. SNP rs28903073 genotyping was performed using standard bi-directional Sanger sequencing.

Data summaries and statistical analysis

The cohort's demographics and clinical characteristics were summarized using appropriate descriptors including mean and standard deviation (SD) for continuous variables and percent frequency for categorical variables and compared between groups (PD and controls) using two-tailed χ^2 test for categorical variables or non-parametric alternative for continuous variables as relevant. A Univariate Analysis of Covariance (ANCOVA), adjusted for age, gender, and disease duration, was conducted to compare UPDRS scores among PD p.Gly2019Ser carriers and non-carriers, and those for patients from lower and upper Egypt. Statistical analysis was performed using the Statistical Package for Social Sciences (SPSS) Statistics for Windows, version 29.0 (IBM Corp, Armonk, NY, USA). The genotyping data for SNPs were analyzed using SNP Viewer software (version 1.99; Hoddesdon, UK) and tested for Hardy-Weinberg equilibrium using the χ^2 test. Results of the sanger sequencing were analyzed using SEQUENCHER (Version 5.4.5, Gene Codes Corporation, Inc).

Verifying the minor allele frequencies (MAF) of the *LRK2* 12 SNPs using the Egyptian genome reference browser

Further, we examined the frequencies of the 12 *LRK2* pathogenic variants in an external dataset comprising WGS data of 110 participants from the Egyptian Genome Reference Project^{5,30}, all of whom were healthy controls aged 40–50 years. For this purpose, we used the web based Variant Query tool, available on the Egyptian Genome Reference website (<https://www.egyptian-genome.org/>). This online tool provides information on variant allele frequencies, as well as the total number of individuals and chromosomes genotyped for each specific variant.

Data availability

Anonymized genotyping and clinical data can be provided to *bone fide* researchers upon request from the corresponding author(s).

Received: 21 May 2024; Accepted: 21 October 2024;

Published online: 06 November 2024

References

- Wolters, E. C., Van Der Werf, Y. D. & Van Den Heuvel, O. A. Parkinson's disease-related disorders in the impulsive-compulsive spectrum. *J. Neurol.* **255**, 48–56 (2008).
- Gupta, S. & Shukla, S. Non-motor symptoms in Parkinson's disease: Opening new avenues in treatment. *Curr. Res. Behav. Sci.* **2**, 100049 (2021).
- Sakowski, S. A., Koubek, E. J., Chen, K. S., Goutman, S. A. & Feldman, E. L. Role of the exposome in neurodegenerative disease: Recent insights and future directions. *Ann. Neurol.* **95**, 635–652 (2024).
- Funayama, M., Nishioka, K., Li, Y. & Hattori, N. Molecular genetics of Parkinson's disease: Contributions and global trends. *J. Hum. Genet.* **68**, 125–130 (2023).
- Wohlers, I. et al. An integrated personal and population-based Egyptian genome reference. *Nat. Commun.* **11**, 4719–4728 (2020).
- Population by Country (2024) – Worldometer. https://www.worldometers.info/world-population/population-by-country/#google_vignette.
- Safiri, S. et al. The burden of Parkinson's disease in the Middle East and North Africa region, 1990–2019: Results from the global burden of disease study. *BMC Public Health* **23**, 107–118 (2023).
- Khedr, E. M., El Fetoh, N. A., Khalifa, H., Ahmed, M. A. & El Beh, K. M. A. Prevalence of non motor features in a cohort of Parkinson's disease patients. *Clin. Neurol. Neurosurg.* **115**, 673–677 (2013).
- El-Tallawy, H. N. et al. Prevalence of Parkinson's disease and other types of Parkinsonism in Al Kharga district, Egypt. *Neuropsychiatr. Dis. Treat.* **9**, 1821–1826 (2013).

10. Blauwendraat, C., Nalls, M. A. & Singleton, A. B. The genetic architecture of Parkinson's disease. *Lancet Neurol.* **19**, 170–178 (2020).
11. Coukos, R. & Krainc, D. Key genes and convergent pathogenic mechanisms in Parkinson disease. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/S41583-024-00812-2> (2024).
12. Bryant, N. et al. Identification of LRRK2 missense variants in the accelerating medicines partnership Parkinson's disease cohort. *Hum. Mol. Genet.* **30**, 454–466 (2021).
13. ClinVar. <https://www.ncbi.nlm.nih.gov/clinvar/>.
14. MDSGene. <https://www.mdsgene.org/>.
15. Wojewska, D. N. & Kortholt, A. LRRK2 targeting strategies as potential treatment of Parkinson's disease. *Biomolecules* **11**, 1101 (2021).
16. Turski, P. et al. Review of the epidemiology and variability of LRRK2 non-p.Gly2019Ser pathogenic mutations in Parkinson's disease. *Front. Neurosci.* **16**, 971270 (2022).
17. Taymans, J. M. et al. Perspective on the current state of the LRRK2 field. *NPJ Park. Dis.* **9**, 104 (2023).
18. Simpson, C. et al. Prevalence of ten LRRK2 variants in Parkinson's disease: A comprehensive review. *Parkinsonism Relat. Disord.* **98**, 103–113 (2022).
19. Shu, L., Zhang, Y., Sun, Q., Pan, H. & Tang, B. A Comprehensive Analysis of Population Differences in LRRK2 Variant Distribution in Parkinson's Disease. *Front. Aging Neurosci.* **11**, 13 (2019).
20. Bouhouche, A. et al. LRRK2 G2019S mutation: Prevalence and clinical features in Moroccans with Parkinson's disease. *Parkinsons. Dis.* **2017**, 2412486 (2017).
21. Simón-Sánchez, J. et al. Parkinson's disease due to the R1441G mutation in Dardarin: a founder effect in the Basques. *Mov. Disord.* **21**, 1954–1959 (2006).
22. Zhang, Y. et al. Genetic Analysis of LRRK2 R1628P in Parkinson's disease in Asian populations. *Parkinsons. Dis.* **2017**, 8093124 (2017).
23. Pan, H. et al. Genome-wide association study using whole-genome sequencing identifies risk loci for Parkinson's disease in Chinese population. *NPJ Park. Dis.* **9**, 22 (2023).
24. Lim, S. Y. et al. Parkinson's disease in the Western Pacific Region. *Lancet Neurol.* **18**, 865–879 (2019).
25. Zimprich, A. et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**, 601–607 (2004).
26. Funayama, M. et al. An LRRK2 mutation as a cause for the parkinsonism in the original PARK8 family. *Ann. Neurol.* **57**, 918–921 (2005).
27. Hashad, D. I., Abou-Zeid, A. A., Achmawy, G. A., Allah, H. M. O. S. & Saad, M. A. G2019S mutation of the leucine-rich repeat kinase 2 gene in a cohort of Egyptian patients with Parkinson's disease. *Genet. Test. Mol. Biomark.* **15**, 861–866 (2011).
28. El Desoky, E. S., Khedr, E. M., Khalil, M. S. & Gasser, T. Genetic analysis of leucine-rich repeat kinase 2 (LRRK2) G2019S mutation in a sample of Egyptian patients with Parkinson's disease, a pilot study. *Br. J. Med. Med. Res.* **5**, 404–408 (2015).
29. Ali, S. & Wszolek, Z. K. LRRK2 R1441C mutation causing Parkinson's Disease in an Egyptian family. *Neurol. Neurochir. Pol.* **56**, 191–192 (2022).
30. Egyptian Genome. <https://www.egyptian-genome.org/>.
31. Trinh, J. et al. Molecular mechanisms defining penetrance of LRRK2-associated Parkinson's disease. *Medizinische Genet* **34**, 103–116 (2022).
32. Khalil, H. et al. Parkinson's Disease in the Middle East, North Africa, and South Asia: Consensus from the International Parkinson and Movement Disorder Society Task Force for the Middle East. *J. Parkinsons. Dis.* **10**, 729–741 (2020).
33. Rizig, M. et al. The international Parkinson disease genomics consortium Africa. *Lancet Neurol.* **20**, 335 (2021).
34. Helmy, A. et al. Baseline predictors of progression of Parkinson's disease in a sample of Egyptian patients: clinical and biochemical. *Egypt. J. Neurol. psychiatry Neurosurg.* **58**, 9 (2022).
35. Shalash, A. S. et al. Non-motor symptoms in essential tremor, akinetic rigid and tremor-dominant subtypes of Parkinson's disease. *PLoS One* **16**, e0245918 (2021).
36. Fahmy, E. M., Elawady, M. E., Sharaf, S., Heneidy, S. & Ismail, R. S. Vitamin D status in idiopathic Parkinson's disease: An Egyptian study. *Egypt. J. Neurol. Psychiatry Neurosurg.* **56**, 1–4 (2020).
37. Khedr, E. M. et al. Epidemiological study and clinical profile of Parkinson's disease in the Assiut Governorate, Egypt: A community-based study. *Neuroepidemiology* **38**, 154–163 (2012).
38. Riboldi, G. M., Frattini, E., Monfrini, E., Frucht, S. J. & Di Fonzo, A. A practical approach to early-onset Parkinsonism. *J. Parkinsons. Dis.* **12**, 1–26 (2022).
39. Rizig, M. et al. Genome-wide association identifies novel etiological insights associated with Parkinson's disease in African and African admixed populations. *medRxiv* <https://doi.org/10.1101/2023.05.05.23289529> (2023).
40. Kolichski, A. et al. Early-onset Parkinson's disease: Creating the right environment for a genetic disorder. *J. Parkinsons. Dis.* **12**, 2353–2367 (2022).
41. Latourelle, J. C. et al. The Gly2019Ser mutation in LRRK2 is not fully penetrant in familial Parkinson's disease: the GenePD study. *BMC Med.* **6**, 32 (2008).
42. Lee, A. J. et al. Penetrance estimate of LRRK2 p.G2019S mutation in individuals of non-Ashkenazi Jewish ancestry. *Mov. Disord.* **32**, 1432–1438 (2017).
43. Healy, D. G. et al. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.* **7**, 583–590 (2008).
44. Lesage, S. et al. Parkinson's disease-related LRRK2 G2019S mutation pathogenic variant results from independent mutation pathogenic variant events in humans. *Hum. Mol. Genet.* **19**, 1998–2004 (2010).
45. Warren, L. et al. A founding LRRK2 haplotype shared by Tunisian, US, European and Middle Eastern families with Parkinson's disease. *Parkinsonism Relat. Disord.* **14**, 77–80 (2008).
46. Watterson, G. A. & Guess, H. A. Is the most frequent allele the oldest? *Theor. Popul. Biol.* **11**, 141–160 (1977).
47. El Haj, R. B. et al. Evidence for prehistoric origins of the G2019S mutation in the North African Berber population. *PLoS One* **12**, e0181335 (2017).
48. Zabetian, C. P. et al. LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. *Am. J. Hum. Genet.* **79**, 752–758 (2006).
49. Kmiecik, M. J. et al. Genetic analysis and natural history of Parkinson's disease due to the LRRK2 G2019S variant. *Brain* **147**, 1996–2008 (2024).
50. Rizig, M. et al. Negative screening for 12 rare LRRK2 pathogenic variants in a cohort of Nigerians with Parkinson's disease. *Neurobiol. Aging* **99**, 101.e15–101.e19 (2021).
51. Okubadejo, N. U. et al. Leucine rich repeat kinase 2 (LRRK2) GLY2019SER mutation pathogenic variant is absent in a second cohort of Nigerian Africans with Parkinson disease. *PLoS One* **13**, e0207984 (2018).
52. Henn, B. M. et al. Genomic ancestry of North Africans supports back-to-Africa migrations. *PLoS Genet.* **8**, e1002397 (2012).
53. Tadmouri, G. O., Sastry, K. S. & Chouchane, L. Arab gene geography: From population diversities to personalized medical genomics. *Glob. Cardiol. Sci. Pract.* **2014**, 394–408 (2014).
54. Hughes, A. J., Daniel, S. E., Kilford, L. & Lees, A. J. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry* **55**, 181–184 (1992).
55. Postuma, R. B. et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* **30**, 1591–1601 (2015).
56. Trinh, J. et al. Genotype-phenotype relations for the Parkinson's disease genes SNCA, LRRK2, VPS35: MDSGene systematic review. *Mov. Disord.* **33**, 1857–1870 (2018).

57. Landrum, M. J. et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018).

Acknowledgements

We would like to thank the patients, their care partners, and the controls for their participation. M.R. received funding from the University College London Grand Challenges Small Grants (Award ID:177813), the MRC neurosciences and mental health research grant. April 2023–April 2026 and The Michael J Fox Foundation Genetic Diversity in Parkinson's Disease 2019 (Grant ID:17483). H.H. received funding from the Michael J Fox Foundation Genetic Diversity in Parkinson's Disease (Grant ID: 17483). M.S. received funding from the Bartlett Fund for Critical Challenges- 2021 (Agreement Number: 2—Cycle 3) and the American University in Cairo Faculty Support Grant-2021 awarded to. The funders were not involved in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization and design: M.R. and M.S. Data acquisition: M.B.W., S.H., A.S., E.M.K., M.H.Y., S.E., G.F., A.H., E.H., M.E., H.L., A.J., M.K., D.M.M., S.E., Y.E., A.S.G., N.S., N.A., T.M.B., N.A.B.E., M.E., S.E., S.R., J.M., L.A., S.N., G.R., M.A.H., M.T.H., H.H., M.S., M.R. Drafting of manuscript: M.B.W. and M.R. Critical revision of manuscript for intellectual content: M.B.W., S.H., A.S., E.M.K., M.H.Y., S.E., G.F., A.H., E.H., M.E., H.L., A.J., M.K., D.M.M., S.E., Y.E., A.S.G., N.S., N.A., T.M.B., N.A.B.E., M.E., S.E., S.R., J.M., L.A., S.N., G.R., M.A.H., M.T.H., H.H., M.S., M.R. Statistical analysis: M.B.W., M.H.Y., and M.R. Administrative, technical, or material support: M.R. and M.S. Supervision of genotyping: M.R. and M.S. Data management: M.R. and M.S. Funding acquisition: M.R., H.H., and M.S. All authors have read and approved the final manuscript version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41531-024-00826-8>.

Correspondence and requests for materials should be addressed to Mohamed Salama or Mie Rizig.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024