Neurobiology of Disease 124 (2019) 555–562

Contents lists available at ScienceDirect

Neurobiology of Disease

journal homepage: www.elsevier.com/locate/ynbdi

Integrative analysis of blood metabolomics and PET brain neuroimaging data for Parkinson's disease

Enrico Glaab[a,*], Jean-Pierre Trezzi[a,b], Andrea Greuel[c], Christian Jäger[a], Zdenka Hodak[a], Alexander Drzezga[d], Lars Timmermann[c], Marc Tittgemeyer[e,g], Nico Jean Diederich[f,1], Carsten Eggers[c,1]

[a] Luxembourg Centre for Systems Biomedicine, Esch-sur-Alzette, Luxembourg
[b] Integrated Biobank of Luxembourg, Luxembourg Institute of Health, Dudelange, Luxembourg
[c] Clinic for Neurology, University Hospital Giessen and Marburg, Marburg, Germany
[d] Department of Nuclear Medicine, University Hospital Cologne, Cologne, Germany
[e] Max Planck Institute for Metabolism Research, Cologne, Germany
[f] Department of Neurology, Centre Hospitalier du Luxembourg, Luxembourg-City, Luxembourg
[g] Cologne Cluster of Excellence in Cellular Stress and Aging-Associated Disease (CECAD), Cologne, Germany

ARTICLE INFO

Keywords:
Parkinson's disease
Metabolomics
Neuroimaging
Machine learning
PET

ABSTRACT

Background: The diagnosis of Parkinson's disease (PD) often remains a clinical challenge. Molecular neuroimaging can facilitate the diagnostic process. The diagnostic potential of metabolomic signatures has recently been recognized.

Methods: We investigated whether the joint data analysis of blood metabolomics and PET imaging by machine learning provides enhanced diagnostic discrimination and gives further pathophysiological insights. Blood plasma samples were collected from 60 PD patients and 15 age- and gender-matched healthy controls. We determined metabolomic profiles by gas chromatography coupled to mass spectrometry (GC–MS). In the same cohort and at the same time we performed FDOPA PET in 44 patients and 14 controls and FDG PET in 51 patients and 16 controls. 18 PD patients were available for a follow-up exam after one year. Both data sets were analysed by two machine learning approaches, applying either linear support vector machines or random forests within a leave-one-out cross-validation scheme and computing receiver operating characteristic (ROC) curves.

Results: In the metabolomics data, the baseline comparison between cases and controls as well as the follow-up assessment of patients pointed to metabolite changes associated with oxidative stress and inflammation. For the FDOPA and FDG PET data, the diagnostic predictive performance (DPP) in the ROC analyses was highest when combining imaging features with metabolomics data (ROC AUC for best FDOPA + metabolomics model: 0.98; AUC for best FDG + metabolomics model: 0.91). DPP was lower when using only PET attributes or only metabolomics signatures.

Conclusion: Integrating blood metabolomics data combined with PET data considerably enhances the diagnostic discrimination power. Metabolomic signatures also indicate interesting disease-inherent changes in cellular processes, including oxidative stress response and inflammation.

1. Introduction

Parkinson's disease (PD) is a complex neurodegenerative disorder, influenced by a wide variety of genetic and environmental risk factors, and characterized by heterogeneous motor and non-motor symptoms (Kalia and Lang, 2015).

Currently, only symptomatic treatments are available for PD. For the development of disease-modifying treatments, a more reliable diagnosis is regarded as an important intermediate step (Michell et al., 2004). However, to date, the commonly used clinical diagnostic criteria for PD (UK Parkinson's Disease Society Brain Bank criteria) only achieve around 76% specificity (Berg et al., 2013). The more recently refined MDS Clinical Diagnostic Criteria for PD (Postuma et al., 2015) were validated for the first time this year with 88.5% specificity, and include

* Corresponding author.
E-mail address: enrico.glaab@uni.lu (E. Glaab).
[1] Contributed equally.

https://doi.org/10.1016/j.nbd.2019.01.003
Received 31 October 2018; Received in revised form 19 December 2018; Accepted 7 January 2019
Available online 09 January 2019

a number of clinical signs that can only be properly evaluated after several years of disease duration (Postuma et al., 2018).

Previous efforts to improve the diagnosis for PD via objective molecular profiling methods (Bogdanov et al., 2008; Molochnikov et al., 2012; Scherzer et al., 2007; Chikina et al., 2015; Shamir et al., 2017; Potashkin et al., 2012; Santiago et al., 2013; Abdi et al., 2006; Ishigami et al., 2012) and neuroimaging approaches (Antonini and DeNotaris, 2004; Brooks et al., 2003; Schindlbeck and Eidelberg, 2018) have led to the identification of new measurable alteration patterns in PD; however, no sufficiently robust, reproducible, and clinically validated diagnostic biomarker signature could be derived from these studies. The challenges that may have prevented a more successful development of diagnostic biomarker models include the high heterogeneity of PD in terms of clinical phenotypes, differences in genetic background, comorbidities, lifestyle and diet; the scarce availability of samples for the early pre-symptomatic stages of PD; and a wide range of pre-analytical confounding factors including sample storage duration and time delay between sample collection, among others (Del Campo et al., 2012). Moreover, previous studies on metabolome changes in PD have mainly focused on systemic metabolic profiles alone. A combination with other modalities, such as PET, has not been performed but could provide important information about the relationship between local and systemic changes, both in terms of diagnostic usefulness and understanding of pathophysiology.

Here, instead of proposing a new biomarker signature, we investigate the added value of integrating two previously explored biomarker approaches for predictive model building: Neuroimaging data from positron emission tomography (PET) measurements using the tracers 3,4-dihydroxy-6-^{18}F-fluoro-L-phenyl-alanine (FDOPA) and 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG), and GC–MS metabolome data derived from blood plasma. PET imaging is used in clinical practice when differential diagnoses of parkinsonism are considered: FDOPA detects degeneration of dopaminergic neurons, and FDG can point towards atypical parkinsonian syndromes (Eshuis et al., 2009; Tang et al., 2010). While it provides significant pathophysiological insights (Antonini and DeNotaris, 2004), PET imaging is expensive and time-consuming. Considered alone without other types of measurements, it may still leave room for uncertainty. By contrast, metabolomics profiling of blood biospecimens is a cost-effective measurement approach requiring limited efforts by the patient. Previous metabolomics studies using blood and cerebrospinal fluid samples have shown significant metabolite alterations in PD patients compared to unaffected controls (Bogdanov et al., 2008; Trezzi et al., 2017a), but the data sets have been insufficient to assess the general predictive value of metabolite profiles for PD diagnosis, and have not been assessed for their complementary predictive value when combined with PET imaging data. So it remains unknown to which extent an integrative approach sheds further light on the pathophysiological cascade in PD.

In order to address these challenges, we present two types of analyses in this study: (1) A comparison of machine learning classification results to distinguish PD patients from controls, when using FDOPA and FDG PET imaging, GC–MS blood metabolomics data, or combination sets of both; (2) a discussion of the reported blood metabolomics group differences and changes over time in PD. The findings indicate that future combinations of PET imaging and metabolomics may enhance clinical decision algorithms.

2. Methods

2.1. Study cohort

60 PD patients were recruited at the University Hospital Cologne (UHC) and collaborating community neurologist offices. All patients were diagnosed by an experienced movement disorder specialist (CE). Additionally, 15 age- and gender-matched neurologically healthy controls were included. Their number was limited due to ethical considerations concerning the use of radioactive tracers in healthy subjects, with restrictions imposed by the local ethics committee, the federal office for radiation protection and German federal law. Criteria for participation in the study were age over 40 years and absence of dementia or conditions directly impairing brain function. In the PD group, Hoehn and Yahr (H&Y) stage I-III and tolerability of complete medical OFF-state (i.e. wash-out of dopaminergic medication, see below) were required. A subset of 18 subjects from the PD group additionally underwent a follow-up examination after one year. Clinical characteristics at baseline were similar between the whole PD group and the follow-up subgroup. The study was approved by the local ethics committee of the Medical Faculty of the University of Cologne (approval number 12–270). Every subject provided informed consent to participate in the study, in accordance with the declaration of Helsinki. Patients and controls were matched by age, gender and body mass index (BMI). Two patients had type 2 diabetes mellitus, one of them insulin-dependent, and one patient had myasthenia gravis treated with pyridostigmine and azathioprine. These patients were not excluded, but checked for potential outliers across all analyses. No other significant metabolic diseases were present. Of note, no subject was suffering from renal insufficiency. Questionnaires about alcohol use and eating behaviour revealed similar scores for both groups. For a more detailed cohort description and group comparison, see Table 1.

2.2. Biospecimen collection and processing

Blood samples were collected at UHC on the morning of PET scanning (either FDG or FDOPA, as they were performed on separate days). Venipuncture sets of the “butterfly” type (Venofix A, 0.8 mm or 0,65 mm (Braun, Germany) were used with K2-EDTA vacutainers. Samples were visually checked for hemolysis, which would result in red colouring of the samples due to increased free haemoglobin concentrations, but no hemolyzed sample was detected. For both PET exams PD patients were in a complete medical OFF-state, defined as a minimum of 12 h without levodopa, amantadine and MAO-inhibitors, and 72 h discontinuation of dopamine agonists. Any other medication was taken as usual (except insulin, used by one patient, where dose was adapted to fasting). Subjects were instructed expressively not to eat after dinner the previous night, but to keep adequate hydration through water or tea without sugar. For the FDOPA scans, 100 mg of carbidopa were administered orally to inhibit peripheral metabolization of the tracer. Of note, the blood samples were collected prior to this. Venous blood was drawn from the cubital vein in a 10 mL EDTA vacuum tube. Immediately after collection, tubes were inverted 8 times and placed on ice. Processing at UHC was performed by the same person for all subjects. Samples were centrifuged at 2000 $\times g$ for 10 min at +4 °C to

Table 1
Cohort characteristics. UPDRS part 3: United Parkinson's Disease Rating Scale, H&Y: Hoehn & Yahr scale, LEDD: levodopa equivalent daily dose, BMI: Body mass index, MMSE: Mini mental state exam, DEBQ: Dutch eating behaviour questionnaire (van Strien et al., 1986), AUDIT: alcohol use disorder identification test (Reinert and Allen, 2007). Group differences were compared using a *t*-test for independent samples, means and standard deviations, Fisher's exact test for gender distribution.

	PD patients	Unaffected controls	*P*-value
N (female/male)	60 (19/41)	15 (8/7)	0.14
Age	65.7 ± 9.0 years	65.1 ± 8.4 years	0.831
UPDRS part 3	25.1 ± 9.7	2.1 ± 2.6	0.000
H&Y stage	2.3 ± 0.4	–	–
Disease duration	5.2 ± 3.9 years	–	–
LEDD	488.0 ± 272.4	–	–
BMI	26.8 ± 4.7	24.6 ± 4.1	0.101
MMSE	28.3 ± 1.9	28.9 ± 1.0	0.288
DEBQ	68.6 ± 20.6	72.9 ± 21.1	0.472
AUDIT	3.4 ± 2.9	3.9 ± 2.5	0.501

separate the plasma, which was then homogenized by pipetting, divided into 10 matrix tubes of 500 µL each and stored at −80 °C. For transfer to the Integrated Biobank of Luxemburg (IBBL) the matrix tubes were packed in dry ice.

2.3. Metabolite extraction and GC–MS data acquisition

An internal standard mixture, consisting of U-13C5 Ribitol (c = 150 µg/mL; Omicron Biochemicals), Pentanedioic-d6 acid (c = 150 µg/mL; C/D/N Isotopes Inc.) and Tridecanoic-d25 acid (c = 100 µg/mL; C/D/N Isotopes Inc.), was prepared beforehand and used for extraction and run evaluation as well as data normalization.

Plasma samples were centrifuged at 15000 × *g* for 3 min at 4 °C, then 100 µL of the supernatant were transferred into a new 1.5 mL Eppendorf tube and 40 µL of the internal standard mix were added. First, a protein precipitation was performed, followed by a liquid-liquid extraction as detailed in the Supplementary Materials. Polar and non-polar metabolite extracts were measured via gas chromatography coupled to mass spectrometry (GC–MS). A detailed description of the metabolite extraction, the run settings and the subsequent GC–MS data acquisition can be found in the Supplementary Materials.

2.4. GC–MS metabolomics data quality control, filtering and pre-processing

All GC–MS runs were quality controlled by internal standard evaluation. In addition, pool samples were generated by mixing equal amounts of each sample, extracted identically to the samples of interest and run every 8th sample within the GC–MS sequence. Sensitivity drops and other measurement deviations are corrected by dividing (for each sample) each metabolite by the average of the corresponding metabolite in the 2 chronologically closest pools (Trezzi et al., 2017b).

All GC-MS chromatograms were processed using MetaboliteDetector, v3.020151231 (Hiller et al., 2011). The software package supports automatic deconvolution of all mass spectra. Compounds were annotated by retention time and mass spectrum using an in-house mass spectral library. The applied deconvolution settings are listed in the Suppl. Materials (Table S1). Retention index calibration was based on an C10–C40 even n-alkane mixture. Finally, the processed GC–MS metabolomics data was scaled to a mean value zero and a standard deviation of one.

2.5. PET imaging data acquisition

FDG- and FDOPA-PET scans were acquired in the same week, 1–2 days apart. Both scans were performed after overnight fasting and, for PD patients, in a medical OFF-state, as described above in the biospecimen collection section. After collection of the blood samples and one hour before injection of FDOPA, 100 mg of Carbidopa were administered orally to inhibit peripheral metabolization of the tracer. Images were recorded on a Siemens high resolution research tomograph (ECAT HRRT, Siemens CTI, Knoxville, TN, USA) with 207 transaxial image planes and a voxel size of 1.219 mm in 3D acquisition mode. Subjects lay comfortably in a supine position in a quiet room with dimmed light. Following a transmission scan for attenuation correction, 185 MBq of the respective tracer were injected. Recording started 20 s after injection and lasted 60 min for FDG and 90 min for FDOPA. An average image of minute 20–60 (FDG), respectively 30–90 (FDOPA) was generated and used for the following analysis.

2.6. PET imaging data pre-processing

Pre-processing of PET data was performed using SPM12 in Matlab (http://www.fil.ion.ucl.ac.uk/spm/software/spm12). Each subject's averaged FDG and FDOPA images were co-registered, centred on the anterior commissure, and aligned horizontally. Spatial normalization to MNI space was performed using tracer-specific templates available on the SPM website. Normalized image dimensions were 91/109/91 voxels (x/y/z) sized 2x2x2 mm. A 6 mm Gaussian smoothing filter was applied to improve the signal-to-noise ratio. To test different levels of denoising, all analyses were also tested with a further successive Gaussian smoothing of 10 mm (resulting in a rounded combined smoothing filter of 12 mm). The pre-processed data in Nifti-format was loaded into the R statistical programming software using the R-package '*Rniftilib*'. For both FDOPA and FDG PET data, a global mean normalization was applied to the pre-processed signal intensities. Due to a large number of voxels with low variance across the measured signal intensities for different subjects in the FDOPA PET dataset, only the top 1000 voxels with the highest variances were retained from this dataset for machine learning analyses. For the FDG PET data, the variances across signal intensities were more evenly spread across different voxels, and therefore only voxels with a variance of zero were filtered out.

2.7. Statistical analyses, machine learning and cross-validation

The empirical Bayes moderated t-statistic (Smyth, 2004) was applied to the pre-processed FDOPA and FDG PET imaging data in order to identify and visualize the most significantly affected brain regions in PD patients compared to controls in both datasets. A whole-brain voxel-based approach was chosen for both methods, rather than the ratio of specific vs. non-specific uptake ratio for FDOPA, thereby including extrastriatal regions. Similarly, for the GC–MS metabolomics data, dedicated statistical tests were applied for supervised feature selection, including the Welch's *t*-test to compare baseline PD vs. control samples, and the paired t-test to compare the second clinical visit for PD patients vs. the baseline visit measurements. These tests were chosen to exploit the approximate normal distribution of the data. For all statistical tests applied, the resulting *p*-values were adjusted for multiple hypothesis testing using the method by Benjamini and Hochberg (Benjamini and Hochberg, 1995).

Machine learning models for supervised sample classification were trained and tested on the PET imaging and metabolomics datasets using both a 50% random training/test set split and a leave-one-out cross-validation. Two different model building algorithms were tested and compared: A linear support vector machine (SVM) (Vapnik, 1995) to assess the separability between PD patients and controls via a linear model, and a random forest algorithm (Breiman, 1999) (RF, with 250 decision trees) to investigate whether a hyper-rectangular partitioning of the input feature space reflects the complexity of the data more adequately than a linear approach. The predictive performance was evaluated using Receiver Operating Characteristic (ROC) curves for the 50% training/test set split, and averaged accuracies for the cross-validation runs. The whole evaluation procedure was applied both to the individual datasets and to combinations of their standardized attributes, with two different Gaussian smoothing filters applied to PET imaging data (see *PET imaging data pre-processing*).

3. Results and Discussion

3.1. Baseline metabolomics differential abundance analysis (PD vs. control)

The metabolomics profiles of 71 detected metabolites in blood were compared at study baseline between 60 PD patients and 15 unaffected controls by Welch's test, and statistics for the top-ranked metabolites are reported in Table 2. Specifically, abundance levels of the unknown metabolite RI1446 showed a significant increase (FDR < 0.05). Suggestive increases (here defined as nominal *p*-value < .05 and FDR < 0.5) were observed for urea, hexadecanoic acid and dodecanoic acid, and the unknown metabolite RI1050 (see box plots in Fig. 1 and Table 2). To rule out an impact of type two subjects with diabetes mellitus on the observed group differences, we repeated the statistical analysis without the two co-morbid diabetic PD patients, obtaining the

Table 2
Fold-changes (FC), *p*-values and false-discovery rates (FDR) for the metabolites with the most significant changes between PD and controls at baseline (threshold p-value < .05).

Metabolite	FC	*P*	FDR
RI 1446	1.270	0.001	0.039
Urea	1.262	0.005	0.140
RI 1050	1.324	0.006	0.140
Hexadecanoic acid	1.256	0.030	0.371
Dodecanoic acid	1.403	0.033	0.371

same qualitative results (Supplementary Table S2 and Supplementary Fig. S3).

Interestingly, urea levels and accumulation of urea cycle related metabolites have previously been linked to oxidative stress (Zhang et al., 1999; Parmeggiani and Vargas, 2018). Specifically, urea treatment of murine renal medullary collecting duct cells in culture was associated with oxidative stress, as assessed by intracellular reduced glutathione content, and led to increased levels of the oxidative stress-responsive transcription factor Gadd153/CHOP (Zhang et al., 1999). (Zhang et al., 1999). Dehydration causing urea increases can be excluded in our study as the recruited subjects maintained normal hydration regimes. The increases of the fatty acids hexadecanoic acid and octadecanoic acid in PD patients could result from an increased fat breakdown into glycerol and fatty acids that could be either diet- or pathology-related. In the latter case, fatty acid accumulation may be associated with alterations in fatty acid oxidation, which have previously been proposed to play a key role in the pathophysiology of PD (Ruipérez et al., 2010). Specifically, the increase in fatty acid oxidation is thought to counteract progressive neuronal cell death by upregulating neuroprotective agents, such as the brain-derived neurotrophic factor (BDNF) (Burté et al., 2017). The unknown metabolites (RI1446, RI1050) cannot be identified with the currently available mass spectral libraries, and will require further elucidation. RI1050 shows only a low Pearson correlation with the levodopa equivalent dose ($r = 0.17$, $p = .15$); however, the correlation between RI1446 and the levodopa equivalent dose is statistically significant ($r = 0.28$, $p = .016$). This could reflect an association with disease severity, since RI1446 is also correlated significantly with the UPDRS part 3 motor score ($r = 0.26$, $p = .02$), but at this point we cannot exclude that RI1446 may be a by-product of dopaminergic treatment.

3.2. Longitudinal metabolomics differential abundance analysis (PD second clinical visit vs. PD at baseline)

In 18 PD patients, metabolomics profiles could be collected one year after the baseline exam. They were compared with the baseline profiles via the paired *t*-test. Out of 71 metabolites, 9 displayed suggestive changes (here defined as $p < .05$, FDR < 0.5) at the follow-up exam. The most significant changes were seen for threonic acid ($p = .0009$), glycolic acid ($p = .0017$) and iminodiacetic acid ($p = .0074$ [Table 3 and Fig. 2]).

Table 3
Fold-changes (FC), *p*-values and false-discovery rates (FDR, adjusted according to Benjamini and Hochberg (Benjamini and Hochberg, 1995)) for the top-ranked metabolites comparing baseline to the second clinical visit. Only suggestive changes with p < .05 and FDR < 0.5 were identified.

Metabolite	FC	*P*	FDR
Threonic acid	1.353	0.001	0.059
Glycolic acid	1.258	0.002	0.059
Iminodiacetic acid	1.154	0.007	0.140
Glycerol	0.642	0.008	0.140
Succinic acid	1.161	0.029	0.317
Mannose	1.148	0.030	0.317
Glyceric acid	1.229	0.031	0.317
Citric acid	1.144	0.046	0.375
RI 1708	0.876	0.048	0.375

Interestingly, recent studies in a subclinical PD cohort have indicated that patients with an ascorbic acid deficiency later on developed PD more frequently (Senarath Yapa, 1992).

Apart from this, minor changes were observed in mannose ($p = .03$, FC = 1.15, FDR = 0.32), which could result from decreased levels of mannose binding lectin (MBL) in response to inflammation in PD as discussed previously (Trezzi et al., 2017a). MBL plays a key role in innate immune system recognition of pathogenic agents. Thus, the increased free mannose found in this study might reflect perturbed immune signals in PD. This hypothesis is in line with previous studies proposing a mechanistic link between reduced MBL levels and the most common neurodegenerative disorder, Alzheimer's disease (Sjölander et al., 2013).

In a previous independent metabolomics-based study involving members of our research group, we had also identified increased threonic acid levels, but in cerebrospinal fluid (CSF) from early-stage PD patients compared to healthy control subjects from the DeNoPa cohort (Trezzi et al., 2017a). Higher levels of threonic acid, an ascorbic acid catabolite, may reflect an increased oxidative stress response in PD associated with lower dehydroascorbic acid levels. Moreover, in the previous analysis of CSF samples we also observed increased levels of mannose, as a marker of an induced inflammatory response. Although the metabolomics study presented here differs from the prior investigations on the DeNoPa cohort both in terms of the studied body fluid (blood plasma vs. CSF) and the disease stage (mid-stage PD vs. early-stage PD), the observed trends in the changes for threonic acid and mannose across the two studies may suggest that these metabolites already have increased levels in early stages of PD, which progressively increase during the course of the disease. Further longitudinal

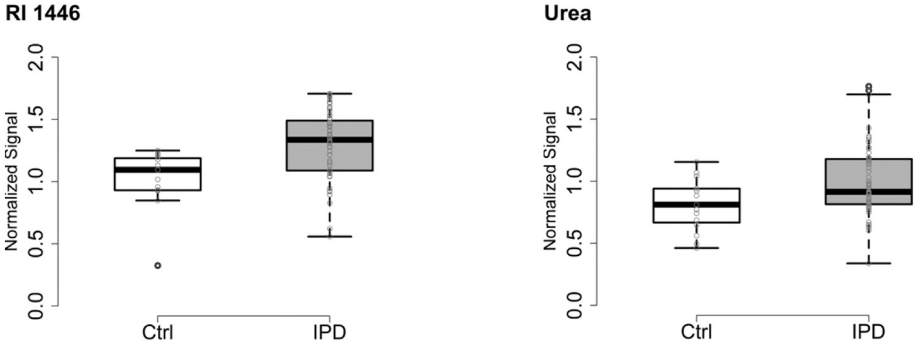

Fig. 1. Box plot visualizations of the normalized signal intensities for top-ranked differentially abundant metabolites in PD vs. unaffected controls.

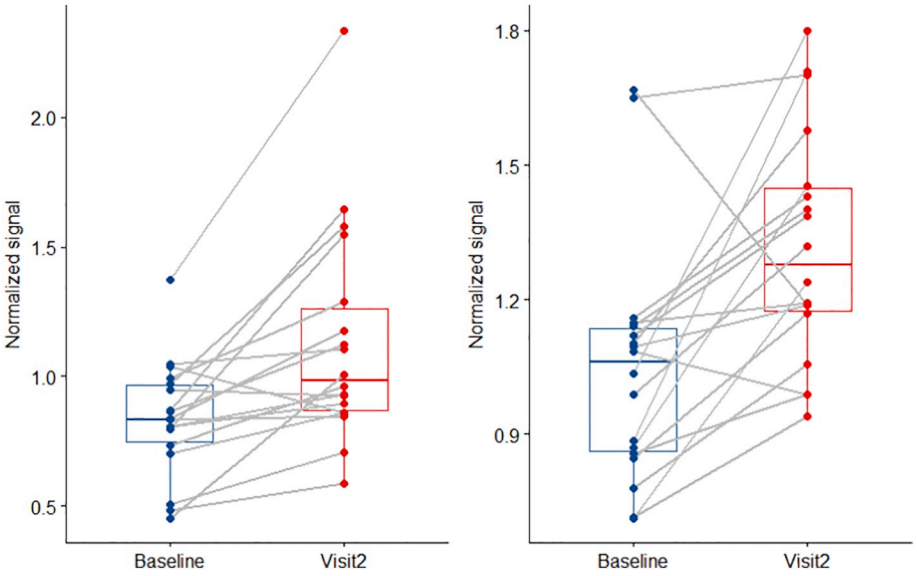

Fig. 2. Box plot visualizations for the top-ranked metabolites with differential abundance in PD patients comparing the second clinical visit to the baseline examination (left: threonic acid, right: glycolic acid). For threonic acid, one potential outlier sample was observed (top left), corresponding to a patient taking a larger than average number of medications for conditions other than PD. For the two patients who displayed decreasing threonic acid levels from baseline to visit 2, no special clinical characteristics were observed. For glycolic acid, two observed potential outlier samples (top right) did not display any specific clinical characteristics and may result from a larger than average biological variation for this metabolite.

measurements on independent cohorts and at shorter time intervals will be required to confirm these alteration trends.

3.3. Differential feature analysis of FDOPA and FDG PET imaging data

A voxel-wise group comparison was first performed between PD patients' and controls' PET signals. We used the empirical Bayes moderated t-statistic across all voxels with non-zero variance. For FDOPA, two voxel clusters located bilaterally in the posterior part of the putamen displayed significant differences after multiple hypothesis testing adjustments (FDR < 0.05, see red areas highlighted in Fig. 3). This finding is a well-known feature of PD as it reflects the loss of dopaminergic terminals projecting from the substantia nigra (SN), pars compacta (SNc) to the posterior putamen subregion of the striatum (Gallagher et al., 2011; Jokinen et al., 2009; O'Brien et al., 2004).

By application of the same method to the FDG-PET dataset, we identified two clusters in the cerebral peduncles of the lower midbrain, where the SNc is located (see green areas highlighted in Fig. 3). As degeneration of the cell bodies in SNc leads to the onset of motor symptoms (Gallagher et al., 2011), it seems intuitive to find hypometabolism in this region at a middle disease stage of our patients (Eggers et al., 2009).

3.4. Machine learning analyses

In order to investigate the predictive power of the collected data for separating PD patients from unaffected controls, machine learning models were built and evaluated. The two considered modelling approaches, support vector machines (SVM) and random forests (RF), were applied separately to the metabolomics data and the FDOPA and FDG PET imaging data, with and without applying the successive 10 mm Gaussian smoothing, as well as to the combined imaging and metabolomics data after standardization.

The average classification accuracies achieved by different models within a leave-one-out cross-validation (LOOCV) are shown in Table 4, and ROC curve visualizations of the performances for different SVM models in a 50% training/test set split of data are provided in Fig. 4.

Importantly, both for the FDOPA and FDG PET data, the predictive performance given by the area under the ROC curve (AUC) was highest when combining standardized imaging features with those from the metabolomics data (SVM AUC for FDOPA + metabolomics: 0.98; SVM AUC for FDG + metabolomics: 0.91). By contrast, the performance was lower when using only the respective PET attributes (AUC for FDOPA: 0.94, AUC for FDG: 0.8) or only the metabolomics data (AUC: 0.66). Similar results were obtained when using LOOCV runs to determine average accuracies for different modelling approaches: In most cases the highest accuracies were obtained for combinations of neuroimaging and metabolomics features as input data (FDOPA PET + metabolomics – avg. accuracy using random forest: 91.4%, FDG PET + metabolomics – avg. accuracy: 77.3%), with the exception of a random forest model for the metabolomics data, providing a slightly higher avg. accuracy (80%, see Table 1) than the combination of metabolomics with the FDG PET data, which may be explained by a larger variation of the performance measures on this dataset. The improvement of discrimination

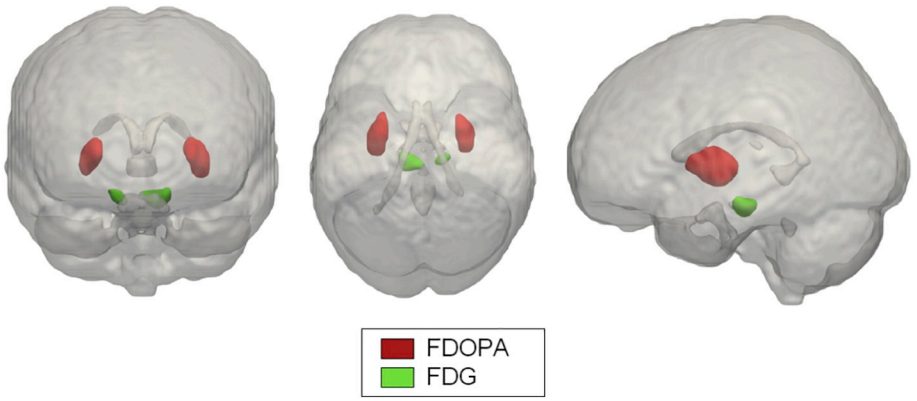

Fig. 3. Glass brain visualization highlighting the brain regions with the most significant changes (false discovery rate (FDR) < 0.05) in signal intensities in FDOPA PET (red regions) and FDG PET (green regions) measurements for PD patients compared to unaffected controls using the empirical Bayes moderated t-statistic. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Predictive performance for PD vs. control classification using FDOPA PET, FDG PET and metabolomics data, or combinations of multiple of these input datasets (see first column). Metabolomics data was available from 60 patients and 15 controls, FDOPA PET data from a subset of 44 patients and 14 controls and FDG PET data from 51 patients and 16 controls (for one control subject no biomaterial was available for metabolomics). Linear Support Vector Machines (SVM) and Random Forest classifiers with 250 trees were used for classification, with or without a successive 10 mm Gaussian smoothing of the PET datasets (see column 2, the additional Gaussian smoothing is highlighted by the abbreviation "GS"). The obtained average classification accuracy using leave-one-out cross-validation is shown in the last column.

Data	Prediction method (GS = with successive 10 mm Gaussian smoothing)	Average accuracy	SENS	SPEC	TP	TN	FP	FN
Metabolomics	SVM	69.3%	78.3%	33.3%	47	5	10	13
Metabolomics	Random Forest	80.0%	98.3%	6.6%	59	1	14	1
FDG	SVM	52.2%	64.7%	12.5%	33	2	14	18
FDG	SVM (GS)	70.1%	78.4%	43.7%	40	7	9	11
FDG	Random Forest	73.1%	94.1%	6.3%	48	1	15	3
FDG	Random Forest (GS)	74.6%	94.1%	12.5%	48	2	14	3
FDG + Metabolomics	SVM	60.6%	70.6%	26.7%	36	4	11	15
FDG + Metabolomics	SVM (GS)	71.2%	80.3%	40.0%	41	6	9	10
FDG + Metabolomics	Random Forest	77.3%	98.0%	6.7%	50	1	14	1
FDG + Metabolomics	Random Forest (GS)	74.2%	94.1%	6.7%	48	1	14	3
FDOPA	SVM	79.3%	88.6%	50%	39	7	7	5
FDOPA	SVM (GS)	81.0%	93.2%	64.3%	41	9	5	3
FDOPA	Random Forest	84.5%	88.6%	71.4%	39	10	4	5
FDOPA	Random Forest (GS)	81.0%	88.6%	57.1%	39	8	6	5
FDOPA + Metabolomics	SVM	84.5%	88.6%	71.4%	39	10	4	5
FDOPA + Metabolomics	SVM (GS)	84.5%	88.6%	71.4%	39	10	4	5
FDOPA + Metabolomics	Random Forest	91.4%	93.1%	85.7%	41	12	2	3
FDOPA + Metabolomics	Random Forest (GS)	91.4%	93.1%	85.7%	41	12	2	3

between healthy controls and IPD patients when using *combined* FDOPA PET imaging and metabolomics in contrast to FDOPA PET alone is limited. This may be due in part to a ceiling effect, since FDOPA PET by itself already reaches an AUC close to the maximum of 1 (AUC: 0.94). However, the improvement obtained by combining FDOPA and metabolomics (AUC: 0.98) shows that metabolomic profiles include independent predictive information. On average 4 misclassifications out of 100 subjects would be avoided with the improved model, which is relevant in clinical practice.

From a methodological perspective, applying the successive 10 mm Gaussian smoothing to the imaging data provided superior results compared to using only the default 6 mm smoothing in most but not all cases. Moreover, using a random forest approach for model building provided better or similar prediction results in comparison to the linear support vector machine in all cases where the same input data was used. Thus, our analyses suggest that the application of a larger Gaussian smoothing filter to neuroimaging data in combination with a random forest classification approach is the method of choice for this type of data.

Models derived from FDOPA PET data generally reached higher average cross-validated accuracies (see Table 4) and AUC values (see Fig. 4) than FDG PET derived models. Moreover, when testing a random forest model using the combined FDOPA and FDG PET features, an average accuracy of 78.9% was obtained, in between the accuracies obtained for the models trained on the separated data sources, suggesting that the predictive features from the FDOPA PET can only partly compensate for the lower predictive information content of the FDG PET data. Since FDOPA accumulates in presynaptic vesicles of dopaminergic axon terminals almost exclusively in the striatum, a clearly weakened signal is seen in the majority of mid-stage PD cases (Eshuis et al., 2009). By contrast, FDG reflects neuronal activity in general, and while there are distinct FDG changes in PD, they occur as a more

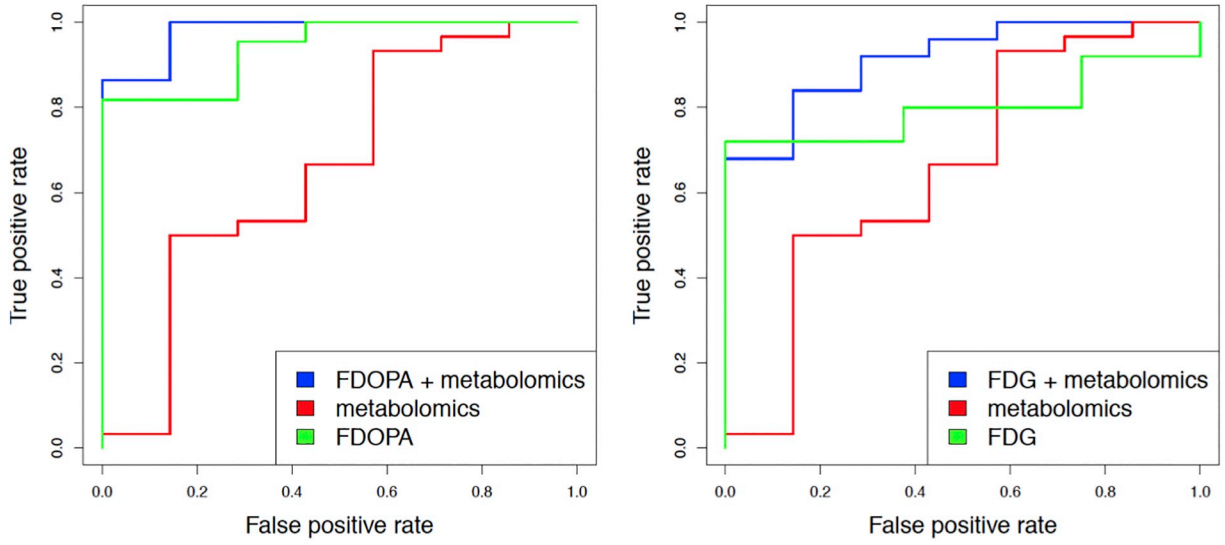

Fig. 4. ROC curve visualizations of the predictive performances for SVM models created using only PET imaging data (green, left: FDOPA PET, right: FDG PET) only metabolomics data (red), and metabolomics data combined with PET imaging data (blue, left: FDOPA PET + metabolomics, right: FDG PET + metabolomics). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

complex covariance pattern rather than absolute local changes (Brooks et al., 2003).

In order to explore the usefulness of the collected data for estimating a quantitative disease outcome, we evaluated to which extent the UPDRS part 3 (motor score), assessed in the off-medication state, could be predicted from the data (Goetz et al., 2008). As the Random Forest algorithm with 250 trees had provided the best prediction results in the previous analyses, regression models were trained to estimate the motor score sum from the FDOPA and FDG PET data. The predicted scores were compared to the real motor scores by Pearson correlation (see Suppl. Fig. S2 and S3). Significant correlations of $r = 0.43$ ($p = .0009$) for FDOPA PET and $r = 0.60$ ($p = 7.3E\text{-}06$) for FDG PET were obtained, showing that the imaging data contains significant informative value for estimating a PD patient's overall motoric performance. While the correlations were significant for both types of PET data, the relative difference between these two correlations was not significant. Moreover, applying the same regression approach to the metabolomics dataset did not result in a significant Pearson correlation with the real motor scores ($p = .49$). Independent analyses on distinct cohorts will be necessary to confirm these indicative trends.

4. Conclusions

The complex, multifactorial nature of Parkinson's disease (PD) is reflected by a wide range of measurable alteration patterns in human tissues and body fluids, as illustrated by the data presented here and in other independent studies. Most of the previous work on metabolic biomarker models for PD has focused on a single biospecimen or type of measurement. While our new analysis is based on a limited set of measurements and will require further validation on other cohorts, its integration of two very diverse data types, blood metabolomics and brain PET imaging data, is unique. The results suggest that capturing different types of PD-associated changes via highly diverse readouts can increase the predictive performance of such models and also sheds further light on the ongoing pathophysiological cascades. Specifically, the observed metabolite alterations provide pointers to the injury-response processes involved in the disease, as we found metabolite changes associated with oxidative stress and inflammation. Although to date predictors derived from metabolomics datasets are less performant than those derived from neuroimaging data, metabolomic profiling could give diagnostic clues when PET is not available or would not be practical. The scope of this study was however limited to investigating the relative difference in cross-validated predictive performance between using PET data in isolation and combining it with metabolomics. While the available sample sizes provided sufficient statistical power to identify significant differences between patients and controls in both neuroimaging and metabolomics analyses, further measurements are warranted in order to identify PD-associated alterations with smaller effect size and to substantiate the generalizability of the predictive models. Thus, in order to develop and validate a new biomarker machine learning model as an improved diagnostic test, the most stable and generalizable predictive features will still need to be identified in independent cohorts. Future extensions of this work should also include samples from subjects with clinical diagnoses similar to PD, e.g. from cohorts focusing on atypical forms of parkinsonism and essential tremor. Furthermore, the added predictive value of metabolomics data should be assessed in integrative analyses with other relevant types of neuroimaging data, e.g. dopamine transporter SPECT imaging, which is more commonly applied in practice than FDOPA PET. Combined with known PD-associated risk factors or prodromal symptoms, such as REM sleep behaviour disorder, hyposmia and constipation (Noyce et al., 2014), refined metabolomics profiling could ultimately be helpful in selecting subjects for disease modifying studies. Thus, while significant progress is still warranted for increasing the accuracy and robustness of predictive machine learning models for PD, our study demonstrates that metabolic signatures can help distinguish PD patients from controls.

Acknowledgements

The bioinformatics analyses presented in this paper were partly conducted using the HPC facilities at the University of Luxembourg (see http://hpc.uni.lu).

Funding

The study was funded by the German Research Association (DFG) in context of the Clinical Research Group 219 (KFO 219, EG350/1–1). EG received support from the Luxembourg National Research Fund (FNR) as part of the National Centre for Excellence in Research in PD (NCER-PD) and the project MitoPD, under the auspices of the bilateral e:Med program by the German Federal Ministry of Education and Research and the FNR (INTER/BMBF/13/04).

Disclosures

NJD has received research support by Fondation Think, Rotary Luxembourg and Rotary International and personal or travelling fees from Alformec (Luxembourg), BioExpert Nieder-Olm and Klinikum Koblenz (both Germany).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nbd.2019.01.003.

References

Abdi, F., Quinn, J.F., Jankovic, J., McIntosh, M., Leverenz, J.B., Peskind, E., Nixon, R., Nutt, J., Chung, K., et al., 2006. Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. J. Alzheimers Dis. 9, 293–348.

Antonini, A., DeNotaris, R., 2004. PET and SPECT functional imaging in Parkinson's disease. Sleep Med. 5, 201–206.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300.

Berg, D., Lang, A.E., Postuma, R.B., Maetzler, W., Deuschl, G., Gasser, T., Siderowf, A., Schapira, A.H., Oertel, W., et al., 2013. Changing the research criteria for the diagnosis of Parkinson's disease: obstacles and opportunities. Lancet Neurol. 12, 514–524.

Bogdanov, M., Matson, W.R., Wang, L., Matson, T., Saunders-Pullman, R., Bressman, S.S., Beal, M.F., 2008. Metabolomic profiling to develop blood biomarkers for Parkinson's disease. Brain 131, 389–396.

Breiman, L., 1999. Random forest. Mach. Learn. 45, 1–35.

Brooks, D.J., Frey, K.A., Marek, K.L., Oakes, D., Paty, D., Prentice, R., Shults, C.W., Stoessl, A.J., 2003. Assessment of neuroimaging techniques as biomarkers of the progression of Parkinson's disease. Exp. Neurol. S68–S79.

Burté, F., Houghton, D., Lowes, H., Pyle, A., Nesbitt, S., Yarnall, A., Yu-Wai-Man, P., Burn, D.J., Santibanez-Koref, M., et al., 2017. Metabolic profiling of Parkinson's disease and mild cognitive impairment. Mov. Disord. 32, 927–932.

Chikina, M.D., Gerald, C.P., Li, X., Ge, Y., Pincas, H., Nair, V.D., Wong, A.K., Krishnan, A., Troyanskaya, O.G., et al., 2015. Low-variance RNAs identify Parkinson's disease molecular signature in blood. Mov. Disord. 30, 813–821.

Del Campo, M., Mollenhauer, B., Bertolotto, A., Engelborghs, S., Hampel, H., Simonsen, A.H., Kapaki, E., Kruse, N., Le Bastard, N., et al., 2012. Recommendations to standardize preanalytical confounding factors in Alzheimer's and Parkinson's disease cerebrospinal fluid biomarkers: an update. Biomark. Med 6, 419–430.

Eggers, C., Hilker, R., Burghaus, L., Schumacher, B., Heiss, W.D., 2009. High resolution positron emission tomography demonstrates basal ganglia dysfunction in early Parkinson's disease. J. Neurol. Sci. 276, 27–30.

Eshuis, S.A., Jager, P.L., Maguire, R.P., Jonkman, S., Dierckx, R.A., Leenders, K.L., 2009. Direct comparison of FP-CIT SPECT and F-DOPA PET in patients with Parkinson's disease and healthy controls. Eur. J. Nucl. Med. Mol. Imaging 36, 454.

Gallagher, C.L., Oakes, T.R., Johnson, S.C., Chung, M.K., Holden, J.E., Bendlin, B.B., Mclaren, D.G., Xu, G., Nickles, R.J., et al., 2011. Rate of 6-[18F]fluorodopa uptake decline in striatal subregions in Parkinson's disease. Mov. Disord. 26, 614–620.

Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M.B., et al., 2008. Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. Mov. Disord. 23, 2129–2170.

Hiller, K., Metallo, C., Stephanopoulos, G., 2011. Elucidation of Cellular Metabolism Via Metabolomics and Stable-Isotope Assisted Metabolomics. Curr. Pharm. Biotechnol. 12, 1075–1086.

Ishigami, N., Tokuda, T., Ikegawa, M., Komori, M., Kasai, T., Kondo, T., Matsuyama, Y., Nirasawa, T., Thiele, H., et al., 2012. Cerebrospinal fluid proteomic patterns

discriminate Parkinson's disease and multiple system atrophy. Mov. Disord. 27, 851–857.
Jokinen, P.A., Rinne, J., Aalto, S., 2009. Impaired cerebral glucose metabolism and atrophy, but not amyloid accumulation is related to cognitive impairment in Parkinson's disease. Parkinsonism Relat. Disord. 15, 122.
Kalia, L.V., Lang, A.E., 2015. Parkinson's disease. Lancet 386, 896–912.
Michell, A.W., Lewis, S.J.G., Foltynie, T., Barker, R.A., 2004. Biomarkers and Parkinson's disease. Brain 127, 1693–1705.
Molochnikov, L., Rabey, J.M., Dobronevsky, E., Bonucelli, U., Ceravolo, R., Frosini, D., Grünblatt, E., Riederer, P., Jacob, C., et al., 2012. A molecular signature in blood identifies early Parkinson's disease. Mol. Neurodegener. 7, 26.
Noyce, A.J., Bestwick, J.P., Silveira-Moriyama, L., Hawkes, C.H., Knowles, C.H., Hardy, J., Giovannoni, G., Nageshwaran, S., Osborne, C., et al., 2014. PREDICT-PD: Identifying risk of Parkinson's disease in the community: methods and baseline results. J. Neurol. Neurosurg. Psychiatry 85, 31–37.
O'Brien, J.T., Colloby, S., Fenwick, J., Williams, E.D., Firbank, M., Burn, D., Aarsland, D., McKeith, I.G., 2004. Dopamine transporter loss visualized with FP-CIT SPECT in the differential diagnosis of dementia with lewy bodies. Arch. Neurol. 61, 919–925.
Parmeggiani, B., Vargas, C.R., 2018. Oxidative stress in urea cycle disorders: findings from clinical and basic research. Clin. Chim. Acta 477, 121–126.
Postuma, R.B., Berg, D., Stern, M., Poewe, W., Olanow, C.W., Oertel, W., Obeso, J., Marek, K., Litvan, I., et al., 2015. MDS clinical diagnostic criteria for Parkinson's disease. Mov. Disord. 30, 1591–1601.
Postuma, R.B., Poewe, W., Litvan, I., Lewis, S., Lang, A.E., Halliday, G., Goetz, C.G., Chan, P., Slow, E., et al., 2018. Validation of the MDS clinical diagnostic criteria for Parkinson's disease. Mov. Disord. https://doi.org/10.1002/mds.27362. in press.
Potashkin, J.A., Santiago, J.A., Ravina, B.M., Watts, A., Leontovich, A.A., 2012. Biosignatures for Parkinson's disease and atypical parkinsonian disorders patients. PLoS One 7, e43595.
Reinert, D.F., Allen, J.P., 2007. The alcohol use disorders identification test: an update of research findings. Alcohol. Clin. Exp. Res. 31, 185–199.
Ruipérez, V., Darios, F., Davletov, B., 2010. Alpha-synuclein, lipids and Parkinson's disease. Prog. Lipid Res. 49, 420–428.
Santiago, J.A., Scherzer, C.R., Potashkin, J.A., 2013. Specific splice variants are associated with Parkinson's disease. Mov. Disord. 28, 1724–1727.
Scherzer, C.R., Eklund, A.C., Morse, L.J., Liao, Z., Locascio, J.J., Fefer, D., Schwarzschild, M.A., Schlossmacher, M.G., Hauser, M.A., et al., 2007. Molecular markers of early Parkinson's disease based on gene expression in blood. Proc. Natl. Acad. Sci. U. S. A. 104, 955–960.
Schindlbeck, K.A., Eidelberg, D., 2018. Network imaging biomarkers: insights and clinical applications in Parkinson's disease. Lancet Neurol. 17, 629–640.
Senarath Yapa, S.C., 1992. Detection of subclinical ascorbate deficiency in early Parkinson's disease. Public Health 106, 393–395.
Shamir, R., Klein, C., Amar, D., Vollstedt, E.-J., Bonin, M., Usenovic, M., Wong, Y.C., Maver, A., Poths, S., et al., 2017. Analysis of blood-based gene expression in idiopathic Parkinson disease. Neurology 89, 1676–1683.
Sjölander, A., Minthon, L., Nuytinck, L., Vanmechelen, E., Blennow, K., Nilsson, S., 2013. Functional mannose-binding lectin haplotype variants are associated with Alzheimer's disease. J. Alzheimers Dis. 35, 121–127.
Smyth, G., 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3, 3.
van Strien, T., Frijters, J.E.R., Bergers, G.P.A., Defares, P.B., 1986. The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. Int. J. Eat. Disord. 5, 295–315.
Tang, C.C., Poston, K.L., Eckert, T., Feigin, A., Frucht, S., Gudesblatt, M., Dhawan, V., Lesser, M., Vonsattel, J.P., et al., 2010. Differential diagnosis of parkinsonism: a metabolic imaging study using pattern analysis. Lancet Neurol. 9, 149–158.
Trezzi, J.P., Galozzi, S., Jaeger, C., Barkovits, K., Brockmann, K., Maetzler, W., Berg, D., Marcus, K., Betsou, F., et al., 2017a. Distinct metabolomic signature in cerebrospinal fluid in early parkinson's disease. Mov. Disord. 32, 1401–1408.
Trezzi, J.-P., Jäger, C., Galozzi, S., Barkovits, K., Marcus, K., Mollenhauer, B., Hiller, K., 2017b. Metabolic profiling of body fluids and multivariate data analysis. MethodsX 4, 95–103.
Vapnik, V., 1995. Support vector machine. Mach. Learn. 20, 273–297.
Zhang, Z., Yang, X.Y., Cohen, D.M., 1999. Urea-associated oxidative stress and Gadd153/CHOP induction. Am. J. Phys. 276, F786–F793.