

## Article

# Plasma Metabolite Signature Classifies Male LRRK2 Parkinson's Disease Patients

Chen Dong <sup>1</sup>, Chandrashekhar Honrao <sup>1</sup>, Leonardo O. Rodrigues <sup>1</sup>, Josephine Wolf <sup>1</sup>, Keri B. Sheehan <sup>1</sup>, Matthew Surface <sup>2</sup>, Roy N. Alcalay <sup>2,3</sup> and Elizabeth M. O'Day <sup>1,\*</sup>

<sup>1</sup> Olaris, Inc., Framingham, MA 01702, USA; cdong@olarisbor.com (C.D.); chonrao@olarisbor.com (C.H.); lrodrigues@olarisbor.com (L.O.R.); jwolf@olarisbor.com (J.W.); ksanborn@gmail.com (K.B.S.)

<sup>2</sup> Department of Neurology, Columbia University Irving Medical Center, New York, NY 10032, USA; surfacematt12@gmail.com (M.S.); rna2104@cumc.columbia.edu (R.N.A.)

<sup>3</sup> Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel

\* Correspondence: eoday@olarisbor.com

**Abstract:** Parkinson's disease (PD) is a progressive neurodegenerative disease, causing loss of motor and nonmotor function. Diagnosis is based on clinical symptoms that do not develop until late in the disease progression, at which point the majority of the patients' dopaminergic neurons are already destroyed. While many PD cases are idiopathic, hereditary genetic risks have been identified, including mutations in the gene for LRRK2, a multidomain kinase with roles in autophagy, mitochondrial function, transcription, molecular structural integrity, the endo-lysosomal system, and the immune response. A definitive PD diagnosis can only be made post-mortem, and no non-invasive or blood-based disease biomarkers are currently available. Alterations in metabolites have been identified in PD patients, suggesting that metabolomics may hold promise for PD diagnostic tools. In this study, we sought to identify metabolic markers of PD in plasma. Using a 1H-13C heteronuclear single quantum coherence spectroscopy (HSQC) NMR spectroscopy metabolomics platform coupled with machine learning (ML), we measured plasma metabolites from approximately age/sex-matched PD patients with G2019S LRRK2 mutations and non-PD controls. Based on the differential level of known and unknown metabolites, we were able to build a ML model and develop a Biomarker of Response (BoR) score, which classified male LRRK2 PD patients with 79.7% accuracy, 81.3% sensitivity, and 78.6% specificity. The high accuracy of the BoR score suggests that the metabolomics/ML workflow described here could be further utilized in the development of a confirmatory diagnostic for PD in larger patient cohorts. A diagnostic assay for PD will aid clinicians and their patients to quickly move toward a definitive diagnosis, and ultimately empower future clinical trials and treatment options.

**Citation:** Dong, C.; Honrao, C.; Rodrigues, L.O.; Sheehan, K.B.; Surface, M.; Alcalay, R.N.; O'Day, E.M. Plasma Metabolite Signature Classifies Male LRRK2 Parkinson's Disease Patients. *Metabolites* **2022**, *12*, 149. <https://doi.org/10.3390/metabo12020149>

Academic Editor: Leonardo Tenori

Received: 23 December 2021

Accepted: 2 February 2022

Published: 5 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Parkinson's disease, biomarker, metabolite, leucine, machine learning

## 1. Introduction

Parkinson's disease (PD) is a chronic, progressive neurodegenerative disease that causes devastating loss of motor and nonmotor function over time. PD is the second-most common neurodegenerative disorder after Alzheimer's dementia, with over 6 million people affected worldwide [1,2]. Currently, the diagnosis is based on clinical assessment of symptoms, and there are no blood-based biomarkers to diagnose the disease.

Postural instability, bradykinesia, rigidity, and resting tremor are the primary clinical indications suggestive of PD. Most patients do not exhibit these symptoms until 5–10 years after the neuropathology develops, at which point it is estimated that 80% of their dopaminergic neurons are lost [3]. Once symptoms are present, PD remains challenging to diagnose due to comorbidities such as cerebrovascular disease, dementia, and other neurological disorders [4]. Although resting tremor is a hallmark clinical feature of PD,

not all patients with pathologically confirmed PD present with it. In clinically uncertain cases, DaTScan imaging, an FDA-approved imaging test, may be used to evaluate dopamine activity in commonly affected areas such as the striatum [5]. A pathologically confirmed PD diagnosis can only be obtained post-mortem by evaluating additional hallmarks such as loss of midbrain dopaminergic (mDA) neurons in the substantia nigra pars compacta (SNpc) and the presence of Lewy bodies containing aggregated alpha-synuclein protein filaments and/or hyperphosphorylated Tau protein [6]. However, post-mortem analyses have revealed that as many as 32% of patients with a pathologically confirmed PD diagnosis were misdiagnosed [7]. At present, therapies for PD, such as dopamine replacement therapy, aim to alleviate motor symptoms, and do not slow or reverse disease progression. Furthermore, patients treated with dopamine replacement therapy such as L-DOPA often develop L-DOPA-induced dyskinesia, an adverse treatment reaction occurring during advanced stages of PD [3,8].

Biomarkers of disease are urgently needed to reduce time to diagnosis and improve diagnostic accuracy, allowing patients to seek treatment while motor skills remain preserved. Further PD biomarkers are needed to enroll the appropriate patient populations in clinical trials and to more accurately assess disease-modifying therapies. As PD is widely recognized as a clinically and pathologically diverse group of disorders rather than a single disease, these markers also have the potential to improve diagnosis of subdisorders, development of precision therapeutics, and identification of underlying diverse disease mechanisms [9,10]. At present, several proposed PD biomarkers have failed to correlate with diagnosis and/or progression, including perturbed dopamine metabolites, decreased levels of alpha-synuclein in the cerebrospinal fluid and elevated levels of oligomeric alpha-synuclein in the plasma [11], gait-cycle irregularities [12], and isotope tracing to the SNpc [13]. Thus, there remains a critical unmet need to identify biomarkers to detect PD at its earliest onset.

PD is a multifactorial disorder; males are 1.4 times more likely to develop PD, and age, environment, and genetic factors contribute to disease development [8]. Although 90% of PD cases are idiopathic, ~10% of cases are related to known genetic risk factors [14], and overall heritability is estimated to be ~20% for some populations [15]. Several genetic risk factors have been identified, including mutations in alpha-synuclein (SNCA), Parkin (PRKN), DJ-1 (PARK7), PTEN-induced kinase (PINK1), and leucine-rich repeat kinase 2 (LRRK2) protein, as well as glucocerebrosidase (GBA) and other lysosomal enzymes [6]. Patients with genetically linked PD who undergo regular screening are an important population to uncover biomarkers associated with disease progression, as the patients are often aware they are at risk for the disease before developing symptoms.

PD-causing LRRK2 mutations are the greatest known cause of heritable PD, representing approximately 10% of autosomal familial dominant PD, and have also been implicated in 1–3% of sporadic PD cases. The prevalence of familial and sporadic Parkinsonism due to LRRK2 mutations ranges from 0.1–41%, varying by ethnic population, with the highest prevalence of cases found in patients from North Africa (30–40%) and Ashkenazi Jews (10–20%) [16–18]. Presentation of clinical symptoms in LRRK2 PD carriers is described as similar, if not indistinguishable, from sporadic PD patients [19,20].

LRRK2 is a large multi-domain enzyme expressed in diverse tissues and cell types that plays a role in many signal transduction pathways within the cell [16,21]. LRRK2 is a member of the ROCO protein family, which is characterized by a GTPase Ras-like G domain (Roc) followed by a C-terminal of Roc (COR) domain. LRRK2 also contains a serine-threonine kinase domain and can regulate several proteins by phosphorylation [22]. LRRK2 signaling has been shown to influence autophagy, mitochondrial function, transcription, molecular structural integrity, the endo-lysosomal system, and the immune response [16,21]. Although its complete function remains unknown, LRRK2 is predicted to play an important role in WNT/beta-catenin, Rab, and WNT/PCP signaling pathways [14]. In addition to PD, LRRK2 mutations also play a role in systemic lupus erythematosus,

Crohn's disease, inflammatory bowel disease, leprosy, and cancer [16]. Thus, better understanding of LRRK2's function has the potential to affect efforts to target a wide range of human diseases.

Glu2019Ser (G2019S), the most prevalent LRRK2 mutation, has been shown to enhance LRRK2 kinase activity, perturbing mitochondrial bioenergetics and dynamics and leading to increased reactive oxygen species (ROS) within the cell [1,8,16]. A recent study demonstrated that PD patients without a LRRK2 mutation also show increased LRRK2 activity, suggesting that LRRK2 may be a common downstream node affected in PD [23]. However, not all persons with a LRRK2 mutation develop PD, and efforts are underway to understand the mutation penetrance [24]. Global metabolic profiling has the potential to highlight the imbalance of metabolic homeostasis driven by LRRK2 mutation and shed light on the role of this key protein in PD development.

Metabolomics is the comprehensive analysis of low-molecular-weight molecules that provide energy and biomass for life to exist [25]. Alterations in metabolism have been linked to several diseases, including cancer, diabetes, metabolic disorders, and neurodegeneration [3]. Understanding differences between the metabolic profiles of individuals or groups of healthy or diseased persons can shed light on the pathways underlying a disease process. Differences between healthy individuals and those who develop a disease are a result of a combination of factors, including genetic background, epigenetic modifications, lifestyle, stress, and environment [26,27]. Metabolic profiling elucidates the complex interaction between genes and the internal and external environment. Evaluating the metabolic profile of samples from PD patients and non-PD controls has the potential to identify candidate biomarkers of PD, creating a "fingerprint" for early diagnosis and prognosis, allowing clinicians to monitor disease progression and the effect of disease-modifying therapy.

Several small studies have identified metabolite changes in different cohorts of patients with PD, including changes to branched amino acids, fatty acids, and antioxidant levels [28]. One study found that bis(monoacylglycerol)phosphate isoforms were increased in the urine of patients with LRRK2 mutation independent of PD status, and that among PD patients, total levels correlated with cognitive decline [29]. In a separate study comparing PD patients with G2019S-LRRK2 mutations, asymptomatic family members with or without G2019S mutations, patients with idiopathic PD, and non-PD controls, patients with PD had significantly lower levels of urate in their blood than non-PD controls or asymptomatic LRRK2 mutation carriers [11], a result later validated by a larger study [24]. Urate is an endogenous purine metabolite with antioxidant and neuroprotectant properties [24]. One hypothesis suggests that elevated urate in non-PD patients serves as a neuroprotectant to reduce the risk and rate of disease [30]. Unfortunately, a randomized clinical trial evaluating urate-elevating inosine treatment in PD was terminated early due to lack of effect, which suggests that other factors may play a role in PD progression [31]. New strategies are needed to comprehensively analyze metabolites differentially expressed in PD and link them back to the underlying biology.

We have previously demonstrated that using a global, unbiased, multi-dimensional <sup>1</sup>H-<sup>13</sup>C heteronuclear single quantum coherence spectroscopy (HSQC) NMR spectroscopy metabolomics platform coupled with machine learning (ML) techniques, we are able to uncover metabolite-based biomarkers that correlate with disease status and response to therapy [32–34]. Further, as metabolomics studies are widely acknowledged to suffer from reproducibility challenges [35–38], we documented the reproducibility of our platform and how to translate our results for *vitro* diagnostic development [25,32,33]. In this study, we used our NMR and ML platform to measure plasma metabolites from approximately age/sex-matched PD patients with G2019S LRRK2 mutations and non-PD controls wild-type for LRRK2. The final model predicted PD with 79.7% accuracy, 81.3% sensitivity, and 78.6% specificity in male patients with LRRK2 mutations. The differential metabolites identified had links to known biological changes occurring during PD development.

Our metabolomics-based biomarker platform has the potential to improve the diagnosis for PD, which is a critical first step towards treating and ultimately curing the disease.

## 2. Results

### 2.1. LRRK2 PD Patient Sample Characteristics

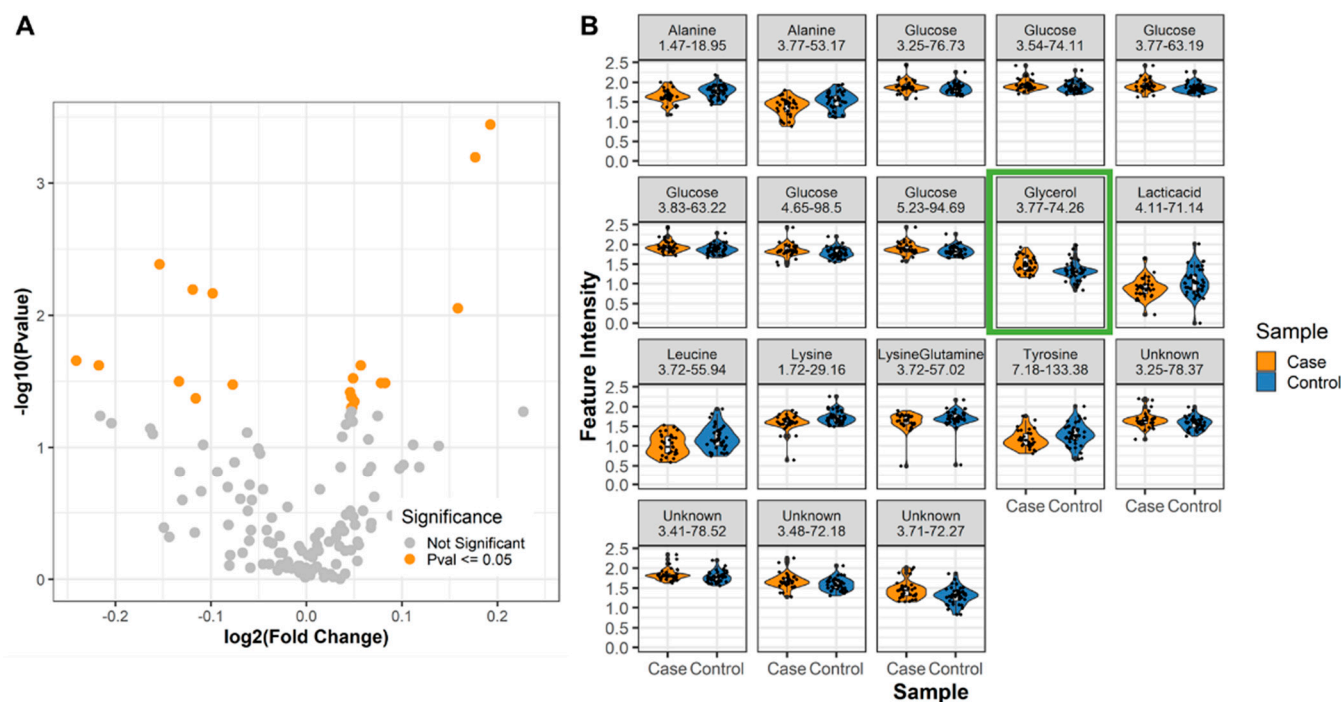
Metabolites were extracted from plasma samples collected from patients diagnosed with PD with a LRRK2 G2019S mutation (N = 46 cases) and non-PD approximately age/sex-matched controls without LRRK2 mutations (N = 65 controls) (Table S1). All samples were assessed for quality assurance and control (QA/QC) by measuring the mean, median, maximum, and minimum intensity values and the number of metabolite resonances detected (Figure S1). One outlier sample (patient ID 15) was excluded, as it had severe streaking and significantly lower mean intensity value than other samples, which can be a sign of sample degradation (Figure S2). The filtered sample set was enriched for male samples (71% LRRK2 PD and 64% control). As sex is known to influence metabolite levels [39], we conducted analysis for both sexes together, as well as males and females separately (Table S2). Based on the small number of differential features for both the unstratified and the female analysis, as well as the reduced size of the female cohort (N = 13 cases and N = 22 controls), all additional analysis focused on male subjects, with N = 32 male LRRK2 PD patients and N = 42 male non-PD controls (Table 1). In the LRRK2 PD cohort, only 2 out of the 32 samples were collected from patients with a post-mortem confirmatory diagnosis (patient ID 47 and 57). The mean age at sample collection for the male LRRK2 PD patients was  $70.0 \pm 8.1$  years, and the non-PD controls were slightly younger, but not statistically significant, with an average age of  $69.5 \pm 7.2$  years ( $p$ -value = 0.780). In the male LRRK2 PD cohort, age at PD onset had a broad range, from 37 years old to 74 years old, with an average of  $58.9 \pm 10.1$  years. In addition, the date at which the samples were collected after PD diagnosis ranged from 0.0 to 28.2 years, with an average of 10.9 years. The Montreal Cognitive Assessment (MoCA) score and body mass index (BMI) for the LRRK2 cohort was lower than, but not statistically significant for, the controls, 25.4 vs. 26.5 ( $p$ -value = 0.206) and 26.7 vs. 28.3 ( $p$ -value = 0.090), respectively. All cases and most controls (92.9%) were white, but not statistically different in ethnicity ( $p$ -value = 0.304). Additional metadata on treatment interventions such as brain surgery were provided (Table S3). In addition to patient clinical data, sample storage conditions, such as length of time in a freezer, can impact metabolomic results [40]. The majority of the samples (85.1%) had freezer time under 5 years (Table S4), and those with longer freezer time were closely monitored.

**Table 1.** Male LRRK2 patient sample characteristics.

	Case	Control
Number of samples	32	42
Age at sample collection in years (SD)	70.0 (8.1)	69.5 (7.2)
Age at onset in years (SD)	58.9 (10.1)	-
Sample collection time after PD Dx in years (SD)	10.9 (8.3)	-
MoCA (SD)	25.4 (4.8)	26.5 (2.5)
BMI (SD)	26.7 (2.6)	28.3 (4.8)
Ethnicity	White	32
	Hispanic	0
	Black	0

## 2.2. Differential Metabolites in Male LRRK2 PD

Using a Kruskal–Wallis (KW) test for significance, we identified 18 metabolite resonances with intensity values that were significantly different ( $p < 0.05$ ) between male cases and controls (Figure 1A,B), in which 11 resonances were increased and 7 decreased in the PD patients compared to the non-PD controls. While not a direct 1:1 relationship due to transfer of magnetization and other factors, the intensity for each resonance can be used as a proxy for relative concentration. Likely due to the limited sample size, only 1 resonance passed the false discovery rate (FDR) correction (FDR  $p$ -value  $< 0.05$ , Figure 1B, green box).



**Figure 1.** Differential metabolites correlated with male LRRK2 PD. (A) Volcano plot of the metabolite resonances extracted from the plasma of male LRRK2 PD patients and non-PD controls, with orange dots indicating significantly different resonances ( $p < 0.05$ ). (B) Violin plots of the significant resonances identified in (A). The resonance at 3.77 and 74.26 ppm in the  $^1\text{H}$  and  $^{13}\text{C}$  dimensions indicated by the green box passed the FDR  $p$ -value of  $< 0.05$ .

The differential metabolite resonances were matched by chemical shift to reference libraries [40] of known metabolites for annotation (Table 2, and Figure S3). Each resonance represents a carbon atom attached to a unique proton (a C–H pair), and thus it is possible that the differential resonances map to the same metabolite. For example, glucose and alanine had 6 and 2 matching resonances, respectively. Other differential resonances were mapped to lysine/glutamine, leucine, tyrosine, lactate, and glycerol. Resonances from differential metabolites that would require experimental validation to distinguish matching-ambiguities (“unknown” features) were kept for further analysis and referenced by their chemical shift position. The average fold change between PD patients and controls was 1.2.

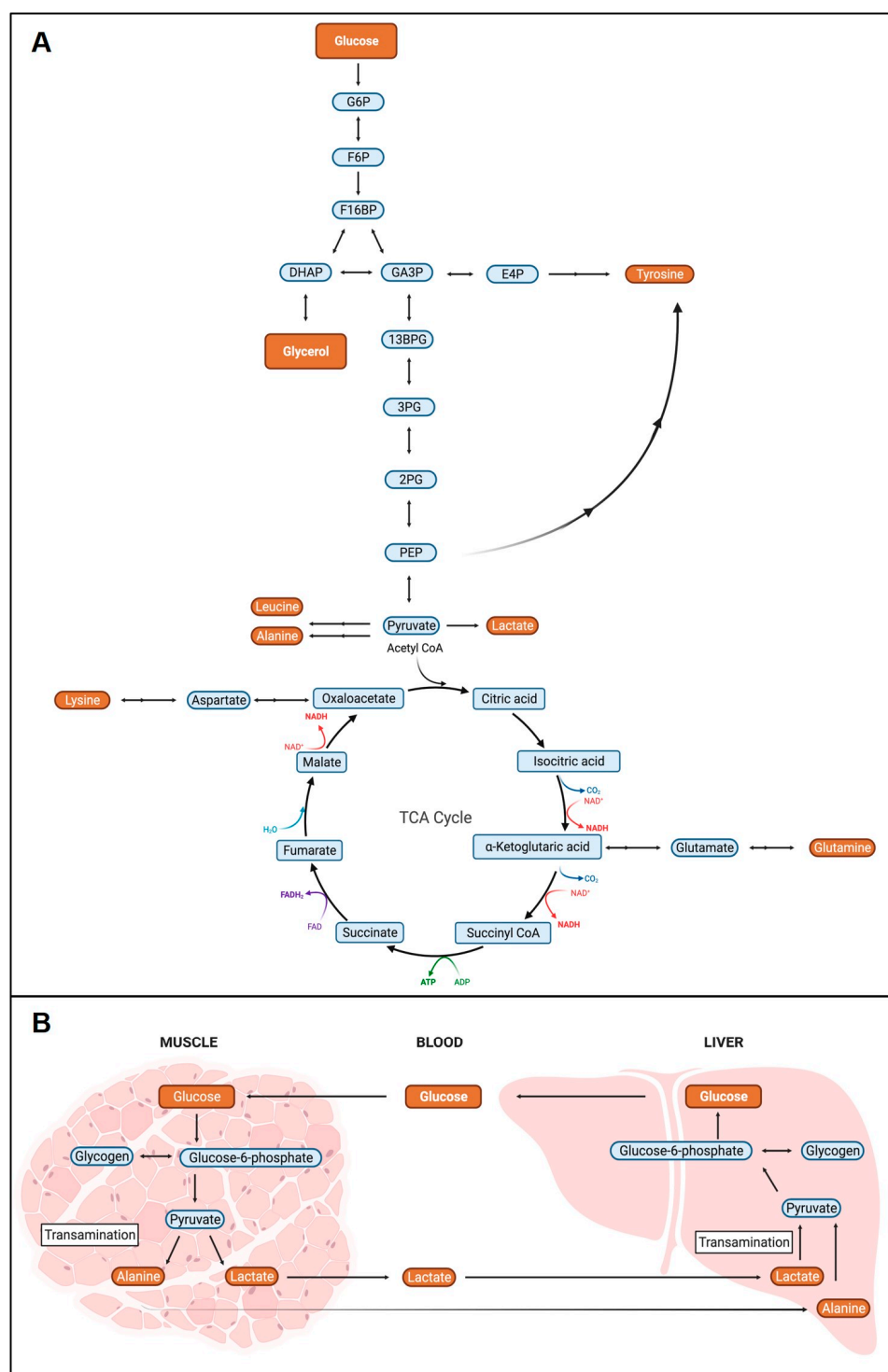
**Table 2.** Top significantly different <sup>1</sup>H-<sup>13</sup>C HSQC metabolite resonances (*p*-value < 0.05) between male LRRK2 PD and controls.

Annotation	<sup>1</sup> H ppm	<sup>13</sup> C ppm	Fold Change (Case/Control)	<i>p</i> -Value	FDR <i>p</i> -Value
Glycerol	3.77	74.26	1.130	0.001	0.042
Alanine	3.77	53.17	0.899	0.004	0.178
Lysine	1.72	29.16	0.921	0.006	0.178
Alanine	1.47	18.95	0.934	0.007	0.178
Unknown	3.71	72.27	1.116	0.009	0.193
Lactic acid	4.11	71.14	0.846	0.022	0.302
Leucine	3.72	55.94	0.860	0.024	0.302
Unknown	3.41	78.52	1.040	0.024	0.302
Glucose	3.77	63.19	1.034	0.030	0.302
Tyrosine	7.18	133.38	0.912	0.032	0.302
Unknown	3.48	72.18	1.056	0.032	0.302
Unknown	3.25	78.37	1.059	0.032	0.302
Lysine/Glutamine	3.72	57.02	0.948	0.033	0.302
Glucose	3.54	74.11	1.032	0.038	0.302
Glucose	4.65	98.5	1.033	0.041	0.302
Glucose	3.25	76.73	1.035	0.045	0.302
Glucose	3.83	63.22	1.033	0.050	0.302
Glucose	5.23	94.69	1.034	0.050	0.302

### 2.3. Metabolite Pathway Analysis

The dysregulated metabolites in the PD patients were enriched for amino acids and metabolites associated with glycolysis and gluconeogenesis. Many of the metabolites can be interconverted through known metabolic pathways (Figure 2A). For example, pyruvate, the end product of glycolysis, can be directly converted to lactate and used to generate alanine or supply the citric acid cycle (TCA) to produce additional amino acids such as glutamine. Glycolysis also generates 3-phosphoglycerate, which is a building block for the production of erythrose 4-phosphate, a precursor for tyrosine biosynthesis. Similarly, glycerol, identified as increased in PD patients, can be used to generate glucose and/or other glycolytic intermediates by conversion into dihydroxyacetone phosphate. Leucine and lysine are essential amino acids and must be obtained through diet [41].

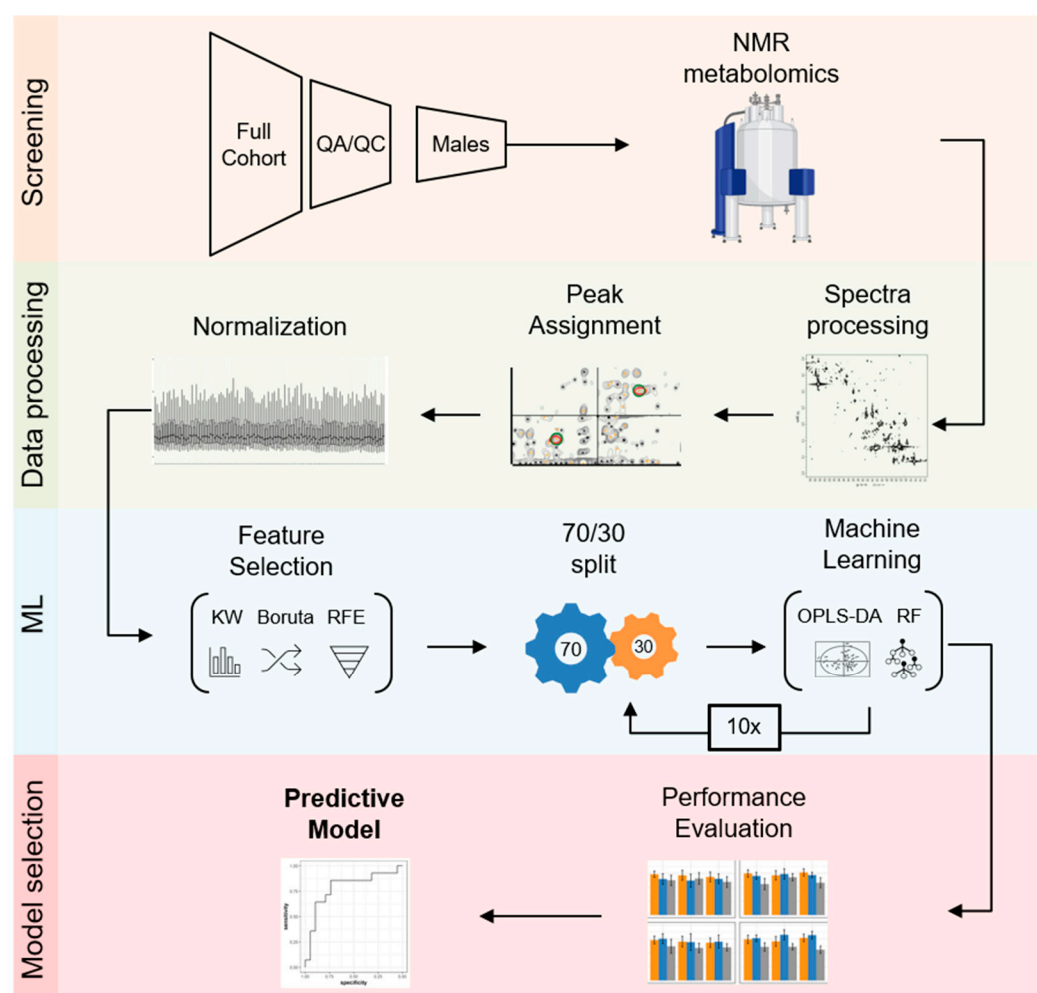
Many of these metabolites have been previously identified as altered in PD [42], suggesting that flux through these pathways could be an important contributor to PD biology. This analysis was performed on plasma; thus, it is also important to consider how circulating metabolite levels affect the PD biology. We identified glucose, alanine, and lactate as altered metabolites in PD patients. These metabolites form part of the glucose–alanine and Cori cycle, wherein in organs like skeletal muscle increase glycolysis and protein catabolism to drive the production of lactate and alanine, both of which easily diffuse to the liver to become substrates for gluconeogenesis (Figure 2B). This helps increase glucose availability in the blood, especially during times of stress [43]. Investigating the impact of these processes in LRRK2 PD and PD warrants further research in additional patient cohorts and will be a focus of future efforts. In addition, we are working with collaborators to test the significance of metabolic flux through these pathways in animal models of the disease, as they could point to new therapeutic targets for intervention.



**Figure 2.** Dysregulated Male LRRK2 PD Metabolic Pathways. **(A)** Altered metabolites in male LRRK2 PD are enriched for amino acids and metabolites associated with glycolysis and gluconeogenesis. **(B)** Altered metabolites glucose, alanine, and lactate are central components of the glucose-alanine and Cori cycle important for supporting gluconeogenesis in the liver. Detected differential metabolites are highlighted in orange boxes. Double arrows represent multiple enzymatic transformational steps. Abbreviations: G6P; Glucose-6-phosphate, F6P; Fructose-6-phosphate, F16BP; Fructose 1,6-biphosphate, DHAP; dihydroxyacetone phosphate, GA3P; Glyceraldehyde-3-phosphate, 13BP; 1,3-Biphosphoglycerate, 3PG; 3-Phosphoglycerate, 2PG; 2-Phosphoglycerate, PEP; Phosphoenolpyruvate.

#### 2.4. Building a Male LRRK2 PD Classifier

The process that generated an accurate model to classify case from control patients comprised generating pre-defined feature sets, evaluating each set in distinct cross-validated machine learning (ML) algorithms, and selecting the final model based on performance in the training and test datasets (Figure 3). The pre-defined feature sets included: “Known”, differential metabolite resonances assigned to a unique metabolite; “Unknowns”, differential metabolite resonances that require experimental clarification; and “Clinical”, available sample-donor clinical information (Table S5). In addition to the KW test, feature sets were further subjected to 2 selection methods, Boruta, that selects features based on the Random Forest (RF) importance of the unchanged vs. shuffled feature [44], and RFE (Recursive Feature Elimination), that recursively eliminates the least important feature of each retrained model. Combinations of the feature sets (“Known”, “Unknown” and “Clinical”) with the feature selection methods (KW-only, KW-Boruta and KW-RFE) led to 12 feature sets for ML analysis (Table S6). The data were then split 70/30 into training and test data, 10x cross-validated, and assessed using orthogonal partial least squares discriminate (OPLS-DA) and RF models (Figure 3).



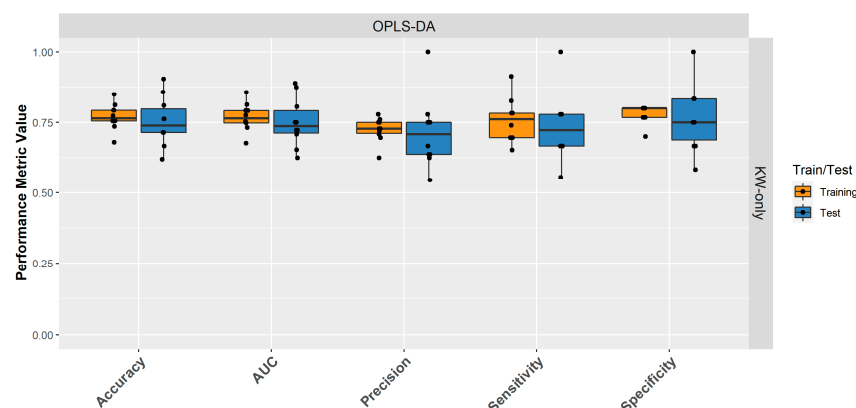
**Figure 3.** Classifier workflow. Following sample preparation and NMR filtering, raw spectra were processed, resonance peaks assigned to the database-matching metabolites, and the resonances across samples normalized. Modeling comprised feature selection and 10x cross-validation of OPLS-DA and RF methods. The model with the highest stability and overall performance was selected as the champion model.

Cross-validation models were assessed based on accuracy, sensitivity, specificity, precision, and area under the curve (AUC) to distinguish LRRK2 PD patient samples from those of non-PD controls. Model selection started with comparing the performance and stability of the cross-validated ML models between the training and test splits (Figure S4 and Table S7). A pre-evaluation of the models excluded 10 unstable, and potentially over-fitting, models because at least one of the five performance metrics was significantly different between the training and test datasets ( $p < 0.05$ , Student's  $t$ -test).

For the remaining 14 models, the average cross-validated accuracy, sensitivity, specificity, precision, and AUC for all models were  $67.6 \pm 5.1\%$ ,  $59.8 \pm 7.1\%$ ,  $73.3 \pm 5.5\%$ ,  $63.8 \pm 5.7\%$ , and  $70.9 \pm 2.2\%$ , respectively. Comparing models using “Known” vs. “Known + Unknown” features, we observed improved testing AUC from  $69.6 \pm 0.5\%$  to  $71.8 \pm 2.4\%$  with the addition of “Unknown” features. Similarly, upon adding clinical data to the “Known” feature set, the testing AUC improved to  $73.1 \pm 3.1\%$ . However, when adding clinical data to the “Known + Unknown” feature sets, the testing AUC dropped to  $70.3 \pm 3.0\%$ , similar to the model performance using “Known” features. On average, OPLS-DA models had better performance than RF models, with  $71.7 \pm 2.1\%$  and  $70.3 \pm 2.2\%$  in AUC, respectively. Boruta and RFE selected features showed improved classification in RF models, but not in OPLS-DA models. In RF analysis, KW-Boruta and KW-RFE increased the AUC from  $67.5 \pm 14.3\%$  (KW-only) to  $69.0 \pm 0.0\%$  (KW-Boruta) and  $71.9 \pm 1.8\%$  (KW-RFE). However, in OPLS-DA models, AUC decreased from  $75.0 \pm 8.6\%$  (KW-only) to  $70.4 \pm 0.8\%$  (KW-Boruta) and  $73.3 \pm 6.1\%$  (KW-RFE).

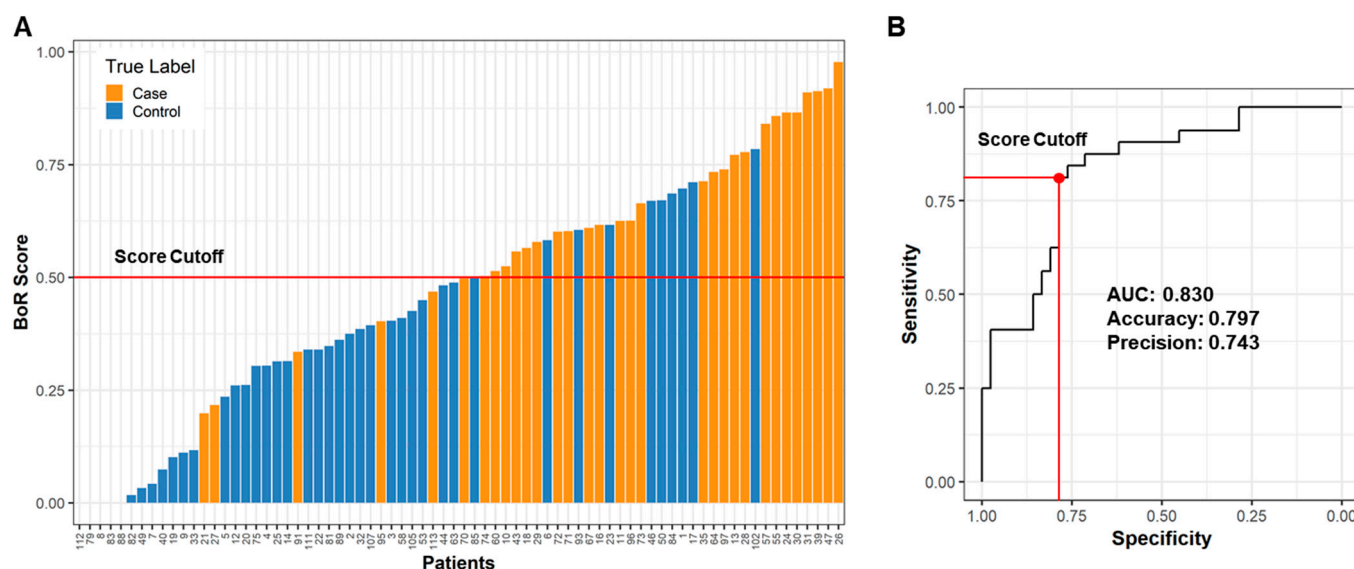
Following model ranking based on cross-validated stability and performance, and model-selection based on the cross-validated AUCs, an OPLS-DA model using KW-only and “Known + Unknown” features was selected as the champion model. The OPLS-DA cross-validation model had  $77.0 \pm 4.6\%$  accuracy,  $75.7 \pm 7.7\%$  sensitivity,  $78.0 \pm 3.2\%$  specificity,  $72.4 \pm 4.3\%$  precision, and  $76.8 \pm 4.9\%$  AUC for the training data, and  $75.2 \pm 8.6\%$  accuracy,  $73.3 \pm 11.9\%$  sensitivity,  $76.7 \pm 11.7\%$  specificity,  $71.4 \pm 12.5\%$  precision, and  $75.0 \pm 8.6\%$  AUC for the test data (Figure 4).

In addition, sensitivity analysis evaluated the effect of each metabolite on the champion model performance (Table S8). Removing any metabolite led to decreased AUC. Removing two resonances from alanine caused the most significant drop in AUC from 85% to 60.8%. This was followed by loss of two lysine resonances and then loss of all six glucose peaks, which dropped AUC to 64.6% and 65.4%, respectively. Withholding any of the single resonance identified from glycerol, tyrosine, lactic acid, and leucine dropped AUC to ~70%. Taken together, the sensitivity analysis demonstrated that all resonances are important for the accuracy of the model. Furthermore, it demonstrates that metabolites like glucose, which contributed six resonances, were not overly biasing the model, as other metabolites (alanine and lysine) more strongly influenced the ML metrics.



**Figure 4.** Cross-validation performance of champion machine learning algorithm for LRRK2 PD classification. The data with KW-only and “Known + Unknown” features was split 70/30 into training and test data, 10x cross-validated using an OPLS-DA model.

Following selection of the champion model, OPLS-DA rankings were used to derive a proprietary Olaris® BoR (biomarker of response) score that ranges from 0 to 1, with PD samples scoring over 0.5 and non-PD controls under 0.5 (Figure 5A). Once the BoR score was applied to the full dataset, the LRRK2 PD classifier had 79.7% accuracy, 81.3% sensitivity, 78.6% specificity, 74.3% precision and 83.0% AUC (Figure 5B).



**Figure 5.** Olaris BoR score has high accuracy to classify male LRRK2 PD. **(A)** Waterfall plot of BoR score for all data with predicted cases scoring over 0.5 and predicted controls scoring under 0.5. True cases are colored orange and true controls are colored blue. **(B)** Receiver operator characteristic (ROC) analysis with 0.830 area under the curve (AUC). Red lines mark corresponding 81.3% sensitivity and 78.6% specificity using a score cutoff at 0.5.

From the non-PD control population, one patient (ID 102) had a high BoR score suggestive of PD (Figure 5A). This patient had a relatively low MoCA score of 22 compared to the average of controls of 26.5, suggesting the patient could have mild cognitive impairment (Table S3). Similarly, there were two PD patients that had low scores suggestive of a non-PD status (ID 21 and 27). There were not any available clinical data supporting the low score, and neither of these two patients had a post-mortem confirmatory PD diagnosis. Even though there were only two patients with post-mortem confirmatory diagnosis, we were very encouraged to see they were predicted correctly and ranked 2nd (BoR 0.92) and 8th (0.84) highest in BoR scores. In addition we applied this model to the female patients, but as expected, it had poor overall accuracy (51.4% Figure S5), suggesting the importance to examine the influence of sex on metabolites.

### 3. Discussion

PD is a growing concern, with a prevalence increasing more rapidly than other neurological disorders [8]. There is an unmet clinical need for improved diagnosis of PD to enable either early detection or greater confidence in diagnosis. While there are currently limited treatments available for PD, efforts are underway to develop new therapies. For these treatments to be successful, improved PD diagnostics is a prerequisite.

PD is influenced by aging, genetic predisposition, and exposure to exogenous and endogenous toxins [45]. While it was previously considered to be primarily a central nervous system disorder, PD is now understood to be multifactorial, with multiple mechanisms contributing to its pathogenesis. As metabolomics has been used to closely examine PD, PD may even be considered a metabolic disorder, with metabolic dysregulation contributing to neurodegeneration [1]. Previous studies have identified metabolites altered in

older adults with PD, including elevation of circulating  $\beta$ -amino butyric acid, cystine, ornithine, phosphoethanolamine, and proline [46], and altered bacterial composition in patients with PD may be linked with changes in bacterial metabolites that affect patient health [47].

In this study, we utilized multi-dimensional NMR spectroscopy and ML to develop a classifier for LRRK2 PD. Focusing on male patients with G2019S LRRK2 mutations, we identified 18 significantly differential metabolite resonances between male LRRK2 PD patients and controls. Although the metabolites identified in our analysis have been previously linked to PD, many of the studies were performed in different biofluids, in both sexes and using different analytical techniques, making it difficult to directly compare results between studies. Nonetheless, there were many consistencies, such as the altered glucose metabolism in patients with PD [48–50]. We also observed an increase in glycerol levels in PD patients with LRRK2 mutations, consistent with previous reports, suggesting that glycerol phospholipid metabolism is upregulated in the plasma of PD patients [51]. Finally, we observed a decrease in lactate, which was similarly observed in the CSF of patients with PD [52]. Many of the altered metabolites, including lactate, alanine, and glutamine, are connected via known metabolic pathways, and differences in their levels may reflect dysregulation of these pathways overall in PD.

We found that several amino acids, including alanine and lysine, were decreased in LRRK2 PD plasma samples. Similarly, in previous studies, alanine was found to be decreased in both serum and CSF of PD patients [52,53], and lysine levels were decreased in the CSF of PD patients [54]. Our data thus provide independent confirmation of these markers in a novel patient cohort, which is encouraging given the reproducibility concerns facing metabolomics studies. In contrast, our analysis showed that leucine and tyrosine were also decreased in LRRK2 PD plasma samples, while previous studies suggested that leucine is increased in the CSF, fecal and urine samples of patients with PD [55–57] and increased tyrosine levels were found in the CSF, urine, and saliva of PD patients [55,58,59]. Dysregulation of the synthesis of dopamine from tyrosine in PD has been well established [60], however it is possible that both tyrosine and leucine levels may be altered in different directions in different biofluids. Further, leucine and tyrosine have both been demonstrated to be elevated in people with metabolically unhealthy obesity [61], and while adding BMI did not improve our model accuracy, it is still possible that obesity and associated comorbidities may act as confounders in our analysis. Additionally, our study was restricted to LRRK2 PD, while previous studies examined levels of these metabolites in idiopathic PD, and it is unclear if genetic forms of PD are a reporter for all cases of PD. Finally, it is not unexpected to find some discordance due to reproducibility challenges in metabolomics, the observed small fold-change between PD patients and controls in our studies and others, and the sensitivity of metabolites to sex, diet, medication, and other factors. This underscores the need for additional controlled validation studies. Finally, while other studies observed decreased urate levels in LRRK2 PD plasma samples, we were unable to measure urate [11,24], as it lacks  $^1\text{H}$ - $^{13}\text{C}$  resonances.

Glucose, alanine, and lactate were the primary metabolites altered in patients with PD. These metabolites form part of the glucose-alanine cycle and the Cori cycle (also known as the lactic acid cycle), which are central to gluconeogenesis, increasing glucose bioavailability and providing ATP to cells during times of stress or exertion. Interestingly, the glucose–alanine cycle (both hits in our analysis) has also been shown to be downregulated in PD patients [62]. This may be particularly relevant in our patient cohort given the established role of LRRK2 in oxidative stress and the effect of the G2019S mutation on ROS production and neuronal cell death in vitro [63,64]. Further research will focus on better understanding the role of these pathways in LRRK2-related PD and PD overall.

Using metabolomics data from this cohort of PD patients with LRRK2 mutations, we developed an OPLS-DA model to derive a biomarker-based score for PD. Although the current BoR score is based on limited data from a heterogeneous patient cohort and must be considered a first step towards a more comprehensive model, the score differentiated LRRK2 PD

patients from controls with 79.7% accuracy, 81.3% sensitivity, 78.6% specificity, 74.3% precision and 83.0% AUC. While there were a few patients and controls who fell outside the score cutoff for non-PD vs. PD status, the majority were correctly predicted using the score. Furthermore, the two patients with a post-mortem biopsy-proven PD diagnosis were classified correctly. In this study, performance stability was used to rank the predictive models and a conservative BoR cutoff (0.5) was selected. Future studies on larger cohorts should also explore models and BoR cutoffs in which the balance of specificity and sensitivity is refined towards maximizing true positive or true negative predictions.

Despite the encouraging first step towards a non-invasive confirmatory diagnostic, appropriate caution must be applied regarding the generalizability of these results to additional PD cohorts due to the limitations of this study. While we were able to build a classifier with high predictive accuracy to classify LRRK2 PD patients from non-PD controls, it remains to be determined if this signature classifies other familial or idiopathic forms of PD, especially as differences in omics results have been observed between familial vs. idiopathic PD [65]. Additionally, it is unclear whether the classifier can distinguish individuals with a LRRK2 mutation that do not manifest symptoms of PD. Moreover, although the analyzed cohort was approximately age- and gender-matched, we were not able to control for additional potential confounding factors such as medications or diet. Hence, it is possible that the PD LRRK2 BoR score may have bias towards some of these confounders. Full access to the patients' clinical information would be necessary to address this limitation.

Finally, this study utilized samples from a single relatively small cohort of LRRK2 PD patients, collected at various time points post diagnosis. The varied interval between the date of PD onset and sample collection is a concern. Some patients provided plasma samples upon diagnosis, while others had been suffering from disease for 28.2 years. This suggests that the "case" cohort had patients at various stages of disease progression, which could obscure differences, and might explain why the addition of clinical data such as MoCA and BMI did not add value to the model performance. Furthermore, analysis of samples collected prior to diagnosis would improve the ability of the BoR score to diagnose asymptomatic and early PD onset. While the model developed was built to account for the large amount of variability, further research will focus on lowering patient and sample variability by recruiting patients within 2–5 years of PD diagnosis; including samples collected prior to diagnosis in analysis; requiring, for example, fasting of patients prior to sample collection; and ensuring access to clinical data. The metabolite and ML pipeline described here provides a foundation to develop a confirmatory diagnostic for PD.

#### 4. Materials and Methods

The NMR reference standard, deuterated 3-(trimethylsilyl)-1-propanesulfonic acid sodium salt (DSS-d<sub>6</sub>, 98%) and deuterium oxide (D<sub>2</sub>O, 99.0%) were purchased from Cambridge Isotope Laboratory, Andover, MA, USA. Potassium dihydrogen phosphate (KH<sub>2</sub>PO<sub>4</sub>) and dipotassium hydrogen phosphate (K<sub>2</sub>HPO<sub>4</sub>) were purchased from Millipore-Sigma (St. Louis, MO, USA).

##### 4.1. Patient Enrollment and Plasma Collection

The study included 46 PD patients and 65 controls who participated in the 'Spot' study [66,67]. In brief, the Spot study included PD patients and genetically unrelated controls (mostly spouses) from the Center for Parkinson's Disease at Columbia University Irving Medical Center in New York, NY, recruited between 2010–2021. All participants have been fully sequenced for GBA mutations and screened for the LRRK2 G2019S mutation. Information on demographics, medical history, medication, family history of PD, the Unified Parkinson's Disease Rating Scale (UPDRS) in the "on" state and the Montreal Cognitive Assessment (MoCA) were collected from all participants. A 10cc EDTA tube was used to collect blood, which was centrifuged and aliquoted to 1cc plasma aliquots

within 60 min of collection. Participants were non-fasting at the time of the blood donation. Samples were stored in a  $-80^{\circ}\text{C}$  freezer until processing. All study procedures were approved by the Columbia University IRB, and all participants signed informed consent.

#### 4.2. NMR Sample Preparation

Metabolites were extracted from 1 mL of human plasma via a methanol and chloroform liquid–liquid extraction. The aqueous phase was transferred to a Falcon tube and freeze-dried. Lipid extracts (0.25 mL) were isolated, dried, and stored at  $-80^{\circ}\text{C}$  for separate lipidomic analysis. The powder was reconstituted in 0.18 mL of 50 mM phosphate buffer at pH 7.4 in  $\text{D}_2\text{O}$  and then immediately transferred to a 3 mm NMR tube for NMR data collection. The NMR standard, DSS- $\text{d}_6$ , was added to each sample for chemical shift referencing.

#### 4.3. NMR Data Collection and Processing

All NMR data were acquired on a Bruker AVANCE II solution-state NMR spectrometer equipped with a liquid nitrogen-cooled Prodigy TXI Cryoprobe operating at a proton frequency of 600 MHz using `hsqcetgpsisp2.2` pulse program with other parameters:  $90^{\circ}$  pulse angle,  $8.53\ \mu\text{s}$  pulse width, 1.5 s relaxation delay, 128 dummy scans, 203 receiver gain and 36 number of scans. The  $1\text{JCH}$  used was 145 Hz (1.72 ms). The number of complex points was set to 512 and 32 (25% NUS) for direct and indirect dimension, respectively. The spectral width of 15.97 ppm (9578.44 Hz) and 160.00 ppm (24131.775 Hz) were used for  $^1\text{H}$  and  $^{13}\text{C}$ , respectively. NUS schedules were generated using Poisson gap distribution [68], and spectral data were processed using the NMRPipe software package as previously described [69]. The NUS data was reconstructed using iterative soft thresholding according to the hmsIST algorithm [68]. Chemical shift queries, metabolite identifications and quantifications were performed using the COLMARm NMR webserver (<http://spin.ccic.osu.edu/index.php/colmarm/index>, accessed on 21 October 2021) [41]. Metabolite resonances with a signal intensity greater than  $2\text{e}6$  ( $24\times \text{S/N}$ ) were dynamically binned into clusters using DBSCAN (Density-Based Spatial Clustering of Applications with Noise) in both  $^1\text{H}$  and  $^{13}\text{C}$  dimensions. Resonances below the signal intensity cut-off were imputed with a two-step process either by matching resonances between  $1\text{e}6$  ( $12\times \text{S/N}$ ) and  $2\text{e}6$  by setting those below the limit of detection to a constant value of  $7\text{e}5$ . Resonance clusters were normalized and filtered using a criterion of being present in at least 50% of all samples in one class.

#### 4.4. Statistical Analysis

The distribution of metabolic resonance intensity was tested by Shapiro–Wilk normality test and Q–Q plot. Due to non-normal distribution, a Kruskal–Wallis (KW) non-parametric one-way analysis of variance (ANOVA) was used to test for significant differences in measured NMR resonances between groups of interest. The test of significance was also adjusted based on false discovery rate (FDR) for multiple hypothesis testing correction. Fold changes (FC) were calculated as the ratio of the mean intensities of the two groups. All statistical analyses and ML modeling were performed using R 4.1.1 (<http://cran.rproject.org>, accessed 19 November 2021) with the following packages: `caret` and `Boruta` for feature selection [44,70], `randomForest` and `ropls` for modeling [71,72], and `ggplot2` and `pROC` for result visualization [73,74].

### 5. Conclusions

The process of diagnosis for Parkinson’s disease can be arduous, with an average time of 7 years from symptom onset until official diagnosis [75]. During this time, patients struggle through a frustrating and costly diagnostic odyssey, only to find limited treatment options. Early diagnostics stand to revolutionize PD care. While further research is needed to expand these results beyond a limited patient cohort, the Olaris® BoR score for

PD represents a substantial step forward in improving diagnostic precision for LRRK2 PD. The model relies on quantifying metabolomic signatures in the plasma of male LRRK2 PD patients, thus requiring only minimally invasive sample collection and providing rapid results. This novel scoring system may help clinicians and their patients to quickly move toward a definitive PD diagnosis, widening the critical window in which therapy is beneficial, allowing time for further research toward curative therapy, and providing PD patients with many more functional years.

**Supplementary Materials:** The following supporting information can be downloaded at: [www.mdpi.com/article/10.3390/metabo12020149/s1](http://www.mdpi.com/article/10.3390/metabo12020149/s1), Table S1, Patient Characteristics; Table S2, Summary of Case vs. Control Analysis; Table S3, Male Patient Metadata; Table S4, Male Sample Freezer Time; Table S5, Pre-defined Feature Sets; Table S6, Feature Sets for ML Analysis; Table S7, Cross-validation ML Model Performance; Table S8, Sensitivity Analysis of Champion Model; Figure S1, NMR data quality control and assurance by measuring the mean, median, maximum and minimum intensity values and number of metabolite resonances detected; Figure S2, NMR spectral comparison; Figure S3, NMR annotations for top differential features; Figure S4, Performance of champion machine learning algorithm for LRRK2 PD classification; Figure S5, Receiver operator characteristic (ROC) analysis with 0.556 area under the curve (AUC) using male BoR score to predict on female cohort.

**Author Contributions:** Conceptualization, E.M.O.; investigation, E.M.O., C.D., C.H., L.O.R., J.W., M.S., R.N.A.; resources, M.S., R.N.A.; data curation, C.D., C.H., L.O.R.; writing—original draft preparation, E.M.O., K.B.S.; writing—review and editing, E.M.O., C.D., C.H., L.O.R., J.W., K.B.S., M.S., R.N.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Olaris, Inc and in part by funding from the Michael J. Fox Foundation grant number “MJFF-010153”. The Columbia University cohort recruitment was supported by the Parkinson’s Foundation, and the National Institutes of Health (grants K02NS080915 and UL1 TR000040).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Columbia University Irving Medical Center (IRBAAAF3108, initial date of approval: 08/18/2021).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data available on request subject to proprietary restrictions.

**Acknowledgments:** The authors would like to thank Shalini Padmanabhan, Bradford Casey, Srihari Raghavendra Rao, Kanchan Sonkar, Robert Powers, Gerhard Wagner, Stan Lapidus, Charles Sang, Tan Chui-Mae and Mahantesh Kothiwale for their many valuable discussions that motivated the writing of this communication.

**Conflicts of Interest:** The authors report that this work (design of the study, collection, analysis, and interpretation of data) was supported by Olaris, Inc. Chen Dong, Chandrashekhara Honrao, Leonardo Rodrigues, Josephine Wolf, and Elizabeth O’Day are employees of Olaris. Keri Sheehan is a consultant at Olaris, Inc.

## References

1. Anandhan, A.; Jacome, M.S.; Lei, S.; Hernandez-Franco, P.; Pappa, A.; Panayiotidis, M.I.; Powers, R.; Franco, R. Metabolic Dysfunction in Parkinson’s Disease: Bioenergetics, Redox Homeostasis and Central Carbon Metabolism. *Brain Res. Bull.* **2017**, *133*, 12–30.
2. Hague, S.M.; Klaffke, S.; Bandmann, O. Neurodegenerative Disorders: Parkinson’s Disease and Huntington’s Disease. *J. Neurol. Neurosurg. Psychiatry* **2005**, *76*, 1058–1063.
3. Havelund, J.F.; Heegaard, N.H.H.; Færgeman, N.J.K.; Gramsbergen, J.B. Biomarker Research in Parkinson’s Disease Using Metabolite Profiling. *Metabolites* **2017**, *7*, 42.
4. Wang, X.; Zeng, F.; Jin, W.S.; Zhu, C.; Wang, Q.H.; Bu, X.L.; Luo, H.B.; Zou, H.Q.; Pu, J.; Zhou, Z.H.; et al. Comorbidity Burden of Patients with Parkinson’s Disease and Parkinsonism between 2003 and 2012: A Multicentre, Nationwide, Retrospective Study in China. *Sci. Rep.* **2017**, *7*, 1671, <https://doi.org/10.1038/s41598-017-01795-0>.
5. Isaacson, J.R.; Brillman, S.; Chhabria, N.; Isaacson, S.H. Impact of DaTscan Imaging on Clinical Decision Making in Clinically Uncertain Parkinson’s Disease. *J. Parkinsons Dis.* **2021**, *11*, 885–889, <https://doi.org/10.3233/JPD-202506>.

6. Lin, M.K.; Farrer, M.J. Genetics and Genomics of Parkinson's Disease. *Genome Med.* **2014**, *6*, 48.
7. Hughes, A.J.; Ben-Shlomo, Y.; Daniel, S.E.; Lees, A.J. What Features Improve the Accuracy of Clinical Diagnosis in Parkinson's Disease: A Clinicopathologic Study. *Neurology* **1992**, *42*, 1142, <https://doi.org/10.1212/wnl.42.6.1142>.
8. Armstrong, M.J.; Okun, M.S. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA—J. Am. Med. Assoc.* **2020**, *323*, 548–560.
9. Selikhova, M.; Williams, D.R.; Kempster, P.A.; Holton, J.L.; Revesz, T.; Lees, A.J. A Clinico-Pathological Study of Subtypes in Parkinson's Disease. *Brain* **2009**, *132*, 2947–2957, <https://doi.org/10.1093/brain/awp234>.
10. Thenganatt, M.A.; Jankovic, J. Parkinson Disease Subtypes. *JAMA Neurol.* **2014**, *71*, 499–504.
11. Johansen, K.K.; Wang, L.; Aasly, J.O.; White, L.O.R.; Matson, W.R.; Henchcliffe, C.; Beal, M.F.; Bogdanov, M. Metabolomic Profiling in LRRK2-Related Parkinson's Disease. *PLoS ONE* **2009**, *4*, e7551, <https://doi.org/10.1371/journal.pone.0007551>.
12. Khoury, N.; Attal, F.; Amirat, Y.; Oukhellou, L.; Mohammed, S. Data-Driven Based Approach to Aid Parkinson's Disease Diagnosis. *Sensors* **2019**, *19*, 242, <https://doi.org/10.3390/s19020242>.
13. Tolosa, E.; vander Borght, T.; Moreno, E. Accuracy of DaTSCAN (123I-Ioflupane) SPECT in Diagnosis of Patients with Clinically Uncertain Parkinsonism: 2-Year Follow-up of an Open-Label Study. *Mov. Disord.* **2007**, *22*, 2346–2351, <https://doi.org/10.1002/mds.21710>.
14. Salašová, A.; Yokota, C.; Potěšil, D.; Zdráhal, Z.; Bryja, V.; Arenas, E. A Proteomic Analysis of LRRK2 Binding Partners Reveals Interactions with Multiple Signaling Components of the WNT/PCP Pathway. *Mol. Neurodegener.* **2017**, *12*, 54, <https://doi.org/10.1186/s13024-017-0193-9>.
15. Blauwendraat, C.; Nalls, M.A.; Singleton, A.B. The Genetic Architecture of Parkinson's Disease. *Lancet Neurol.* **2020**, *19*, 170–178.
16. Price, A.; Manzoni, C.; Cookson, M.R.; Lewis, P.A. The LRRK2 Signalling System. *Cell Tissue Res.* **2018**, *373*, 39–50.
17. Khan, N.L.; Jain, S.; Lynch, J.M.; Pavese, N.; Abou-Sleiman, P.; Holton, J.L.; Healy, D.G.; Gilks, W.P.; Sweeney, M.G.; Ganguly, M.; et al. Mutations in the Gene LRRK2 Encoding Dardarin (PARK8) Cause Familial Parkinson's Disease: Clinical, Pathological, Olfactory and Functional Imaging and Genetic Data. *Brain* **2005**, *128*, 2786–2796, <https://doi.org/10.1093/brain/awh667>.
18. Lesage, S.; Brice, A. Parkinson's Disease: From Monogenic Forms to Genetic Susceptibility Factors. *Hum. Mol. Genet.* **2009**, *18*, R48–R59.
19. Adams, J.R.; van Netten, H.; Schulzer, M.; Mak, E.; McKenzie, J.; Strongosky, A.; Sossi, V.; Ruth, T.J.; Lee, C.S.; Farrer, M.; et al. PET in LRRK2 Mutations: Comparison to Sporadic Parkinson's Disease and Evidence for Presymptomatic Compensation. *Brain* **2005**, *128*, 2777–2785, <https://doi.org/10.1093/brain/awh607>.
20. Kay, D.M.; Zabetian, C.P.; Factor, S.A.; Nutt, J.G.; Samii, A.; Griffith, A.; Bird, T.D.; Kramer, P.; Higgins, D.S.; Payami, H. Parkinson's Disease and LRRK2: Frequency of a Common Mutation in U.S. Movement Disorder Clinics. *Mov. Disord.* **2006**, *21*, 519–523, <https://doi.org/10.1002/mds.20751>.
21. Erb, M.L.; Moore, D.J. LRRK2 and the Endolysosomal System in Parkinson's Disease. *J. Parkinsons Dis.* **2020**, *10*, 1271–1291.
22. Seegobin, S.P.; Heaton, G.R.; Liang, D.; Choi, I.; Blanca Ramirez, M.; Tang, B.; Yue, Z. Progress in LRRK2-Associated Parkinson's Disease Animal Models. *Front. Neurosci.* **2020**, *14*, 674.
23. di Maio, R.; Hoffman, E.K.; Rocha, E.M.; Keeney, M.T.; Sanders, L.H.; de Miranda, B.R.; Zharikov, A.; van Laar, A.; Stepan, A.F.; Lanz, T.A.; et al. LRRK2 Activation in Idiopathic Parkinson's Disease. *Sci. Transl. Med.* **2018**, *10*, eaar5429, <https://doi.org/10.1126/scitranslmed.aar5429>.
24. Bakshi, R.; Macklin, E.A.; Logan, R.; Zorlu, M.M.; Xia, N.; Crotty, G.F.; Zhang, E.; Chen, X.; Ascherio, A.; Schwarzschild, M.A. Higher Urate in LRRK2 Mutation Carriers Resistant to Parkinson Disease. *Ann. Neurol.* **2019**, *85*, 593–599, <https://doi.org/10.1002/ana.25436>.
25. Hoyt, L.; O'Day, E.M. Perspective: A Potential Role for NUS in Metabolite-Based in Vitro Diagnostics. *Magn. Reson. Chem.* **2021**, *59*, 257–263, <https://doi.org/10.1002/mrc.5104>.
26. Patti, G.J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: The Apogee of the Omics Trilogy. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263–269.
27. Shao, Y.; Le, W. Recent Advances and Perspectives of Metabolomics-Based Investigations in Parkinson's Disease. *Mol. Neurodegener.* **2019**, *14*, 3, <https://doi.org/10.1186/s13024-018-0304-2>.
28. Okuzumi, A.; Hatano, T.; Ueno, S.I.; Ogawa, T.; Saiki, S.; Mori, A.; Koinuma, T.; Oji, Y.; Ishikawa, K.I.; Fujimaki, M.; et al. Metabolomics-Based Identification of Metabolic Alterations in PARK2. *Ann. Clin. Transl. Neurol.* **2019**, *6*, 525–536, <https://doi.org/10.1002/acn3.724>.
29. Fraser, K.B.; Moehle, M.S.; Alcalay, R.N.; West, A.B. Urinary LRRK2 Phosphorylation Predicts Parkinsonian Phenotypes in G2019S LRRK2 Carriers. *Neurology* **2016**, *86*, 994–999, <https://doi.org/10.1212/WNL.0000000000002436>.
30. Cipriani, S.; Chen, X.; Schwarzschild, M.A. Urate: A Novel Biomarker of Parkinsons Disease Risk, Diagnosis and Prognosis. *Biomark. Med.* **2010**, *4*, 701–712.
31. The Parkinson Study Group SURE-PD3 Investigators. Effect of Urate-Elevating Inosine on Early Parkinson Disease Progression: The SURE-PD3 Randomized Clinical Trial. *JAMA* **2021**, *326*, 926–939, <https://doi.org/10.1001/jama.2021.10207>.
32. Zhang, B.; Warner, J.; Pinto, C.; Juric, D.; ODay, E. NMR-Metabolite-Resonance Signature to Predict HR+ Breast Cancer Patient Response to CDK4/6 Inhibitors. *J. Clin. Oncol.* **2019**, *37*, 3043, [https://doi.org/10.1200/jco.2019.37.15\\_suppl.3043](https://doi.org/10.1200/jco.2019.37.15_suppl.3043).
33. Honrao, C.; Rao, S.R.; Teissier, N.; Call, S.G.; ODay, E.M.; Janku, F. Abstract LB031: Plasma Based Metabolic Profiling in Metastatic Gastrointestinal Stromal Tumors (GIST). In Proceedings of the AACR Annual Meeting 2021, Philadelphia, PA, USA, 17–21 May 2021.

34. O'Day, E.; Leitzel, K.; Ali, S.M.; Zhang, B.; Dong, C.; Gu, H.; Shi, X.; Drabick, J.J.; Cream, L.; Vasekar, M.; et al. Abstract P4-10-25: Pretreatment serum metabolome predicts PFS in first-line trastuzumab-treated metastatic breast cancer. In *SABCS, Proceedings of the 2019 San Antonio Breast Cancer Symposium, San Antonio, TX, USA, 10–14 December 2019*; AACR Publications: Philadelphia, PA, USA, 2020.
35. Seol, W.; Kim, H.; Son, I. Urinary Biomarkers for Neurodegenerative Diseases. *Exp. Neurobiol.* **2020**, *29*, 325–333.
36. Tzoulaki, I.; Ebbels, T.M.D.; Valdes, A.; Elliott, P.; Ioannidis, J.P.A. Design and Analysis of Metabolomics Studies in Epidemiologic Research: A Primer on -Omic Technologies. *Am. J. Epidemiol.* **2014**, *180*, 129–139.
37. Johnson, C.H.; Gonzalez, F.J. Challenges and Opportunities of Metabolomics. *J. Cell. Physiol.* **2012**, *227*, 2975–2981.
38. Huang, S.; Chong, N.; Lewis, N.E.; Jia, W.; Xie, G.; Garmire, L.X. Novel Personalized Pathway-Based Metabolomics Models Reveal Key Metabolic Pathways for Breast Cancer Diagnosis. *Genome Med.* **2016**, *8*, 34, <https://doi.org/10.1186/s13073-016-0289-9>.
39. Brennan, L.; Gibbons, H. Sex Matters: A Focus on the Impact of Biological Sex on Metabolomic Profiles and Dietary Interventions. *Proc. Nutr. Soc.* **2020**, *79*, 205–209, <https://doi.org/10.1017/S002966511900106X>.
40. Wagner-Golbs, A.; Neuber, S.; Kamlage, B.; Christiansen, N.; Bethan, B.; Rennefahrt, U.; Schatz, P.; Lind, L. Effects of Long-Term Storage at  $-80^{\circ}\text{C}$  on the Human Plasma Metabolome. *Metabolites* **2019**, *9*, 99, <https://doi.org/10.3390/metabo9050099>.
41. Bingol, K.; Li, D.W.; Zhang, B.; Brüscheweiler, R. Comprehensive Metabolite Identification Strategy Using Multiple Two-Dimensional NMR Spectra of a Complex Mixture Implemented in the COLMARm Web Server. *Anal. Chem.* **2016**, *88*, 12411–12418, <https://doi.org/10.1021/acs.analchem.6b03724>.
42. Ahmed, S.S.; Santosh, W.; Kumar, S.; Christlet, H.T.T. Metabolic Profiling of Parkinson's Disease: Evidence of Biomarker from Gene Expression Analysis and Rapid Neural Network Detection. *J. Biomed. Sci.* **2009**, *16*, 63, <https://doi.org/10.1186/1423-0127-16-63>.
43. Argilés, J.M.; Campos, N.; Lopez-Pedrosa, J.M.; Rueda, R.; Rodriguez-Mañas, L. Skeletal Muscle Regulates Metabolism via Interorgan Crosstalk: Roles in Health and Disease. *J. Am. Med. Dir. Assoc.* **2016**, *17*, 789–796.
44. Kurs, M.B.; Jankowski, A.; Rudnicki, W.R. Boruta—A System for Feature Selection. *Fundam. Inform.* **2010**, *101*, 271–285, <https://doi.org/10.3233/FI-2010-288>.
45. Rajput, C.; Sarkar, A.; Sachan, N.; Rawat, N.; Singh, M.P. Is Gut Dysbiosis an Epicenter of Parkinson's Disease? *Neurochem. Res.* **2021**, *46*, 425–438.
46. Picca, A.; Calvani, R.; Landi, G.; Marini, F.; Biancolillo, A.; Gervasoni, J.; Persichilli, S.; Primiano, A.; Urbani, A.; Bossola, M.; et al. Circulating Amino Acid Signature in Older People with Parkinson's Disease: A Metabolic Complement to the EXosomes in PARKINSON Disease (EXPAND) Study. *Exp. Gerontol.* **2019**, *128*, 110766, <https://doi.org/10.1016/j.exger.2019.110766>.
47. van Kessel, S.P.; el Aidy, S. Bacterial Metabolites Mirror Altered Gut Microbiota Composition in Patients with Parkinson's Disease. *J. Parkinsons Dis.* **2019**, *9*, S359–S370.
48. Peppard, R.F.; Martin, W.R.W.; Schulzer, M.; Guttman, M.; Tsui, J.K.C.; Carr, G.D.; Phillips, A.G.; McGeer, P.L.; Grochowski, E.; Calne, D.B. Cerebral Glucose Metabolism in Parkinson's Disease With and Without Dementia. *Arch. Neurol.* **1992**, *49*, 1262–1268, <https://doi.org/10.1001/archneur.1992.00530360060019>.
49. Shao, Y.; Li, T.; Liu, Z.; Wang, X.; Xu, X.; Li, S.; Xu, G.; Le, W. Comprehensive Metabolic Profiling of Parkinson's Disease by Liquid Chromatography-Mass Spectrometry. *Mol. Neurodegener.* **2021**, *16*, 4, <https://doi.org/10.1186/s13024-021-00425-8>.
50. Wuolikainen, A.; Jonsson, P.; Ahnlund, M.; Antti, H.; Marklund, S.L.; Moritz, T.; Forsgren, L.; Andersen, P.M.; Trupp, M. Multi-Platform Mass Spectrometry Analysis of the CSF and Plasma Metabolomes of Rigorously Matched Amyotrophic Lateral Sclerosis, Parkinson's Disease and Control Subjects. *Mol. Biosyst.* **2016**, *12*, 1287–1298, <https://doi.org/10.1039/c5mb00711a>.
51. Zhao, H.; Wang, C.; Zhao, N.; Li, W.; Yang, Z.; Liu, X.; Le, W.; Zhang, X. Potential Biomarkers of Parkinson's Disease Revealed by Plasma Metabolic Profiling. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2018**, *1081–1082*, 101–108, <https://doi.org/10.1016/j.jchromb.2018.01.025>.
52. Öhman, A.; Forsgren, L. NMR Metabonomics of Cerebrospinal Fluid Distinguishes between Parkinson's Disease and Controls. *Neurosci. Lett.* **2015**, *594*, 36–39, <https://doi.org/10.1016/j.neulet.2015.03.051>.
53. Mally, J.; Szalai, G.; Stone, T.W. Changes in the Concentration of Amino Acids in Serum and Cerebrospinal Fluid of Patients with Parkinson's Disease. *J. Neurol. Sci.* **1997**, *151*, 159–162, [https://doi.org/10.1016/S0022-510X\(97\)00119-6](https://doi.org/10.1016/S0022-510X(97)00119-6).
54. Molina, J.A.; Jiménez-Jiménez, F.J.; Gómez, P.; Vargas, C.; Navarro, J.A.; Ortí-Pareja, M.; Gasalla, T.; Benito-León, J.; Bermejo, F.; Arenas, J. Decreased Cerebrospinal Fluid Levels of Neutral and Basic Amino Acids in Patients with Parkinson's Disease. *J. Neurol. Sci.* **1997**, *150*, 123–127, [https://doi.org/10.1016/S0022-510X\(97\)00069-5](https://doi.org/10.1016/S0022-510X(97)00069-5).
55. Wu, J.; Wuolikainen, A.; Trupp, M.; Jonsson, P.; Marklund, S.L.; Andersen, P.M.; Forsgren, L.; Öhman, A. NMR Analysis of the CSF and Plasma Metabolome of Rigorously Matched Amyotrophic Lateral Sclerosis, Parkinson's Disease and Control Subjects. *Metabolomics* **2016**, *12*, 101, <https://doi.org/10.1007/s11306-016-1041-6>.
56. Vascellari, S.; Palmas, V.; Melis, M.; Pisanu, S.; Cusano, R.; Uva, P.; Perra, D.; Madau, V.; Sarchioto, M.; Oppo, V.; et al. Gut Microbiota and Metabolome Alterations Associated with Parkinson's Disease. *mSystems* **2020**, *5*, e00561-20, <https://doi.org/10.1128/msystems.00561-20>.
57. Luan, H.; Liu, L.F.; Tang, Z.; Zhang, M.; Chua, K.K.; Song, J.X.; Mok, V.C.T.; Li, M.; Cai, Z. Comprehensive Urinary Metabolomic Profiling and Identification of Potential Noninvasive Marker for Idiopathic Parkinson's Disease. *Sci. Rep.* **2015**, *5*, srep13888, <https://doi.org/10.1038/srep13888>.

58. Kumari, S.; Goyal, V.; Kumaran, S.S.; Dwivedi, S.N.; Srivastava, A.; Jagannathan, N.R. Quantitative Metabolomics of Saliva Using Proton NMR Spectroscopy in Patients with Parkinson's Disease and Healthy Controls. *Neurol. Sci.* **2020**, *41*, 1201–1210, <https://doi.org/10.1007/s10072-019-04143-4>.
59. Kumari, S.; Kumaran, S.S.; Goyal, V.; Sharma, R.K.; Sinha, N.; Dwivedi, S.N.; Srivastava, A.K.; Jagannathan, N.R. Identification of Potential Urine Biomarkers in Idiopathic Parkinson's Disease Using NMR. *Clin. Chim. Acta* **2020**, *510*, 442–449, <https://doi.org/10.1016/j.cca.2020.08.005>.
60. Segura-Aguilar, J.; Paris, I.; Muñoz, P.; Ferrari, E.; Zecca, L.; Zucca, F.A. Protective and Toxic Roles of Dopamine in Parkinson's Disease. *J. Neurochem.* **2014**, *129*, 898–915.
61. Cheng, D.; Zhao, X.; Yang, S.; Cui, H.; Wang, G. Metabolomic Signature between Metabolically Healthy Overweight/Obese and Metabolically Unhealthy Overweight/Obese: A Systematic Review. *Diabetes Metab. Syndr. Obes. Targets Ther.* **2021**, *14*, 991–1010, <https://doi.org/10.2147/DMSO.S294894>.
62. Powers, R.; Lei, S.; Anandhan, A.; Marshall, D.D.; Worley, B.; Cerny, R.L.; Dodds, E.D.; Huang, Y.; Panayiotidis, M.I.; Pappa, A.; et al. Metabolic Investigations of the Molecular Mechanisms Associated with Parkinson's Disease. *Metabolites* **2017**, *7*, 22.
63. Heo, H.Y.; Park, J.M.; Kim, C.H.; Han, B.S.; Kim, K.B.S.; Seol, W. LRRK2 Enhances Oxidative Stress-Induced Neurotoxicity via Its Kinase Activity. *Exp. Cell Res.* **2010**, *316*, 649–656, <https://doi.org/10.1016/j.yexcr.2009.09.014>.
64. Hur, E.M.; Lee, B.D. Lrrk2 at the Crossroad of Aging and Parkinson's Disease. *Genes* **2021**, *12*, 505, <https://doi.org/10.3390/genes12040505>.
65. Redenšek, S.; Dolžan, V.; Kunej, T. From Genomics to Omics Landscapes of Parkinson's Disease: Revealing the Molecular Mechanisms. *OMICS J. Integr. Biol.* **2018**, *22*, 1–16.
66. Chan, R.B.; Perotte, A.J.; Zhou, B.; Liong, C.; Shorr, E.J.; Marder, K.B.S.; Kang, U.J.; Waters, C.H.; Levy, O.A.; Xu, Y.; et al. Elevated GM3 Plasma Concentration in Idiopathic Parkinson's Disease: A Lipidomic Analysis. *PLoS ONE* **2017**, *12*, e0172348, <https://doi.org/10.1371/journal.pone.0172348>.
67. Alcalay, R.N.; Levy, O.A.; Waters, C.C.; Fahn, S.; Ford, B.; Kuo, S.H.; Mazzoni, P.; Pauciulo, M.W.; Nichols, W.C.; Gan-Or, Z.; et al. Glucocerebrosidase Activity in Parkinson's Disease with and without GBA Mutations. *Brain* **2015**, *138*, 2648–2658, <https://doi.org/10.1093/brain/awv179>.
68. Hyberts, S.G.; Milbradt, A.G.; Wagner, A.B.; Arthanari, H.; Wagner, G. Application of Iterative Soft Thresholding for Fast Reconstruction of NMR Data Non-Uniformly Sampled with Multidimensional Poisson Gap Scheduling. *J. Biomol. NMR* **2012**, *52*, 315–327, <https://doi.org/10.1007/s10858-012-9611-z>.
69. Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6*, 277–293, <https://doi.org/10.1007/BF00197809>.
70. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; The R Core Team; et al. Caret: Classification and Regression Training. Available online: <https://cran.r-project.org/package=caret> (accessed on 13 December 2021).
71. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R News* **2002**, *2/3*, 18–22.
72. Thévenot, E.A.; Roux, A.; Xu, Y.; Ezan, E.; Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **2015**, *14*, 3322–3335, <https://doi.org/10.1021/acs.jproteome.5b00354>.
73. Wickham, H. Ggplot2: Elegant Graphics for Data Analysis. Available online: <https://ggplot2.tidyverse.org> (accessed on 13 December 2021).
74. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.C.; Müller, M. PROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinform.* **2011**, *12*, 77, <https://doi.org/10.1186/1471-2105-12-77>.
75. Darweesh, S.K.L.; Verlinden, V.J.A.; Stricker, B.H.; Hofman, A.; Koudstaal, P.J.; Ikram, M.A. Trajectories of Prediagnostic Functioning in Parkinson's Disease. *Brain* **2017**, *140*, 429–441, <https://doi.org/10.1093/brain/aww291>.