In-depth mass-spectrometry reveals phospho-RAB12 as a blood biomarker of G2019S LRRK2-driven Parkinson's disease

Adriana Cortés,[1,†] Toan K. Phung,[2,†] Lorena de Mena,[3] Alicia Garrido,[3] Jon Infante,[4] Javier Ruíz-Martínez,[5] Miquel À. Galmés-Ordinas,[6] Sophie Glendinning,[2] Jesica Pérez,[3] Ana Roig,[3] Marta Soto,[3] Marina Cosgaya,[3] Valeria Ravasi,[3] Manel Fernández,[3] Alejandro Rubiano-Castro,[3] Ramón Díaz,[1] Haizea Hernández-Eguiazu,[5] Coro Sánchez-Quintana,[4] Ana Vinagre-Aragón,[5] Elisabet Mondragón,[5] Ioana Croitoru,[5] María Rivera-Sánchez,[4] Andrea Corrales-Pardo,[4] María Sierra,[4] Eduardo Tolosa,[3] Cristina Malagelada,[7] Raja S. Nirujogi,[2] Joaquín Fernández-Irigoyen,[1] Enrique Santamaría,[1] Dario R. Alessi,[2] María J. Martí,[3] Mario Ezquerra[3,‡] and Rubén Fernández-Santiago[3,‡]

[†,‡]**These authors contributed equally to this work.**

Abstract

Leucine-rich repeat kinase 2 (LRRK2) inhibition is a promising disease-modifying therapy for LRRK2-associated Parkinson's disease (L2PD) and idiopathic PD (iPD). However, pharmacodynamic readouts and progression biomarkers for clinical trials aiming for disease modification are insufficient since no endogenous marker reflecting enhanced kinase activity of the most common LRRK2 G2019S mutation has been reported yet in L2PD patients.

Employing phospho-/proteomic analyses we assessed the impact that LRRK2 activating mutations had in peripheral blood mononuclear cells (PBMCs) from a LRRK2 clinical cohort from Spain (n=174). The groups of study encompassed G2019S L2PD patients (n=37), non-manifesting LRRK2 mutation carriers of G2019S, here, G2019S L2NMCs (n=27), R1441G L2PD patients (n=14), R1441G L2NMCs (n=11), iPD patients (n=40), and healthy controls (n=45).

We identified 207 differential proteins in G2019S L2PD compared to controls (39 up/ 168 down) and 67 in G2019S L2NMCs (10 up/ 57 down). G2019S down-regulated proteins affected the endolysosomal pathway, proteostasis, and mitochondria, e.g., ATIC, RAB9A, or LAMP1. At the phospho-proteome level, we observed increases in endogenous phosphorylation levels of pSer106 RAB12 in G2019S carriers, which were validated by immunoblotting after 1 year of follow-up (n=48). Freshly collected PBMCs from 3 G2019S L2PD, 1 R1441G L2PD, 1 iPD, and 5 controls (n=10) showed strong diminishment of pSer106 RAB12 phosphorylation levels after in-vitro administration of the MLi-2 LRRK2 inhibitor. Using machine learning, we identified an 18-feature G2019S phospho-/protein signature discriminating G2019S L2PD, L2NMCs, and controls with 96% accuracy that correlated with disease severity, i.e., UPDRS-III motor scoring.

Using easily accessible PBMCs from a LRRK2 clinical cohort, we identified elevated levels of pSer106 RAB12 as an endogenous biomarker of G2019S carriers. Our data suggest that monitoring pSer106 RAB12 phosphorylation could be a relevant biomarker for tracking LRRK2 activation, particularly in G2019S carriers. Future work may determine whether pSer106 RAB12 could help with patient enrichment and monitoring drug efficacy in LRRK2 clinical trials.

Author affiliations:

1 Proteored-ISCIII, Proteomics Platform, Clinical Neuroproteomics Unit, Navarrabiomed, Departamento de Salud, UPNA, IdiSNA, ES 31008 Pamplona, Spain

2 Medical Research Council Protein Phosphorylation and Ubiquitylation Unit, University of Dundee, Dundee, DD1 5EH, UK

3 Lab of Parkinson's & Other Movement Disorders, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS); Parkinson's Disease and Movement Disorders Unit, Neurology Service, Hospital Clínic de Barcelona; Institut de Neurociències, Universitat de Barcelona; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED) CB06/05/0018-ISCIII; ES 08036 Barcelona, Spain

4 Neurology Service, University Hospital Marqués de Valdecilla-IDIVAL; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED); ES 39008 Santander, Spain

5 Department of Neurology, Donostia University Hospital; Biogipuzkoa Health Research Institute; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED); San Sebastián ES 20014, Spain

6 In Silico Medicinal Chemistry, Division of Cancer Therapeutics, The Institute of Cancer Research, London, SW7 3RP, UK

7 Department of Biomedicine, Faculty of Medicine, Universitat de Barcelona; Institut de Neurociències, Universitat de Barcelona; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED); ES 08036 Barcelona, Spain

Correspondence to: Dr Rubén Fernández-Santiago, PhD

Lab for Parkinson's & Other Movement Disorders

IDIBAPS-Hospital Clínic de Barcelona

CELLEX building (Faculty of Medicine)

Universitat de Barcelona

Casanova 143, Floor 3B

08036, Barcelona, Spain

E-mail: ruben.fernandez.santiago@gmail.com

Correspondence may also be addressed to: Dr Mario Ezquerra, PhD

E-mail: ezquerra@recerca.clinic.cat

Running Title: Elevated pSer106 RAB12 levels in LRRK2 G2019S

Keywords: Parkinson’s disease (PD); leucine-rich repeat kinase 2 (LRRK2); peripheral blood mononuclear cells (PBMCs); phospho-/proteomics; non-manifesting carriers; biomarker

Introduction

Activating mutations in the leucine-rich repeat kinase 2 (*LRRK2*), e.g., G2019S at the kinase or R1441G at the GTPase domains, increase LRRK2 kinase activity[1–4] causing autosomal-dominant LRRK2 Parkinson's disease (L2PD)[5,6]. By converging pathways, LRRK2 kinase activity appears to be also enhanced in patients with idiopathic PD (iPD)[7–9], which is clinically undistinguishable from L2PD[10,11]. Thus, ongoing clinical trials of small-molecule type-I inhibitors targeting active LRRK2 protein conformation represent a promising disease-modifying strategy for a broad spectrum of patients[12,13]. LRRK2 non-manifesting carriers (L2NMCs) are at high risk of PD in an age-dependent progressive manner[14–16], encompassing a candidate population for the continued clinical follow-up and disease course modification by early neuroprotective interventions.[13]

A subset of G-proteins from the Ras-related small GTPase superfamily[17] was reported as phosphorylation substrates of the LRRK2 Ser/Thr kinase[2,3]. Among these, pThr73 RAB10 was validated as an LRRK2 substrate[18] showing elevated endogenous phosphorylation levels in a large set of R1441G carriers, PD-manifesting, and non-manifesting, yet not in G2019S subjects[19]. Moreover, pThr73 RAB10 represents a readout for LRRK2 pharmacological inhibition using MLi-2 or DNL201[20,21]. In addition, RAB29[22,23], and more recently, RAB12[24,25] and RAB32[26], have been described as critical upstream LRRK2 activators. Despite significant progress, there is an urgent need of lack of clinical progression biomarkers and robust pharmaco-dynamic readouts useful for disease modification clinical trials.

By data-independent acquisition (DIA) mass-spectrometry (MS), we have screened the LRRK2 phospho-/proteome using peripheral blood mononuclear cells (PBMCs) from an extense LRRK2 clinical cohort (n=174) including G2019S L2PD (n=37), G2019S L2NMCs (n=27), R1441G L2PD (n=14), R1441G L2NMCs (n=11), iPD (n=40), and controls (n=45). We identified differential phospho-/proteins in G2019S and R1441G carriers, PD-manifesting and non-manifesting. More specifically we detected elevated pSer106 RAB12 phosphorylation levels in G2019S carriers. Our results suggest that pSer106 RAB12 is an endogenous biomarker of G2019S, which can be similarly elevated in G2019S L2PD and L2NMCs. Consistent with RAB12 being phosphorylated by LRRK2, we found that pSer106 RAB12 levels strongly diminished after MLi-2 LRRK2 inhibition in all kinds of subjects, regardless of disease or mutation status. We propose that pSer106 RAB12 could be exploited as a marker of LRRK2 activity in clinical trials[13].

Following FAIR[27] and through interactive Curtain weblinks for non-MS experts [28], we provide full open access to all data generated in this study.

Materials and methods

Subjects

Probands participated in the study after ethics approval and signed informed consent. Subjects included symptomatic and asymptomatic LRRK2 mutation carriers, iPD patients, and controls, which were healthy spouses and companions of Spanish descent. Patient inclusion criteria were a clinical diagnosis of PD by a movement disorders specialist based on the MDS criteria for Parkinson's[29]. Exclusion criteria were chronic inflammatory and autoimmune diseases, e.g., Crohn's (CD), inflammatory bowel disease (IBD), rheumatoid arthritis, systemic lupus erythematosus (SLE), chronic neurological diseases such as myasthenia gravis, chronic use of nonsteroidal anti-inflammatory drugs (NSAIDs) or corticosteroid anti-inflammatory medication, and viral or bacterial infection during the week precedent to blood donation. Subjects were recruited at three centres from Spain: Hospital Clínic de Barcelona (n=76) ('B')[30], Hospital Marqués de Valdecilla in Santander (n=55) ('S')[31], and Hospital de Donostia in San Sebastian (n=43) ('D')[32] (**Table 1**). By cohort and subject type, the sample included G2019S L2PD (n=37) (16 from B, 20 from S, and 1 from D), G2019S L2NMCs (n=27) (11 B, 15 S, and 1 D), R1441G L2PD (n=14) (1 B, and 13 D), R1441G L2NMCs (n=11) (3 B, and 8 D), iPD (n=40) (20 B, 10 S, and 10 D), and controls (n=45) (25 B, 10 S, and 10 D). We also collected sex, age at sampling, age-at-onset (AAO), LRRK2 mutation status, kinship to index cases, UPDRS-III[33], MoCA[34], autoimmune and environmental questionnaires, and COVID-19 history. PD patients had a mean age-at-sampling of 63.5 years for G2019S L2PD, 67.1 for R1441G L2PD, and 67.3 for iPD. Asymptomatic blood relatives of L2PD patients, i.e., L2NMCs, were younger than PD patients with an average of 56.7 years for G2019S L2NMCs and 61.1 for R1441G L2NMCs. The AAO was similar for G2019S and R1441G L2PD, with 55.1 and 55.8 years, respectively, whereas iPD had 62.1 years. Mean disease duration was 8.4 years for G2019S L2PD, 12.3 for R1441G L2PD, and 5.2 for IPD. Average disease severity, UPDRS-III motor scoring, was similar (mild) in all three patient groups, i.e., 16.0 in G2019S L2PD, 19.8 R1441G L2PD, and 19.7 for iPD. Mean

MoCA scores were also mild and similar in all patients: 24.3 for G2019S L2PD, 23.2 for R1441G L2PD, and 25.6 for IPD. Average L-DOPA equivalent daily dose (LEDD) was 635.8 mg for G2019S L2PD, 711.5 mg for R1441G L2PD, and 584.7 mg for iPD.

Genotyping

We genotyped the most common LRRK2 mutations in our population using Taqman SNP assays-on-demand for *LRRK2* G2019S (Thermo Fisher Sci. #C-63498123-10) and a commercial TaqMan assay for *LRRK2* R1441G[35] on a Step-One Plus Real-time PCR System (Life Tech. Inc.)

PBMC isolation

40 ml of peripheral blood were drawn early in the morning in fasting, and PBMCs were isolated by density gradient using Sodium-Citrate tubes (BD Vacutainer CPT, EAN30382903627821) following manufacturer's instructions. As usual in large-scale proteomic studies, dry PBMC pellets were isolated at every patient visit, immediately snap-frozen into liquid N_2, stored at -80°C overnight, and cryopreserved at -196°C in liquid N_2 for long storage (half a year on average) until DIA-MS.

PBMC preparation

PBMC samples from the three cohorts were processed in parallel. Blind experimental groups to the operator were balanced and randomised in runs to avoid manipulation bias. Briefly, PBMCs were homogenised in lysis buffer (7 M urea, 2 M thiourea, 50 mM dithiothreitol) supplemented with cOmplete Mini protease (Roche #11836153001) and PhosSTOP phosphatase (Roche #4906845001) inhibitors. Lysates were centrifuged at 20,000g, 1h, 15°C, and the resulting supernatant was quantified by the Bradford assay (Bio-Rad #5000201). Above 400 µg of protein were separated for protein digestion to obtain phosphorylated fractions. Proteins were reduced with DTT (final concentration of 20 mM; 30 min; room temperature), alkylated with

iodoacetamide (final concentration of 30 mM; 30 min in dark; room temperature), diluted to 0.9 M with ABC, and digested with trypsin (Promega #V5280) (1:20 w/w enzyme protein ratio, 18h, 37°C). Protein digestion was interrupted by acidification (acetic acid, pH<6), and the resulting peptides were cleaned up using Pierce Peptide Desalting Spin Columns (Thermo Fisher Sci. #89851). Phospho-peptide enrichment was performed using the High-Select TiO_2 Phospho-peptide enrichment Kit (Thermo Fisher Sci. #A32993) according to the manufacturer's instructions. Lastly, the enriched phospho-enriched fractions were cleaned up as described above and dried down in a Speed Vacuum system. Aliquots of 10 µg cleaned-up peptides from protein digestions were set aside for total protein analyses.

Data-independent acquisition (DIA) mass-spectrometry (MS)

Dried-down peptide samples were reconstituted with 2% ACN-0.1% FA (Acetonitrile-Formic acid), spiked with internal retention time peptide standards (iRT, Biognosys), and quantified by NanoDropTM spectrophometer (ThermoFisher Sci.) before LC-MS/MS in an EASY-1000 nanoLC system coupled to an EZ-Exploris 480 mass spectrometer (Thermo Fisher Sci.). Peptides were resolved using C18 Aurora column (75µm x 25cm, 1.6 µm particles; IonOpticks) at a flow rate of 300 nL/min using a 60-min gradient (50°C): 2% to 5% B in 1 min, 5% to 20% B in 48 min, 20% to 32% B in 12 min, and 32% to 95% B in 1 min (A = FA, 0.1%; B = 100% ACN:0.1% FA). Peptides were ionised using 1.6 kV spray voltage at a capillary temperature of 275 °C. We used data-independent acquisition (DIA) with full MS scans (scan range: 400 to 900 m/z; resolution: 60,000; maximum injection time: 22 ms; normalised AGC target: 300%) and 24 periodical MS/MS segments applying 20 Th isolation windows (0.5 Th overlap: Resolution: 15000; maximum injection time: 22 ms; normalised AGC target: 100%). Peptides were fragmented using a normalised HCD collision energy of 30%. MS data files were analysed using Spectronaut (Biognosys) by direct DIA analysis (dDIA). MS/MS spectra were searched against the Uniprot proteome reference from the *Homo sapiens* database UP000005640 using standard settings. The enzyme was set to trypsin in a specific mode. On the one hand, Carbamidomethyl (C) was set as a fixed modification, and oxidation (M), acetyl (protein N-term), deamidation (N), and Gln to pyroGlu as variable modifications for total protein analysis. On the other hand, Carbamidomethyl

(C) was set as a fixed modification, and oxidation (M), acetyl (protein N-term), and Phospho (STY) were used as variable modifications for phospho-proteome analysis. Identifications were filtered by a 1% Q-value. After MS, samples that did not pass QC were omitted from the study, resulting in a sample of G2019S L2PD (n=32) (15 from B and 17 from S), G2019S L2NMCs (n=22) (9 B and 13 S), R1441G L2PD (n=13) (1 B, and 12 D), R1441G L2NMCs (n=7) (2 B, and 5 D), iPD patients (n=39) (19 B, 10 S, and 10 D), and healthy controls (n=42) (23 B, 10 S, and 9 D). Lastly, to disambiguate peptide IDs into gene names we used the Uniprot online database (https://uniprotparser.proteo.info/). The resulting number of proteins (3,815) and phospho-peptides (10,288) identified by DIA-MS in human PBMCs attest to the optimal experimental quality of the LRRK2 clinical samples.

Proteome differential analysis

Proteome MS output data was exported from .SNE files from Spectronaut in a pivot table text format. For the differential analyses between groups, MS data was processed using QFeatures (doi: 10.18129/B9.bioc.QFeatures) in R (QFeatures v1.13.1). We applied the following R workflow: (i) Data was filtered to remove proteins identified by only 1 peptide sequence. (ii) Data selection was done based on condition and sub-group labels, with overall analysis containing all samples, G2019S analyses containing Barcelona and Santander samples labelled with the prefix 'B' or 'S', and R1441G analyses containing samples labelled with the prefix 'D' from Donostia-San Sebastian and 'B' from Barcelona if carrying R1441G. For each analysis, we provided a separate Rscript file with a customised group selection, as well as a single collapsing file with all scripts available at *Brain* online (**Suppl. Material**) and as a cloud weblink (doi.org/10.5281/zenodo.13774022). (iii) A protein ID column was assigned as an identification column for the analysis at QFeatures. (iv) We filtered out any row with 70% or more missing data. Here, with a 70% missing data cut-off, a meta-analysis would have 3,815 rows, while a more common 30% missing data cut-off would result in 3,789 rows. Since there was only about a 0.71% difference between the cut-off threshold, we chose the 70% cut-off to keep entries potentially found in only one group without affecting the statistical power of the entire analysis. (v) Imputation of missing data was done using the kNN method (QFeatures v1.13.1). Subsequently, (vi) we

performed a log2 transformation of the imputed data matrix, and (vii) designed a contrast matrix for differential analysis using limma[36]. (viii) For each contrast matrix, we performed a limma analysis using the Benjamini-Hochberg false discovery rate (FDR) multiple-testing adjustment under an statistical significance of an FDR adjusted $P<0.05$ (1.12 $\log_{10}$) and a $\log_2$ fold-change (FC) above |0.6| (|1.5| in lineal values). Scripts for proteome raw data download and re-analysis are available online (**Suppl. Material**). For ANOVA analysis, we used the normalised data from above as a starting point. The data from each row was grouped depending on the criteria used for grouping. Then, for each comparison, we applied a Python script using one-way ANOVA analysis on the grouped data within the comparison and returned the P-value output as a new column[37]. Lastly, we perform the same FDR correction from above to obtain the multiple-testing adjusted P-values using the Statsmodels Python package[38] with Python scripts also available online (**Suppl. Material**).

Phospho-proteome differential analysis

Phospho-proteome MS data was exported from .SNE files from Spectronaut in a long-form table format using a Spectronaut param export file available online (**Suppl. Material**). Data was imputed using a modified version of a collapsing R script (Perseus Plugin Peptide Collapse)[39] with phosphorylation as target modification at a confidence cut-off above 0.75. Modified collapsing.R and Perseus parameter.xml files are available online (**Suppl. Material**). We applied the following R workflow: (i) Columns with more than 70% blank cells were removed to meet the kNN requirement of less than 80% blank columns. (ii) Data selection for QFeatures input was based on condition and sub-group labels using all samples for overall analysis or specific group combination for location-specific and mutation-specific group combination, with overall analysis containing all samples, G2019S analyses containing Barcelona and Santander samples labelled with the prefix 'B' or 'S', and R1441G analyses containing samples labelled with the prefix 'D' from Donostia-San Sebastian and "B" from Barcelona when carrying R1441G. For each phospho-analysis group, we provide a separate Rscript file with customisation to the selection group (**Suppl. Material**). Subsequently, (iii) we performed imputation by removing any row with 30% or more empty data, similar to the proteome analysis using the kNN method, and (vi) performed $\log_2$ transformation

normalisation of the data using the quantile normalisation method. (v) The statistical significance criteria were set at an FDR multiple-testing adjusted P<0.05 (1.12 in $\log_{10}$) and a $\log_2$ fold-change (FC) above |0.6| (|1.5| in lineal values). (vii) In each differential analysis, we matched the protein and its original sequence using protein UniProt ID and extracted PTM position in protein and peptide, the peptide sequence, and the sequence window for visualisation at the Curtain tool[28]. Scripts for phospho-proteome data re-analysis are available at *Brain* online (**Suppl. Material**) and as a cloud weblink (doi.org/10.5281/zenodo.13774022). For phospho-proteome ANOVA analysis, we followed the same methodology as the proteome analysis, using the normalised phospho-proteome datasets from above. Data belonging to each group was identified from their column name. One-way ANOVA was applied on each row of cell groups from their respective comparison. The final statistically significant output values were adjusted using the Statsmodels package under the same FDR multiple-testing adjusted P<0.05. Python scripts for ANOVA phospho-proteome analyses are available online (**Suppl. Material**).

Data visualisation

Aligning to FAIR principles[27] of data findability, accessibility, interoperability, and reusability, we used Curtain and Curtain PTM[28], as free open-source tools for MS phospho-/proteomics data mining and exploitation by MS non-experts. Visualisation of each of the differential analysis results from limma was done in volcano plot representation using the default cut-off settings of a fold-change (FC) above |1.5| (|0.6| log2) and an FDR multiple-testing adjusted P<0.05 (1.12 in $\log_{10}$). The Curtain tools enable interactively perusing volcano plots, deconvoluting primary experimental data to individual replicates that can be visualised in bar charts or violin plots, allowing statistical analysis and export of plots in .SVG format (Curtain tutorials). For each analysis, we also provide web links in the Figure legends. From each link, users can view the data associated with each data point on the volcano plot as bar charts and violin plots. The magnitude of the data within these plots represents the relative intensity of the protein (total proteome) or phospho-site (phospho-proteome) before normalisation. Beyond simple visualisation of the numerical data, Curtain tools aggregate data for different knowledgebases, including UniProt, AlphaFold, PhosphoSitePlus, ProteomicsDB, and StringDB.

Machine learning modelling of G2019S differential phospho-/proteins

The normalised and imputed datasets comprising differentially expressed peptides and phospho-peptides were employed to train a multi-class classifier to distinguish between Controls, G2019S L2PD, and G2019S L2NMCs. Three distinct candidate models were considered, including Support Vector Machine (SVM), Random Forest (RF), and Gradient Boosting (GB) classifiers as described in other studies[40]. Parameter optimisation of the models was done through a grid search with a 5-fold cross-validation. To mitigate potential performance degradation due to unbalanced group sizes, we applied the Synthetic Minority Over-sampling Technique (SMOTE)[41] to the training split. We used the balanced accuracy score[42] defined as the average recall across each class to evaluate model performances. Implementation of the models was done using the Scikit-learn[43] v1.3.1 library within Python[44] programming language v3.9.18.

Classifier selection by a comparative performance of machine learning models

In the G2019S proteome dataset, we included 32 G2019S L2PD, 22 G2019S L2NMCs, and 42 controls that overpassed the QC criteria described above. Similarly, the phospho-proteome dataset comprised 29 G2019S L2PD, 19 G2019S L2NMCs, and 35 controls. First, we assessed comparative model performances for each dataset considering an initial number of features 3,816 peptides and 10,180 phospho-peptides respectively (**Suppl. Table 1**). Notably, in the proteome dataset, the SVM classifier demonstrated a substantial enhancement in balanced accuracy score following redundant feature elimination, achieving 0.91. This outcome indicates that the selective elimination of features contributed to obtaining a more discriminative model. Contrarily, the RF classifier showed limited improvement, implying that feature elimination methods were less effective for this specific model. Consistently, we obtained similar results for the phospho-proteome dataset where, after feature elimination, SVM achieved a balanced accuracy of 0.95, again highlighting the efficacy of feature selection in enhancing model performance. Furthermore, GB demonstrate significant improvement with only 43 features. This result indicates that the

model performance can be enhanced with only a small subset of features. After comparative evaluation and parameter optimisation, we identified SVM as the most optimal model to derive informative LRKK2 signatures using the minimum subset of relevant features that maximise the discrimination between classes.

Identification of a differential G2019S phospho-/protein signature

After the SVM model selection, an initial set of relevant features was determined by incorporating only statistically significant features ($P<0.05$) identified by the ANOVA test. Subsequently, we applied backwards Recursive Feature Elimination with Cross-Validation (RFECV)[45], i.e., we eliminated features with relative lower importance to iteratively reduce the number of features while maximising the balanced accuracy score. To obtain the LRKK2 signature, we employed the Monte Carlo Tree Search (MCTS)[46] method. The MCTS strategy involved selecting the minimum combination of features that maximise the score in an additive manner. Considering that the combinatorial features scale rapidly, the depth of the tree was fixed to five to manage computational complexity. The reward at each tree node was computed as the balanced accuracy score obtained through model training with cross-validation, utilising the selected subset of features. At each iteration, the number of trees evaluated was set to 10 times the number of features. After evaluating all the trees, the MCTS identified the best feature to add, maximising the reward. A stop node was introduced to halt the algorithm when no further improvement could be achieved. In summary, the procedure comprised: (i) selection of the first feature, (ii) MTCS evaluation of all possible trees and reward calculation, (iii) selection of the best feature to be added, (iv) iteration from the second step until the model stops, (v) repetition from the first step until all features were screened. After the screening of all features, we selected the combinations of features with a balanced accuracy score above 0.90. The most prominently represented features were used as initial features for refinement by MCTS. Discriminant LRKK2 signatures were defined as the subset with the highest score after the refinement. Feature selection and refinement were implemented in Python v3.9.18 using Scikit-learn v1.3.1 and MCTS v2.0.4 libraries (https://pypi.org/project/monte-carlo-tree-search).

Phospho-/protein gene ontology enrichment

Differential phospho-/proteins gene ontology (GO) was assessed using Metascape[47] cell component term using default settings (min. overlap: 3, min. enrichment: |1.5|, P<0.05), and a Benjamini-Hochberg false discovery rate (FDR) multiple-testing adjusted P<0.05. Specifically, for signature phospho-/proteins, we used a combination of cell component and biological processes, KEEGs, Reactome, and wiki pathways under the same statistical significance cut-off.

pSer106 RAB12 immunoblotting of 1-year follow-up PBMCs and MLi-2 LRRK2 inhibition assessment

Further details on pSer106 RAB12 validation by immunoblot in 1-year follow-up PBMC samples, as well as pSer106 RAB12 response to the LRRK2 MLi-2 inhibitor in freshly collected PBMCs can be found in the Supplementary Material.

Clinical correlation of LRRK2 differential phospho-/proteins and disease severity

We performed a Spearman's association analysis between the differential proteins and phospho-proteins across different comparisons (log2FC>|0.6|, adj. P<0.05) and UPDRS-III motor scores from PD patients and healthy controls. To this end, we used the "cor.test" function from R (stats v4.3.1) to calculate Rho coefficients and the EnhancedVolcano package (v1.20.0) to represent correlation outputs. Statistical significance was set a Spearman's correlation coefficient Rho>|0.5| and an FDR multiple-testing adjusted P<0.05.

Results

G2019S proteome analyses show endolysosomal pathway deregulation

We succeeded in quantifying the levels of 3,798 unique proteins using DIA-MS in our LRRK2 clinical cohort (**Fig. 1**). Pairwise analysis, under a cut-off of ≥2 peptide mapping, <0.30 imputation, $log_2FC>|0.60|$, and adj. $P<0.05$ revealed that G2019S L2PD was the most distinct group displaying a set of 207 proteins whose levels differed vs controls, with 85% down-regulated proteins (168 down/ 39 up) (**Fig. 2**). Specifically, G2019S L2PD showed a number of proteins which had reduced expression, among others, ATIC, which can repress LRRK2 and rescue neurodegeneration[48] ($log_2FC=-0.97$, adj. $P=1.92x10^{-13}$); RAB9A, involved in phagocytic vesicle trafficking and lysosomal function ($log_2FC=-1.17$, adj. $P=3.97x10^{-10}$); or LAMP1, a lysosome biogenesis and autophagy regulator ($log_2FC=-1.32$, adj. $P=1.63x10^{-9}$). G2019S L2NMCs vs controls showed 67 differential hits, also involving 85% down-regulated proteins (57 down/ 10 up), which were mostly common and with the same FC direction as in G2019S L2PD (42 of 67), e.g., ATIC or LAMP1 (**Suppl. Fig. 1**). G2019S L2PD vs L2NMCs differed in only 2 proteins, which were down-regulated in G2019S L2PD, i.e., RAB9A ($log_2FC=-0.77$, adj. $P=0.038$) and SCLY, a Selenocysteine lyase involved in peptide elongation ($log_2FC=-1.58$, adj. $P=0.038$). These results indicate proteome changes associated with the G2019S mutation, common to all G2019S carriers.

Proteome pathway deficits of R1441G are similar to G2019S

Regarding the R1441G proteome, R1441G L2PD vs controls revealed 80 hits (45 down/ 35 up) (**Suppl. Fig. 2**). Of these, 44% proteins (30 down/ 3 up) overlapped with G2019S L2PD and had the same FC direction, including down-regulation of NDUFB8, a mitochondrial Complex I subunit; PDCD6, a calcium sensor involved vesicle trafficking and apoptosis; RPL11, a component of the 60S ribosomal subunit; and other hits such as ATIC, RAB9A, LAMP1, and SLCY. Similarly, R1441G L2NMCs vs controls showed 5 down-regulated proteins, all common

to R1441G L2PD, including NDUFB8 and PDCD6. Between R1441G L2PD and L2NMCs, 2 proteins were up-regulated in R1441G L2PD, i.e., ATG3, an E2 ubiquitin-like conjugating enzyme, and MAGT2, which is essential for Golgi protein N-glycosylation. In addition, iPD vs controls, despite their larger sample, had only 3 differential hits, all down-regulated and common to L2PD, i.e., SRSF1, an RNA splicing factor; UQCRB, a mitochondrial Complex III subunit; and LAMP1 (**Suppl. Fig. 1** and **Suppl. Fig. 4**). Such findings can be related to the clinical heterogeneity of iPD with diverse genetic and environmental aetiology. Functionally, proteome changes in G2019S and R1441G L2PD, even iPD, revealed a shared biological enrichment affecting endolysosomal trafficking, protein homeostasis (i.e., proteostasis), and mitochondrial function (**Suppl. Fig. 3**).

LRRK2 phospho-proteome analyses uncovers elevated pSer106 RAB12 in G2019S carriers

Regarding the G2019S phospho-proteome, we found 10,288 phospho-sites mapping to 2,657 proteins. Using the same stringent cut-off as above, G2019S L2PD vs controls displayed a single differential phospho-site, pSer106 RAB12, which was hyper-phosphorylated in G2019S L2PD vs controls ($log_2FC=0.97$; adj. P=0.036) as well as in L2NMCs (log2FC=0.92; adj. P=0.057) (**Fig. 3**). Remarkably, pSer106 RAB12 was shown as a key physiological LRRK2 substrate of higher expression than other RABs including pThr73 RAB10 in brain from PD models[49,50]. G2019S carriers as a whole also showed elevated levels of pSer106 RAB12 ($log_2FC=0.95$; adj. P=0.003) along with pTyr334 SKAP2 ($log_2FC=1.05$; adj. P=0.003), a protein involved in immune response at peripheral tissues that regulates neural functions in the CNS,[51] including a-synuclein phosphorylation[52]. G2019S L2PD vs L2NMCs showed down-regulated pSer205 MON2 levels ($log_2FC=1.25$; adj. P=0.05), a regulator of endosome to Golgi trafficking. Lastly, we found no differential hit in G2019S L2NMCs compared to controls. Collectively, these results identify elevated pSer106 RAB12 levels in a large clinical cohort of G2019S carriers, pinpointing for the first time pSer106 RAB12 as an endogenous biomarker in PBMCs from G2019S carriers.

Phospho-proteome analyses in R1441G carriers and iPD

As for the R1441G phospho-proteome, R1441G L2PD vs controls showed no hit overpassing the multiple-testing adjustment of P-values. In addition, R1441G L2NMCs vs controls had 25 differential phospho-sites (20 down/ 5 up), but none of these included pSer106 RAB12 (**Suppl. Fig. 2**). Moreover, R1441G carriers as a whole vs controls also showed no hit overpassing the statistical significance cut-off. Altogether, these findings indicate that enhancement of pSer106 RAB12 phosphorylation is a specific effect in G2019S PBMCs and suggest distinct phospho-signalling preferences occurring for different pathogenic LRRK2 activating mutations such as G2019S and R1441G. Regarding iPD, at the phospho-proteome level, we found no phospho-peptide change vs controls (**Suppl. Fig. 4**). However, iPD revealed significant phospho-peptide differences compared to G2019S L2PD (84 down/ 9 up), including pSer106 RAB12, whose levels were elevated in G2019S L2PD, and also to R1441G L2PD (409 down/ 225 up). Altogether, these findings indicate that phospho-protein derangements are more prominent in L2PD due to phospho-signalling effects of LRRK2 activating mutations than in iPD, being pSer106 RAB12 a preferred LRRK2 substrate in PBMCs from G2019S carriers.

pSer106 RAB12 immunoblot validation and LRRK2 inhibition assessment

By immunoblot, we assessed pSer106 RAB12 levels as a pSer106 RAB12 / Total RAB12 ratio using >1-year follow-up PBMC samples from the G2019S cohort recruited at Clínic-Barcelona (n=48). These encompassed G2019S L2PD (n=12), G2019S L2NMCs (n=6), iPD (n=15), and healthy controls (n=15) (**Table 1**). Consistent with DIA-MS data, we found phosphorylation differences across the different groups (Kruskal-Wallis P=0.01), with increased pSer106 RAB12 phosphorylation levels in G2019S L2PD (Dunn's adj. P=0.069) and L2NMCs (Dunn's adj. P=0.118) vs controls. Similarly, G2019S carriers as a whole also showed elevated pSer106 RAB12 levels compared to controls (Kruskal-Wallis P=0.003; Dunn's adj. P=0.027), but not in iPD (**Fig. 4**, and **Suppl. Fig. 5**). However, by immunoblot[53] we did not observe down-regulation of proteome hits such as RAB9A in G2019S L2PD or iPD (Kruskal-Wallis P=0.08) nor LAMP1 except in iPD

(Kruskal-Wallis P=0.03; Dunn's adj. P=0.046). Lastly, we assessed pSer106 RAB12 response to LRRK2 pharmacological inhibition by MLi-2 using technical replicates from freshly collected PBMC pellets from an additional set of probands (n=10), including 3 G2019S L2PD, 1 R1441G L2PD, 1 iPD, and 5 controls treated with MLi-2 (200 nM; 30 min) or DMSO (**Suppl. Fig. 6**). In all subjects, we observed a substantial diminishment of pSer106 RAB12 phosphorylation levels after MLi-2 treatment, confirming pSer106 RAB12 as a pharmaco-dynamic readout of LRRK2 inhibition in PBMCs.

Phospho-/protein signatures define PD manifesting and non-manifesting G2019S carriers

Next, we interrogated G2019S phospho-/protein signatures. We applied a supported vector machine (SVM) classifier, adjusting for unbalanced group sizes, using 5-fold cross-validations as overfitting control. After recursive feature elimination, we obtained 510 peptides and 204 phospho-proteins as multi-class informative items. By Montecarlo Tree Search (MCTS), we refined combinations to the minimal numbers of features yielding the maximal balanced accuracy. We identified an 18-feature signature of 15 proteins and 3 phospho-proteins (**Fig. 5** and **Suppl. Fig. 7**) including pSer106 RAB12 and pSer205 MON2, ATIC, RAB9A, LAMP1, NDUFB8, and SCLY which yielded a balanced accuracy of 96% to discriminate G2019S carrier groups and controls. Specifically, receiver-operating curve (ROC) analysis revealed an area under the curve (AUC) of 1.00 for G2019S L2PD vs controls, 0.99 for G2019S L2NMCs vs controls, and 0.98 between G2019S L2PD and L2NMCs. The top gene ontology term of the 18 features was vesicle transport, thus supporting a biological plausibility. Altogether, the 18-feature phospho-/protein signature correctly classified 96% of G2019S L2PD, G2019S L2NMCs, and healthy controls, thus holding a potential to assess disease progression.

Differential phospho-/proteins correlate with disease severity

Lastly, we assessed the relation between deregulated phospho-/proteins and disease severity. Under a Spearman's Rho>|0.5| and a P<0.05, 16% of the differential proteins between G2019S L2PD vs controls (34 of 207) had an inverse association with UPDRS-III motor scores whereas pSer106 RAB12 and pSer205 MON2 had a direct correlation. Moreover, 55% of the 18 features at the G2019S phospho-/protein signature correlated inversely with UPDRS-III (ATIC, PDCD6, RAB9A, PSMC5, LAMP1, HSD13B10, ARHGAP45, NDUFB8, and SCLY) whereas pSer106 RAB12 correlated positively (Rho = 0.49, adj. P = $1.60x10^{-4}$), i.e., the higher the pSer106 RAB12 levels, the higher the UPDRS-III scores (**Fig. 6**). In R1441G L2PD vs controls, 81% of the differential proteins (65 of 80) correlated with UPDRS-III, both inversely (59%) or positively (41%). Remarkably, several of these R1441G proteins are part of the 18-feature G2019S signature (PDCD6, ARHGAP45, NDUFB8, RAB9A, ATIC, SCLY, and LAMP1), whereas other proteins were specific of R1441G, e.g., the mitochondrial UBQLN4 (Rho=-0.89, P=$1.64x10^{-6}$) or the cytoskeletal PLEC (Rho=0.84, P=$3.50x10^{-5}$) proteins. As an example, PDCD6, a top common correlating protein between G2019S and R1441G (Rho=-0.75, P=$5.51x10^{-10}$), participates in vesicle trafficking, mediates mitochondrial cytochrome c release and apoptosis[54], and has been linked to PD[55]. In summary, although correlation does not mean causality, differential phospho-/proteins at the 18-feature G2019S classifier are associated with disease severity, therefore holding clinical interest.

Discussion

Following FAIR principles[27], we employed an interactive tool called Curtain[28] in which the raw and differential phospho-/proteomic data from all analyses are saved as weblinks that non-MS experts can readily explore for public data mining. Overall, the G2019S L2PD proteome showed the highest number of changes, 207 proteins, most of which were down-regulated (85%). The G2019S L2NMCs displayed fewer protein differences, 67, which were also mostly down-regulated (85%). There was a substantial overlap between proteins that changed in both groups (60%). The comparison between G2019S L2PD and L2NMCs revealed two proteins, RAB9A and SCLY, which were down-regulated in the symptomatic carriers. Our findings indicate prominent protein

deficits associated with LRRK2 pathogenic mutations such as G2019S, which begin at G2019S L2NMCs premotor stages[56] and progress to G2019S L2PD stages.

Endolysosomal and proteostasis defects in G2019S LRRK2 blood

Gene ontology analysis annotated the protein changes in the G2019S carriers as participating in the endolysosomal pathway, i.e., vesicle trafficking. For example, G2019S L2PD showed down-regulation of RAB9A, which controls phagocytosis and lysosomal biology[57,58]. In G2019S carriers, we also observed down-regulated levels of LAMP1, a canonical lysosomal marker involved in lysosome biogenesis, which supports an enhanced LRRK2 activity in G2019S carriers[7–9]. A previous study also noted that LAMP1 levels were reduced in CSF of L2PD[59]. Our findings are consistent with the current understanding of the LRRK2 pathway, indicating that the LRRK2 protein plays a crucial role in controlling the endolysosomal pathway[50,60]. Beyond that, we also observed protein changes related to G2019S affecting ribosomal function, protein homeostasis, mitochondrial function, and alternative splicing[61]. For instance, ATIC, the top protein down-regulated in G2019S carriers, catalyses the last two steps of mitochondria purine biosynthesis[62,63]. Another study has also linked ATIC to LRRK2 toxicity[48]. Other protein deficits included KARS1, a tRNA synthetase; PSMC5, the proteasomal 26S subunit; or SCLY, seleno-cysteine lyase, an enzyme involved in peptide elongation that has been related to neurodegeneration[64]. These findings align with studies reporting transcriptional repression of proteostasis regulators in G2019S L2PD[65] and proteostasis defects in PD substantia nigra[66].

Similar proteomic deficits in R1441G carriers as in G2019S

R1441G L2PD showed 80 differential proteins, 44% of which were shared with G2019S L2PD. Enrichment analysis showed that the functions of the proteins deregulated in R1441G carriers also affected the endolysosomal pathway, protein homeostasis, and mitochondrial function. Indeed, the R1441G L2PD top down-regulated protein, NDUFB8, is a subunit of the mitochondrial Complex I (NADH to Ubiquinone oxidoreductase) whose activity is deficient in PD[67]. In addition, R1441G L2PD and L2NMCs displayed few protein differences, notably ATG3, which is involved in

autophagy, and MGAT2, a Golgi glycosyl transferase. In summary, the proteomic effects of the R1441G mutation in our LRRK2 clinical cohort were largely similar to G2019S[68,69]. Furthermore, the iPD proteome, despite being the largest group, displayed only 3 differential proteins, which were commonly decreased in G2019S L2PD, R1441G L2PD, and iPD. These encompassed LAMP1, which further supports endolysosomal dysfunction occurring in iPD[60,70]; SRSF1, a Serine/Arginine-rich splicing factor; and UQCRB, a mitochondrial Complex III subunit (Ubiquinol-cytochrome c oxidoreductase). Beyond the etiopathological heterogeneity of iPD[71,72], the fewer protein changes detected in iPD than in G2019S and R1441G L2PD indicate stronger signal transduction derangements due to pathogenic LRRK2 mutations in L2PD than in iPD. However, proteome changes in G2019S and R1441G L2PD, even iPD, similarly affected the same biological processes.

pSer106 RAB12 as an endogenous G2019S biomarker

At the phospho-proteome level, a single hit, pSer106 RAB12, was elevated specifically in G2019S carriers but not in R1441G. Excitingly, this phospho-site comprises a critical physiological substrate of LRRK2[2]. Overall, the roles that RAB12 plays and its phosphorylation by LRRK2 are poorly understood. Phosphorylation of RAB12 is prominent in the brain in PD models and observed to be higher than other RAB substrates such as RAB10 in this organ[49,50]. Functionally, other studies showed that RAB12 is located in phagosomes, lysosomes, and late endosomes, where it may regulate endosome to trans-Golgi trafficking and exocytosis[73,74]. Ours is the first report of hyper-phosphorylated RAB12 in PBMCs from a large clinical cohort of G2019S carriers. In addition, we analysed n=48 follow-up PBMC samples after 1 year of follow-up by immunoblotting. Despite the lower sample than DIA-MS, we found an increase of pSer106 RAB12 in G2019S L2PD and L2NMCs. Previous studies in neutrophils probing for RAB10 but not RAB12 phosphorylation revealed elevated pThr73 RAB10 in R1441G but not G2019S carriers[19]. In our study, by DIA-MS, pThr73 RAB10 did not pass the QC cut-offs in all three cohorts, only in PBMCs from Barcelona (n=76), which did not show pRAB10 differences between G2019S L2PD and controls ($log_2FC=0.71$, adj. P=0.999) or between G2019S carriers and controls ($log_2FC=0.68$, adj. P=0.999). Such results in G2019S PBMCs suggest that either RAB12 is a

preferred substrate for LRRK2 - indeed, distinct mutation effects cannot be ruled out[75] - or that pThr73 RAB10 phosphatases, e.g., PPM1H[76], could dephosphorylate RAB10 more efficient than RAB12. Mechanistic studies on how G2019S and other LRRK2 variants preferentially phosphorylate different RABs in various cell types and using larger cohorts are warranted. Our study identifies pSer106 RAB12 as an endogenous biomarker in easily accessible PBMCs from carriers of the most prevalent G2019S mutation, either PD-manifesting or non-manifesting, suggesting that pSer106 RAB12 can be used as a marker of LRRK2 activity in G2019S carriers.

pSer106 RAB12 as a marker of LRRK2 activity

Upstream of LRRK2, PD cell models showed LRRK2 activation by VPS35/ RAB29 (RAB7L1) binding to a region on the Armadillo (ARM) domain termed 'Site-1'[22,23]. More recently, non-phosphorylated RAB12 was shown as a key LRRK2 activator that binds to a distinct site at the ARM domain termed 'Site-3'[24,25]. One study showed that RAB12 plays a role in recruiting LRRK2 to damaged or stressed lysosomes[25]. These studies suggested that ARM domain Site-1 or Site-3 inhibitors that block RAB binding could serve as novel therapeutic targets for allosteric inhibitors of LRRK2 kinase activity[24]. The biological effect of pSer106 RAB12 phosphorylation on LRRK2 regulation has not been well characterised, and our results in G20919S carriers emphasise that additional work is warranted to investigate this. Specifically, it is key to investigate whether pSer106 RAB12 binding to the ARM Site-3 can create feedback loops modulating LRRK2 activity through activation or inhibition depending on the cellular context and, importantly, how this translates to LRRK2 clinical cohorts. Downstream of LRRK2, MLi-2 phospho-proteomics identified RAB3A, RAB8A, RAB10, RAB12, RAB29, and RAB43 as LRRK2 substrates[2,3,20,77]. In line with these studies, using freshly collected PBMCs (n=10), we found that pSer106 RAB12 levels strongly diminished after MLi-2 LRRK2 inhibition in all subjects, regardless of disease or mutation status. In the clinical setting, only pThr73 RAB10 has been validated as an LRRK2 substrate[19] and exploited as a readout of LRRK2 activity in previous studies[18], including in LRRK2 inhibitor clinical trials[12,21]. As mentioned above, there has not been a specific way of assessing elevated LRRK2 activity in G2019S carriers yet due to the lack of effect in pThr73 RAB10 phosphorylation. Thus, monitoring pSer106 RAB12 phosphorylation levels could be useful for

assessing G2019S selective inhibitors that have been newly developed in clinical studies[78–81], as these would be expected to preferentially reduce pSer106 RAB12 phosphorylation in patients with heterozygous G2019S mutations.

Dual role of RAB12 upstream and downstream of LRRK2 signalling

RAB12 was shown to play a dual role both in downstream and upstream signalling of LRRK2. Current evidence points to a mechanism by which lysosomal stress and dysfunction lead to the GTP loading and activation of dephosphorylated RAB12 at the lysosome membrane. This, in turn, recruits LRRK2 to the lysosome, where RAB12 directly interacts with Site-3 on the ARM domain of LRRK2.[82] LRRK2 is then activated at the lysosomal membrane, although the exact mechanism remains poorly understood. Once activated, LRRK2 phosphorylates nearby RAB proteins, including RAB12 at Ser106. In this context, our data suggest that monitoring pSer106 RAB12 phosphorylation could be a relevant biomarker for tracking LRRK2 activation, particularly in LRRK2 G2019S PBMCs. Further research is needed to fully understand the biological roles of phosphorylated RAB12 and to identify the proteins and downstream pathways it regulates.

G2019S phospho-/protein signatures can reflect disease progression

In G2019S carriers, we identified a signature of 15 proteins and 3 phospho-sites, including pSer106 RAB12, that was found to provide a 96% accuracy in discriminating G2019S L2PD, L2NMCs, and controls. Although correlation does not imply causality, 55% of the signature features correlated with PD motor severity as determined by UPDRS-III scores, including pSer106 RAB12, pSer2015 MON2, ATIC, PDCD6, RAB9A, PSMC5, LAMP1, HSD13B10, ARHGAP45, NDUFB8, and SCLY. These results suggest that this phospho-signature can be related to PD progression. However, further work in larger LRRK2 clinical cohorts would be required to assess this clinically[56]. Altogether, as a proof-of-principle, we identified the first phospho-/protein signature in G2019S PBMCs based on DIA-MS data, which complements previous G2019S signatures reported in blood[83] and urine[84,85].

Study limitations

Despite the exciting findings, our study has limitations. Inherent variation in humans markedly affects differential protein expression and phosphorylation. Slightly different PBMCs preparation and storage procedures at different centres can also affect the results. To minimise this variation, we undertook DIA-MS analyses simultaneously for all subcohorts and blind to study groups. Differential enrichment of phospho-peptides on titanium dioxide beads can result in further variety. Indeed, we discarded one of the phospho-peptide batches due to not passing quality control. Based on our phospho-peptide enrichment approach, the detection of phospho-Tyrosines was under-represented. We used stringent significance cut-off criteria filtering in only hits mapped by at least 2 peptides, and we cannot rule out that other important proteins have been excluded. The number of R1441G carriers, especially L2NMCs, was smaller than G2019S, and it was limited for phospho analyses and insufficient to assess signatures by machine learning. Yet, pSer106 RAB12 did not show significant differences nor trends in R1441G groups. The validation of pRAB12 by immunoblot in 1-year follow-up clinical samples strengthens the robustness of this as an endogenous biomarker for G2019S carriers. Lastly, other phospho-/protein candidates identified by DIA-MS should be validated in additional studies.

Concluding remarks

Aligning with urine[85], in PBMCs, we found elevated pSer106 RAB12 levels as an endogenous biomarker for G2019S carriers. This finding holds clinical implications, suggesting that pSer106 RAB12 can be a marker of LRRK2 activity in G2019S carriers. Other studies should also assess pSer106 RAB12 levels in CSF and brain tissue of LRRK2 patients carrying the G2019S mutation. In addition, given that RAB12 was shown as a key LRRK2 activator in PD models able to increase pThr73 RAB10 levels[24,25], future studies ought to investigate the effect of pSer106 RAB12 phosphorylation on LRRK2 activation. Moreover, in line with findings from PD models[50,86], in human LRRK2 PBMCs, we also found that pSer106 RAB12 represents a pharmaco-dynamic readout of LRRK2 inhibition. In addition, we found an 18-feature signature, including pSer106

RAB12, with a 96% accuracy in discriminating symptomatic, asymptomatic G2019S carriers and controls. Future large-scale studies need to assess pSer106 RAB12 in other G2019S clinical cohorts. Moreover, developing novel assays able to quantify pSer106 RAB12 in LRRK2 clinical samples such as blood cells and CSF, e.g., reaction monitoring or ELISA-based assays, are needed to translate our findings to clinical settings. If validated, pSer106 RAB12 can aid patient enrichment and target engagement in clinical trials of novel LRRK2 inhibitors targeting the G2019S mutation[78–81].

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[87] partner repository with the dataset identifiers PXD050865 for the proteome and PXD050944 phospho-proteome analyses. Following FAIR principles,[27] through the interactive tool called Curtain[28], raw and differential phospho-/proteomic data from all comparisons are also provided as weblinks to be readily explored by non-MS experts. Programming Rscripts for data analyses are publicly available at *Brain* online (**Suppl. Material**) and as a cloud weblink (doi.org/10.5281/zenodo.13774022).

Web resources

Curtain weblinks of LRRK2 phospho-/proteomic analyses as presented in the article. Expanded analyses can be found at *Brain* online (**Suppl. Figures** and **Suppl. Material**).

Proteome G2019S cohort

G2019S L2PD vs controls

G2019S carriers vs controls

G2019S L2NMCs vs controls

G2019S L2NMCs vs L2PD

iPD vs controls

G2019S L2PD vs iPD

G2019S carriers vs iPD

Proteome R1441G cohort

R1441G L2PD vs controls

R1441G carriers vs controls

R1441G L2NMCs vs controls

R1441G L2NMCs vs L2PD

iPD vs controls

R1441G L2PD vs iPD

R1441G carriers vs iPD

Phospho G2019S cohort

G2019S L2PD vs controls

G2019S carriers vs controls

G2019S L2NMCs vs controls

G2019S L2NMCs vs L2PD

iPD vs controls

G2019S L2PD vs iPD

G2019S carriers vs iPD

Phospho R1441G cohort

R1441G L2PD vs controls

R1441G carriers vs controls

R1441G L2NMCs vs controls

R1441G L2NMCs vs L2PD

iPD vs controls

R1441G L2PD vs iPD

R1441G carriers vs iPD

Acknowledgements

We thank the patients and their relatives for their continued and essential collaboration. LDM was funded by the Beatriu-de-Pinòs programme (#BP00176) from the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR). AR was funded by the PFIS programme (#FI21/00104) from the Instituto de Salud Carlos III (ISCIII) co-funded by the European Union. MF was funded by the María-de-Maeztu programme (#MDM-2017-0729) to the Parkinson's disease and Movement Disorders group of the Institut de Neurociències (Universitat de Barcelona). ARC was funded by the EU Next-Generation 2022 Investigo programme from the European Commission (EC) / Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR). Research in the Alessi lab were funded by the UK Medical Research Council (grant number MC_UU_00018/1) and the pharmaceutical companies supporting the Division of Signal Transduction Therapy Unit (Boehringer Ingelheim, GlaxoSmithKline, and Merck KGaA). RFS was supported by a Miguel Servet grant (#CP19/00048), a FIS grant (#PI20/00659), a PFIS grant (#FI21/00104), and a M-AES grant (#MV22/00041) from the Instituto de Salud Carlos III (ISCIII) co-funded by the European Union. IDIBAPS receives support from the CERCA program of Generalitat de Catalunya.

Funding

This study was supported by a grant from the Michael J. Fox Foundation for Parkinson's Research (MJFF) (#MJFF-000858) to RFS, ME, CM, ES, JI, and JRM.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Sheng Z, Zhang S, Bustos D, et al. Ser1292 autophosphorylation is an indicator of LRRK2 kinase activity and contributes to the cellular effects of PD mutations. *Sci Transl Med*. 2012;4(164):164ra161. doi:10.1126/scitranslmed.3004485

2. Steger M, Tonelli F, Ito G, et al. Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases. *Elife*. 2016;5(JANUARY2016):e12813. doi:10.7554/eLife.12813.001

3. Steger M, Diez F, Dhekne H, et al. Systematic proteomic analysis of LRRK2-mediated Rab GTPase phosphorylation establishes a connection to ciliogenesis. *Elife*. 2017;6:e31012. doi:10.7554/ELIFE.31012

4. Taylor M, Alessi DR. Advances in elucidating the function of leucine-rich repeat protein kinase-2 in normal cells and Parkinson's disease. *Curr Opin Cell Biol*. 2020;63:102-113. doi:10.1016/j.ceb.2020.01.001

5. Zimprich A, Biskup S, Leitner P, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron*. 2004;44(4):601-607.

doi:S0896627304007202 [pii]10.1016/j.neuron.2004.11.005

6. Paisán-Ruíz C, Jain S, Evans EW, et al. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron*. 2004;44(4):595-600. doi:S0896627304006890 [pii]10.1016/j.neuron.2004.10.023

7. Di Maio R, Hoffman EK, Rocha EM, et al. LRRK2 activation in idiopathic Parkinson's disease. *Sci Transl Med*. 2018;10(451):eaar5429. doi:10.1126/scitranslmed.aar5429

8. Fraser KB, Rawlins AB, Clark RG, et al. Ser(P)-1292 LRRK2 in urinary exosomes is elevated in idiopathic Parkinson's disease. *Mov Disord*. 2016;31(10):1543-1550. doi:10.1002/mds.26686

9. Wang X, Negrou E, Maloney MT, et al. Understanding LRRK2 kinase activity in preclinical models and human subjects through quantitative analysis of LRRK2 and pT73 Rab10. *Sci Rep*. 2021;11(1):12900. doi:10.1038/s41598-021-91943-4

10. 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;7(7):583-590. doi:10.1016/S1474-4422(08)70117-0

11. Marras C, Alcalay RN, Caspell-Garcia C, et al. Motor and nonmotor heterogeneity of LRRK2-related and idiopathic Parkinson's disease. *Mov Disord*. Published online 2016. doi:10.1002/mds.26614

12. Jennings D, Huntwork-Rodriguez S, Vissers MFJM, et al. LRRK2 Inhibition by BIIB122 in Healthy Participants and Patients with Parkinson's Disease. *Mov Disord*. 2023;38(3):386-398. doi:10.1002/mds.29297

13. Tolosa E, Vila M, Klein C, Rascol O. LRRK2 in Parkinson disease: challenges of clinical trials. *Nat Rev Neurol*. 2020;16(2):97-107. doi:10.1038/s41582-019-0301-2

14. 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;83(6):568-569. doi:10.1212/WNL.0000000000000675

15. Marder K, Wang Y, Alcalay RN, et al. Age-specific penetrance of LRRK2 G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 Consortium. *Neurology*. 2015;85(1):89-95.

doi:10.1212/WNL.0000000000001708

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;32(10):1432-1438. doi:10.1002/mds.27059

17. Homma Y, Hiragi S, Fukuda M. Rab family of small GTPases: an updated view on their regulation and functions. *FEBS J*. 2021;288(1):36-55. doi:10.1111/febs.15453

18. Karayel, Ö, Tonelli F, Virreira Winter S, et al. Accurate MS-based Rab10 Phosphorylation Stoichiometry Determination as Readout for LRRK2 Activity in Parkinson's Disease. *Mol Cell Proteomics*. 2020;19(9):1546-1560. doi:10.1074/MCP.RA120.002055

19. Fan Y, Nirujogi RS, Garrido A, et al. R1441G but not G2019S mutation enhances LRRK2 mediated Rab10 phosphorylation in human peripheral blood neutrophils. *Acta Neuropathol*. 2021;142(3):475-494. doi:10.1007/s00401-021-02325-z

20. Thirstrup K, Dächsel JC, Oppermann FS, et al. Selective LRRK2 kinase inhibition reduces phosphorylation of endogenous Rab10 and Rab12 in human peripheral mononuclear blood cells. *Sci Rep*. 2017;7(1):10300. doi:10.1038/s41598-017-10501-z

21. Jennings D, Huntwork-Rodriguez S, Henry AG, et al. Preclinical and clinical evaluation of the LRRK2 inhibitor DNL201 for Parkinson's disease. *Sci Transl Med*. 2022;14(648):eabj2658. doi:10.1126/scitranslmed.abj2658

22. Purlyte E, Dhekne HS, Sarhan AR, et al. Rab29 activation of the Parkinson's disease-associated LRRK2 kinase. *EMBO J*. 2018;37(1):1-18. doi:10.15252/embj.201798099

23. Mir R, Tonelli F, Lis P, et al. The Parkinson's disease VPS35[D620N] mutation enhances LRRK2-mediated Rab protein phosphorylation in mouse and human. *Biochem J*. 2018;475(11):1861-1883. doi:10.1042/BCJ20180248

24. Dhekne HS, Tonelli F, Yeshaw WM, et al. Genome-wide screen reveals Rab12 GTPase as a critical activator of Parkinson's disease-linked LRRK2 kinase. *Elife*. 2023;12:e87098. doi:10.7554/elife.87098

25. Wang X, Bondar V V, Davis OB, et al. Rab12 is a regulator of LRRK2 and its activation

by damaged lysosomes. *Elife*. 2023;12:12:e87255. doi:10.7554/elife.87255

26. Gustavsson EK, Follett J, Trinh J, et al. RAB32 Ser71Arg in autosomal dominant Parkinson's disease: linkage, association, and functional analyses. *Lancet Neurol*. Published online 10 April 2024. doi:10.1016/S1474-4422(24)00121-2

27. Wilkinson M, Dumontier M, Aalbersberg I. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (2016). *Sci data*. 2016;3:1-9. doi:10.1038/sdata.2016.18

28. Phung TK, Berndsen K, Shastry R, et al. CURTAIN-A unique web-based tool for exploration and sharing of MS-based proteomics data. *Proc Natl Acad Sci U S A*. 2024;121(7):e2312676121. doi:10.1073/pnas.2312676121

29. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30(12):1591-1601. doi:10.1002/mds.26424

30. Gaig C, Marti MJ, Ezquerra M, Rey MJ, Cardozo A, Tolosa E. G2019S LRRK2 mutation causing Parkinson's disease without Lewy bodies. *J Neurol Neurosurg Psychiatry*. 2007;78(6):626-628. doi:jnnp.2006.107904 [pii]10.1136/jnnp.2006.107904

31. Sierra M, González-Aramburu I, Sánchez-Juan P, et al. High frequency and reduced penetrance of LRRK2 G2019S mutation among Parkinson's disease patients in Cantabria (Spain). *Mov Disord*. 2011;26(13):2343-2346. doi:10.1002/mds.23965

32. Ruiz-Martínez J, Gorostidi A, Ibañez B, et al. Penetrance in Parkinson's disease related to the LRRK2 R1441G mutation in the Basque country (Spain). *Mov Disord*. 2010;25(14):2340-2345. doi:10.1002/mds.23278

33. Goetz CC. The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. *Mov Disord*. 2003;18(7):738-750. doi:10.1002/mds.10473

34. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x

35. Simon-Sanchez J, Marti-Masso JF, Sanchez-Mut J V, et al. Parkinson's disease due to the R1441G mutation in Dardarin: a founder effect in the Basques. *Mov Disord*.

2006;21(11):1954-1959. doi:10.1002/mds.21114

36. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007

37. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2

38. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. In: *Proceedings of the 9th Python in Science Conference*. ; 2010:92-96. doi:10.25080/majora-92bf1922-011

39. Bekker-Jensen DB, Bernhardt OM, Hogrebe A, et al. Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries. *Nat Commun*. 2020;11(1):787. doi:10.1038/s41467-020-14609-1

40. Ogutu JO, Piepho HP, Schulz-Streeck T. A comparison of random forests, boosting and support vector machines for genomic selection. In: *BMC Proceedings*. Vol 5. ; 2011. doi:10.1186/1753-6561-5-S3-S11

41. Chawla N V, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-357. doi:10.1613/jair.953

42. Brodersen KH, Ong CS, Stephan KE, Buhmann JM. The balanced accuracy and its posterior distribution. In: *Proceedings - International Conference on Pattern Recognition*. ; 2010:3121-3124. doi:10.1109/ICPR.2010.764

43. Pedregosa F, Weiss R, Brucher M, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res*. 2011;12:2825-2830. http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html%5Cnhttp://arxiv.org/abs/1201.0490

44. Van Rossum G, Drake FL, Harris CR, et al. *Python 3 Reference Manual*. Vol 585.; 2009.

45. Chen XW. Gene selection for cancer classification using bootstrapped genetic algorithms and support vector machines. In: *Proceedings of the 2003 IEEE Bioinformatics*

Conference, CSB 2003. Vol 46. Springer; 2003:504-505. doi:10.1109/CSB.2003.1227389

46. Gaudel R, Sebag M. Feature selection as a one-player game. In: *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*. ; 2010:359-366.

47. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523. doi:10.1038/s41467-019-09234-6

48. Liu Q, Zhu D, Li N, et al. Regulation of LRRK2 mRNA stability by ATIC and its substrate AICAR through ARE -mediated mRNA decay in Parkinson's disease. *EMBO J*. 2023;42(15). doi:10.15252/embj.2022113410

49. Kalogeropulou AF, Freemantle JB, Lis P, Vides EG, Polinski NK, Alessi DR. Endogenous Rab29 does not impact basal or stimulated LRRK2 pathway activity. *Biochem J*. 2020;477(22):4397-4423. doi:10.1042/BCJ20200458

50. Kluss JH, Mazza MC, Li Y, et al. Preclinical modeling of chronic inhibition of the Parkinson's disease associated kinase LRRK2 reveals altered function of the endolysosomal system in vivo. *Mol Neurodegener*. 2021;16(1):17. doi:10.1186/s13024-021-00441-8

51. Ghelman J, Grewing L, Windener F, Albrecht S, Zarbock A, Kuhlmann T. SKAP2 as a new regulator of oligodendroglial migration and myelin sheath formation. *Glia*. 2021;69(11):2699-2716. doi:10.1002/glia.24066

52. Takahashi T, Yamashita H, Nagano Y, et al. Identification and Characterisation of a Novel Pyk2/Related Adhesion Focal Tyrosine Kinase-associated Protein That Inhibits α-Synuclein Phosphorylation. *J Biol Chem*. 2003;278(43):42225-42233. doi:10.1074/jbc.M213217200

53. Ayoubi R, Ryan J, Biddle MS, et al. Scaling of an antibody validation procedure enables quantification of antibody performance in major research applications. *Elife*. 2023;12:RP91645. doi:10.1101/2023.06.01.543292

54. Suzuki K, Dashzeveg N, Lu ZG, Taira N, Miki Y, Yoshida K. Programmed cell death 6, a novel p53-responsive gene, targets to the nucleus in the apoptotic response to DNA damage. *Cancer Sci*. 2012;103(10):1788-1794. doi:10.1111/j.1349-7006.2012.02362.x

55. Erekat NS. Apoptosis and its Role in Parkinson's Disease. In: *Parkinson's Disease: Pathogenesis and Clinical Aspects*. Codon Publications; 2018:65-82. doi:10.15586/codonpublications.parkinsonsdisease.2018.ch4

56. Tolosa E, Garrido A, Scholz SW, Poewe W. Challenges in the diagnosis of Parkinson's disease. *Lancet Neurol*. 2021;20(5):385-397. doi:10.1016/S1474-4422(21)00030-2

57. Ao X, Zou L, Wu Y. Regulation of autophagy by the Rab GTPase network. *Cell Death Differ*. 2014;21(3):348-358. doi:10.1038/cdd.2013.187

58. Hirota Y, Yamashita S ichi, Kurihara Y, et al. Mitophagy is primarily due to alternative autophagy and requires the MAPK1 and MAPK14 signaling pathways. *Autophagy*. 2015;11(2):332-343. doi:10.1080/15548627.2015.1023047

59. Hirschberg Y, Valle-Tamayo N, Dols-Icardo O, et al. Proteomic comparison between non-purified cerebrospinal fluid and cerebrospinal fluid-derived extracellular vesicles from patients with Alzheimer's, Parkinson's and Lewy body dementia. *J Extracell Vesicles*. 2023;12(12):12383. doi:10.1002/jev2.12383

60. Vidyadhara DJ, Lee JE, Chandra SS. Role of the endolysosomal system in Parkinson's disease. *J Neurochem*. 2019;150(5):487-506. doi:10.1111/jnc.14820

61. Erb ML, Moore DJ. LRRK2 and the Endolysosomal System in Parkinson's Disease. *J Parkinsons Dis*. 2020;10(4):1271-1291. doi:10.3233/JPD-202138

62. Bulock KG, Peter Beardsley G, Anderson KS. The kinetic mechanism of the human bifunctional enzyme ATIC (5-amino-4-imidazolecarboxamide ribonucleotide transformylase/inosine 5′-monophosphate cyclohydrolase). A surprising lack of substrate channeling. *J Biol Chem*. 2002;277(25):22168-22174. doi:10.1074/jbc.M111964200

63. Vergis JM, Beardsley GP. Catalytic Mechanism of the Cyclohydrolase Activity of Human Aminoimidazole Carboxamide Ribonucleotide Formyltransferase/Inosine Monophosphate Cyclohydrolase. *Biochemistry*. 2004;43(5):1184-1192. doi:10.1021/bi035139b

64. Byrns CN, Pitts MW, Gilman CA, Hashimoto AC, Berry MJ. Mice lacking selenoprotein P and selenocysteine lyase exhibit severe neurological dysfunction, neurodegeneration, and audiogenic seizures. *J Biol Chem*. 2014;289(14):9662-9674. doi:10.1074/jbc.M113.540682

65. Flinkman D, Hong Y, Gnjatovic J, et al. Regulators of proteostasis are translationally repressed in fibroblasts from patients with sporadic and LRRK2-G2019S Parkinson's disease. *npj Park Dis*. 2023;9(1):20. doi:10.1038/s41531-023-00460-w

66. Jang Y, Pletnikova O, Troncoso JC, et al. Mass Spectrometry–Based Proteomics Analysis of Human Substantia Nigra From Parkinson's Disease Patients Identifies Multiple Pathways Potentially Involved in the Disease. *Mol Cell Proteomics*. 2023;22(1):100452. doi:10.1016/j.mcpro.2022.100452

67. Grünewald A, Rygiel KA, Hepplewhite PD, Morris CM, Picard M, Turnbull DM. Mitochondrial DNA Depletion in Respiratory Chain-Deficient Parkinson Disease Neurons. *Ann Neurol*. 2016;79(3):366-378. doi:10.1002/ana.24571

68. Marchand A, Drouyer M, Sarchione A, Chartier-Harlin MC, Taymans JM. LRRK2 Phosphorylation, More Than an Epiphenomenon. *Front Neurosci*. 2020;14:527. doi:10.3389/fnins.2020.00527

69. Harvey K, Outeiro TF. The role of LRRK2 in cell signalling. *Biochem Soc Trans*. 2018;47(1):197-207. doi:10.1042/BST20180464

70. Rocha EM, Keeney MT, Di Maio R, De Miranda BR, Greenamyre JT. LRRK2 and idiopathic Parkinson's disease. *Trends Neurosci*. 2022;45(3):224-236. doi:10.1016/j.tins.2021.12.002

71. Graham JM, Sagar HJ. A data-driven approach to the study of heterogeneity in idiopathic Parkinson's disease: Identification of three distinct subtypes. *Mov Disord*. 1999;14(1):10-20. doi:10.1002/1531-8257(199901)14:1<10::AID-MDS1005>3.0.CO;2-4

72. Lewis SJG, Foltynie T, Blackwell AD, Bobbins TW, Owen AM, Barker RA. Heterogeneity of Parkinson's disease in the early clinical stages using a data driven approach. *J Neurol Neurosurg Psychiatry*. 2005;76(3):343-348. doi:10.1136/jnnp.2003.033530

73. Bae E-J, Lee S-J. The LRRK2-RAB axis in regulation of vesicle trafficking and α-synuclein propagation. *Biochim Biophys Acta - Mol Basis Dis*. 2020;1866(3):165632. doi:10.1016/j.bbadis.2019.165632

74. Matsui T, Fukuda M. Rab12 regulates mTORC1 activity and autophagy through

controlling the degradation of amino-acid transporter PAT4. *EMBO Rep*. 2013;14(5):450-457. doi:10.1038/embor.2013.32

75. Xenias HS, Chen C, Kang S, et al. R1441C and G2019S LRRK2 knockin mice have distinct striatal molecular, physiological, and behavioral alterations. *Commun Biol*. 2022;5(1):1-14. doi:10.1038/s42003-022-04136-8

76. Berndsen K, Lis P, Yeshaw WM, et al. PPM1H phosphatase counteracts LRRK2 signaling by selectively dephosphorylating rab proteins. *Elife*. 2019;8:e50416. doi:10.7554/eLife.50416

77. Liu Z, Bryant N, Kumaran R, et al. LRRK2 phosphorylates membrane-bound Rabs and is activated by GTP-bound Rab7L1 to promote recruitment to the trans-Golgi network. *Hum Mol Genet*. 2018;27(2):385-395. doi:10.1093/hmg/ddx410

78. Lesniak RK, Nichols RJ, Montine TJ. Development of mutation-selective LRRK2 kinase inhibitors as precision medicine for Parkinson's disease and other diseases for which carriers are at increased risk. *Front Neurol*. 2022;13:1016040. doi:10.3389/fneur.2022.1016040

79. Leśniak RK, Nichols RJ, Schonemann M, et al. Discovery of azaspirocyclic 1H-3,4,5-Trisubstitued pyrazoles as novel G2019S-LRRK2 selective kinase inhibitors. *Eur J Med Chem*. 2022;242:114693. doi:10.1016/j.ejmech.2022.114693

80. Lesńiak RK, Nichols RJ, Schonemann M, et al. Discovery of 1H-Pyrazole Biaryl Sulfonamides as Novel G2019SLRRK2 Kinase Inhibitors. *ACS Med Chem Lett*. 2022;13(6):981-988. doi:10.1021/acsmedchemlett.2c00116

81. Leśniak RK, Nichols RJ, Schonemann M, et al. Discovery of G2019S-Selective Leucine Rich Repeat Protein Kinase 2 inhibitors with in vivo efficacy. *Eur J Med Chem*. 2022;229:114080. doi:10.1016/j.ejmech.2021.114080

82. Li X, Zhu H, Huang BT, et al. RAB12-LRRK2 Complex Suppresses Primary Ciliogenesis and Regulates Centrosome Homeostasis in Astrocytes. *Nat Commun*. 2024;15(1):8434. doi:10.1038/s41467-024-52723-6

83. Garrido A, Santamaría E, Fernández-Irigoyen J, et al. Differential Phospho-Signatures in Blood Cells Identify LRRK2 G2019S Carriers in Parkinson's Disease. *Mov Disord*.

2022;37(5):1004-1015. doi:10.1002/mds.28927

84. Virreira Winter S, Karayel O, Strauss M, et al. Urinary proteome profiling for stratifying patients with familial Parkinson's disease. *EMBO Mol Med*. 2021;13(3):e13257. doi:10.15252/EMMM.202013257

85. Hadisurya M, Li L, Kuwaranancharoen K, et al. Quantitative proteomics and phosphoproteomics of urinary extracellular vesicles define putative diagnostic biosignatures for Parkinson's disease. *Commun Med*. 2023;3(1). doi:10.1038/s43856-023-00294-w

86. Kluss JH, Beilina A, Williamson CD, Lewis PA, Cookson MR, Bonet-Ponce L. Lysosomal positioning regulates Rab10 phosphorylation at LRRK2+ lysosomes. *Proc Natl Acad Sci U S A*. 2022;119(43):e2205492119. doi:10.1073/pnas.2205492119

87. Perez-Riverol Y, Bai J, Bandla C, et al. The PRIDE database resources in 2022: A hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res*. 2022;50(D1):D543-D552. doi:10.1093/nar/gkab1038

Figure legends

Figure 1 Experimental workflow using PBMCs from a Spanish LRRK2 clinical cohort. (**A**) Peripheral blood mononuclear cells (PBMCs) processing for different applications. 40 ml of blood were drawn from subjects of a LRRK2 clinical cohort from Spain (n=174) encompassing G2019S L2PD patients (n=37), G2019S L2NMCs (n=27), R1441G L2PD patients (n=14), R1441G L2NMCs (n=11), iPD (n=40), and controls (n=45). (**B**) After PBMCs isolation, homogenisation, and protein digestion, DIA-MS identified a total of 3,815 proteins on an EZ-Exploris 480 mass-spectrometer (Thermo), and 10,288 phospho-sites after phospho-enrichment. For the group differential analysis, we only considered proteins and phospho-sites mapped by ≥2 different peptides (Spetronaut), and with <30% imputation, with a significance cut-off of $log_2FC>|0.6|$ and an FDR multiple-testing adjusted P<0.05. Data deconvolution and interactive representation of findings were done using the Curtain / Curtain PTM Tool, and gene ontology was assessed by Metascape. Using machine learning, we identified an 18-feature G2019S phospho-/protein signature able to discriminate G2019S L2PD, G2019S L2NMCs, and controls. By immunoblot, we assessed pSer106 RAB12 / total RAB12 levels in PBMCs from a subject of subjects (n=48) after 1 year of follow-up, including G2019S L2PD (n=12), G2019S L2NMCs (n=6), iPD (n=15) and controls (n=15). Lastly, in freshly isolated PBMCs from a second subset of subjects (n=10) encompassing G2019S L2PD (n=3), R1441G L2PD (n=1), iPD (n=1) and healthy controls (n=5), treated with DMSO or the MLi-2 LRRK2 inhibitor, we performed an LRRK2 kinase assay measuring pSer106 RAB12 / total RAB12 levels.

Figure 2 Proteome overview and differential analyses in G2019S carriers. (**A**) Barplots showing the numbers of differential proteins in different pairwise comparisons involving G2019S carriers, R1441G carriers, iPD, and controls, with up-regulated proteins in dark grey, and down-regulated in light grey. All cohorts were run in parallel, with balanced study groups per run, blind to the operator, and using 1 quantile normalisation (Limma). The significance cut-off was set at a $log_2FC>|0.6|$ and an FDR multiple-testing adjusted P<0.05. (**B**) Volcano plot of the proteome differential analysis in G2019S L2PD vs healthy controls, with Curtain weblinks to access raw and differential analysis data, showing proteins up-regulated in G2019S L2PD as red dots on the right,

and proteins up-regulated in controls (i.e., down-regulated in G2019S L2PD) as red dots on the left (Curtain). A legend colour code applying to all panels is shown at the bottom of the Figure, depicting statistically significant hits as red dots. (**C**) Volcano plot of the proteome differential analysis in G2019S carriers as a whole, i.e., L2PD and L2NMCs, vs healthy controls (Curtain). (**D**) Volcano plot showing the proteome differential analysis between G2019S L2NMCs and healthy controls (Curtain). (**E**) Volcano plot representing the proteome comparison between G2019S L2NMCs and G2019S L2PD. A Venn diagram at the bottom of the Figure shows the overlap of differential hits in PD-manifesting and non-manifesting G2019S carriers (Curtain). Curtain weblinks provide access to the differential analyses.

Figure 3 Phospho-proteome differential analyses of G2019S carriers. (**A**) Volcano plot of the phospho-proteome differential analysis of G2019S L2PD vs controls, and Curtain weblinks to raw and differential analysis data, representing hyper-phosphorylated proteins in G2019S L2PD as red dots on the right with a single hit, elevated pSer106 RAB12 levels in G2019S L2PD, emerging as a differential phospho-peptide at a $\log_2$FC>|0.6| and an FDR multiple-testing adjusted P<0.05 (Curtain PTM). A legend colour code applying to all the panels shows hits categorisation by statistical significance. (**B**) Volcano plot showing phospho-protein hits in G2019S carriers as a whole, PD-manifesting and non-manifesting, compared to controls (Curtain PTM). (**C**) Phospho-proteome differences in G2019S L2NMCs vs controls (Curtain PTM). (**D**) Volcano showing phospho-proteome differences in G2019S L2NMCs vs G2019S L2PD (Curtain PTM). (**E**) QC crude non-imputed (lower bar plot), non-normalised (upper violin plot) mass-spectrometry data from pSer106 RAB12 levels across all study groups showing higher pSer106 phosphorylation levels in G2019S L2PD and G2019S L2NMCs respect to the rest of the groups. The adj. P-values and FC on top of the violin plot correspond to those from the differential analysis. (**F**) A similar analysis to the previous panel with G2019S L2PD and G2019S L2NMCs grouped into a single group of G2019S carriers. Curtain weblinks provide access to the differential analyses.

Figure 4 One-year follow-up of pSer106 RAB12 by immunoblot and MLi-2 response. Immunoblot assessment of pSer106 RAB12 phosphorylation levels in >1-year follow-up PBMC samples from part of the LRRK2 subcohort from Clínic-Barcelona (n=48), including G2019S

L2PD (n=12), G2019S L2NMCs (n=6), iPD (n=15), and controls (n=15). (**A**) Schematic workflow of immunoblot assessment and representative blot from 5 different blots shown in the Supplement. (*) Denotes intergel control. (**B**) dot plots comparing pSer106 RAB12 / Total RAB12 levels obtained by DIA-MS at the entire LRRK2 clinical cohort (n=174) on the left and by immunoblot of part of the Clínic-Barcelona cohort after 1-year of follow-up (n=48) in G2019S carriers on the right. In each plot, overall intergroup differences were assessed using the Kruskal-Wallis test followed by post-hoc Dunn's test to evaluate for pSer106 RAB12 / Total RAB12 differences in G2019S carriers. (**C**) Representative immunoblot analysis of pSer106 RAB12 / Total RAB12 and pThr73 RAB10 / Total RAB10 using technical replicates from additional freshly collected PBMCs from one R1441G L2PD, one G2019S L2PD, one iPD, and 3 controls (expanded to a total n=10 subjects in the Supplement), treated with DMSO or the MLi-2 LRRK2 inhibitor (200 nM, 30 min), showing a diminishment of pSer106 RAB12 phosphorylation levels after LRRK2 inhibition by MLi-2 treatment.

Figure 5 Identification of an 18-feature phospho-/protein classifier for G2019S carriers. After comparing the performance of several models, we applied supported vector machine (SVM) learning, adjusted by unbalanced groups using the Synthetic Minority Over-sampling Technique (SMOTE), corrected from overfitting with 5-fold cross-validation, identified cross-group differential proteins and phospho-proteins by ANOVA and Recursive Feature Elimination with Cross-Validation (RFECV), and refined informative combinations to the minimal numbers of features yielding the maximal balanced accuracy by the Montecarlo Tree Search (MCTS) method. (**A**) 18-feature G2019S phospho-/protein best classifier identified in G2019S carriers, PD-manifesting and non-manifesting subjects, and healthy controls. Red dots indicate individual features correlating with disease severity (UPDRS-III) (See next Figure). (**B**) Relative contribution of the different proteins (n=15) and phospho-sites (n=3), including pSer106 RAB12, from the 18-feature G2019S classifier on the upper bar plot; Metascape gene ontology enrichment analysis of the 18-features G2019S signature lower bar plot. (**C**) Receiver Operating Curve (ROC) analysis of the 18-feature G2019S phospho-/protein signature showing an overall balanced accuracy of 0.957 to discriminate G2019S L2PD, G2019S L2NMCs, and controls, specifically with an area under the curve (AUC) of 1.00 between G2019S L2PD and controls, 0.99 between G2019S L2NMCs and controls, and 0.98 between G2019S L2PD and G2019S L2NMCs. (**D**) Principal

component analysis (PCA) based on the 18-feature G2019S phospho-/protein classifier in G2019S carriers and healthy controls showing distinct group profiles based on LRRK2 mutation and disease status, with G2019S L2NMCs in between G2019S L2PD and controls, consistent with their disease status.

Figure 6 Association between differential LRRK2 phospho-/proteins and disease severity. Correlation analysis of differential proteins and phospho-proteins (log2FC>|0.6|, adj. P<0.05) and UPDRS-III motor scores from L2PD patients and healthy controls with statistical significance set at a Spearman's correlation coefficient Rho>|0.5| and an FDR multiple-testing adj. P<0.05. (**A**) Correlation plots between differential proteins in G2019S L2PD vs controls on the left and R1441G L2PD vs controls on the right, showing differential hits correlating with UPDRS-III in red. (**B**) Scatter plot of 10 hits from the 18-feature G2019S phospho-/protein signature correlating with UPDRS-III in G2019S L2PD patients represented as orange dots and healthy controls as blue dots, including PDCD6, ARHGAP45, ATIC, SCLY, PSMC5, NDUFB8, LAMP1, HSD17B10, RAB9A, and pSer106 RAB12.

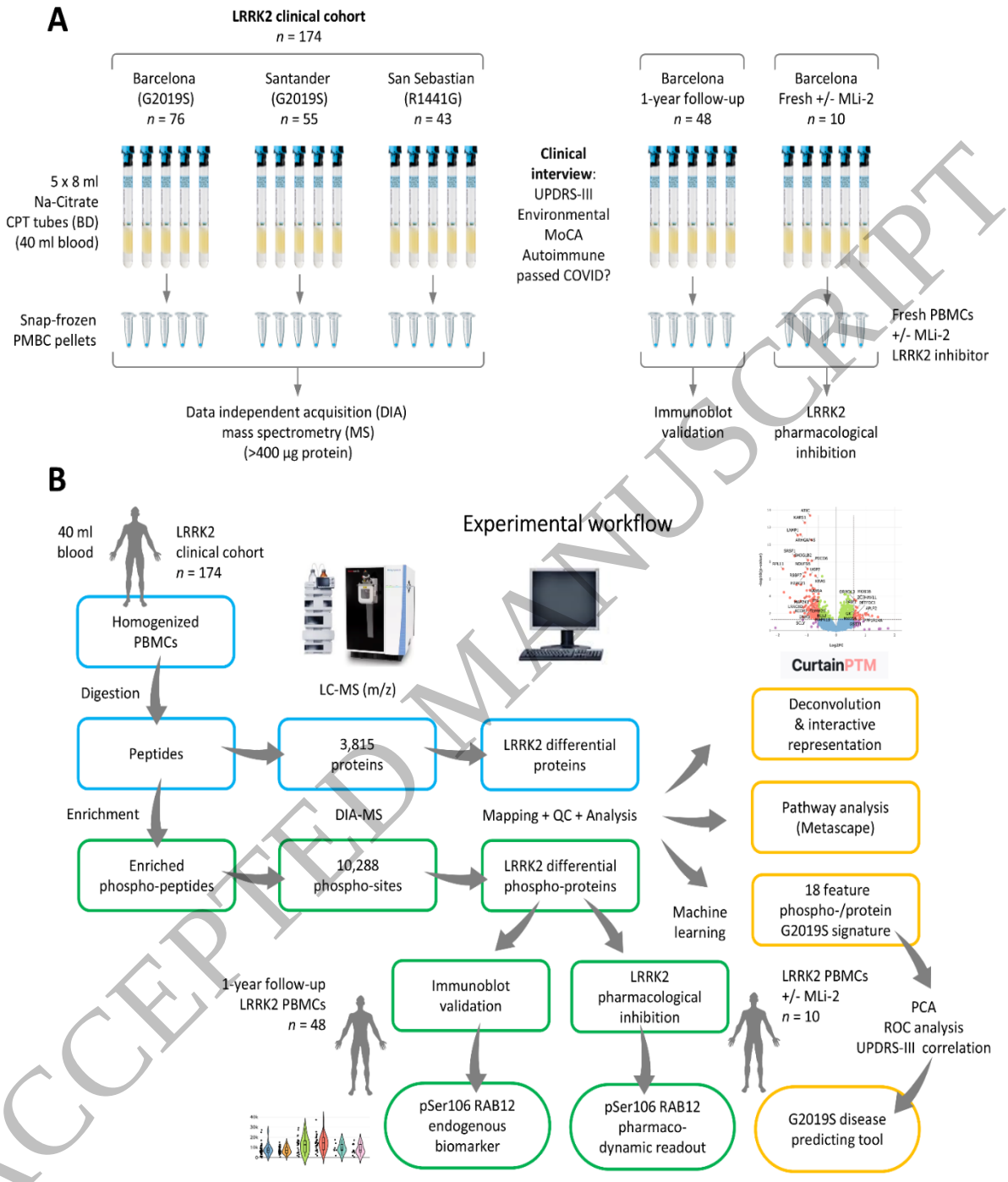

Figure 1
183x201 mm (x DPI)

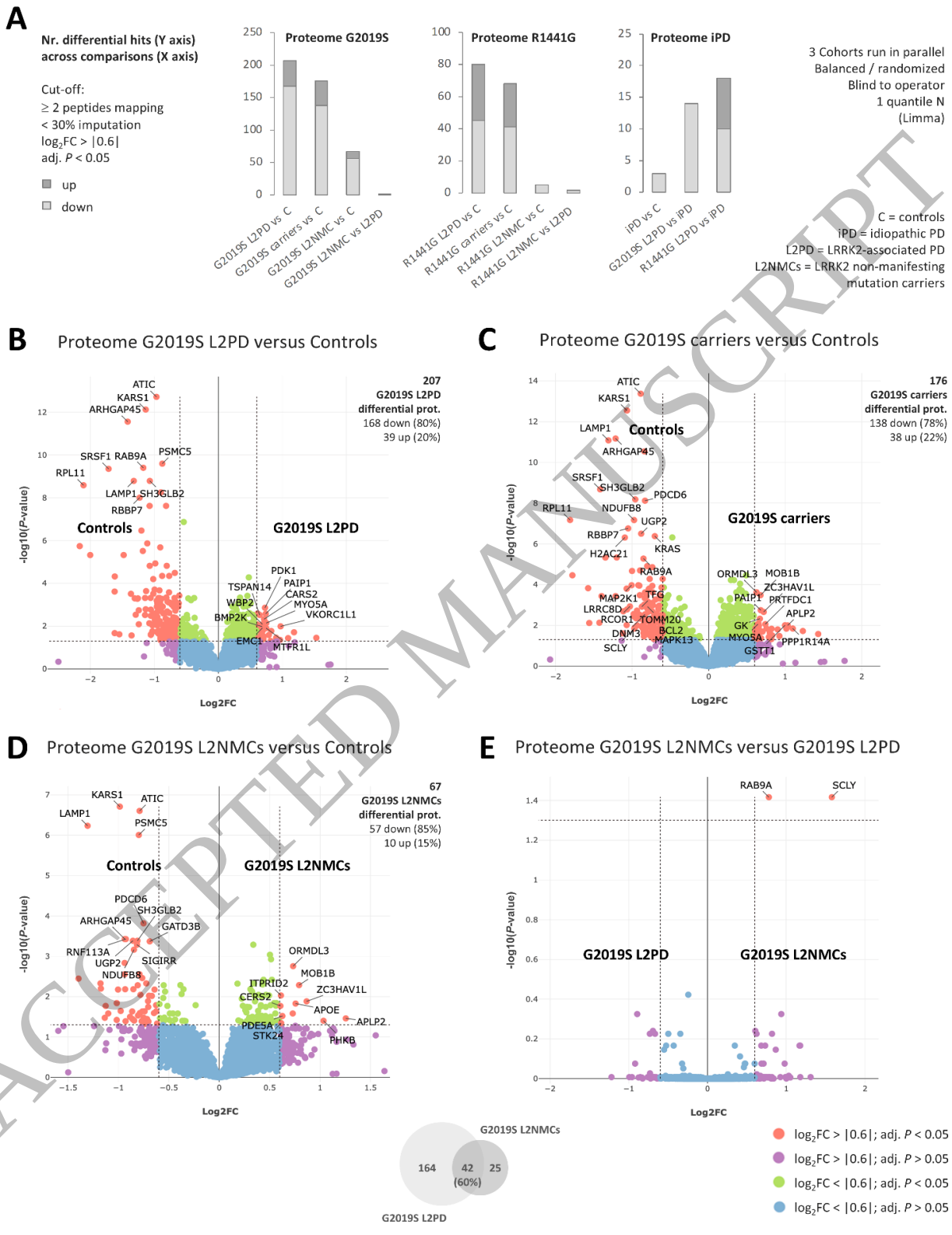

Figure 2
175x225 mm (x DPI)

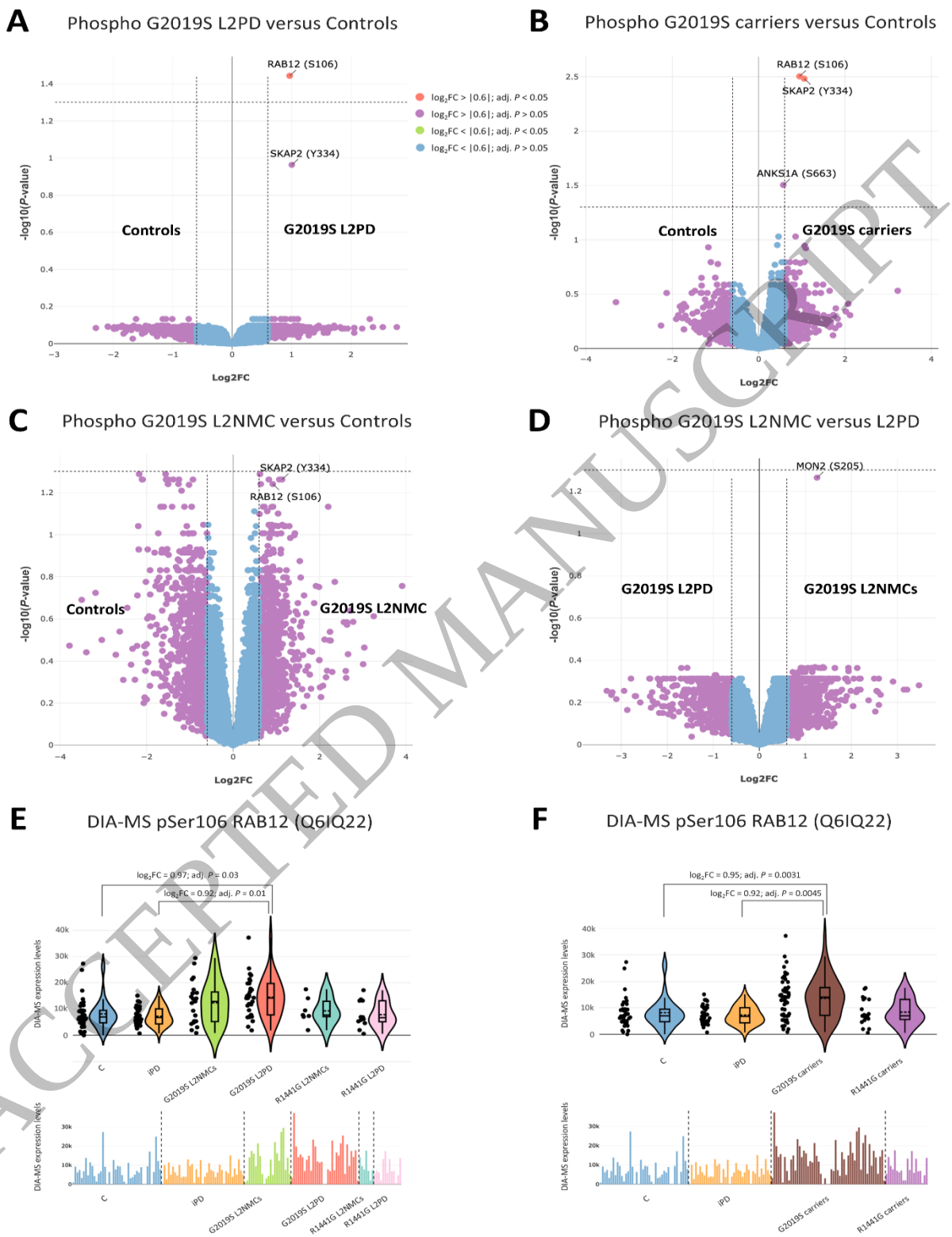

Figure 3
162x229 mm (x DPI)

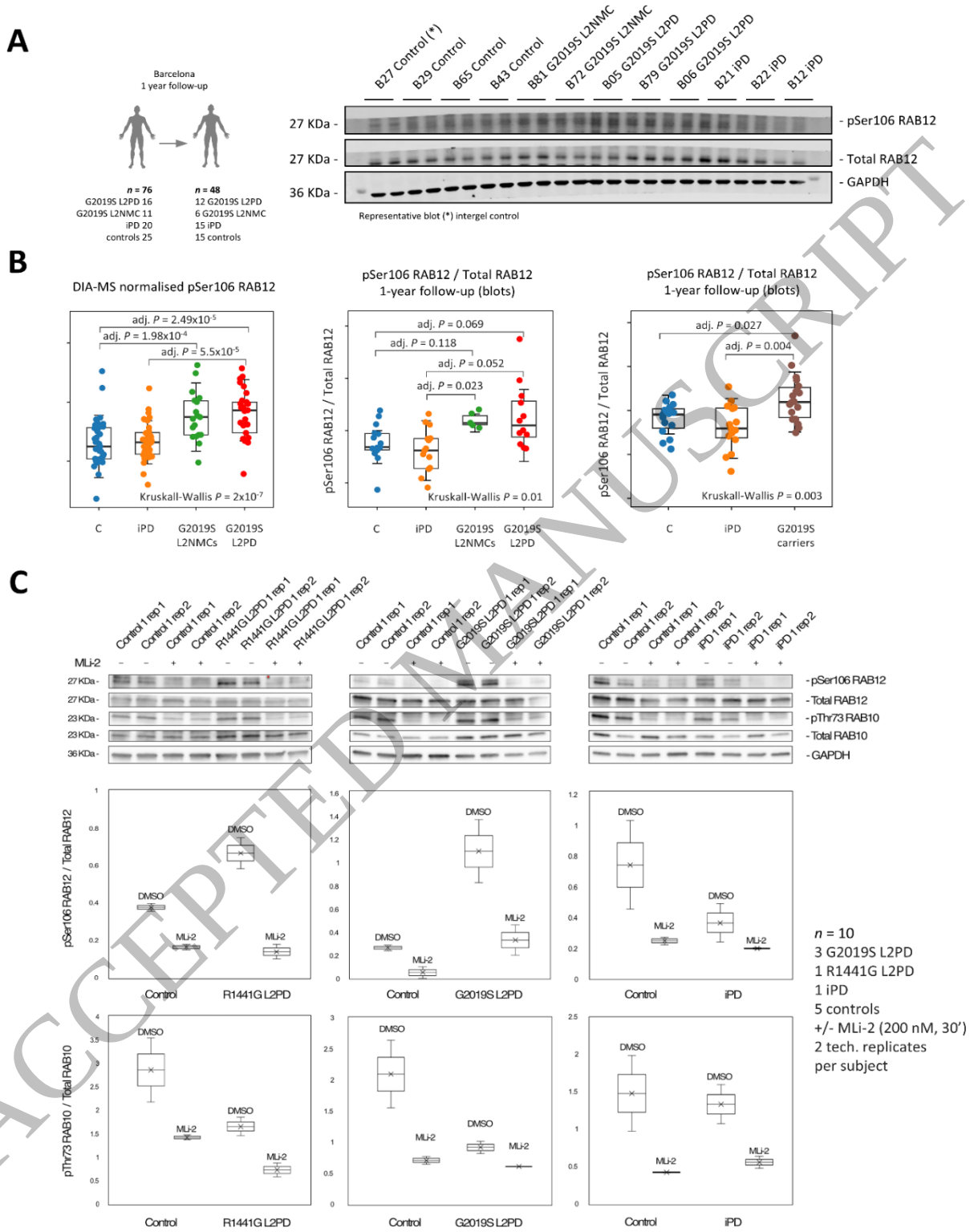

Figure 4
171x225 mm (x DPI)

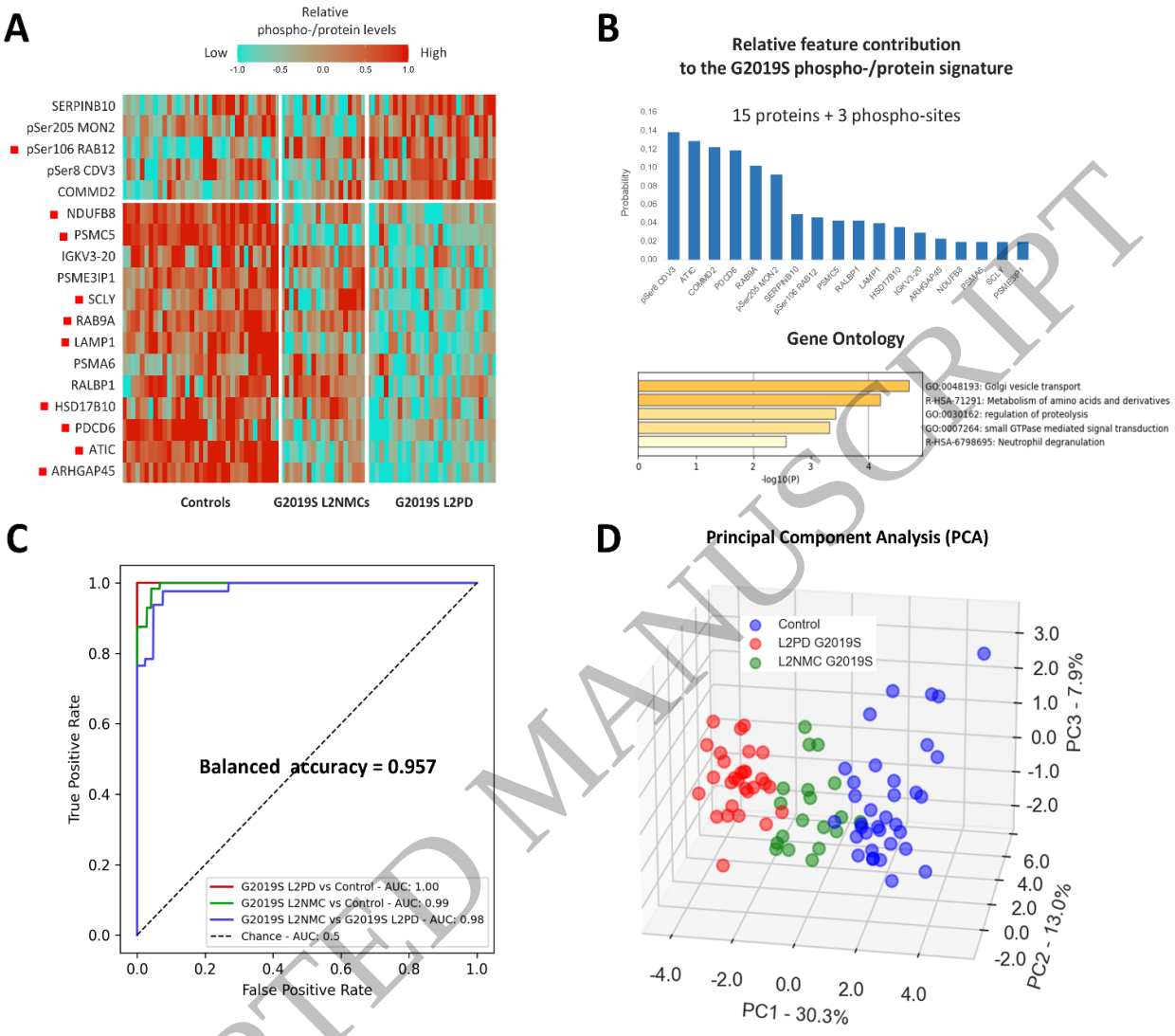

Figure 5
179x151 mm (x DPI)

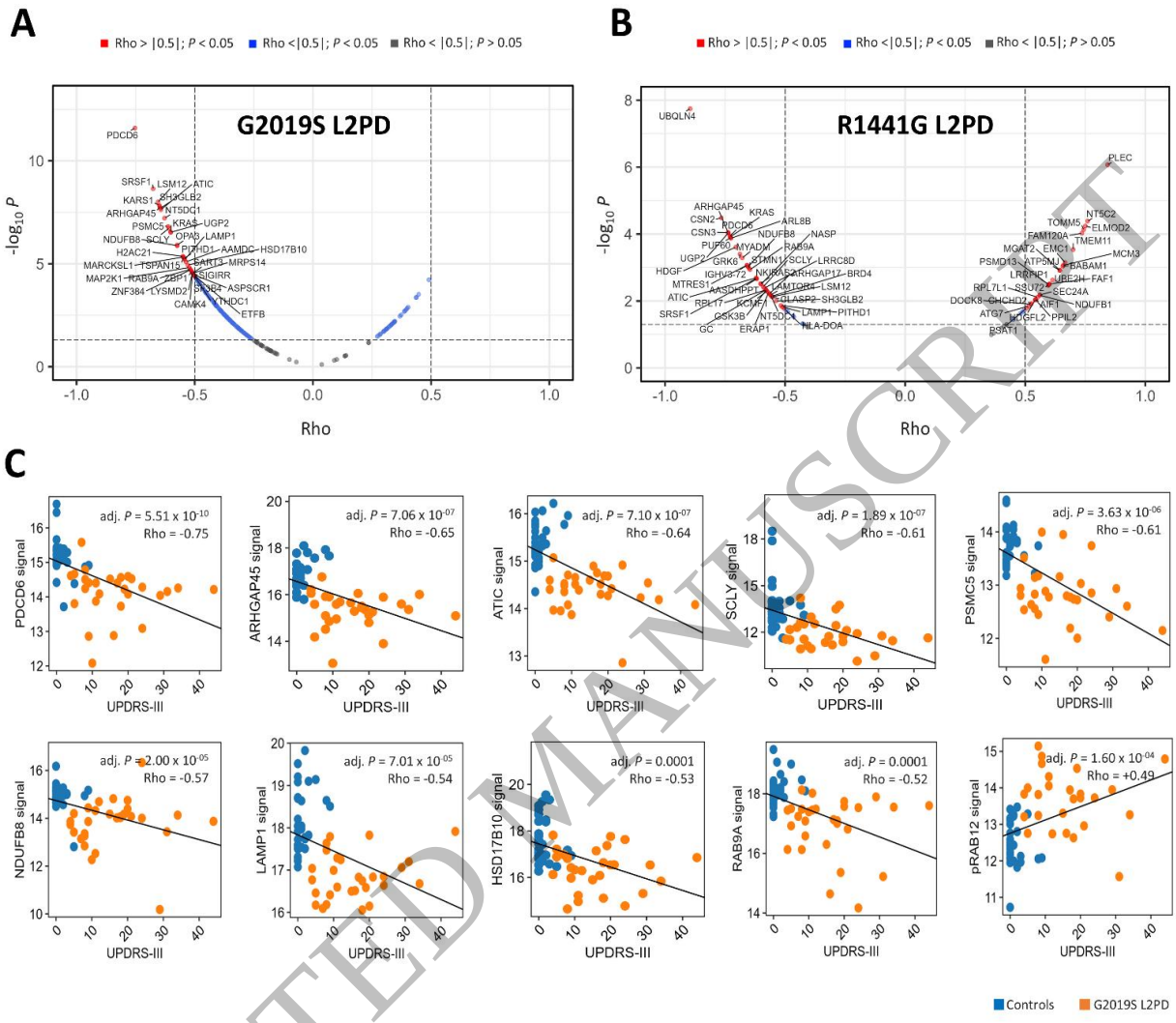

Figure 6
184x155 mm (x DPI)

Table 1 Participant clinic-demographics

Cohort	N (males/females)	Age at sampling (years)	PD AAO (years)	Disease duration (years)	UPDRS-III	H&Y	MoCA	LEDD (mg)	Passed COVID-19 (yes/no)
Entire cohort	**174**								
G2019S L2PD	37 (20/17)	63.5 ± 9.1 (37/37)	55.1 ± 10.2 (33/37)	8.4 ± 6.3 (33/37)	16.0 ± 9.7 (34/37)	2.0 ± 0.6 (21/37)	24.3 ± 4.5 (35/37)	635.8 ± 438.8 (29/39)	4/30 (34/37)
G2019S L2NMC	27 (18/9)	56.7 ± 14.1 (26/27)	–	–	1.0 ± 1.6 (22/27)	–	25.4 ± 6.6 (25/27)	–	6/19 (25/27)
R1441G L2PD	14 (7/7)	67.1 ± 9 5 (14/14)	55.8 ± 11.4 (14/14)	12.3 ± 5.5 (14/14)	19.8 ± 12.0 (14/14)	2.2 ± 0.9 (13/14)	23.2 ± 5.5 (10/14)	711.5 ± 355.7 (14/14)	4/10 (14/14)
R1441G L2NMC	11 (4/7)	61.1 ± 55 (11/11)	–	–	1.2 ± 2.1 (11/11)	–	28.6 ± 2.0	–	1/10 (11/11)

							(11/11)		
iPD	40 (30/10)	67.3 ± 7.7 (40/40)	62.1 ± 8.5 (40/40)	5.2 ± 4.3 (40/40)	19.7 ± 13.2 (40/40)	2.2 ± 0.6 (31/40)	25.6 ± 3.7 (33/40)	584.7 ± 373.6 (37/40)	1/37 (38/40)
C	45 (18/27)	60.0 ± 10.9 (45/45)	–	–	1.2 ± 2.2 (17/27)	–	27.5 ± 3.2 (27/27)	–	14/30 (44/45)
B - Barcelona	**76**								
G2019S L2PD	16 (7/9)	65.5 ± 8.3 (16/16)	53.5 ± 11.3 (14/16)	11.4 ± 7.1 (14/16)	13.0 ± 7.2 (13/16)	2.0 ± 0.5 (11/16)	25.0 ± 4.3 (14/16)	596.5 ± 269.7 (13/18)	2/12 (14/16)
G2019S L2NMC	11 (7/4)	47.8 ± 15.5 (10/11)	–	–	0.3 ± 0.7 (10/11)	–	28.2 ± 2.0 (10/11)	–	4/5 (9/11)
R1441G L2PD	1 (0/1)	44.0 (1/1)	32.0 (1/1)	12.0 (1/1)	16.0 (1/1)	NA (0/1)	30.0 (1/1)	400 (1/1)	0/1 (1/1)
R1441G L2NMC	3 (2/1)	65.3 ± 15 (3/3)	–	–	3.0 ± 3.6 (3/3)	–	29.3 ± 0.6 (3/3)	–	1/2 (3/3)
iPD	20 (16/4)	68.3 ± 7.9 (20/20)	64.3 ± 7.8 (20/20)	4.0 ± 3.0 (20/20)	14.7 ± 4.2 (20/20)	1.9 ± 0.2 (20/20)	27.2 ± 3.2 (20/20)	453.8 ± 279.8 (19/20)	1/19 (20/20)
C	25 (9/16)	63.9 ± 10.6 (25/25)	–	–	1.7 ± 2.4 (16/25)	–	27.9 ± 2.1 (25/25)	–	8/17 (25/25)
S - Santander	**55**								
G2019S L2PD	20 (13/0)	61.3 ± 9.0 (20/20)	55.6 ± 9.1 (18/20)	6.1 ± 4.7 (18/20)	18.1 ± 11.0 (20/20)	1.9 ± 0.6 (9/20)	24.2 ± 4.6 (20/20)	652.7 ± 562.3 (15/20)	2/17 (19/20)
G2019S L2NMC	15 (11/4)	62.6 ± 10.2 (15/15)	–	–	1.8 ± 1.9 (11/15)	–	23.3 ± 8.1 (14/15)	–	2/13 (15/15)
iPD	10 (6/4)	67.2 ± 7.6 (10/10)	62.0 ± 6.3 (10/10)	5.2 ± 4.0 (10/10)	17.9 ± 8.3 (10/10)	2.5 (1/10)	23 ± 3.5 (10/10)	479.9 ± 220.2 (8/10)	0/8 (8/10)
C	10 (4/6)	56.9 ± 11.1 (10/10)	–	–	NA	–	25.0 ± 3.6 (6/6)	–	1/8 (9/10)
D - Donostia	**43**								
G2019S L2PD	1 (0/1)	78.0 (1/1)	68.0 (1/1)	10.0 (1/1)	15.0 (1/1)	3.0 (1/1)	19.0 (1/1)	893 (1/1)	0/1 (1/1)
G2019S L2NMC	1 (0/1)	58.0 (1/1)	–	–	NA	–	29 (1/1)	–	0/1 (1/1)
R1441G L2PD	13 (7/6)	68.9 ± 7.1 (13/13)	56.5 ± 9.7 (13/13)	12.4 ± 5.7 (13/13)	20.1 ± 12.5 (13/13)	2.2 ± 0.9 (13/13)	22.4 ± 5.2 (9/13)	735.5 ± 358.3 (13/13)	4/9 (13/13)
R1441G L2NMC	8 (2/6)	59.5 ± 5.6 (8/8)	–	–	0.5 ± 0.8 (6/6)	–	28.3 ± 2.3 (6/6)	–	0/8 (8/8)
iPD	10 (8/2)	65.4 ± 7.8 (10/10)	57.9 ± 10.6 (10/10)	7.5 ± 5.9 (10/10)	31.7 ± 20.7 (10/10)	2.7 ± 0.7 (10/10)	23.7±2.9 (10/10)	917.3 ± 441.8 (10/10)	0/10 (10/10)
C	10 (5/5)	53.5 ± 8.1 (10/10)	–	–	NA	–	29.1 ± 3.9 (5/5)	–	5/5 (10/10)

Data are expressed as a mean ± standard deviation (S.D.) with the number of available subjects/totals in brackets. L2PD = LRRK2-associated PD patients; L2NMC = LRRK2 non-manifesting carriers; iPD = idiopathic PD; C = controls; AAO = age-at-onset; UPDRS-III = Unified Parkinson's Disease Rating Scale; MoCA = Montreal cognitive assessment; LEDD = levodopa equivalent daily dose; "–" = not applicable; NA = not available.