

LRRK2 variant spectrum and association study in a multi-ethnic cohort of Malaysian Parkinson's disease patients

Kai Shi Lim

Universiti Malaya

Jia Lun Lim

Universiti Malaya

Maria Teresa Perinán

Universidad de Sevilla

Yi Wen Tay

Universiti Malaya

Tzi Shin Toh

Universiti Malaya

Lei-Cheng Lit

Universiti Malaya

Anis Nadhirah Khairul Anuar

Universiti Malaya

Hans Xing Ding

Universiti Malaya

Khairul Azmi Ibrahim

Hospital Sultanah Nur Zahirah

Ahmad Shahir Mawardi

Hospital Kuala Lumpur

Yuen Kang Chia

Queen Elizabeth Hospital

Joshua Chin Ern Ooi

Queen Elizabeth Hospital

Thien Thein Lim

Island Hospital

Irene Looi

Hospital Seberang Jaya

Jie Ping Schee

Universiti Malaya

Yuan Ye Beh

Universiti Malaya

Wee Kooi Cheah

Hospital Taiping

Wan Chung Law

Sarawak General Hospital

Siaw Cheng Wong

Sarawak General Hospital

Tien Lee Ong

Hospital Sungai Buloh

Mann Leon Chin

Hospital Raja Permaisuri Bainun

Pei Chiek Teh

Tengku Ampuan Rahimah Hospital

Yue Hui Lau

Tengku Ampuan Rahimah Hospital

Christina Lai Ling Lee

Universiti Malaya

Laurel Screven
The Global Parkinson's Genetics Program (GP2)

Sara Bandres-Ciga
National Institutes of Health

Shen-Yang Lim
Universiti Malaya

Ai Huey Tan
Universiti Malaya

Azlina Ahmad-Annuar
`azlina_aa@um.edu.my`

Universiti Malaya

Article

Keywords: Parkinson's disease, LRRK2, Malaysia

Posted Date: September 4th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9924440/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

The variant spectrum of *LRRK2* in the multi-ancestry Malaysian population remains largely uncharacterized. In this study, 2,034 Parkinson's disease (PD) patients and 1,059 controls (Malays, Chinese, Indians and Indigenous) were screened. Seventy-seven heterozygous non-synonymous variants were identified, including four known pathogenic/likely pathogenic variants (p.R1067Q, p.R1325Q, p.R1441C/H) and three risk alleles (p.A419V, p.R1628P, p.G2385R). There were 51 variants of uncertain significance (VUS), including two high-interest VUS (p.A1413T, p.P1446L) and five novel VUS (p.L162F, p.D523N, p.F877S, p.E1566D, p.D1858A). In the Chinese cohort, the p.G2385R variant was significantly associated with increased PD risk by 3.2-fold. After Bonferroni correction, the p.R1628P risk allele and the p.N551K-p.R1398H-p.K1423K protective haplotype associations were not significant. Double heterozygous p.R1628P-p.G2385R patients were observed to have a mean age at diagnosis (AAD) of 54.4 ± 12.1 years while non-*LRRK2* carriers had a mean AAD of 59.0 ± 12.0 years. Domain-based low-frequency and rare variant tests highlighted the WD40 domain, largely driven by the p.G2385R variant, as disease-relevant for the Chinese subgroup. Overall, this study offers the most comprehensive catalog of *LRRK2* variation in the Malaysian population to date, providing a valuable resource for future genetic studies of Southeast Asian ancestries, and aiding in the development of relevant *LRRK2*-targeted therapies.

Introduction

Variability in the Leucine-rich repeat kinase 2 (*LRRK2*) gene is unequivocally linked with Parkinson's disease (PD), whereby rare pathogenic variants (p.R1067Q, p.N1437H, p.R1441C/G/H, p.Y1699C, p.L1795F, p.G2019S, p.I2020T, and others) account for monogenic forms of PD with reduced penetrance¹. Genome-wide association studies (GWAS) and other large-scale genetic studies have identified several *LRRK2* variants (p.A419V, p.R1628P, p.G2385R) as important risk factors in East Asian (EAS) and Central Asian (CAS) populations^{2–8}. However, these variants are very rare in South Asian (SAS) populations, in which no other significant *LRRK2* risk alleles have been consistently identified to date^{2,9}. Known pathogenic and risk alleles are found across the multidomain *LRRK2* protein but cluster mainly in the catalytic core (ROC, COR, and kinase domains), where they are thought to increase kinase activity¹⁰. Apart from the more common risk alleles, the cumulative burden of rare coding variants (minor allele frequency, MAF < 1%) has been reported to increase PD risk by approximately 1.9-fold¹¹. Several case-control studies have also identified a common protective haplotype, p.N551K-p.R1398H-p.K1423K, with an odds ratio (OR) for PD of 0.6–0.9 in Asians^{12–14} and Europeans^{15,16}.

Studies in multiethnic Malaysian cohorts have largely focused on specific *LRRK2* variants, including p.N551K–p.R1398H, p.G2385R, and p.R1628P, providing important insights into population risk and clinical correlates^{2,4,14,17}. Interestingly, we have uncovered a relatively high frequency of the p.R1628P variant in Malay and indigenous populations - groups that remain underrepresented in global genetic studies². However, expanding variant discovery beyond known hotspots is critical to better define the regional genetic architecture of PD and to improve understanding of *LRRK2* biology, population-specific disease risk, and genotype–phenotype relationships in Southeast Asian populations. These efforts are increasingly important in the era of *LRRK2*-targeted clinical trials and precision therapeutic strategies⁴.

Utilising SNP arrays, targeted sequencing, and whole genome sequencing (WGS), this study sought to catalogue pathogenic/likely pathogenic, common, and rare coding *LRRK2* variants present in Malaysian (Chinese, Malay, Indian, and Indigenous) populations, identify novel variants, and perform association, haplotype, and cumulative rare variant burden analyses.

Results

Cohort demographics and ancestry prediction

After removing overlapping samples in the NeuroBooster Array (NBA)¹⁸, WGS, and CENTOGENE next-generation sequencing (NGS)-based PD gene panel datasets (Supplementary Table 1), the cohort consisted of 2,034 patients and 1,059 controls (Demographics: Table 1). The mean age of PD patients was 63.8 ± 10.7 years, while controls had a mean age of 63.1 ± 8.2 years. The mean age of diagnosis (AAD) for PD patients was 59.0 ± 12.0 years, with 24.3% reporting a positive family history ($n = 495$) (family history data missing for $n = 43$). Early-onset PD (EOPD) patients (AAD ≤ 50 years old) accounted for 25.4% ($n = 516/2,034$) of the overall cohort.

Table 1
Cohort demographics for the variant spectrum analysis

Ancestries	CHINESE	MALAY	INDIAN	INDIGENOUS [#]
Total, n	1770	981	274	68
PD, n (%)	1413 (79.83)	412 (42.0)	184 (67.15)	25 (36.76)
Control, n (%)	357 (20.17)	569 (58.0)	90 (32.85)	43 (63.24)
Sex				
Male, n (%)	936 (52.88)	580 (59.12)	153 (55.84)	34 (50.0)
Female, n (%)	834 (47.12)	401 (40.88)	121 (44.16)	34 (50.0)
Mean age, years $\pm$ SD				
All	64.99 $\pm$ 9.64	62.09 $\pm$ 9.75	61.22 $\pm$ 10.15	57.81 $\pm$ 9.8
PD	65.24 $\pm$ 10.05	60.76 $\pm$ 11.23	61.3 $\pm$ 11.11	52.04 $\pm$ 10.16
Control	63.99 $\pm$ 7.76	63.05 $\pm$ 8.41	61.03 $\pm$ 7.88	61.16 $\pm$ 7.93
Family History of PD				
Presence (%)	331 (23.93)	73 (18.07)	81 (45.0)	10 (41.67)
Mean AAD, years $\pm$ SD	60.53 $\pm$ 11.41	55.85 $\pm$ 12.53	55.93 $\pm$ 12.19	45.32 $\pm$ 10.34
	N = 570	N = 176	N = 134	N = 25
Mean AAO, years $\pm$ SD	53.77 $\pm$ 12.26	47.85 $\pm$ 11.66	53.55 $\pm$ 12.27	44.12 $\pm$ 11.05
	N = 1413	N = 412	N = 184	N = 25
PD (Parkinson's disease), SD (standard deviation), N (number of samples with available information), AAD (Age at Diagnosis), AAO (Age at Onset), *Mean age at onset for cases where information was available.				
[#] Indigenous are individuals that self-report as indigenous from Peninsular Malaysia				

Based on ancestry prediction using the SG10K reference panel (Fig. 1b), our cohort consisted of 1,770 Chinese (1,413 PD and 357 controls), 981 Malays (412 PD and 569 controls), and 274 Indians (184 PD and 90 controls). In addition, we included 68 self-reported Indigenous participants (25 PD and 43 controls). There was a 98.2% concordance between genetically-inferred ancestry and self-reported ancestry.

LRRK2 variant spectrum in Malaysian PD patients

After quality control (QC) filtering, 86 heterozygous non-synonymous variants were identified across the three platforms (NBA, WGS, and CENTOGENE). Variants detected in the NBA dataset were cross-validated against WGS and CENTOGENE data, while variants unique to NBA/CENTOGENE were validated using Sanger sequencing (Supplementary Table 2). After this validation step, 77 variants remained for further analysis; 30 were present in both patients and controls, and 47 only in the patients. ACMG classifications were applied to the variants according to pathogenic/likely pathogenic/variants of uncertain significance/benign/likely benign classifications. As ACMG is more applicable to Mendelian inheritance, the ClinGen Low-Penetrance/Risk Allele Working Group recommended classifications¹⁹ were also applied for the variants that have previously been associated with increased/reduced risk of PD. The p.G2385R variant was classed as an 'Established risk allele' (a variant associated with increased risk of disease based on very strong, consistent evidence from association studies with multiple replications), while p.A419V and p.R1628P were classed as 'Likely risk alleles' (variants with strong evidence from association studies in the risk same direction as the 'Established risk allele' but with fewer replication studies). In addition, using the same guidelines, variants in the reported protective haplotype (p.N551K-R1398H) were classed as 'Established likely protective alleles'.

Details of the variants including their location within LRRK2 domains, allele frequencies in patients and controls, population allele frequencies from gnomAD v4.1 and SG10K databases, Clinvar/ACMG/ClinGen Low-Penetrance/Risk Allele classifications¹⁹, kinase activity and RAB10 phosphorylation status (where available)^{10,20}, CADD scores, and relevant published studies are provided in Supplementary Tables 3 and 4. Based on this classification approach, out of the 77 variants, four were classed as pathogenic (P) or likely pathogenic (LP) (4/77; 5.2%), three as risk alleles (3/77; 3.9%), 17 benign/likely benign (17/77; 22.1%), two as likely protective alleles (2/77; 2.6%) and 51 as uncertain significance (VUS)/conflicting classifications of pathogenicity (Franklin/ClinVar), (51/77; 66.2%) (Fig. 2a).

Pathogenic/Likely pathogenic variants

Four pathogenic/likely pathogenic variants (p.R1067Q, p.R1325Q, p.R1441C/H) were identified in nine patients (9/2,034; 0.4%). Key clinico-demographic details of these P/LP patients are presented in Table 2a (mean AAD 59.0 ± 10.5 [range 37–70] years; 25.0% with EOPD; 44.4% with positive family history). The p.R1441C variant was seen exclusively in Chinese patients; $MAF_{Chinese PD}=0.14\%$ ($n = 4$), while p.R1441H was detected in a Malay patient ($MAF_{Malay PD}=0.14\%$, $n = 1$). Meanwhile, p.R1067Q was observed in two Chinese patients and one Indian patient ($MAF_{Chinese PD}=0.07\%$, $n = 2$; $MAF_{Indian PD}=0.27\%$, $n = 1$). The p.R1325Q variant was identified in a single Chinese patient ($MAF_{Chinese PD}=0.04\%$, $n = 1$).

Table 2

PD patients carrying pathogenic/likely pathogenic, double heterozygous risk, high-interest VUS and novel *LRRK2* variants

Sample ID	Gender	Ethnicity	Age at: Symptom onset / Diagnosis / Last FU	Family history	Clinical features
Pathogenic/Likely pathogenic variants					
p.R1067Q					
PD-0330* UMKLM_002042	M	Chinese	64 / 66 / Died age 79	No	MF+ (S); Dysk-; Dementia+; VH-; RBD-
PD-1785* UMKLM_000456	M	Chinese	62 / 62 / 67	No	MF+ (S); Dysk+ (S); Dementia-; VH?; RBD?
PD-2560 KUL_000939	F	Indian	60 / 60 / 68	Paternal aunt with LOPD	MF+; Dysk+; Dementia-; VH-; RBD-
p.R1325Q					
PD-2585 UMKLM_001978	F	Chinese	67 / 69 / 76	No	MF-; Dysk-; Dementia+(S); VH-; RBD?
p.R1441C					
PD-0218** KUL_000642	F	Chinese	37 / 37 / Died age 77	Sister with PD (PD-219)	MF+; Dysk+; Dementia+; VH+; RBD-
PD-0219** KUL_000643	F	Chinese	58 / 58 / Died age 73	Sister with PD (PD-218)	MF+; Dysk+; Dementia-; VH+; RBD-
PD-1374 UMKLM_000309	M	Chinese	60 / 61 / 70	Brother suspected to have PD	MF+ (S); Dysk+; Dementia-; VH+; RBD+
PD-3089 KUL_000631	M	Chinese	48 / 48 / 54	No	N/A
p.R1441H					
PD-1997 UMKLM_002084	M	Malay	70 / 70 / 73	No	MF-; Dysk-; Dementia-; VH-; RBD?
Double heterozygous risk allele carriers					
p.A419V-p.R1628P					
PD-2861 UMKLM_001996	F	Chinese	51 / 51 / 63	No	MF+; Dysk+; Dementia-; VH-; RBD-
p.R1628P-p.G2385R					
* Previously published in Lim SY et al. Clinical and functional evidence for the pathogenicity of the <i>LRRK2</i> p.Arg1067Gln variant. <i>NPJ Parkinsons Dis.</i> 2025 Feb 23;11(1):34. doi: 10.1038/s41531-025-00884-6.					
** Previously published in Lim SY et al. Clinical Phenotype of <i>LRRK2</i> R1441C in 2 Chinese Sisters. <i>Neurodegener Dis.</i> 2020;20(1):39–45. doi: 10.1159/000508131.					
*** Published in Goh et al. <i>LRRK2</i> p.G2385R and p.R1628P variants in a multi-ethnic Asian Parkinson's Cohort: epidemiology and clinical insights. <i>NPJ Parkinsons Dis.</i> 2025 Nov 18;11(1):320. doi: 10.1038/s41531-025-01166-x.					
^^Individual sequenced via CENTOGENE, no GP2 ID available					
Dx=Diagnosed; Dysk=Dyskinesias; FU=Follow-up; LOPD=Late-onset Parkinson's disease; MF=Motor fluctuations; N/A = Not available; RBD = REM sleep behaviour disorder symptoms; VH=Visual hallucinations; WM=White Matter; M=Male, F=Female; ?=Unknown.					

Sample ID	Gender	Ethnicity	Age at: Symptom onset / Diagnosis / Last FU	Family history	Clinical features
Pathogenic/Likely pathogenic variants					
p.R1067Q					
PD-1042 KUL_000237 Plus <i>GCH1</i> p.S80N likely pathogenic/risk allele	F	Chinese	44 / 44 / 64	No	MF+ (S); Dysk+ (S); Dementia-; VH-; RBD?
PD-2496 KUL_000105	M	Chinese	77 / 78 / 81	Sister with EOPD (PD-0390)	MF- ; Dysk-; Dementia+(S)*; VH+; RBD-; *(also had prominent WM T2-hyperintensities on brain MRI)
PD-0271*** KUL_000162	M	Chinese	45 / 45 / Died 60 years	Uncertain	MF+ ; Dysk-; Dementia+; VH+; RBD+
PD-0390*** KUL_000104	F	Chinese	49 / 49 / Died 84 years	Brother with LOPD (PD-2496)	MF+; Dysk+; Dementia+(S)*; VH+; RBD+
PD-0569*** UMKLM_000117	F	Chinese	59 / 59 / 61	No	MF- ; Dysk-; Dementia-; VH-; RBD?
PD-1350*** KUL_000682	M	Chinese	65 / 65 / 76	Multiple (4) 2nd- or 3rd-degree paternal and maternal relatives with PD, mostly LOPD	MF+; Dysk+; Dementia-; VH+; RBD+
PD-1610*** KUL_000821	M	Chinese	54 / 54 / 66	No	MF+; Dysk+; Dementia-; VH-; RBD-
PD-2219*** KUL_000517	M	Chinese	45 / 45 / 49	No	MF-; Dysk-; Dementia-; VH?; RBD?
PD-2524*** UMKLM_000727	F	Chinese	64 / 67 / 69	No	MF+; Dysk-; Dementia-; VH+; RBD-
PD-2587*** KUL_000942	M	Chinese	55 / 55 / 66	Father and brother with tremor in older age, but not formally Dx with PD	MF+; Dysk+; Dementia+; VH+; RBD+
PD-2862*** KUL_000116 Plus 'severe' <i>GBA1</i> RecNcil variant	M	Chinese	63 / 66 / 72	Distant relative with LOPD (but who is negative for both <i>LRRK2</i> risk alleles)	MF+; Dysk+; Dementia-; VH+; RBD+
PD-2969*** KUL_000313	F	Chinese	36 / 37 / 48	Mother and grandmother apparently with similar symptoms	MF+; Dysk-; Dementia-; VH?; RBD-
High interest VUS					
p.A1413T					
PD-2016	M	Malay	66 / 68 / 79	No	MF+; Dysk-; Dementia-; VH+; RBD+
* Previously published in Lim SY et al. Clinical and functional evidence for the pathogenicity of the <i>LRRK2</i> p.Arg1067Gln variant. <i>NPJ Parkinsons Dis.</i> 2025 Feb 23;11(1):34. doi: 10.1038/s41531-025-00884-6.					
** Previously published in Lim SY et al. Clinical Phenotype of <i>LRRK2</i> R1441C in 2 Chinese Sisters. <i>Neurodegener Dis.</i> 2020;20(1):39–45. doi: 10.1159/000508131.					
*** Published in Goh et al. <i>LRRK2</i> p.G2385R and p.R1628P variants in a multi-ethnic Asian Parkinson's Cohort: epidemiology and clinical insights. <i>NPJ Parkinsons Dis.</i> 2025 Nov 18;11(1):320. doi: 10.1038/s41531-025-01166-x.					
^^Individual sequenced via CENTOGENE, no GP2 ID available					
Dx=Diagnosed; Dysk=Dyskinesias; FU=Follow-up; LOPD=Late-onset Parkinson's disease; MF=Motor fluctuations; N/A = Not available; RBD = REM sleep behaviour disorder symptoms; VH=Visual hallucinations; WM=White Matter; M=Male, F=Female; ?=Unknown.					

Sample ID	Gender	Ethnicity	Age at: Symptom onset / Diagnosis / Last FU	Family history	Clinical features
Pathogenic/Likely pathogenic variants					
p.R1067Q					
KUL_000878					
Plus <i>LRRK2</i> p.R1628P risk allele					
PD-0502	M	Malay	67 / 67 / 72	No	MF?; Dysk?; Dementia?; VH+; RBD?
KUL_000662					
As reported above for the novel p.E1566D variant (Table 2a)					
PD-0249^	M	Malay	60 / 60 / 75	No	MF+; Dysk+; Dementia-; VH+; RBD+
PD-1196	M	Malay	50 / 50 / 59	Maternal aunt with LOPD	MF+; Dysk+ (S); Dementia+; VH?; RBD?
KUL_000406					
Also likely pathogenic, severe <i>GBA1</i> variant p.N227K					
p.P1446L					
PD-1948	M	Chinese	76 / 77 / 86	No	MF-; Dysk-; Dementia-; VH-; RBD-
UMKLM_000528					
PD-0228^	F	Chinese	75 / 75 / 88	No	MF-; Dysk-; Dementia?; VH-; RBD-
Novel variants					
p.L162F	M	Chinese	75 / 76 / Died 78 years	Father had hand shakes and walking difficulty, but was not formally Dx with PD	MF+; Dysk-; Dementia-; VH+; RBD+
PD-2430					
KUL_000922					
(double heterozygous with the <i>LRRK2</i> p.R1628P risk allele)					
p.D523N	M	Chinese	72 / 72 / 80	No	MF-; Dysk-; Dementia-; VH-; RBD+
PD-2050					
UMKLM_000562					
p.F877S	M	Malay	60 / 61 / 73	No	MF+; Dysk?; Dementia-; VH?; RBD?
PD-2002					
UMKLM_002088					
p.E1566D	M	Malay	66 / 67 / 72	No	MF?; Dysk?; Dementia?; VH+; RBD?
PD-0502					
KUL_000662					
* Previously published in Lim SY et al. Clinical and functional evidence for the pathogenicity of the <i>LRRK2</i> p.Arg1067Gln variant. <i>NPJ Parkinsons Dis.</i> 2025 Feb 23;11(1):34. doi: 10.1038/s41531-025-00884-6.					
** Previously published in Lim SY et al. Clinical Phenotype of <i>LRRK2</i> R1441C in 2 Chinese Sisters. <i>Neurodegener Dis.</i> 2020;20(1):39–45. doi: 10.1159/000508131.					
*** Published in Goh et al. <i>LRRK2</i> p.G2385R and p.R1628P variants in a multi-ethnic Asian Parkinson's Cohort: epidemiology and clinical insights. <i>NPJ Parkinsons Dis.</i> 2025 Nov 18;11(1):320. doi: 10.1038/s41531-025-01166-x.					
^^Individual sequenced via CENTOGENE, no GP2 ID available					
Dx=Diagnosed; Dysk=Dyskinesias; FU=Follow-up; LOPD=Late-onset Parkinson's disease; MF=Motor fluctuations; N/A = Not available; RBD = REM sleep behaviour disorder symptoms; VH=Visual hallucinations; WM=White Matter; M=Male, F=Female; ?=Unknown.					

Sample ID	Gender	Ethnicity	Age at: Symptom onset / Diagnosis / Last FU	Family history	Clinical features
Pathogenic/Likely pathogenic variants					
p.R1067Q					
<i>(double heterozygous with a LRRK2 VUS p.A1413T)</i>					
p.D1858A PD-2927 KUL_000977	F	Chinese	64 / 65 / 74	Mother with LOPD	MF-; Dysk-; Dementia+; VH-; RBD+
* Previously published in Lim SY et al. Clinical and functional evidence for the pathogenicity of the <i>LRRK2</i> p.Arg1067Gln variant. <i>NPJ Parkinsons Dis.</i> 2025 Feb 23;11(1):34. doi: 10.1038/s41531-025-00884-6.					
** Previously published in Lim SY et al. Clinical Phenotype of <i>LRRK2</i> R1441C in 2 Chinese Sisters. <i>Neurodegener Dis.</i> 2020;20(1):39–45. doi: 10.1159/000508131.					
*** Published in Goh et al. <i>LRRK2</i> p.G2385R and p.R1628P variants in a multi-ethnic Asian Parkinson's Cohort: epidemiology and clinical insights. <i>NPJ Parkinsons Dis.</i> 2025 Nov 18;11(1):320. doi: 10.1038/s41531-025-01166-x.					
^^Individual sequenced via CENTOGENE, no GP2 ID available					
Dx=Diagnosed; Dysk=Dyskinesias; FU=Follow-up; LOPD=Late-onset Parkinson's disease; MF=Motor fluctuations; N/A = Not available; RBD = REM sleep behaviour disorder symptoms; VH=Visual hallucinations; WM=White Matter;M=Male,F=Female;?=Unknown.					

Established risk alleles/protective alleles

Association analyses for the established Asian risk alleles were identified in this cohort (p.A419V, p.R1628P, and p.G2385R)^{2,4–8,21} and the protective haplotype p.N551K-p.R1398H-p.K1423K^{12,14,22} are described below.

Double heterozygous carriers of risk alleles

Thirteen patients carrying double heterozygous risk alleles (p.A419V-p.R1628P and p.R1628P-p.G2385R) were identified in our cohort (Table 2). Only one p.A419V-p.R1628P patient was identified (AAD 59 years, no family history) while the rest were double heterozygous for p.R1628P-p.G2385R (n = 12/13) (mean AAD: 55.3 ± 12.0 [range: 37–67] years, 45.5% EOPD and 45.5% with a positive family history). Ten of the p.R1628P-p.G2385R patients have previously been reported by our group². Two p.R1628P-p.G2385R patients harbored variants in other PD-associated genes; one with a ‘severe’ *GBA1* RecNcil variant, with an AAD of 66 years and a positive family history of PD, while the other patient had a heterozygous likely pathogenic *GCH1* p.S80N variant with AAD of 44 and negative family history.

Variants of uncertain significance (VUS)

A total of 51 VUS were identified in 89 patients and 37 Controls (Supplementary Table 3), with fourteen observed only in patients. Twenty-five VUS have been previously reported: 14 are listed in the MDSGene database from family-based or association studies, and 11 in other PD studies (Supplementary Table 4). Functional data were available for 10 VUS, including evidence from animal models (p.P1446L) and *in vitro* kinase activity assays (p.L153W, p.A211V, p.A312V, p.T776M, p.A1413T, p.S1443A, p.P1446L, p.R1725Q, p.I2012T, p.V2390M)^{10,20,23–26}.

To further evaluate the VUS, we first investigated VUS with increased kinase activity whereby only the p.A1413T variant had increased kinase activity (> 1.5). This variant was previously reported in a Filipino PSP patient²⁷ and has a very low frequency in control databases (MAF gnomAD_EAS = 0.00%, MAF SG10K_Chinese/Malay < 0.5%). It is classed as deleterious by multiple prediction tools and thus this variant was considered as a ‘high-interest VUS’.

The remaining 47 VUS had no increase in kinase activity (< 1.5, n = 6) or no data available (n = 41). We investigated the VUS for further genetic evidence, in which thirteen VUS are reported in other PD patients in MDSgene and had MAF < 1% in control databases. Eight VUS had CADD scores > 20 while seven of them had in vitro models which were inconclusive for PD-relevance. A knock-in mouse model recapitulating PD-like features of age-related motor impairment and dopaminergic neurodegeneration was reported for the p.P1446L variant. We acknowledge that this animal study does not provide LRRK2 kinase activity/Rab10 phosphorylation information, however, taking together all the evidence including its location within the conserved ROC domain and supported by computational predictions, these findings provide justification to highlight the p.P1446L as a high-interest VUS for further investigation.

The p.A1413T variant was found in four Malay PD patients (mean AAD 61.3 [range 50–68] years, no family history) and seven controls (age at collection 50–68 years). One p.A1413T patient (AAD 50 years) also harboured the ‘severe’ *GBA1* p.N277K variant, while another p.A1413T patient (AAD 68 years) co-carried the *LRRK2* p.R1628P risk alleles. Meanwhile, the p.P1446L variant was found in two unrelated Chinese PD patients (AAD 77 and 75 years), both with no reported family history of PD. The motor and non-motor features of these high-interest *LRRK2* VUS patients are presented in Table 2.

Evidence of potential pathogenicity from the literature for the remaining 15 variants (Supplementary Tables 4 and 5) was inconclusive, as they lacked segregation analyses, functional validation, or consistent association in case-control studies.

Five of the VUS (p.L162F, p.D523N, p.F877S, p.E1566D, p.D1858A; Table 2) were novel variants, absent in gnomAD v4.1, SG10K, and in-house control datasets, and had CADD scores ≥ 20 . All had ConSurf conservation scores ≥ 7 , except for p.F877S which had a lower conservation score of 3. All the variants were found in late-onset PD (LOPD) patients, and the motor and non-motor features of these patients are presented in Table 2. The p.L162F, p.D523N, and p.D1858A variants were found in Chinese patients while the p.F877S and p.E1566D variants were in Malays. Positive family history of PD was reported in the p.L162F and p.D1858A patients. The p.L162F patient additionally harbored the p.R1628P risk alleles, while the p.E1566D patient also co-carried one of the high-interest VUS (p.A1413T); phasing was unknown for both carriers. None of the patients harbored pathogenic or likely pathogenic variants in other monogenic PD genes or *GBA1* (Table 2).

LRRK2 association study

Association analysis was conducted using NBA genotyping data which had the largest sample size compared to the other datasets (Supplementary Tables 1 and 5, $n_{\text{patients}}=1,851$ and $n_{\text{controls}}=1,051$). Power calculation showed that the Chinese cohort achieved $\geq 80\%$ power to detect variants with OR 3.0 at MAF ≥ 0.01 , OR 1.7 at MAF ≥ 0.04 , and OR 1.5 at MAF ≥ 0.08 , while the Malay cohort reached $\geq 80\%$ power for OR 3.0 at MAF ≥ 0.02 , OR 1.7 at MAF ≥ 0.06 , and OR 1.5 at MAF ≥ 0.12 (Supplementary Table 6). In contrast, the Indian and Indigenous cohorts were underpowered at the current sample sizes.

Sixteen coding variants (MAF $> 1\%$) that were present in both patients and controls were included for association analysis in the Chinese and Malays (Table 3). In the Chinese cohort, the p.G2385R variant conferred a 3.2-fold increased risk of PD (MAF_{patients} 4.53% vs MAF_{controls} 1.59%; OR 3.25, 95% CI: 1.71–6.14, $p = 0.0003$, Table 3) and the p.R1628P variant was nominally associated with an increased PD risk by 1.7-fold (MAF_{patients} 4.61% patients vs MAF_{controls} 3.17%; OR 1.71, 95% CI: 1.06–2.75, $p = 0.028$). The known protective haplotype p.N551K-p.R1398H-p.K1423K was nominally associated with lower PD risk (MAF_{patients} 8.20% vs MAF_{controls} 11.10%; OR 0.68, 95% CI: 0.5–0.9, $p = 0.0003$) (Table 3). After Bonferroni correction, only the p.G2385R variant remained statistically significant ($P_{\text{BF threshold}}=0.003$). The rarer risk allele, p.A419V (MAF_{patients} 0.59% vs MAF_{controls} 0.43%), was not significantly associated with PD (OR 1.41, 95% CI: 0.40–4.98, $p = 0.60$).

Table 3

Case-control association results for *LRRK2* variants in Chinese and Malay ancestries (16 variants, MAF > 1%, highest - lowest p-values)

CHINESE				MALAY							
SNP (HG38)	A1	AA change	Genotyped/Imputed	MAF PD	MAF CTRL	P adj	ADJUSTED (SEX, AGE, 3 PCs) OR (95% CI)	MAF PD	MAF CTRL	P adj	ADJUSTED (SEX, AGE, 10 PCs) OR (95% CI)
12:40363526	A	p.G2385R	Genotyped	4.53%	1.59%	0.0003	3.245 (1.714–6.144)	0.28%	0.36%	0.283	0.385 (0.068–2.195)
12:40263898	G	p.N551K	Genotyped	8.20%	11.24%	0.007	0.679 (0.514–0.898)	4.65%	6.18%	0.178	0.734 (0.468–1.151)
12:40309185	A	p.K1423K	Genotyped	8.23%	11.10%	0.010	0.693 (0.524–0.917)	4.65%	6.09%	0.219	0.754 (0.481–1.183)
12:40309109	A	p.R1398H	Genotyped	8.24%	11.10%	0.010	0.694 (0.524–0.918)	4.66%	6.09%	0.191	0.741 (0.472–1.161)
12:40320043	C	p.R1628P	Genotyped	4.61%	3.17%	0.028	1.705 (1.058–2.747)	3.81%	2.42%	0.172	1.497 (0.839–2.669)
² 12:40320099	A	p.S1647T	Genotyped	36.29%	33.14%	0.098	1.163 (0.972–1.390)	37.89%	37.37%	0.703	1.040 (0.850–1.273)
12:40348452	A	p.E2108E	Genotyped	36.45%	33.29%	0.103	1.160 (0.971–1.387)	37.89%	37.19%	0.676	1.044 (0.852–1.280)
12:40237989	C	p.L153L	Genotyped	36.91%	34.15%	0.140	1.143 (0.957–1.365)	42.68%	41.65%	0.605	1.055 (0.861–1.293)
12:40294893	C	p.L953L	Genotyped	16.90%	18.73%	0.202	0.866 (0.694–1.080)	9.15%	8.42%	0.915	0.981 (0.695–1.386)
12:40363528	G	p.G2385G	Genotyped	1.78%	2.31%	0.226	0.694 (0.384–1.253)	1.13%	1.25%	0.747	0.859 (0.341–2.163)
12:40283897	T	p.P755L	Genotyped	.	.	.	.	1.69%	1.52%	0.451	1.355 (0.615–2.983)
12:40320032	A	p.G1624G	Genotyped	49.72%	48.99%	0.784	1.024 (0.863–1.215)	38.03%	39.50%	0.413	0.918 (0.748–1.126)
12:40320071	G	p.K1637K	Genotyped	48.11%	48.85%	0.803	0.978 (0.824–1.162)	39.41%	40.56%	0.527	0.937 (0.764–1.147)
12:40322458	C	p.G1819G	Genotyped	48.19%	48.85%	0.835	0.982 (0.827–1.166)	39.30%	40.86%	0.402	0.917 (0.749–1.123)
12:40364850	C	p.M2397T	Genotyped	48.19%	47.84%	0.945	1.006 (0.848–1.193)	46.48%	47.76%	0.703	0.962 (0.788–1.174)
Minor allele frequency (MAF); Odds ratios (ORs); 95% confidence intervals (CIs); p adj (p-values after adjusted for covariates). p-values < 0.05 were indicated in bold. A1 = Effect allele.											
The variant in bold reached Bonferroni adjusted significance.											
“.” Dot indicate association was not carried out due to variant did not pass MAF cut off or it only present in either cases or controls only											

CHINESE								MALAY			
SNP (HG38)	A1	AA change	Genotyped/Imputed	MAF PD	MAF CTRL	ADJUSTED (SEX, AGE, 3 PCs)		MAF PD	MAF CTRL	ADJUSTED (SEX, AGE, 10 PCs)	
						P adj	OR (95% CI)			P adj	OR (95% CI)
12:40364850	C	p.M2397M	Imputed ($R^2 = 1$)	48.14%	47.84%	0.981	1.002 (0.845–1.189)	.	.	.	.
Minor allele frequency (MAF); Odds ratios (ORs); 95% confidence intervals (CIs); p adj (p-values after adjusted for covariates). p-values < 0.05 were indicated in bold. A1 = Effect allele.											
The variant in bold reached Bonferroni adjusted significance.											
“.” Dot indicate association was not carried out due to variant did not pass MAF cut off or it only present in either cases or controls only											

Table 4

Domain-based SKAT and SKAT-O analyses for *LRRK2* variants with MAF < 5% using WGS data. (A) Domain-based rare-variant association tests performed using SKAT and SKAT-O. NumVar indicates the number of variants identified within each domain. FDR_BF denotes the Bonferroni-corrected P-value. $P < 0.05$; $*P < 0.01$. (B) Leave-one-out analysis for SKAT and SKAT-O tests to evaluate the contribution of individual variants to the observed domain-level associations

(A) SKAT and SKAT-O test per domain for <i>LRRK2</i> variants found in the Malaysian Chinese ancestry							
Chinese							
SKAT				SKAT-O			
Domain	NumVar	P-value	FDR_BF	rho	P-value	FDR_BF	Variants
ARM	16	0.876	1	1	0.872	1	p.K53R, p.V62V, p.Q116R, p.L153L, p.L153W, p.L162F, p.S231P, p.I266T, p.Y302Y, p.A356V, p.V366M, p.D392D, p.A419V, p.V458V, p.S476S, p.D523N, p.N551K, p.F676S
ANK	6	0.140	0.982	0	0.219	1	p.I723V, p.P755L, p.A774A, p.L775W, p.T776M, p.I786F
LRR	13	0.015	0.105	0	0.029	0.204	p.E854G, p.C925Y, p.L953L, p.S955L, p.M968K, p.S971S, p.I984V, p.S1001fs, p.I1006V, p.K1015K, p.L1040S, p.R1067Q, p.Q1111H, p.R1223R
ROC	9	0.309	1	0	0.462	1	p.L1388I, p.L1390L, p.R1398H, p.Y1402Y, p.R1412X, p.K1423K, p.R1441C, p.P1446L, p.V1450I, p.T1491I, p.I1504L
COR	14	0.207	1	1	0.144	1	p.E1537E, p.R1538H, p.P1542S, p.L1556L, p.C1618R, p.G1624G, p.R1628P, p.R1628H, p.K1637K, p.S1647T, p.L1694L, p.E1745E, p.E1760A, p.C1774Y, p.E1789Q, p.G1819G, p.H1822R, p.D1858A
KINASE	6	0.610	1	1	0.176	1	p.R1915R, p.R1943W, p.I2012T, p.Y2018Y, p.N2047N, p.E2108E, p.I2131S
WD40	10	0.002	0.015*	0	0.004**	0.028*	p.K2148K, p.V2205V, p.H2206Y, p.T2310A, p.N2328K, p.G2385R, p.G2385G, p.M2397T, p.R2446C, p.S2454S, p.W2496L
(B) Leave-one-out analysis for SKAT and SKAT-O test per domain for Chinese ancestry							
SKAT		SKAT-O		Variant dropped out			
Domain	p-value	rho	p-value				
WD40	0.212	0.5	0.309	p.G2385R			

None of the coding and non-coding variants in the Chinese cohort were found to be in LD ($r^2 > 0.8$) with the significant/nominally-significant variants, suggesting independent association signals for p.N551K-p.R1398H-p.K1423K, p.R1628P, and p.G2385R (Supplementary Table 7a, b). No coding variants reached statistical significance in the Malays, and no non-coding variants with MAF > 1% surpassed the Bonferroni-corrected threshold in either ancestry (Table 3, Supplementary Tables 7c,d, 8).

Patients were grouped as carriers of P/LP variants, single risk allele, double heterozygous risk alleles, protective alleles, and high-interest VUS, and compared with non-carriers for differences in AAD (Fig. 3). None of the group comparisons reached statistical significance. We observed that the

double-heterozygous risk allele carriers *LRRK2* p.R1628P-p.G2385R had a mean AAD of 54.4 ± 12.1 years (range: 37–78 years) compared to the non-*LRRK2* carriers who had a mean AAD of 59.0 ± 12.0 years (range from 19–90 years).

LD blocks and haplotype analysis

A total of seven *LRRK2* LD blocks (size range: 1.75–58.21 kb) were identified in the Chinese group, 4 in the Malays (size range: 11.68–79.25 kb), and 8 in the Indians (size range: 3.47–45.42 kb) (Supplementary Table 9). None of the Chinese, Malay and Indian haplotypes derived from variants with MAF > 5% showed statistically significant association with PD after multiple testing corrections (Supplementary Table 9).

Domain and MAF-based low-frequency and rare variant burden testing

Low-frequency and rare coding variants (MAF < 5%) were grouped by functional domain. A significant association for the WD40 domain ($n_{\text{variant}}=10$) was seen in the Chinese cohort (Table 3a, Supplementary Table 10a,b; $P_{\text{SKAT_BF}} = 0.021$, $P_{\text{SKAT-O_BF}}=0.04$). It was observed that when the ten variants were consecutively excluded during the systematic leave-one-out analysis, the significant association of the WD40 domain was lost only when the p.G2385R variant was excluded ($P_{\text{SKAT}}=0.212$, $P_{\text{SKAT-O}}=0.309$) (Table 3b, Supplementary Table 10c). Rare non-coding variants were grouped according to MAF thresholds (MAF < 5% and < 1%), with no significant association was observed in either group (Supplementary Table 10d).

Discussion

This study represents the most comprehensive *LRRK2* variant spectrum analysis in multi-ancestry Malaysian PD patients to date and, to the best of our knowledge, the largest in Southeast Asia. With ongoing precision medicine initiatives targeting *LRRK2* variants and limited information available for under-represented populations in this region, this study provides a genetic resource of *LRRK2* variants in Southeast Asian ancestries. We observed pathogenic/likely pathogenic variants (p.R1067Q, p.R1325Q, p.R1441C/H) in 0.4% (9/2,034) patients across various ancestries (Chinese, Malay and Indian). This study extends earlier reports of Malaysian p.R1067Q and p.R1441C PD patients^{28,29}, and adds a new, diverse ancestries for *LRRK2* variants reported in the current literature – representing first reports of an Malay p.R1441H PD patient and only the second South Asian p.R1067Q PD carrier^{30–34}. We also add a supportive case of likely pathogenicity for the rare *LRRK2* p.R1325Q variant^{11,15,35–38} in a Malaysian Chinese PD patient.

Approximately 14% patients carried at least one risk allele (p.A419V/p.R1628P/p.G2385R), with double heterozygous risk allele patients accounting for 0.7% of the cohort. The p.A419V (MAF = 0.53%) and p.G2385R (MAF = 4.39%) variants were seen mainly in those of Chinese ancestry, while p.R1628P was seen across Chinese, Malay, and Indigenous groups (combined MAF = 4.16%). In line with previous reports^{6,7,17,39}, the 'Asian' risk allele p.G2385R was significantly associated with PD by 3.2-fold while the p.R1628P allele is nominally associated with PD risk by 1.7-fold in the Malaysian Chinese population. The rarer p.A419V risk allele was not significantly associated with PD in this study, likely due to limited statistical power. However, our Malaysian Chinese cohort was part of a recent, large multi-ancestry GP2 study on p.A419V which validated p.A419V as a significant EAS risk allele (OR = 2.9; 95% CI: 1.66–5.10; $p = 0.0002$)⁴⁰.

Similar to a recent study by our group², single p.R1628P or p.G2385R patients did not have an earlier AAD compared to non-*LRRK2* variant carriers. While our previous study with 14 double heterozygous p.R1628P-G2385R patients indicated a significantly earlier AAD (by 6.4 years), this difference was not statistically significant in the group of eleven double heterozygous p.R1628P-G2385R patients here, likely due to the smaller sample size in the current study.

One of the main aims for this study was to catalogue the *LRRK2* VUS in our cohort, and to investigate their relevance to disease pathogenicity/risk. A considerable proportion of patients (4.4%, 89/2,034) were found to carry *LRRK2* VUS. A comprehensive literature search was performed for all the variants, for which kinase activity data/reported functional studies were available for 19.6% (10/51) variants. Kinase activity and RAB10 phosphorylation data for selected *LRRK2* VUS^{20,41} have been highly valuable in variant interpretation, assisting to distinguish certain VUS as 'high interest'. For example, the *LRRK2* p.V1447L variant was recently reclassified from VUS to likely pathogenic when extensive functional data, including elevated Rab10 phosphorylation, became available⁴¹. Emerging data from animal models and cell-based studies²⁴ on the 'high-interest' p.P1446L variant, seen in two Malaysian Chinese LOPD patients (and previously reported in Japanese and Korean EOPD)^{42,43}, also suggest a likely pathogenic role in PD, although further validation studies are needed to more fully assess its disease-relevance. Evidence of pathogenicity for the other high-interest VUS p.A1413T comes from the increased kinase activity and *in silico* prediction, but it remains inconclusive due to the lack of supportive genetics data.

Five novel VUS (p.L162F, p.D523N, p.F877S, p.E1566D, and p.D1858A) were identified in this study and their relevance to PD is currently unclear. The p.L162F and p.D523N variants are located in the ARM domain, recognised as an important binding site for RAB proteins^{44,45} thus these variants may affect the critical *LRRK2*-RAB signalling pathway. Two variants, p.E1566D and p.D1858A, are located in the COR A and B domains, which are sites for *LRRK2* dimerization and interaction with the ROC domain for GTPase^{46,47}, and represent hotspot sites for pathogenic PD

variants. Efforts to obtain additional family pedigree information and clinical history, together with functional studies including kinase activity and RAB10 phosphorylation assays⁴⁸, are currently ongoing to further clarify their potential role in PD.

Previous domain-based burden analyses of *LRRK2* have reported enrichment predominantly within the kinase domain⁴⁹, although the study focused on rare *LRRK2* variants in European populations ($P_{FDR} = 1.61 \times 10^{-22}$). Similar burden analysis in EAS PD cohorts investigated a combined effect of *LRRK2* 'gene-based' burden but had limited resolution to identify enrichment in a particular domain¹¹. In this study, burden testing by domain in the Chinese cohort showed a significant signal that was confined to the WD40 domain, which was largely driven by the p.G2385R variant. Whether different ancestral backgrounds may have distinct *LRRK2* domains that contribute to PD susceptibility is an area that should be explored in larger cohorts with diverse ancestral backgrounds.

We utilised the SG10K dataset, a Southeast Asian reference panel⁵⁰ which improved population stratification and helped reduce heterogeneity confounders in the association studies. This underscores the importance of population-matched reference panels for genetic studies in Southeast Asia, where underrepresentation in global reference panels may obscure disease-associated signals. Encouragingly, several regional resources within Southeast Asia and Asia to generate population genomic databases are currently underway, including Genome Asia 100K⁵¹, South and East Asian reference Database (SEAD) panel⁵², as well as national projects in Malaysia, Indonesia, Thailand, and Vietnam.

There are several limitations to our study. Despite this being the largest *LRRK2* genetics study in Malaysia so far, an even larger and balanced sample in cases and controls is needed to detect rarer variants and to provide increased statistical power to test for new risk and protective associations. Although previous studies have reported the p.R1628P variant as a risk allele and the p.N551K-p.R1398H-p.K1423K haplotype as protective^{13,15,53,54}, these associations did not remain significant after the multiple testing correction in our cohort. Furthermore, the uneven sizes of the ancestral groups also limited a true comparison of allele frequencies across groups, and efforts are underway to obtain greater representation e.g., by including more Malay and indigenous participants^{55,56}. Unavailability of family members' DNAs and clinical histories hampered efforts to perform segregation studies for the VUS including the novel variants, and the lack of functional cellular data, which are resource-intensive to obtain^{48,57}, impedes interpretation of the pathogenic effects of these variants on *LRRK2* kinase activity.

Despite these limitations, the study builds on existing data and currently stands as the most comprehensive characterization of *LRRK2* coding variation across a multi-ethnic Malaysian PD cohort, encompassing Chinese, Malay, Indian, and Indigenous populations. Our study is relevant for countries regionally in Southeast Asia who share closely matched ancestry such as Singapore, Indonesia, and the Philippines. Future large-scale studies in Southeast Asian populations, combined with functional characterization of rare variants, will be essential for refining *LRRK2*-related risk and advancing equitable precision medicine in PD.

Methodology

Participants recruitment

Ethical clearance for the study was approved by University of Malaya Medical Centre (UMMC) Medical Research Ethics Committee (201910107917) and the Malaysian Ministry of Health (MOH) (NMRR-19-3762-52429). Written informed consent was obtained from all participants prior to enrolment. PD patients were recruited from fourteen participating centers across Malaysia: UMMC and University of Malaya Specialist Centre (UMSC), Sultanah Nur Zahirah Hospital Kuala Terengganu, Kuala Lumpur General Hospital, Queen Elizabeth Hospital Kota Kinabalu, Island Hospital Penang, Penang General Hospital, Seberang Jaya Hospital, Taiping Hospital, Sarawak General Hospital Kuching, Miri General Hospital, Sungai Buloh Hospital, Raja Permaisuri Bainun Hospital, Hospital Tengku Ampuan Rahimah Klang (HTAR) and Tuanku Ja'afar Hospital Seremban.

All PD patients were recruited by neurologists/physicians with an interest in movement disorders, and diagnoses were based on the United Kingdom PD Brain Bank diagnosis criteria (Hughes et al. 1992). PD patients with age at diagnosis (AAD) ≤ 50 years were classified as early-onset PD (EOPD) (Tay et al., 2023). Individuals diagnosed with alternative forms of parkinsonism, including Parkinson-plus syndromes, drug-induced parkinsonism, and vascular parkinsonism, were excluded. Age-matched healthy controls with no family history of PD and no self-reported neurological disease were recruited. Self-reported ancestry across three generations was obtained for both patients and controls for initial ancestry assignment.

Sample processing, data generation and quality control

Genomic DNA was extracted from peripheral blood using FavorPrep Genomic DNA Extraction Maxi Kit, as previously described according to manufacturer protocol (Catalogue number: FABGK003-1, <https://www.favorgen.com/>). Samples underwent either targeted sequencing using the CENTOGENE panel with data processing as previously described⁵⁸, NeuroBooster Array v1.0 (NBA) genotyping¹⁸, and/or whole-genome sequencing (WGS) through the Global Parkinson's Genetics Programme (GP2) from release 11 (DOI: 10.5281/zenodo.17753486). All reads were aligned to the human reference genome assembly GRCh38 (hg38). NBA and WGS data processing and QC are detailed in Vitale et al., 2025⁵⁹.

Ancestry classification

Ancestral background was inferred using a two-stage reference panel approach. First, reference panels from the 1000 Genomes Project (<https://www.internationalgenome.org/data-portal/data-collection/phase-1>)⁶⁰, Human Genome Diversity Project⁶¹, and an Ashkenazi Jewish population dataset were used⁶² (Fig. 1a). To improve ancestry inference and account for population stratification in our cohort, the Singapore Whole Genome Reference pilot data (SG10K)⁵⁰, comprising 2657 Chinese, 927 Malay, and 1125 Indians, was subsequently used as a reference panel (Fig. 1b). Due to the lack of available reference panels for indigenous, self-reported identification was utilised where individuals were from the Bajau, Bidayuh, Bisaya, Bugis, Dusun, Iban, Kadazan, Murut, Rawa and Rungus tribes. For detailed insights into the ancestry estimation method, refer to the GenoTools GitHub repository (<https://github.com/GP2code/GenoTools>). Samples exceeding ± 6 standard deviations from the cluster mean, were removed from the analysis.

Imputation

Variants with a MAF < 0.005 and Hardy Weinberg Equilibrium p-value < 1×10^{-4} were excluded prior to submission to the TOPMed Imputation server. Further details about the TOPMed Imputation Server and Minimac can be accessed at <https://imputation.biodatacatalyst.nhlbi.nih.gov>. Following imputation, variants were filtered using a minor allele count (MAC) threshold of 10 and an imputation quality (Rsq) of 0.8.

Mutational spectrum of LRRK2

LRRK2 variants were annotated using ANNOVAR⁶³. Concordance analysis was performed for individuals with both NBA and WGS data, while variants identified from a single platform (NBA or CENTOGENE panel) were validated by Sanger sequencing. Variant pathogenicity was classified as (pathogenic/likely pathogenic/VUS/benign/likely benign) according to American College of Medical Genetics and Genomics (ACMG) recommendations^{64–66} using the Franklin tool (<https://franklin.genoox.com/clinical-db/home>) and cross-checked against ClinVar. Pathogenic/likely pathogenic candidate variants were validated using Sanger sequencing (primers available upon request). Variants were screened against gnomAD v4.1 and the SG10K Health dataset to retrieve population allele frequencies. Additionally, literature review for each variant (up to January 2026) was performed.

Recent recommendations by the ClinGen Low-Penetrance/Risk Allele Working Group¹⁹ to expand ACMG classifications to recognise other clinically relevant categories were also utilised. Based on the proposed criteria, the p.G2385R variant is classed as an ‘established risk allele’, while p.A419V and p.R1628P are classed as ‘likely risk alleles’. In addition, variants in the reported protective haplotype (p.N551K-R1398H) are classed as ‘established likely protective alleles’.

Identified *LRRK2* patients were further screened for variants in monogenic PD genes, including *GBA1*, *PRKN*, *PINK1*, *PARK7*, *SNCA*, and *ATP13A2* using the Centogene PD Panel, WGS, and/or Multiplex Ligation-dependent Probe Amplification (MLPA), as previously described⁵⁸.

Power Calculation

Power calculations were performed using the GAS power calculator (https://csg.sph.umich.edu/abecasis/cats/gas_power_calculator/). The additive model was used with a significance level set at $p = 0.05$, assuming a PD general population prevalence of 0.5%⁶⁷ across multiple ORs of 1.2, 1.5, 1.7 and 3.0.

Association study

The association between *LRRK2* variants and PD risk was assessed using logistic regression (glm) under an additive genetic model in PLINK 2.0⁶⁸. Covariates including age, sex, and the appropriate number of PCs for each ancestry, were included in the model as shown in Supplementary Table 11. Allele and genotype frequencies, and HWE, were calculated using PLINK 2.0, while pairwise linkage disequilibrium (LD; r^2) was calculated to assess whether observed associations were independent.

Age at Diagnosis (AAD) comparison across LRRK2 patients

Differences in AAD were compared between non-*LRRK2* patients and *LRRK2* patients groups using one-way analysis of variance (ANOVA). Post hoc Dunnett tests were performed to account for multiple testing. Participants that co-harbor any known PD related pathogenic/likely pathogenic variants for other genes were dropped out to prevent bias in the analysis.

Linkage disequilibrium (LD) block and haplotype analysis

Variants with MAF < 0.05 were excluded from this analysis. LD blocks were defined using the `--blocks` function in PLINK1.9, restricting variants to a 200-kb window defined using the ⁶⁹ criteria. Haplotype-based association analysis was performed for each defined block using haplo.stats in R (<https://CRAN.R-project.org/package=haplo.stats>).

Low frequency and rare variant burden testing

For low-frequency and rare variant (MAF < 5%) association testing, WGS data was used. The cumulative effects of rare variants were assessed using the Rare Variant Test (RVtest) software package ⁷⁰, employing the Sequence Kernel Association Test (SKAT), and the Sequence Kernel Association Test-Optimized (SKAT-O) test adjusted for age, sex, and PCs. Coding variants were grouped according to their respective LRRK2 functional domains, whereas non-coding variants were stratified by MAF thresholds of < 5% and < 1%, following established protocols from previous studies ^{11,71}.

Declarations

Data availability

Data used in the preparation of this article were obtained from the Global Parkinson's Genetics Program (GP2; <https://gp2.org>). Specifically, we used Tier 2 data from GP2 release 11 (DOI 10.5281/zenodo.17753486, release 11). All GP2 data are hosted in collaboration with the Accelerating Medicines Partnership in Parkinson's disease and available via application on the website (<https://amp-pd.org>). The 1000 Genome, HGDP and AJ reference panel used for ancestry inference were part of the GenoTools pipeline and can be obtained from <https://github.com/dvitale199/GenoTools>. The SG10K pilot dataset is available from the European Genome-phenome Archive (EGA) under accession number EGAD00001005337 (<https://ega-archive.org/datasets/EGAD00001005337>). The SG10K_Health dataset is available through the Singapore National Precision Medicine (NPM) programme data access framework (<https://www.npm.sg/data-access-to-sg10k-health/>). Access to the data is controlled and governed by the National Precision Medicine Data Access Committee (NPM DAC).

Code Availability

Genotyping imputation, quality control, ancestry prediction, and processing were performed using GenoTools v1.0, publicly available on GitHub (<https://github.com/GP2code/GenoTools>). All code generated for this article, and the identifiers for all software programs and packages used, are available on GitHub https://github.com/GP2code/Malaysian_LRRK2 and were given a persistent identifier via Zenodo [10.5281/zenodo.19859782](https://doi.org/10.5281/zenodo.19859782).

Acknowledgements

We gratefully acknowledge all the participants for their contributions, without whom this research would not have been possible. This work received funding from the Ministry of Higher Education Malaysia Fundamental Research Grant Scheme (FRGS/1/2018/SKK08/UM/02/3) awarded to AAA, and the Universiti Malaya Parkinson's Disease and Movement Disorders Research Program (PV035-2017) awarded to SYL and AHT, and carried out as part of the Global Parkinson's Genetics Program (GP2; <https://gp2.org>). This project was supported by GP2, which is funded by the Aligning Science Across Parkinson's (ASAP) (<https://ror.org/03zj4c476>) initiative and implemented by The Michael J. Fox Foundation for Parkinson's Research (MJFF) (<https://ror.org/03arq3225>). For a complete list of GP2 members see <https://doi.org/10.5281/zenodo.7904831>. We thank the "SG10K_Pilot Investigators" for providing the SG10K_Pilot data (EGAD00001005337). The data from the "SG10K_Pilot Study" reported here were obtained from European Genome-phenome Archive (EGA). This manuscript was not prepared in collaboration with the "SG10K_Pilot Study" and does not necessarily reflect the opinions or views of the "SG10K_Pilot Study. This study made use of data generated as part of the Singapore National Precision Medicine program funded by the Industry Alignment Fund (Pre-Positioning) (IAF-PP:H17/01/a0/007). This study made use of data / samples collected in the following cohorts in Singapore: (1) The Health for Life in Singapore (HELIOS) study at the Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore (supported by grants from a Strategic Initiative at Lee Kong Chian School of Medicine, the Singapore Ministry of Health under its Singapore Translational Research Investigator Award (NMRC/StaR/0028/2017) and the IAF-PP: H18/01/a0/016; (2) The Growing up in Singapore Towards Healthy Outcomes (GUSTO) study, which is jointly hosted by the National University Hospital (NUH), KK Women's and Children's Hospital (KKH), the National University of Singapore (NUS) and the Singapore Institute for Clinical Sciences (SICS), Agency for Science Technology and Research (A*STAR) (supported by the Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Programme and administered by the Singapore Ministry of Health's National Medical Research Council (NMRC), Singapore - NMRC/TCR/004-NUS/2008; (3) The Singapore Epidemiology of Eye Diseases (SEED) cohort at Singapore Eye Research Institute (SERI) (supported by NMRC/CIRG/1417/2015; NMRC/CIRG/1488/2018; NMRC/OFLCG/004/2018); (4) The Multi-Ethnic Cohort (MEC) cohort (supported by NMRC grant 0838/2004; BMRC grant 03/1/27/18/216; 05/1/21/19/425; 11/1/21/19/678, Ministry of Health, Singapore, National University of Singapore and National University HealthSystem, Singapore); (5) The SingHealth Duke-NUS Institute of Precision Medicine (PRISM) cohort (supported by NMRC/CG/M006/2017_NHCS; NMRC/StaR/0011/2012, NMRC/StaR/0026/2015, Lee Foundation and Tanoto Foundation); (6) The TTSH Personalised Medicine Normal Controls (TTSH) cohort (supported by NMRC/CG12AUG17 and CGAug16M012).

The views expressed are those of the author(s) are not necessarily those of the National Precision Medicine investigators, or institutional partners. We thank all investigators, staff members and study participants who made the National Precision Medicine Project possible. We thank the Director-General of the Ministry of Health Malaysia for his permission to publish this manuscript.

Author contributions

AAA, AHT, and KSL conceptualized the manuscript. AAA, KSL, SBC, MTP, JLL, and YWT conducted the formal analysis. KSL, AAA, and MTP wrote the first draft of the manuscript. LS provided logistical support of project management. SYL, KAI, and JCEO provided the details of the patients' clinical phenotypes. AHT, SYL, TST, LLC, ANKA, HXD, JPS, YYB, KAI, ASM, YKC, JCEO, TTL, IL, WKC, WCL, SCW, TLO, MLC, PCT, YHL, CLLL were involved in patient recruitment, sample processing, and contributed data. All authors reviewed, edited, and approved the final version of the manuscript for submission.

Competing Interest

The authors declare no competing financial or non-financial interests.

References

1. Krüger, C. *et al.* Updated MDSGene review on the clinical and genetic spectrum of LRRK2 variants in Parkinson's disease. *NPJ Parkinsons Dis* 11, 30 (2025). <https://doi.org/10.1038/s41531-025-00881-9>
2. Goh, J. W. *et al.* LRRK2 p.G2385R and p.R1628P variants in a multi-ethnic Asian Parkinson's Cohort: epidemiology and clinical insights. *npj Parkinson's Disease* 11, 320 (2025). <https://doi.org/10.1038/s41531-025-01166-x>
3. Kaiyrzhanov, R. *et al.* LRRK2 Mutations and Asian Disease-Associated Variants in the First Parkinson's Disease Cohort from Kazakhstan. *Parkinsons Dis* 2020, 2763838 (2020). <https://doi.org/10.1155/2020/2763838>
4. Lim, S.-Y. *et al.* Uncovering the genetic basis of Parkinson's disease globally: from discoveries to the clinic. *The Lancet Neurology* 23, 1267–1280 (2024). [https://doi.org/10.1016/S1474-4422\(24\)00378-8](https://doi.org/10.1016/S1474-4422(24)00378-8)
5. Pan, H. *et al.* Genome-wide association study using whole-genome sequencing identifies risk loci for Parkinson's disease in Chinese population. *npj Parkinson's Disease* 9, 22 (2023). <https://doi.org/10.1038/s41531-023-00456-6>
6. Foo, J. N. *et al.* Identification of Risk Loci for Parkinson Disease in Asians and Comparison of Risk Between Asians and Europeans: A Genome-Wide Association Study. *JAMA Neurol* 77, 746–754 (2020). <https://doi.org/10.1001/jamaneurol.2020.0428>
7. 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). <https://doi.org/10.3389/fnagi.2019.00013>
8. Wu, X. *et al.* Quantitative assessment of the effect of LRRK2 exonic variants on the risk of Parkinson's disease: a meta-analysis. *Parkinsonism Relat Disord* 18, 722–730 (2012). <https://doi.org/10.1016/j.parkreldis.2012.04.013>
9. Kishore, A. *et al.* Deciphering the Genetic Architecture of Parkinson's Disease in India. *medRxiv*, 2025.2002.2017.25322132 (2025). <https://doi.org/10.1101/2025.02.17.25322132>
10. Alessi, D. R. & Pfeffer, S. R. Leucine-Rich Repeat Kinases. *Annu Rev Biochem* 93, 261–287 (2024). <https://doi.org/10.1146/annurev-biochem-030122-051144>
11. Foo, J. N. *et al.* Analysis of non-synonymous-coding variants of Parkinson's disease-related pathogenic and susceptibility genes in East Asian populations. *Hum Mol Genet* 23, 3891–3897 (2014). <https://doi.org/10.1093/hmg/ddu086>
12. Wu, Y.-R. *et al.* Genetic Variants of LRRK2 in Taiwanese Parkinson's Disease. *PLOS ONE* 8, e82001 (2013). <https://doi.org/10.1371/journal.pone.0082001>
13. Tan, E.-K. *et al.* Multiple LRRK2 variants modulate risk of Parkinson disease: a chinese multicenter study. *Human Mutation* 31, 561–568 (2010). <https://doi.org/https://doi.org/10.1002/humu.21225>
14. Gopalai, A. A. *et al.* LRRK2 N551K and R1398H variants are protective in Malays and Chinese in Malaysia: A case-control association study for Parkinson's disease. *Mol Genet Genomic Med* 7, e604 (2019). <https://doi.org/10.1002/mgg3.604>
15. Ross, O. A. *et al.* Association of < em>LRRK2 exonic variants with susceptibility to Parkinson's disease: a case–control study. *The Lancet Neurology* 10, 898–908 (2011). [https://doi.org/10.1016/S1474-4422\(11\)70175-2](https://doi.org/10.1016/S1474-4422(11)70175-2)
16. Paisán-Ruiz, C., Washecka, N., Nath, P., Singleton, A. B. & Corder, E. H. Parkinson's Disease and Low Frequency Alleles Found Together Throughout LRRK2. *Annals of Human Genetics* 73, 391–403 (2009). <https://doi.org/https://doi.org/10.1111/j.1469-1809.2009.00524.x>
17. Gopalai, A. A. *et al.* LRRK2 G2385R and R1628P mutations are associated with an increased risk of Parkinson's disease in the Malaysian population. *Biomed Res Int* 2014, 867321 (2014). <https://doi.org/10.1155/2014/867321>
18. Bandres-Ciga, S. *et al.* NeuroBooster Array: A Genome-Wide Genotyping Platform to Study Neurological Disorders Across Diverse Populations. *Mov Disord* 39, 2039–2048 (2024). <https://doi.org/10.1002/mds.29902>

19. Schmidt, R. J. et al. Recommendations for risk allele evidence curation, classification, and reporting from the ClinGen Low Penetrance/Risk Allele Working Group. *Genet Med* 26, 101036 (2024). <https://doi.org/10.1016/j.gim.2023.101036>
20. Kalogeropoulou, A. F. et al. Impact of 100 LRRK2 variants linked to Parkinson's disease on kinase activity and microtubule binding. *Biochem J* 479, 1759–1783 (2022). <https://doi.org/10.1042/bcj20220161>
21. Lim, K. S. et al. The LRRK2 p.A419V variant associates with risk of Parkinson's disease in the East Asian population and an evaluation on age of onset. *medRxiv* (2025). <https://doi.org/10.1101/2025.08.28.25333987>
22. Tan, E. K. et al. Multiple LRRK2 variants modulate risk of Parkinson disease: a Chinese multicenter study. *Hum Mutat* 31, 561–568 (2010). <https://doi.org/10.1002/humu.21225>
23. Zhang, P. et al. Crystal structure of the WD40 domain dimer of LRRK2. *Proceedings of the National Academy of Sciences* 116, 201817889 (2019). <https://doi.org/10.1073/pnas.1817889116>
24. Ding, L. et al. The LRRK2 P1446L mutation triggers dopaminergic neurodegeneration via DAPK1-mediated microglial neuroinflammation and neuronal apoptosis. *npj Parkinson's Disease* (2025). <https://doi.org/10.1038/s41531-025-01234-2>
25. Nichols, R. J. et al. 14-3-3 binding to LRRK2 is disrupted by multiple Parkinson's disease-associated mutations and regulates cytoplasmic localization. *Biochem J* 430, 393–404 (2010). <https://doi.org/10.1042/bj20100483>
26. Muda, K. et al. Parkinson-related LRRK2 mutation R1441C/G/H impairs PKA phosphorylation of LRRK2 and disrupts its interaction with 14-3-3. *Proc Natl Acad Sci U S A* 111, E34-43 (2014). <https://doi.org/10.1073/pnas.1312701111>
27. Sanchez-Contreras, M. et al. Study of LRRK2 variation in tauopathy: Progressive supranuclear palsy and corticobasal degeneration. *Mov Disord* 32, 115–123 (2017). <https://doi.org/10.1002/mds.26815>
28. Lim, S.-Y. et al. Clinical and functional evidence for the pathogenicity of the LRRK2 p.Arg1067Gln variant. *npj Parkinson's Disease* 11, 34 (2025). <https://doi.org/10.1038/s41531-025-00884-6>
29. Lim, S. Y. et al. Clinical Phenotype of LRRK2 R1441C in 2 Chinese Sisters. *Neurodegener Dis* 20, 39–45 (2020). <https://doi.org/10.1159/000508131>
30. Kishore, A. et al. Understanding the role of genetic variability in LRRK2 in Indian population. *Mov Disord* 34, 496–505 (2019). <https://doi.org/10.1002/mds.27558>
31. Punia, S. et al. Absence/rarity of commonly reported LRRK2 mutations in Indian Parkinson's disease patients. *Neurosci Lett* 409, 83–88 (2006). <https://doi.org/10.1016/j.neulet.2006.04.052>
32. Vijayan, B., Gopala, S. & Kishore, A. LRRK2 G2019S mutation does not contribute to Parkinson's disease in South India. *Neurology India* 59, 157–160 (2011). <https://doi.org/10.4103/0028-3886.79125>
33. Vishwanathan Padmaja, M., Jayaraman, M., Srinivasan, A. V., Srikumari Srisailapathy, C. R. & Ramesh, A. The SNCA (A53T, A30P, E46K) and LRRK2 (G2019S) mutations are rare cause of Parkinson's disease in South Indian patients. *Parkinsonism & Related Disorders* 18, 801–802 (2012). <https://doi.org/10.1016/j.parkreldis.2012.02.012>
34. Rath, S. et al. Clinico-genetic analysis of patients with LRRK2-related Parkinson's disease from a referral centre in India. *Parkinsonism & Related Disorders* 132 (2025). <https://doi.org/10.1016/j.parkreldis.2025.107299>
35. Nuytemans, K. et al. Founder mutation p.R1441C in the leucine-rich repeat kinase 2 gene in Belgian Parkinson's disease patients. *Eur J Hum Genet* 16, 471–479 (2008). <https://doi.org/10.1038/sj.ejhg.5201986>
36. Lesage, S. et al. Molecular analyses of the LRRK2 gene in European and North African autosomal dominant Parkinson's disease. *J Med Genet* 46, 458–464 (2009). <https://doi.org/10.1136/jmg.2008.062612>
37. Zhang, J.-R. et al. Genetic analysis of LRRK2 in Parkinson's disease in Han Chinese population. *Neurobiology of Aging* 72, 187.e185-187.e110 (2018). <https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2018.06.036>
38. Landoulsi, Z. et al. Genetic landscape of Parkinson's disease and related diseases in Luxembourg. *Frontiers in Aging Neuroscience* Volume 15–2023 (2023). <https://doi.org/10.3389/fnagi.2023.1282174>
39. Pan, H. et al. Genome-wide association study using whole-genome sequencing identifies risk loci for Parkinson's disease in Chinese population. *NPJ Parkinsons Dis* 9, 22 (2023). <https://doi.org/10.1038/s41531-023-00456-6>
40. Lim, K. S. et al. Association of LRRK2 p.A419V with Parkinson's Disease in East Asians and analysis of age at onset. *npj Parkinson's Disease* 12, 51 (2026). <https://doi.org/10.1038/s41531-026-01265-3>
41. Pratuskeviciute, N. et al. Functional and Structural Characterization of LRRK2 p.V1447L in Parkinson's Disease. *Movement Disorders* 40, 2251–2256 (2025). <https://doi.org/https://doi.org/10.1002/mds.30284>
42. Youn, J. et al. Genetic variants of PARK genes in Korean patients with early-onset Parkinson's disease. *Neurobiology of Aging* 75, 224.e229-224.e215 (2019). <https://doi.org/https://doi.org/10.1016/j.neurobiolaging.2018.10.030>
43. Li, Y. et al. Clinical characterization of patients with leucine-rich repeat kinase 2 genetic variants in Japan. *J Hum Genet* 65, 771–781 (2020). <https://doi.org/10.1038/s10038-020-0772-4>

44. McGrath, E., Waschbüsch, D., Baker, B. M. & Khan, A. R. LRRK2 binds to the Rab32 subfamily in a GTP-dependent manner via its armadillo domain. *Small GTPases* 12, 133–146 (2021). <https://doi.org/10.1080/21541248.2019.1666623>
45. Vides, E. G. *et al.* A feed-forward pathway drives LRRK2 kinase membrane recruitment and activation. *Elife* 11 (2022). <https://doi.org/10.7554/eLife.79771>
46. Deniston, C. K. *et al.* Structure of LRRK2 in Parkinson's disease and model for microtubule interaction. *Nature* 588, 344–349 (2020). <https://doi.org/10.1038/s41586-020-2673-2>
47. Myasnikov, A. *et al.* Structural analysis of the full-length human LRRK2. *Cell* 184, 3519–3527.e3510 (2021). <https://doi.org/10.1016/j.cell.2021.05.004>
48. Toh, T. S. *et al.* Ex Vivo LRRK2 Activation in Asian G2385R and R1628P Variant Carriers and Idiopathic Parkinson's Disease. *Movement Disorders* n/a (2026). <https://doi.org/https://doi.org/10.1002/mds.70314>
49. Parlar, S. C. *et al.* LRRK2 rare-variant per-domain genetic burden in Parkinson's Disease: association confined to the kinase domain. *npj Parkinson's Disease* 11, 102 (2025). <https://doi.org/10.1038/s41531-025-00934-z>
50. Wu, D. *et al.* Large-Scale Whole-Genome Sequencing of Three Diverse Asian Populations in Singapore. *Cell* 179, 736–749.e715 (2019). <https://doi.org/10.1016/j.cell.2019.09.019>
51. Wall, J. D. *et al.* The GenomeAsia 100K Project enables genetic discoveries across Asia. *Nature* 576, 106–111 (2019). <https://doi.org/10.1038/s41586-019-1793-z>
52. Yang, M.-Y. *et al.* SEAD reference panel with 22,134 haplotypes boosts rare variant imputation and genome-wide association analysis in Asian populations. *Nature Communications* 15, 10839 (2024). <https://doi.org/10.1038/s41467-024-55147-4>
53. Chen, L. *et al.* LRRK2 R1398H polymorphism is associated with decreased risk of Parkinson's disease in a Han Chinese population. *Parkinsonism Relat Disord* 17, 291–292 (2011). <https://doi.org/10.1016/j.parkreldis.2010.11.012>
54. Heckman, M. G. *et al.* Protective effect of LRRK2 p.R1398H on risk of Parkinson's disease is independent of MAPT and SNCA variants. *Neurobiol Aging* 35, 266.e265-214 (2014). <https://doi.org/10.1016/j.neurobiolaging.2013.07.013>
55. Tan, A. H. *et al.* PINK1 p.Leu347Pro mutations in Malays: Prevalence and illustrative cases. *Parkinsonism & Related Disorders* 79, 34–39 (2020). <https://doi.org/https://doi.org/10.1016/j.parkreldis.2020.08.015>
56. Tay, Y. W. *et al.* Very High Frequency of Early-Onset Parkinson's Disease and PRKN Mutations among Indigenous Patients in Sabah, Malaysia. *Mov Disord* 40, 2407–2418 (2025). <https://doi.org/10.1002/mds.70015>
57. Lim, S. Y. *et al.* Uncovering the genetic basis of Parkinson's disease globally: from discoveries to the clinic. *Lancet Neurol* 23, 1267–1280 (2024). [https://doi.org/10.1016/s1474-4422\(24\)00378-8](https://doi.org/10.1016/s1474-4422(24)00378-8)
58. Tay, Y. W. *et al.* Genetic study of early-onset Parkinson's disease in the Malaysian population. *Parkinsonism Relat Disord* 111, 105399 (2023). <https://doi.org/10.1016/j.parkreldis.2023.105399>
59. Vitale, D. *et al.* GenoTools: an open-source Python package for efficient genotype data quality control and analysis. *G3 (Bethesda)* 15 (2025). <https://doi.org/10.1093/g3journal/jkae268>
60. Auton, A. *et al.* A global reference for human genetic variation. *Nature* 526, 68–74 (2015). <https://doi.org/10.1038/nature15393>
61. Siva, N. 1000 Genomes project. *Nature Biotechnology* 26, 256–256 (2008). <https://doi.org/10.1038/nbt0308-256b>
62. Bray, S. M. *et al.* Signatures of founder effects, admixture, and selection in the Ashkenazi Jewish population. *Proc Natl Acad Sci U S A* 107, 16222–16227 (2010). <https://doi.org/10.1073/pnas.1004381107>
63. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38, e164 (2010). <https://doi.org/10.1093/nar/gkq603>
64. Abou Tayoun, A. N. *et al.* Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat* 39, 1517–1524 (2018). <https://doi.org/10.1002/humu.23626>
65. Biesecker, L. G. & Harrison, S. M. The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genet Med* 20, 1687–1688 (2018). <https://doi.org/10.1038/gim.2018.42>
66. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405–424 (2015). <https://doi.org/10.1038/gim.2015.30>
67. Nalls, M. A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* 18, 1091–1102 (2019). [https://doi.org/10.1016/s1474-4422\(19\)30320-5](https://doi.org/10.1016/s1474-4422(19)30320-5)
68. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4, s13742-13015-10047-13748 (2015). <https://doi.org/10.1186/s13742-015-0047-8>
69. Gabriel, S. B. *et al.* The Structure of Haplotype Blocks in the Human Genome. *Science* 296, 2225–2229 (2002). <https://doi.org/10.1126/science.1069424>

70. Zhan, X., Hu, Y., Li, B., Abecasis, G. R. & Liu, D. J. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 32, 1423–1426 (2016). <https://doi.org/10.1093/bioinformatics/btw079>
71. Lee, B. D., Dawson, V. L. & Dawson, T. M. Leucine-rich repeat kinase 2 (LRRK2) as a potential therapeutic target in Parkinson's disease. *Trends Pharmacol Sci* 33, 365–373 (2012). <https://doi.org/10.1016/j.tips.2012.04.001>

Figures

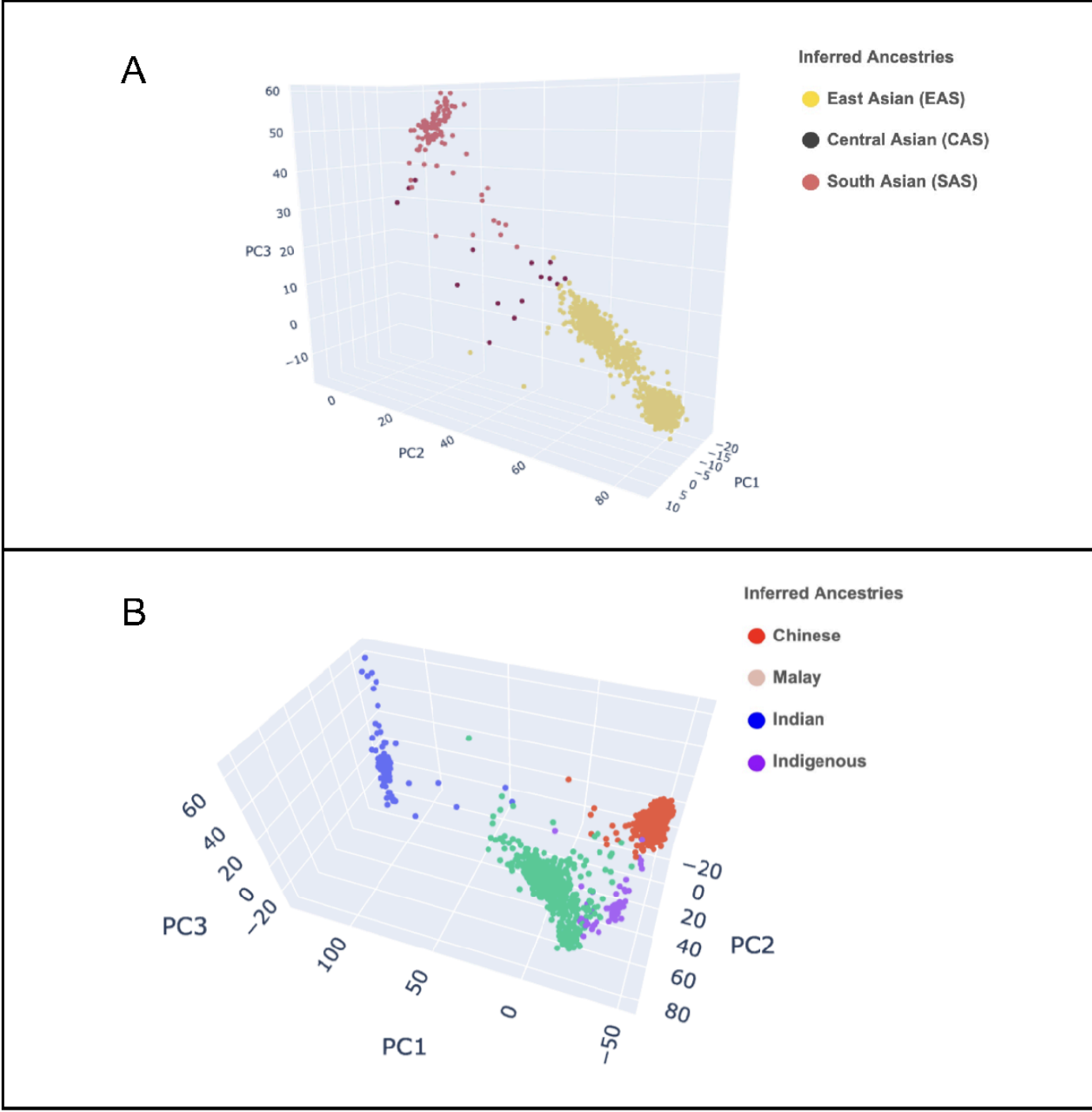


Figure 1

PCA plot of the Malaysian cohort

(A): Ancestry prediction using the 1K Genome, Human Genome Diversity Project (HGDP) and Ashkenazi Jewish population reference panels.

Overall variant spectrum of *LRRK2* in the Malaysian cohort

(A). Schematic diagram of the *LRRK2* protein and the position of the variants found in Malaysian cases and controls

Schematic representation of the *LRRK2* protein showing the positions of variants identified in Malaysian PD cases only or in cases and controls across the major protein domains (ARM, ANK, LRR, ROC, COR, MAPKKK, and WD40). Variants previously reported as pathogenic are highlighted in bold red, while known risk alleles are shown in blue and those in yellow highlights are the novel variants.

(B). The distribution of *LRRK2* genetic variants across the Malaysian Chinese, Malay, Indian and Indigenous populations

Distribution of *LRRK2* variants across the four major ethnic groups in Malaysia (Chinese, Malay, Indian, and Indigenous). The Venn diagram illustrates shared and ancestry-specific variants detected among PD cases and controls within each population group. Sample sizes for each ancestry group are indicated. Variants in bold are pathogenic or risk alleles. Variants with the asterix (*) are the novel variants. n = Number of samples

AAD comparison between non-LRRK2 carrier and other carrier groups

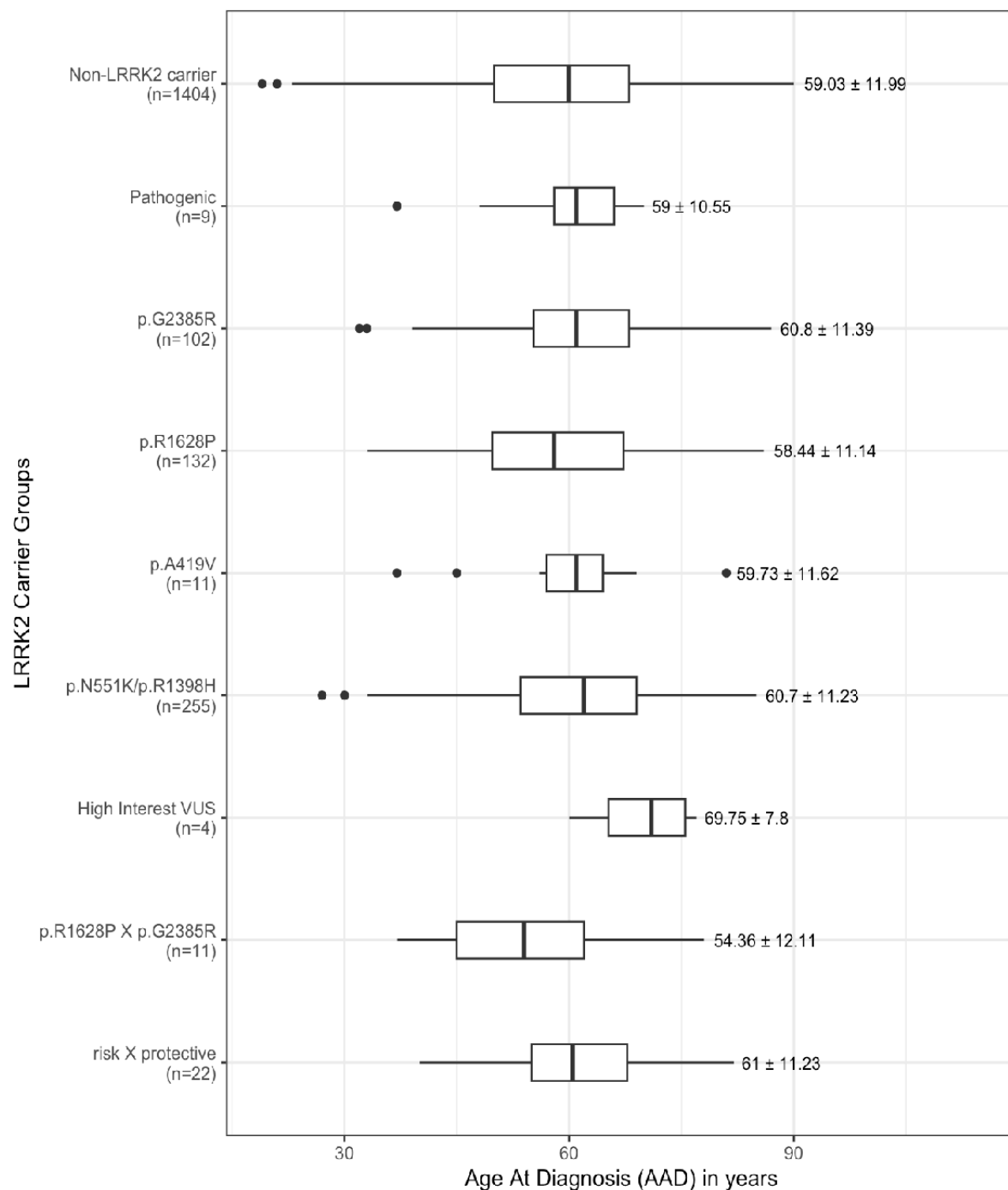


Figure 3

Mean AAD comparison between the non-*LRRK2* patients and subgroups of *LRRK2* pathogenic/likely pathogenic (p.R1067Q/p.R1325Q/p.R1441C/H) patients, single p.A419V, p.R1628P, p.G2385R heterozygous patients, double heterozygous p.R1628P-p.G2385R patients, protective p.N551K-p.R1398H-K1423K patients and high-interest VUS p.A1413T, p.P1446L patients.

The 'Non-*LRRK2* group' includes patients who do not carry any of these variants. The double heterozygous p.A419V-p.R1628P patient was the only sample and is not plotted.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementarytable.xlsx](#)