Analysis of *DNM3* and *VAMP4* as genetic modifiers of *LRRK2* Parkinson's disease

Brown EE[1] MRes, Blauwendraat C[3] PhD, Trinh J[2] PhD, Rizig M[1] MD PhD, Nalls AM[2,5] PhD, Leveille E[6,7], Ruskey JA[6,7] MSc, Jonvik H[1] MSc, Tan MMX[1] BSc, Bandres-Ciga S[2,24] PhD, Hassin-Baer S[8,9,10] MD, Brockmann K[11] MD, Hardy J[1] PhD, Singleton AB[2] PhD, Alcalay RN[12,13] MD MSc, Gasser T[14,15] MD PhD, Grosset D[16,17] MD, Williams NM[18] PhD, Pittman A[19] PhD, Gan-Or Z[6,7,20] MD PhD, Fernandez-Santiago R[21] PhD, Brice A[22] MD MSc, Lesage S[22] PhD, Farrer M[23] PhD, Wood N[1] MD PhD, Morris HR[1] MD PhD* on behalf of the International Parkinson Disease Genomics Consortium (IPDGC)

[1]Department of Clinical and Movement Neurosciences, Institute of Neurology, University College London, WC1N 3BG

[2]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA

[3]Institute of Neurogenetics, University of Luebeck, Lübeck, Germany

[5]Data Tecnica International, Glen Echo, MD, USA

[6]Montreal Neurological Institute, McGill University, Montréal, QC, Canada

[7]Department of Neurology and Neurosurgery, McGill University, Montréal, QC, Canada

[8]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[9]Department of Neurology, Sheba Medical Center, Tel Hashomer, Israel

[10]Movement Disorders Institute, Sheba Medical Center, Tel Hashomerf, Israel

[11]Hertie Institute for Clinical Brain Research and German Center for Neurodegenerative Diseases, University Clinic Tuebingen, Tuebingen, Germany

[12]Department of Neurology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

[13]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York, NY, USA

[14]Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

[15]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

[16]Department of Neurology, Institute of Neurological Sciences, Queen Elizabeth University Hospital, Glasgow, United Kingdom

[17]Institute of Neuroscience & Psychology, University of Glasgow, Glasgow, United Kingdom

[18]Division of Psychological Medicine & Clinical Neuroscience, School of Medicine, Cardiff University, Cardiff, UK

[19]Department of Clinical Genetics, St George's University of London, London SW17 0RE

[20]Department of Human Genetics, McGill University, Montreal, Quebec, Canada

[21]Laboratory of Neurodegenerative Disorders, Department of Neurology, Hospital Clínic of Barcelona Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) University of Barcelona (UB), Barcelona, Spain

[22] Research Unit U1127 at INSERM, Research Unit UMR 7225 at the French National Centre for Scientific Research (CNRS)

[23]Research Unit UMR_1127 at Sorbonne Université, Institutet du Cerveau et de la Moëlle épinière (ICM), Paris, France

[23]Department of Medical Genetics, University of British Columbia, Vancouver, Canada

[24]Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, 18016, Spain

*Corresponding author: Professor Huw Morris, Department of Clinical and Movement Neurosciences, Royal Free Hospital, Rowland Hill Street, London NW3 2PF. Email: h.morris@ucl.ac.uk. Tel: +44 020 7830 2951.

Abbreviations. PD = Parkinson's disease, *LRRK2* = Leucine-rich repeat kinase 2, AAO = age at onset, *DNM3* = Dynamin 3, KASP = 'Kompetitive' allele-specific polymerase chain reaction, MAF = minor allele frequency, GWAS = Genome-wide association study, SNP = single nucleotide polymorphism

Abstract

Objective: To assess the reported *LRRK2* age-at-onset (AAO) modifier *DNM3* rs2421947, and the proximate genome wide association signal *VAMP4* rs11578699 on AAO of Parkinson's disease (PD), with rare (rs34637584; p.G2019S) or common risk factors (rs10878226) in the *LRRK2* locus.

Methods: We assessed the extent of linkage disequilibrium between *DNM3* rs2421947 and *VAMP4* rs11578699, and between *LRRK2* rs34637584 and *LRRK2* rs10878226. We analysed *DNM3* rs2421947 in 724 *LRRK2* p.G2019S heterozygotes using the Cox proportional hazards model and linear regression of AAO. We then meta-analysed this with previously published data (n=754). We analysed the rs11578699 variant in the nearby *VAMP4* gene in 786 *LRRK2* p.G2019S heterozygotes. We evaluated the impact of *VAMP4* variants using AAO regression in 4882 patients with PD carrying a common *LRRK2* risk variant (rs10878226).

Results: There was no evidence for linkage disequilibrium between *DNM3* rs2421947 and *VAMP4* rs11578699, or between *LRRK2* rs34637584 and *LRRK2* rs10878226. Our survival analysis of 724 p.G2019S carriers showed no relationship between *DNM3* rs2421947 and AAO (hazard ratio [HR] 1.09, 95% CI 0.95-1.25, p=0.20) or PD affected status. However, meta-analysis with previously published *LRRK2* p.G2019S data indicates an effect on AAO (HR 1.14, 95% CI 1.02-1.27, p=0.025), and furthermore reveals a significant relationship between *DNM3* rs2421947 genotype and parkinsonism in people with *LRRK2* p.G2019S ($p=3.1x10^{-5}$). The *VAMP4* rs11578699 variant was not associated with AAO (HR 1.02, 95% CI 0.88-1.17, p=0.84) overall. However, *VAMP4*

rs11578699 was associated with AAO in patients dichotomized by *LRRK2* rs10878226 genotype status (TT versus CC and TC (beta=1.68, se=0.81 p=0.037)).

Interpretation: Our analysis suggests that *DNM3* and VAMP4 play a modest and independent role in determining AAO in *LRRK2* PD; there is some evidence for ethnic-specific effects, and for effects on case-control status in *LRRK2* p.G2019S heterozygotes. Analysis of sporadic PD patients stratified by the *LRRK2* PD risk allele rs10878226 indicates a potential interaction between *LRRK2* and *VAMP4.*

Introduction The p.G2019S coding variant in the *LRRK2* gene is the most common high penetrance mutation causing parkinsonism. The mutation occurs in 1-40% of PD cases, varying by ethnicity (1). *LRRK2* parkinsonism is broadly similar to “idiopathic” disease in clinical manifest-ations and age at onset (AAO), generating interest in its potential as a therapeutic target with broader application to PD (2). In addition to rare pathogenic mutations, genome wide association studies (GWAS) show that common variation in *LRRK2* is a risk factor for sporadic PD (3) (23).

There is a wide range in PD AAO among p.G2019S carriers. It is estimated that the heritability of PD AAO (~0.11) is lower than the overall heritability of risk for PD (~0.27) (4). Currently the strongest genetic association with AAO is through genetic risk score (GRS). There is overlap between the heritability of risk, as quantified in the GRS, and AAO. As *LRRK2* is a functionally complex protein with many protein-protein interaction domains, there may be diverse genetic modifiers that vary in effect size and frequency. Trinh, Gustavsson and colleagues (5) presented evidence of modification by a common variant haplotype tag in *DNM3* (rs2421947; minor allele frequency (C) = 0.4), for which the median onset of *LRRK2* parkinsonism for GG homozygotes was 12·5 years younger than that of CC homozygotes (direction of effect GG < GC < CC).

Fernandez-Santiago, Garrido (6) observed a consistent trend for the rs2421947 G allele on AAO in *LRRK2* p.G2019S carriers, with median onset 3 years younger in patients with the G allele, though the result was not significant. In the largest GWAS meta-analysis of idiopathic PD the nearby *VAMP4* gene variant rs11578699, located 113,325

bp from *DNM3* rs2421947, emerged as a statistically significant PD risk association, raising the possibility the effect DNM3 may relate to linkage disequilibrium with VAMP4 (7). *LRRK2* has a role in both monogenic and idiopathic forms of PD and overlap in genetic modifiers may exist.

We investigated the impact of *DNM3* rs2421947 and *VAMP4* rs11578699 on AAO in a large multi-ethnic cohort of European, North African, and Ashkenazi Jewish *LRRK2* p.G2019S carriers and a larger cohort of European patients with and without the *LRRK2* common risk allele.

Methods

Data collection

Patient cohorts: *LRRK2* p.G2019S heterozygotes were identified from cohorts in the International Parkinson's Disease Genomics Consortium (IPDGC) and other collaborative centres. Data from 724 *LRRK2* p.G2019S heterozygotes was contributed by the National Institute of Health (NIH), University College London (UCL), McGill University (MU) (samples were collected in Columbia University (CU) and Sheba Medical Centre (SMC)), Sorbonne University (SU) (in collaboration with Habib Bourguiba hospital, Sfax, Tunisia, and Blida Hospital, Blida, Algeria), National Institute of Neurological Disorders and Stroke (NINDS), and the University of Tübingen (TU). Data from two published studies analysing *DNM3* rs2421947 and *LRRK2* p.G2019S parkinsonism with AAO was also included comprising 754 participants (5, 6). All studies willing to participate and currently holding minimum required data for *LRRK2* p.G2019S

carriers were included: PD onset age or sampling age for asymptomatic carriers; and *DNM3* rs2421947 or a high RSQ proxy SNP. Where these studies had also genotyped *VAMP4* rs11578699 or a high RSQ proxy, these data were also included in the analysis of the *VAMP4* gene (n=786). Patients with PD without a known Mendelian or high risk genetic cause of disease were collated from the UK-wide *Tracking Parkinson's* study (8) and from IPDGC datasets. Summary statistics for *DNM3* rs2421947 and *VAMP4* rs11578699 were obtained from the most recent large-scale PD GWAS (7) and PD AAO GWAS (4). All studies were approved by each respective institutional ethics review committee and all participants provided written informed consent. All studies were carried out in accordance with the Declaration of Helsinki (9).

Genetic analyses: Genotyping was performed on different platforms (Neurochip array, NeuroX array, TaqMan assays, Infinium OmniExpress-24, and Illumina HumanCore Exome array) at participating centres. The *LRRK2* p.G2019S mutation was directly genotyped. Sanger sequence or KASPar verification of the *DNM3* rs2421947 variant was carried out on data subsets at study specific centres. Where the variant *DNM3* rs2421947 was imputed, only RSQ values above 0.85 were used, and only RSQ values above 0.75 were used for *VAMP4* rs11578699 (table-e1), as the RSQ quality of available data for variants *VAMP4* rs11578699 was lower. 1478 *LRRK2* p.G2019S carriers were identified across study cohorts. 2052 samples from patients with idiopathic PD from the Tracking Parkinson's disease study passed quality control and had no identified genetic cause of disease (8). 96 patients with PD were excluded due to missing AAO data; therefore 1956 patients with PD and without *LRRK2* p.G2019S were included in further analysis. Summary statistics and patient data from the most

recent PD GWAS included over 37.7K cases, 18.6K 'proxy-cases' and 1.4M controls. The most recent PD AAO GWAS included data from 28.6K patients (4,7).

Statistical analyses: Data from the studies were pooled. We calculated Hardy-Weinberg Equilibrium (HWE) for *VAMP4* rs11578699 and *DNM3* rs2421947 in p.G2019S carriers. QC HWE for iPD and controls is summarised in the recent large-scale PD GWAS. We assessed linkage disequilibrium (LD) measures RSQ and D' between *DNM3* rs2421947 and *VAMP4* rs11578699 in different populations to evaluate the possibility that an extended haplotype block might explain the relationship between *DNM3* and *LRRK2* (table e-2). The allele based Fisher's exact test was used to compare *DNM3* rs2421947 and *VAMP4* rs11578699 minor allele frequencies (MAF) between *LRRK2* p.G2019S heterozygotes of different ethnic backgrounds and with idiopathic PD (table 1). Student's t-tests and ANOVA were used to assess differences in AAO by genotype (table 2). We then analysed the effect of *DNM3* rs2421947 (figure 1) on AAO in idiopathic PD in the *Tracking Parkinson's* dataset using the Cox proportional hazards model and Kaplan Meier survival analysis. Next, we analysed patients and asymptomatic individuals with *LRRK2* p.G2019S of Ashkenazi Jewish, North African and European ethnicity using survival methods (figure 1 and 2). We then analysed AAO for patients with *LRRK2* p.G2019S by linear regression with available covariates of sex, ethnicity, relatedness, and study centre of origin (figure 2). AAO was regressed by *DNM3* rs2421947 genotypes, and separately for rs11578699 *VAMP4* genotypes. We meta-analysed *DNM3* rs2421947 data from all previously unpublished datasets using Cox proportional hazards (figure 2A) and linear regression models (figure 2B). We pooled these data with previously published data and meta-analysed

again with the same methods (figure 2C and D). We then analysed *VAMP4* rs11578699 impact on AAO with Cox proportional hazards and linear regression (figure 2E and F). Finally we analysed the impact of *VAMP4* rs11578699 on AAO in idiopathic PD carrying the *LRRK2* rs10878226 variant. *LRRK2* rs10878226 has also been implicated in PD risk (odds ratio [OR], 1.20; 95% confidence interval [CI], 1.08–1.33; $p = 6.3\times^{10-4}$, n=6129) (3).

Gene expression: Gene expression profiles of *DNM3* rs2421947 and *VAMP4* rs11578699 were assessed using publicly available Web-based resources: BRAINEAC (www.braineac.org) (11); GTEx (www.gtexportal.org); and Allen Brain Atlas (www.brain-map.org) (12), and evaluated based on the significance level required for inclusion in colocalization (COLOC) analysis (appendix e-1).

Results

DNM3 rs2421947 and *AMP4* rs11578699 and *VAMP4* rs11578699 genotypes were in Hardy-Weinberg equilibrium in controls.

There was no LD between *DNM3* rs2421947 and *VAMP4* rs11578699, as assessed through RSQ and D', within or across population samples, indicating that these are independent variants (table e-2)

We investigated the possibility that ethnic variation in allele frequency might explain a variable effect of *DNM3* on affected *LRRK2* penetrance. Aside from the Norwegian cohort *DNM3* rs2421947 MAF (C) of 0.69 (vs. other Europeans p=0.00012), all *DNM3* rs2421947 MAF in the study were between 0.43-0.51 (table 1). There was no difference between Ashkenazi Jews and European frequencies (p=0.94); or between Ashkenazi Jews and North Africans (p=0.28); nor between Europeans and North Africans (p=0.54)

at this locus. Publicly available allele frequencies (http://gnomad.broadinstitute.org) were similar (table 1). *VAMP4* rs11578699 MAF were between 0.15-0.30 in different populations.

We assessed possible association between these variants and *LRRK2* p.G2019S patients, and between these variants and idiopathic PD. There was no association between *DNM3* rs2421947 and p.G2019S genotype in all PD cases (p=0.42). There was no difference in allele frequencies between patients with PD and asymptomatic heterozygotes (p=0.056). When the original discovery data was meta-analysed with data from this study there was a significant association between *DNM3* rs2421947 and PD affected status in p.G2019S heterozygotes ($p=3.1x10^{-5}$). In the most recent PD GWAS there was no genome-wide significant association between *DNM3* rs2421947 and PD (p=0.0051). There was no association between *VAMP4* rs11578699 and *LRRK2* p.G2019S status in European PD (p=0.60) and no difference in *VAMP4* rs11578699 MAF between *LRRK2* p.G2019S patients and asymptomatic carriers (p=0.64). Mean sampling age for p.G2019S heterozygotes without PD was 56 years across ethnicities; 63 years in AJ subjects; and 56 years for Europeans.

Table 1. Minor allele frequency (MAF) for *DNM3* rs2421947 and *VAMP4* rs11578699 in PD cases and controls with and without p.G2019S

		gnomad[a]	p.G2019S PD[b]	p.G2019S asymptomatic [c]	Idiopathic PD	Total
***DNM3* rs2421947**	**Ashkenazi Jewish**	0.47 (n=145)	0.41 (n=146)	0.79 (n=7)	N/A	0.45 (n=298)
	African/North African	0.42 (n=4349)	0.39 (n=480)	N/A	N/A	0.42 (n=4829)
	Spanish/Latino	0.41 (n=424)	0.49 (n=195)	0.49 (n=117)	N/A	0.44 (n=736)
	European (non-Finnish)	0.41 (n=7,700)[e]	0.47 (n=483)	0.68 (n=50)	0.44 (n=14,508)	0.43 (n=22,734)
	Total	0.41 (n=12,618)	0.44 (n=1304)	0.56 (n=174)	0.44 (n=14,508)	N/A
***VAMP4* rs11578699**	**Ashkenazi Jewish**	0.20 (n=138)	0.27 (n=24)	0.5 (n=8)	N/A	0.22 (n=170)
	African/North African	0.18 (n=3979)	0.26 (n=484)	N/A	N/A	0.19 (n=4463)
	Spanish/Latino	0.16 (n=393)	N/A	N/A	N/A	0.16 (n=393)
	European (non-Finnish)	0.19 (n=6994)	0.20 (n=248)	0.18 (n=22)	0.19 (n=21,242)	N/A
	Total	0.18 (n=11,504)	0.23 (n=752)	0.27 (n=30)	0.19 (n=21,242)	N/A

[a]The gnomad database features unrelated individuals sequenced as part of disease-specific and population based genetic studies. Individuals affected by severe paediatric disease and their first-degree relatives have not been included.

[b]*LRRK2* p.G2019S parkinsonism

[c]*LRRK2* p.G2019S asymptomatic carriers

[d]North African *LRRK*2 p.G2019S parkinsonism

[e]This series includes ethnicity sub-categories: Estonian, Other non-Finnish European, North-western European, Southern European

[f]All samples from the following categories were combined: gnomAD, idiopathic PD, *LRRK2* p.G2019S patients and asymptomatic carriers.

We then assessed the relationship between AAO and *DNM3* rs2421947 and *VAMP4* rs11578699 (table 2) using t-tests and ANOVA. There was no effect of *DNM3* rs2421947 on AAO as assessed through ANOVA (p=0.55, F=0.59, df=2, n=708). When meta-analysed with data from previous cohorts there was a small effect of *DNM3* on AAO (2 years difference between G and C genotypes), (p=0.021, F=3.88, df=2, n=1304).

Table 2 Age at onset by genotype in *LRRK2* p.G2019S carriers in this study

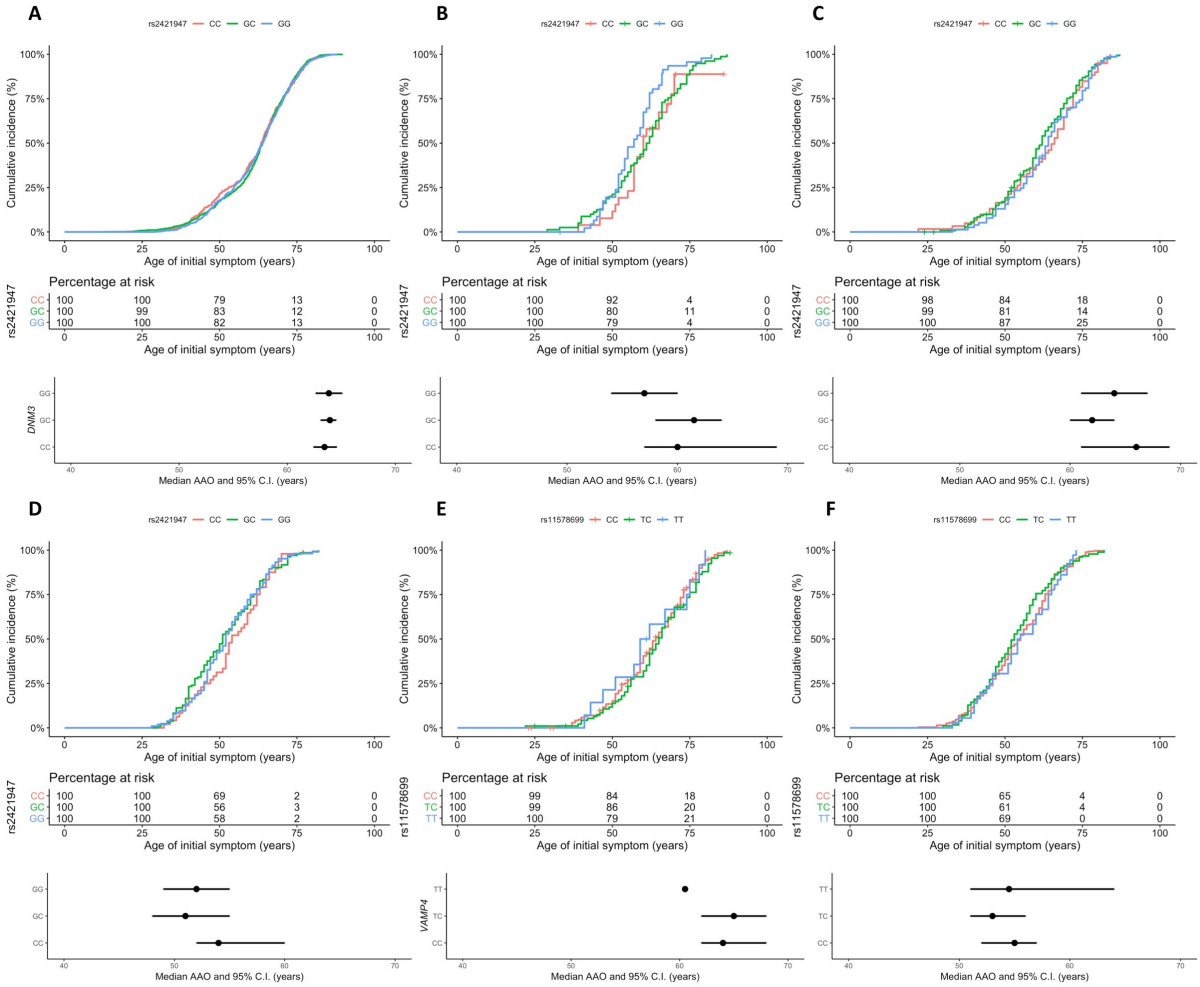
***DNM3* rs2421947**							
Population series	**Mean age at onset (sd)**			**ANOVA p value**	**GG vs CG p[c]**	**CG vs CC p[c]**	**GG vs CC p[c]**
	GG	**CG**	**CC**				
Ashkenazi Jewish (n=146)	57.20 (8.96)	60.11 (12.21)	58.41 (8.01)	0.34	0.13	0.45	0.58
North African (n=285)	52.64 (11.21)	51.95 (11.99)	54.27 (11.05)	0.49	0.65	0.23	0.40
European[a] (n=277)	63.26 (12.14)	61.20 (12.45)	62.17 (13.48)	0.52	0.24	0.64	0.63
Combined[b] (n=708)	57.11 (12.02)	57.44 (12.93)	58.59 (12.30)	0.55	0.76	0.37	0.28
***VAMP4* rs11578699**							
Population series	**CC**	**TC**	**TT**	**ANOVA p value**	**CC vs TC p[c]**	**TC vs TT p[c]**	**CC vs TT p[c]**
Ashkenazi Jewish (n=24)	63.92 (9.57)	57.45 (11.93)	57.00 (N/A)	0.91	0.17	N/A	N/A
North African (n=484)	54.38 (11.89)	53.35 (11.58)	55.22 (11.48)	0.55	0.36	0.38	0.68

European[a] (n=248)	62.44 (12.38)	63.89 (12.21)	61.33 (13.75)	0.66	0.41	0.55	0.79
Combined[b] (n=756)	57.39 (12.61)	56.41 (12.62)	56.76 (12.10)	0.60	0.32	0.85	0.73

Combined subgroup mean and standard deviation age at onset in years of *LRRK2* p.G2019S parkinsonism for *DNM3* rs2421947 and *VAMP4* rs11578699 genotypes; One-way ANOVA between *DNM3* rs2421947 and *VAMP4* rs11578699 genotypes; Student's t test between genotype
[a]The following population series subgroups from were combined: European unspecific, and British.
[b]The following population series subgroups from were combined: Ashkenazi Jewish, North African, European unspecific, and British.
[c]Uncorrected for multiple testing

We then carried out survival analysis. *DNM3* did not influence time to the development of disease in idiopathic PD (hazard ratio [HR] 0.98, 95% CI 0.89-1.07, p = 0·60, n=1956) for GG versus CC and CG carriers. This is consistent with the latest PD AAO GWAS, p=0.39, effect=-0.11, se=0.13. The Kaplan-Meier method was used for visualising risk across *DNM3* rs2421947 genotypes for idiopathic PD (Figure 1A), and *LRRK2* p.G2019S carriers (Figure 1B-D). The same method was also used to visualise risk for *VAMP4* rs11578699 genotypes in individuals with *LRRK2* p.G2019S (Figure 1E and F).

Figure 1. Survival analysis of AAO in *DNM3* rs2421947 (A-D) and *VAMP4* rs11578699 genotype groups (E-F)



A-C *DNM3* rs2421947 Kaplan Meier and Kaplan Meier median plots of **A** European idiopathic PD (n=1956); **B** Ashkenazi Jewish p.G2019S carriers (n=153) Kaplan Meier plot and median plot; **C** European p.G2019S carriers (n=286); **D** North African p.G2019S carriers (n=285).

E-F *VAMP4* rs2421947 Kaplan Meier and Kaplan Meier median plots **E** European and Ashkenazi Jewish p.G2019S carriers (n=302); **F** North African p.G2019S carriers (n=484).

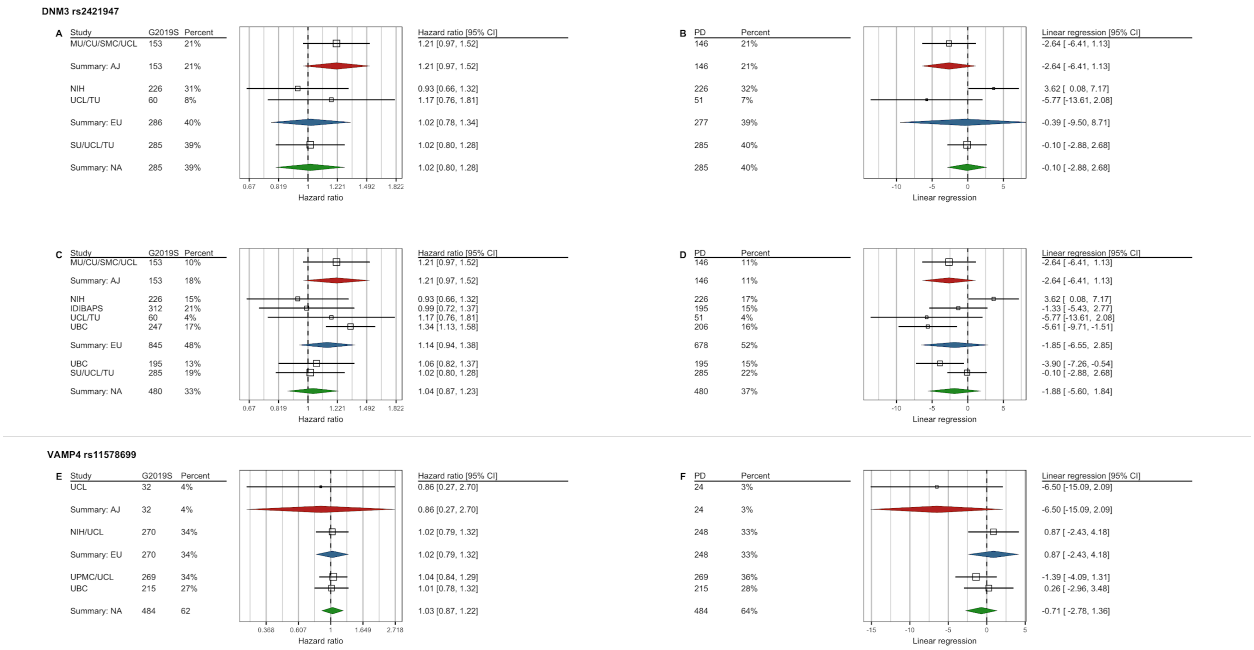
We carried out a multi-ethnic meta-analysis of newly contributed data of *DNM3* rs2421947 GG versus CC and CG genotypes against onset age in 724 *LRRK2* p.G2019S carriers, using a random effects model on disease-free survival. *DNM3* rs2421947 did not significantly affect the age associated hazards of developing parkinsonism (hazard ratio [HR] 1.09, 95% CI 0.95-1.25, p = 0.20, n=724), as shown in Figure 2 . I^2 heterogeneity was <0.01% (p = 0.42). Linear regression meta-analysis on PD AAO was not significant (beta = -1.19, p = 0.55, n =1304), though I^2 heterogeneity was 86.9%, p < 0.01. The impact of GG versus CC and CG was not significant in sub-group analyses (figure 2).

When our new data was pooled with the two previous studies and meta-analysed, there was a significant but small effect of *DNM3* rs2421947 on AAO in *LRRK2* p.G2019S parkinsonism ([HR] 1.14, 95% CI, 1.02-1.27, p = 0.025, n = 1478). I^2 total heterogeneity was 22.8%, p=0.37. However, linear regression meta-analysis of *LRRK2* p.G2019S AAO was not significant (beta = -2.21, p = 0.083, n = 1304). I^2 total heterogeneity for linear regression meta-analysis was 82.3%, p<0.0001 (figure 2).

Meta-analysis of *VAMP4* rs11578699 impact on AAO i*n LRRK*2 p.G2019S parkinsonism was not significant for CC versus TT and TC genotypes (hazard ratio [HR] 1.02, 95% CI 0.88-1.17, p = 0.84, n=786); (beta = -0.25, p=0.75, n=756).

We evaluated the possibility of ethnic specific effects using linear regression of AAO and Cox proportional hazards (figure 2). Though in each subgroup analysis confidence intervals overlapped zero, there appeared to be a trend towards earlier onset in G allele carriers of Ashkenazi Jewish ethnicity in both survival and regression analysis, and in the discovery data effects were strongest in Arab Berbers suggesting ethnic specific effects.

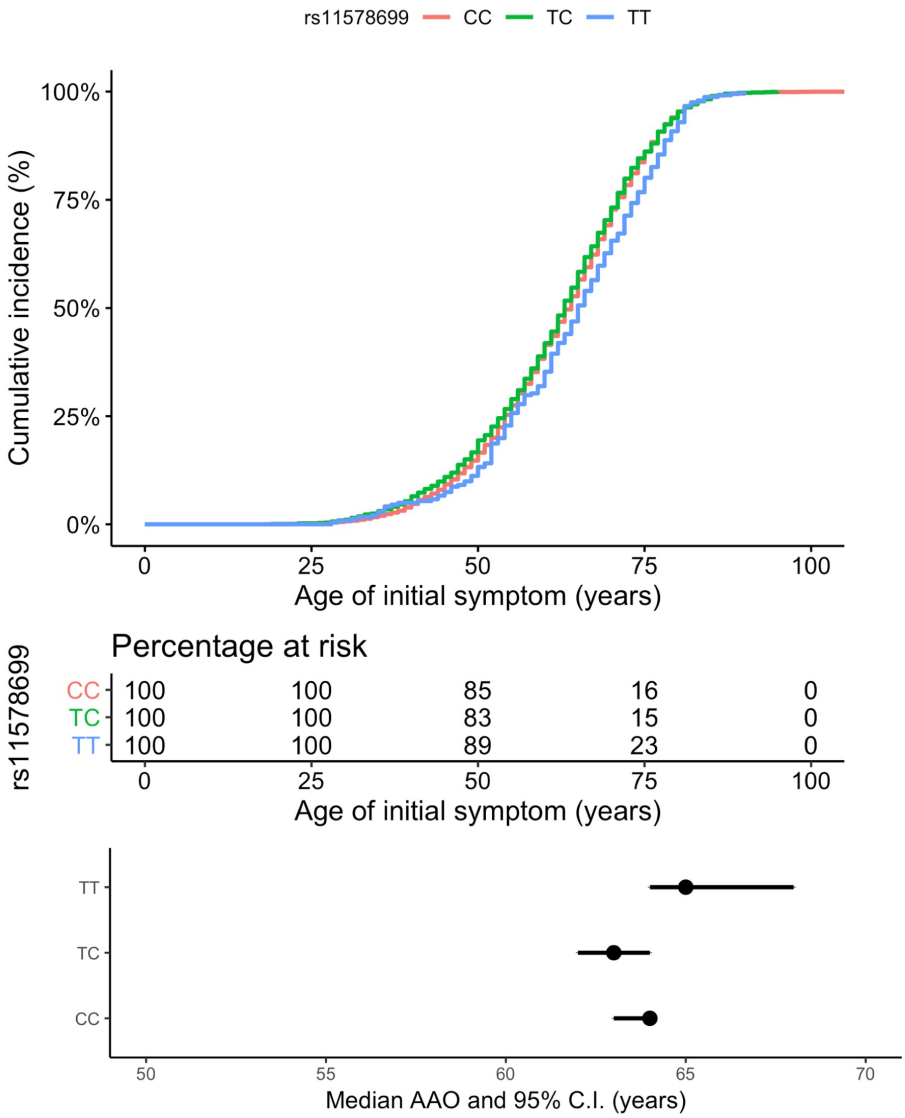
Figure 2. Forest plot of meta-analyses for *DNM3* rs2421947 and *VAMP4* rs11578699



(**A**, **C**, and **E**) Cox proportional hazards model meta-analyses *of LRRK*2 p.G2019S carriers, (**B**, **D**, and **F**) linear regression meta-analyses o*f LRRK*2 p.G2019S patients. **A** and **B**- meta-analyses of *DNM3* rs2421947 GG versus CG and CC genotypes of novel data from this manuscript from 724 *LRRK2* p.G2019S carriers; **C** and **D**- meta-analyses of *DNM3* rs2421947 GG versus CG and CC genotypes from 1478 *LRRK2* p.G2019S carriers including 754 previously published; **E** and **F**- *VAMP4* rs11578699 meta-analyses of CC versus TT and TC. Percentage contribution and numbers of individuals included in each analysis are indicated. Analyses are carried out on ethnicity subgroups: Ashkenazi Jewish (Summary: AJ), European (Summary: EU), and North African (Summary: NA).

Finally, we investigated the possibility that might be an interaction between the GWAS defined common *LRRK2* risk allele (rs10878226), and *DNM3* and *VAMP4*. We analysed 4882 cases with PD carrying the *LRRK2* risk variant minor allele using linear regression of AAO and plotting of the Kaplan Meier curve, as shown in Figure 3. TT versus TC and CC *VAMP4* rs11578699 genotype was significantly associated with AAO risk in this cohort (beta=1.68, se=0.81, p=0.037, n=4882*). VAMP4* rs11578699 was not significantly associated with AAO risk in idiopathic PD cases without the *LRRK2* rs10878226 risk variant (beta=-0.28, se=0.48, p=0.56, n=14970). Far fewer patients with idiopathic PD had been genotyped for *DNM3* rs2421947 in the PD GWAS cohort (table 1), meaning we were not able to assess interaction between *DNM3* and the *LRRK2* risk allele.

Figure 3. Kaplan Meier analysis by *VAMP4* rs11578699 genotype in PD cases carrying *LRRK2* risk variant (rs10878226)



We then reviewed *VAMP4* and *DNM3* human brain gene expression and co-expression data from the BRAINEAC[23] (11), GTEx, and Allen Atlas databases (12) (figure e-3). *VAMP4* rs11578699 is an eQTL for *VAMP4* expression: NES: 0.29, $p = 3.5 \times 10^{-9}$ (cerebellum, GTEx database); NES: 0.32, $p = 6.5 \times 10^{-7}$ (cerebellar hemisphere, GTEx

database). Cerebellum expression data for rs11578699 exceeds the significance level required for inclusion in COLOC analysis: $p = 5x10^{-8}$. *DNM3* rs2421947 did not appear to be an eQTL for *DNM3* expression in the brain in the GTex database, though *VAMP4* and *DNM3* gene expression in brain were high (figure e-3).

Discussion

We have analysed the effect of *DNM3* rs2421947 on AAO in *LRRK2* p.G2019S parkinsonism. In the original discovery analysis, there was a difference of 12.5 years between *DNM3* rs2421947 GG and CC genotypes. When all available data was analysed together there was a small effect of *DNM3* rs2421947 on AAO. However, (meta-)analysis of independent multi-ethnic series did not identify a significant difference between genotypes. Consistent with the recent AAO GWAS (4), our study indicated that *DNM3 rs2421947* does not affect AAO in non-p.G2019S disease. Analysis of carriers of a common *LRRK2* risk SNP provided support for a potential interaction between *VAMP4* and *LRRK2,* which requires replication in larger sample sizes. Gene expression analysis indicated that *VAMP4* rs11578699 is an eQTL for *VAMP4* expression. One possibility for the lack of replication of *DNM3* modification of p.G2019S AAO is that heterogeneity in the sample may limit the observed effect in the meta-analysis; this has previously led to non-replication in independent samples in genome-wide association studies (13) (14). Significant heterogeneity was observed in the linear regression meta-analysis.

Minor allele frequencies across the cohort in different populations were similar, making it less likely that spurious association due to population stratification had occurred. The

impact of *DNM3* may vary between ethnicities and the strongest effects were seen in Ashkenazi Jewish p.G2019S carriers (MAF was also slightly higher in AJ than in European p.G2019S cases), although not reaching significance (Figure 2) . Further studies in large numbers of Ashkenazi Jewish and Arab patients are needed.

Due to its role in innate immunity *LRRK2* may have been under different selective pressures in human evolution, relating to differences in environmental pathogens; Importantly, *LRRK2* p.G2019S penetrance varies across ethnic groups. Hentati, Trinh and colleagues (15) have presented data showing a later AAO in *LRRK2* p.G2019S carriers from Norway as compared to Tunisian patients. Lee, Wang (16) indicated later AAO penetrance for Ashkenazi Jewish PD p.G2019S carriers compared to non-Jewish carriers through plotting of estimated cumulative risk, though this was not significant; penetrance was 42.5% (95% CI: 26.3-65.8%) in non-Ashkenazi Jewish relatives compared to 25% (95% CI: 16.7%-34.2%) in Ashkenazi Jewish heterozygous relatives. There appears to be a reduction in penetrance for Ashkenazi Jewish LRRK2 p.G2019S carriers as compared to European *LRRK2* p.G2019S carriers, which is also apparent in our data. These data imply that there may be potential protective factors in the relatively homogeneous Norwegian and Ashkenazi Jewish population and that ethnic and population effects are likely to be very important in analysing this variant. Our analysis and comparison of background allele frequencies in this study suggest that this is unlikely to primarily relate to the *VAMP4* and *DNM*3 variants that we have studied. However there may be other rare variant effects that will emerge with fine mapping of this region.

The *LRRK2* common risk variant analysis is of interest; the variant is not a proxy for p.G2019S in Europeans, so this represents an independent marker of PD risk in this population. Around a quarter (24.6%) of sporadic PD cases carry the common *LRRK2* risk variant rs10878226, which is associated with PD (combined odds ratio [OR], 1.20, 95% CI, 1.08-1.33, $p=6.3\times10^{-4}$, n=6129) (3). Our common *LRRK2* variant analysis indicates an interaction between common variation in *LRRK2* and *VAMP4,* which requires further study. *VAMP4* and *LRRK2* are both involved in synaptic vesicle dynamics, which has relevance to the etiology of PD. Underscoring this possibility, analysis of *VAMP4* and *DNM3* expression indicated that both are highly expressed in the brain, though these are not in LD and likely represent independent signals. Variation in this genetic region may have important implications for the treatment of PD.

Acknowledgement

We would like to thank all of the subjects who donated their time and biological samples to be part of this study.

This work was supported by Parkinson's UK, University College London (UCL), and the Medical Research Council (MRC). Research ethics approval was provided by West of Scotland Research Ethics Service (reference 11/AL/0163) and multiple regional ethics committees. The study is registered with ClinicalTrials.gov under the identifier NCT02881099.

We would like to particularly thank Dr Benromdhan, Dr Mhiri, Dr Benmahdjoub, and Dr Arezki. We would also like to thank all members of the International Parkinson Disease Genomics Consortium (IPDGC). For a complete overview of IPDGC members, acknowledgements and funding, please see: http://pdgenetics.org/partners. This work was supported in part by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on Aging (NIA). The study was funded by the Medical Research Council and the NIHR Rare Diseases Translational Research Consortium.

Funding for the PFP study was provided by University College London (UCL), Parkinson's UK and the Medical Research Council. Research ethics approval was provided by the London Camden and King's Cross ethics committee (reference 15/LO/0097) and the Health Research Authority (HRA). The study is registered with ClinicalTrials.gov under the identifier NCT02760108.

The authors thank the PD patients and unaffected relatives who participated in this study.

Potential Conflicts of Interest

Dr Nalls reported receiving support from a consulting contract between Data Tecnica International and the National Institute on Aging (NIA), National Institutes of Health (NIH), and consulting for the Michael J. Fox Foundation, Vivid Genomics, Lysosomal Therapeutics Inc., and Neuron23, Inc, among others. Dr. Gan-Or has received consultancy fees from Lysosomal Therapeutics Inc., Idorsia, Denali, Prevail Therapeutics and Inception Sciences. Dr Morris is employed by UCL. In the last 12 months he reports paid consultancy from Biogen, UCB, Abbvie, Denali, Biohaven; lecture fees/honoraria from Biogen, UCB, C4X Discovery, GE-Healthcare, Wellcome Trust, Movement Disorders Society; Research Grants from Parkinson's UK, Cure Parkinson's Trust, PSP Association, CBD Solutions, Drake Foundation, Medical Research Council. Dr Morris is a co-applicant on a patent application related to C9ORF72 - Method for diagnosing a neurodegenerative disease (PCT/GB2012/052140)

References

1. Healy DG, Falchi M, O'Sullivan SS, et al. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. Lancet Neurol. 2008 Jul;7(7):583-90.
2. Di Maio R, Hoffman EK, Rocha EM, et al. LRRK2 activation in idiopathic Parkinson's disease. Sci Transl Med. 2018 Jul 25;10(451).
3. Mata IF, Checkoway H, Hutter CM, et al. Common Variation in the LRRK2 Gene is a Risk Factor for Parkinson's Disease. Movement Disorders. 2012 Dec;27(14):1822-5
4. Cornelis Blauwendraat KH, Costanza L. Vallerga, Sara Bandres-Ciga, Rainer von Coelln, Lasse Pihlstrom, Javier Simon-Sanchez. Parkinson disease age of onset GWAS: defining heritability, genetic loci and a-synuclein mechanisms. 2018.
5. Trinh J, Gustavsson EK, Vilarino-Guell C, et al. DNM3 and genetic modifiers of age of onset in LRRK2 Gly2019Ser parkinsonism: a genome-wide linkage and association study. Lancet Neurol. 2016 Nov;15(12):1248-56.
6. Fernandez-Santiago R, Garrido A, Infante J, et al. alpha-synuclein (SNCA) but not dynamin 3 (DNM3) influences age at onset of leucine-rich repeat kinase 2 (LRRK2) Parkinson's disease in Spain. Mov Disord. 2018 Apr;33(4):637-41.
7. Mike A Nalls CB, Costanza L Vallerga, Karl Heilbron, Sara Bandres-Ciga, Diana Chang, Manuela Tan, Demis A Kia,. Parkinson's disease genetics: identifying novel risk loci, providing causal insights and improving estimates of heritable risk. https://www.biorxiv.org/content/early/2018/08/09/3881652018.
8. Malek N, Swallow DM, Grosset KA, et al. Tracking Parkinson's: Study Design and Baseline Patient Data. J Parkinsons Dis. 2015;5(4):947-59.
9. Williams-Gray CH, Mason SL, Evans JR, et al. The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. J Neurol Neurosurg Psychiatry. 2013 Nov;84(11):1258-64.
10. Trabzuni D, Ryten M, Emmett W, et al. Fine-mapping, gene expression and splicing analysis of the disease associated LRRK2 locus. PLoS One. 2013;8(8):e70724.
11. Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci*. 2014; 17-1418-1428.

12. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 2012; 489:391-399.
13. Dunckley T, Huentelman MJ, Craig DW, et al. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. N Engl J Med. 2007 Aug 23;357(8):775-88.
14. Fernandez-Santiago R, Sharma M, Berg D, et al. No evidence of association of FLJ10986 and ITPR2 with ALS in a large German cohort. Neurobiol Aging. 2011 Mar;32(3):551 e1-4.
15. Hentati F, Trinh J, Thompson C, Nosova E, Farrer MJ, Aasly JO. LRRK2 parkinsonism in Tunisia and Norway: a comparative analysis of disease penetrance. Neurology. 2014 Aug 5;83(6):568-9.
16. Lee AJ, Wang Y, Alcalay RN, et al. Penetrance estimate of LRRK2 p.G2019S mutation in individuals of non-Ashkenazi Jewish ancestry. Mov Disord. 2017 Oct;32(10):1432-8.
17. Seber GAFW, C. J. . Nonlinear Regression. Hoboken, NJ1989.
18. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
19.. RStudioTeam. RStudio: Integrated Development for R. Boston, MA: RStudio, Inc; 2015.
20.. Terry M Therneau. A Package for Survival Analysis in S. 2015.
21.. Wolfgang Viechtbauer. Conducting meta-analyses in R with the metafor package. 2010;36(3):1—48.
22.. Kossmeier M. Visualising meta-analytic data with R package metaviz. 2019 [cited 2019]; Available from: https://cran.r-project.org/web/packages/metaviz/vignettes/metaviz.html.
23. Ross OA, Soto-Ortolaza AI, Heckman MG, et al. Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: a case-control study. Lancet Neurol. 2011 Oct;10(10):898-908.

Supplementary material

Table e-1. Imputation quality per subgroup

RSQ imputation score	
iPD *DNM3* rs2421947	
>0.99	100%
p.G2019S carriers *DNM3* rs2421947	
1.00	70%
>0.99	73%
>0.90	98%
p.G2019S carriers *VAMP4* rs11578699[1]	
1.00	79%
>0.98	86%
>0.75	100%
iPD *VAMP4* rs11578699	
Previously published PD GWAS data[1]	100%

appendix e-1

Supplementary methods

For the meta-analyses data was divided into the same ethnicity subgroups (Ashkenazi Jewish, North African, and European), and study centre specific groups. Where there were <60 datapoints from a particular study centre for *DNM3* rs2421947, data was pooled, and study centre of origin was used as a covariate. Cox proportional hazards model analyses and linear regressions for p.G2019S carriers were carried out on GG

versus CC and CG genotypes, as exploratory data analysis indicated a dominant protective effect of the minor C allele. For the *VAMP4* rs11578699 variant these survival methods were carried out on CC versus TC and TT due to the smaller number of TT genotypes available. Family relatedness, sex, and principal components of ancestry were used as covariates, where available, in each analysis.

A random-effects model was used in meta-analyses. Natural logs of hazard ratios for GG versus CC and CG genotypes were taken for *DNM3* rs2421947, and for CC versus TC and TT for *VAMP4* rs11578699. The log of standard errors was calculated using outputted hazard ratio confidence intervals, using the Taylor expansion with leading term (also known as the delta method) (17).

RSQ and D' were calculated in PLINK 1.9 (18). The allele based Fisher's Exact test, ANOVA, Student's t tests, and linear regressions (PD onset age regressed on *DNM3* rs2421947 genotype, and separately *VAMP4* rs11578699 genotype) were performed in base R RStudio (19). Cox proportional hazards survival analysis and Kaplan Meier curves were carried out using the Terry Therneau "Survival" package (20) in RStudio (19). Time-to-event in survival analyses was the self-reported age at which a participant developed motor symptoms of PD. For p.G2019S carriers without PD time-to-event was right-censored from sampling age. Meta-analysis was carried out in metafor (21) and metaviz R packages (22).

We supplemented the analysis of *LRRK2* p.G2019S by analysing time-to-event (by *DNM3* rs2421947 and *VAMP4* rs11578699) genotypes in PD cases carrying the *LRRK2* risk variant rs10878226 (4). We identified 4882 PD cases heterozygous or homozygous for the minor risk allele from the 21,242 PD cases in IPDGC datasets. Kaplan Meier and Cox proportional hazards analyses were then performed on *VAMP4* rs11578699 genotypes against age at onset. These methods were carried out on TT versus TC and CC genotypes as exploratory data analysis indicated an effect of the homozygous T allele on AAO. Covariates of sex, ethnicity (PC1-PC10) and dataset were used. We carried out the same analysis on 14970 PD cases not carrying the *LRRK2* rs10878226 variant.

Gene Expression

The BRAINEAC database hosts data on brain tissues (frontal cortex, temporal cortex, parietal cortex, occipital cortex, hippocampus, thalamus, putamen, substantia nigra, medulla, cerebellum, and white matter) from 134 healthy controls. The GTEx database consists of 8555 samples from 53 tissues (including 13 brain regions) from 544 donors. The Allen Human Brain Atlas database microarray data is from 8 neuropathologically normal individuals of varying ethnicity, and covers ~150 brain regions.

table e-2. Population specific D' and RSQ values between *DNM3* rs2421947 and *VAMP4* rs11578699

	Controls	PD cases	p.G2019S PD cases
D'			
Ashkenazi Jewish	N/A	N/A	0.087 (n=29)
African/North African	0.089 (n=1322)[a]	N/A	0.061 (n=482)
Spanish/Latino	0.30 (n=214)[b]	N/A	N/A
European non-Finnish	0.46 (n=594)[c]	0.21 (n=14508)	0.27 (n=186)
Total	0.22 (n=2130)	0.21 (n=14508)	0.021 (n=697)
RSQ			
Ashkenazi Jewish	N/A	N/A	0.0035 (n=29)
African/North African	0.0013 (n=1322)[a]	N/A	0.0022 (n=482)
Spanish/Latino	0.037 (n=214)[b]	N/A	N/A
European non-Finnish	0.039 (n=594)[c]	0.0086 (n=14508)	0.017 (n=186)
Total	0.0085 (n=2130)	0.0086 (n=14508)	0.000098 (n=697)

[a]All (AFR) Populations: (YRI, LWK, GWD, MSL, ESN, ASW, ACB)

[b]Selected (EUR) population: (IBS)

[c]Selected (EUR) population: (CEU, GBR, TSI)

figure e-1. Regional association plot of PD cases versus control from the largest PD GWAS, identifying rs11578699 as the lead SNP

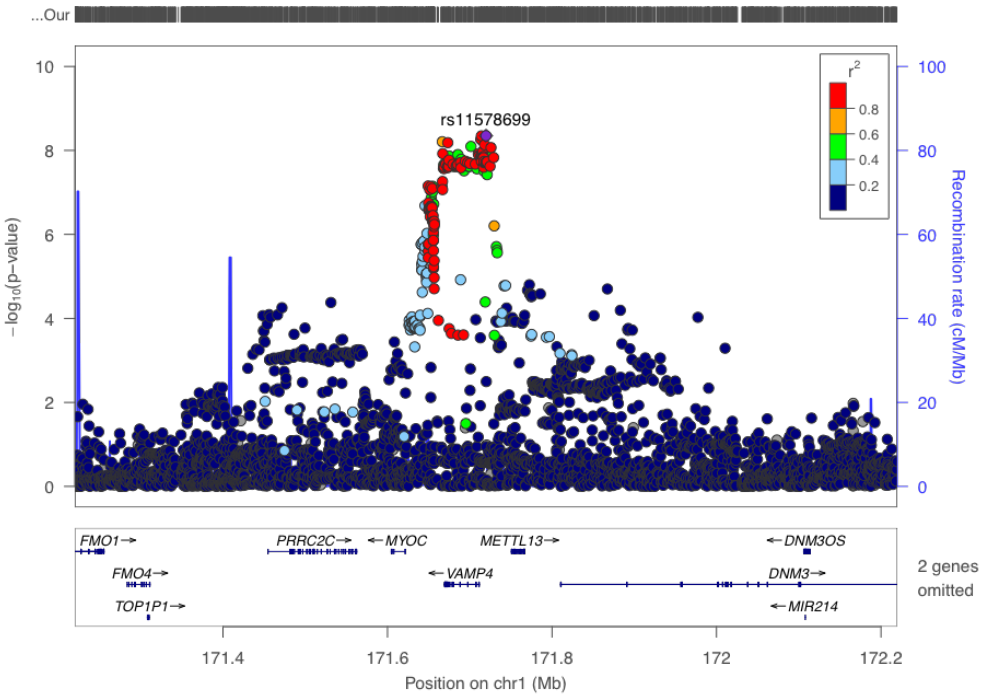


table e-3. Population series mean and standard deviation age at onset in years of p.G2019S PD cases for rs2421947 genotypes; One-way ANOVA between rs2421947 genotypes; Student's t test between rs2421947 genotypes

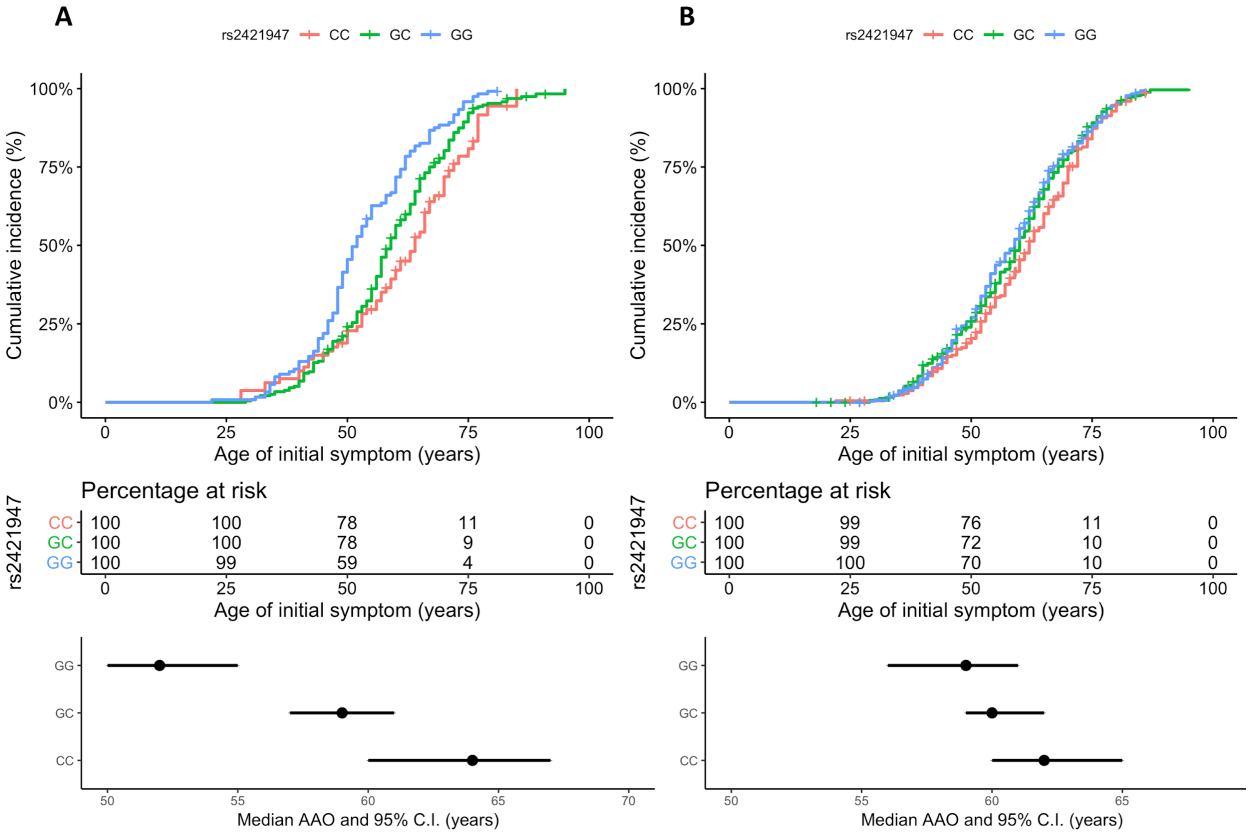
Population series	Mean age at onset (sd)			ANOVA p value	t test GG versus CG p value[b]	t test CG versus CC p value[b]	t test GG versus CC p value[b]
	GG	CG	CC				
European unspecific (n=239)[a]	65.35 (11.15)	62.19 (12.19)	62.51 (14.00)	0.23	0.075	0.88	0.24
American (n=87)	56.30 (7.53)	61.76 (10.14)	64.55 (10.98)	0.16	0.061	0.44	0.058
French (n=98)	52.08 (9.96)	50.93 (13.48)	49.94 (10.51)	0.81	0.66	0.77	0.49
Norwegian (n=18)	44.00 (N/A)	72.55 (15.12)	61.67 (20.81)	0.21	Insufficient count[c]	0.29	Insufficient count[c]
Spanish (n=195)	58.70 (14.01)	59.93 (12.59)	60.24 (12.78)	0.81	0.60	0.89	0.56
British (n=36)	49.68 (9.54)	53.78 (11.20)	60.00 (9.99)	0.13	0.32	0.18	0.042
Ashkenazi Jewish (n=146)	57.20 (8.96)	60.11 (12.21)	58.41 (8.01)	0.34	0.13	0.45	0.58
North African (n=480)	53.02 (11.89)	54.16 (11.41)	55.75 (12.11)	0.23	0.33	0.32	0.10

[a]Samples defined as from a non-specific European background (through PCA or self-identification during clinician interview).
[b]Uncorrected for multiple testing
[c]There are insufficient p.G2019S carriers in these categories for the t test to be calculated.

Comparison of datasets

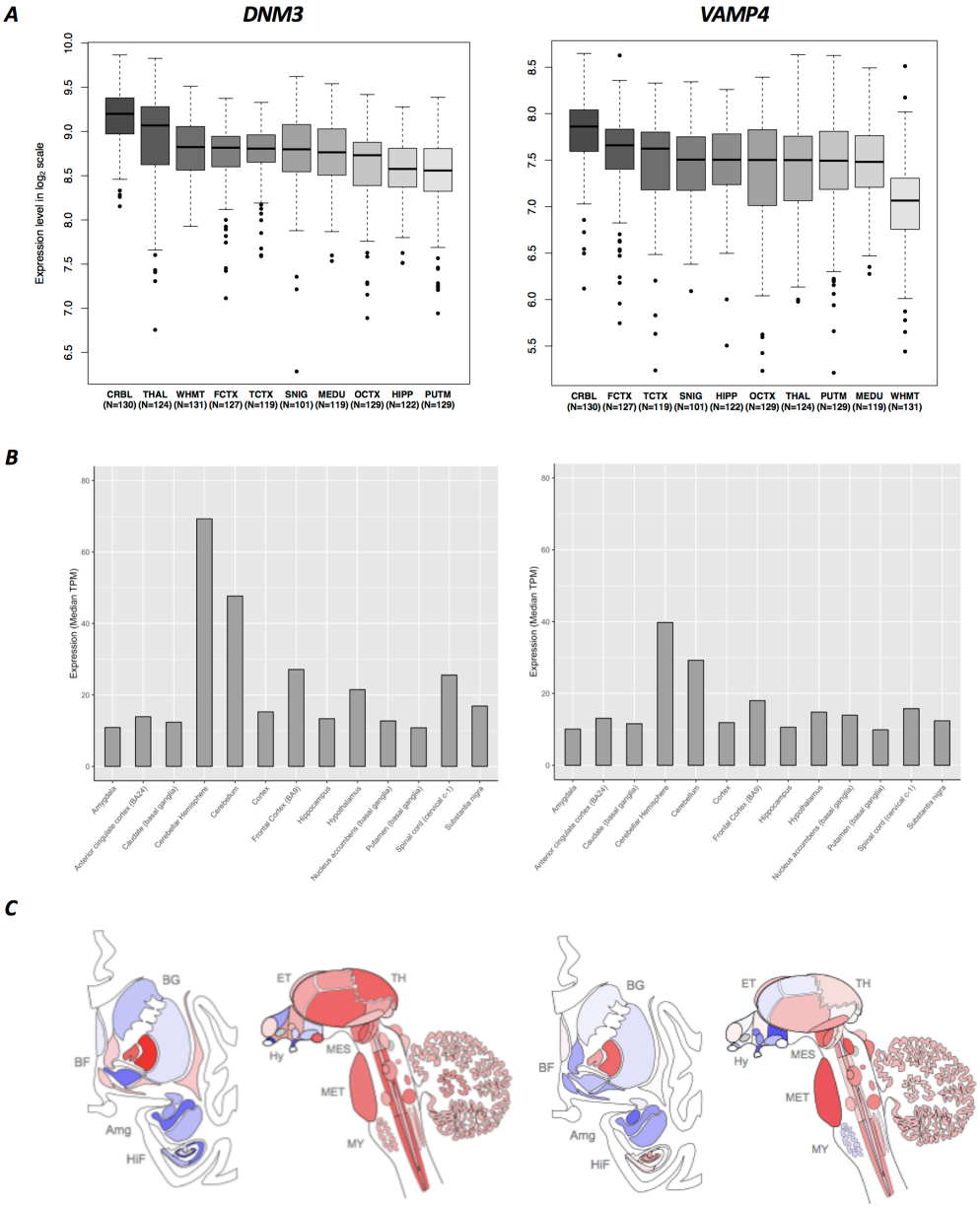
We analysed the discovery data and subsequent cohorts separately (GG versus CC and CG as described previously). There was a strong association between rs2421947 and PD AAO in discovery data (hazard ratio [HR] 1.59, 95% CI 1.28-1.97, $p<0.001$). This was also present through linear regression analysis (beta = -4.93, $p = 0.00019$). The association was absent in subsequent cohorts when analysed together (hazard ratio [HR] 1.09, 95% CI 0.95-1.26, $p = 0.22$). Linear regression analysis of subsequent datasets was not significant (beta = -0.85, $p = 0.35$).

figure e-2 Kaplan Meier plots by rs2421947 genotypes- comparison of discovery data and subsequent cohorts.



(A) Discovery data (n = 440) p.G2019S carriers Kaplan Meier plot and median plot (B) Subsequent cohorts (n = 1038) p.G2019S carriers Kaplan Meier plot and median plot.

figure e-3 Regional variation in the brain of *DNM3* and *VAMP4*



DNM3 and VAMP4 brain expression in: **A.** BRAINEAC database – CRBL = Cerebellum, PUTM = Putamen, HIPP = Hippocampus, TCTX = Temporal cortex, OCTX = Occipital cortex, FCTX = Frontal cortex, THAL = Thalamus, SNIG = Substantia nigra, MEDU = Medulla, WHMT = White matter; **B.** GTEx database; **C.** Allen Atlas database (European subjects) – Amg = Amygdala, BF = Basal Forebrain, BG = Basal Ganglia, CAU = Caudate, ET = Epithalamus, HiF = Hippocampal Formation, Hy = Hypothalamus, MES = Mesencephalon, MET = Metencephalon, MY = Myelencephalon, PUT = Putamen, TH = Thalamus. Image credit: Allen Institute.