ORIGINAL ARTICLE

# A novel multi-marker discovery approach identifies new serum biomarkers for Parkinson's disease in older people: an EXosomes in PArkiNson Disease (EXPAND) ancillary study

**Riccardo Calvani** · **Anna Picca** · **Giovanni Landi** · **Federico Marini** · **Alessandra Biancolillo** · **Hélio José Coelho-Junior** · **Jacopo Gervasoni** · **Silvia Persichilli** · **Aniello Primiano** · **Annalaura Arcidiacono** · **Andrea Urbani** · **Maurizio Bossola** · **Anna Rita Bentivoglio** · **Matteo Cesari** · **Roberto Bernabei** · **Maria Rita Lo Monaco** · **Emanuele Marzetti**

**Abstract** Dopaminergic nigrostriatal denervation and widespread intracellular α-synuclein accumulation are neuropathologic hallmarks of Parkinson's disease (PD). A constellation of peripheral processes, including metabolic and inflammatory changes, are thought to contribute to neurodegeneration. In the present study, we sought to obtain insight into the multifaceted pathophysiology of PD through the application of a multi-marker discovery approach. Fifty older adults aged 70+, 20 with PD and 30 age-matched controls were enrolled as part of the EXosomes in PArkiNson Disease (EXPAND) study. A panel of 68 circulating mediators of inflammation, neurogenesis and neural plasticity, and amino acid metabolism was assayed. Biomarker selection was accomplished through sequential and orthogonalized covariance selection (SO-CovSel), a multi-platform

Riccardo Calvani and Anna Picca contributed equally to this work.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s11357-020-00192-2) contains supplementary material, which is available to authorized users.

R. Calvani · A. Picca (✉) · G. Landi · J. Gervasoni ·
S. Persichilli · A. Arcidiacono · A. Urbani · M. Bossola ·
A. R. Bentivoglio · R. Bernabei · M. R. L. Monaco (✉) ·
E. Marzetti
Fondazione Policlinico Universitario "Agostino Gemelli" IRCCS,
Rome, Italy
e-mail: anna.picca@guest.policlinicogemelli.it
e-mail: mariarita.lomonaco@policlinicogemelli.it

F. Marini · A. Biancolillo
Department of Chemistry, Sapienza Università di Roma, Rome,
Italy

A. Biancolillo
Department of Physical and Chemical Sciences, Università degli
Studi dell'Aquila, L'Aquila, Italy

H. J. Coelho-Junior · J. Gervasoni · S. Persichilli ·
A. Primiano · A. Urbani · M. Bossola · A. R. Bentivoglio ·
R. Bernabei · E. Marzetti
Università Cattolica del Sacro Cuore, Rome, Italy

H. J. Coelho-Junior
Applied Kinesiology Laboratory–LCA, School of Physical
Education, Universidade Estadual de Campinas, Campinas, SP,
Brazil

M. Cesari
Department of Clinical Sciences and Community Health,
Università di Milano, Milan, Italy

M. Cesari
Geriatric Unit, Fondazione IRCCS Ca' Granda Ospedale
Maggiore Policlinico, Milan, Italy

regression method developed to handle highly correlated variables organized in multi-block datasets. The SO-CovSel model with the best prediction ability using the smallest number of variables was built with seven biomolecules. The model allowed correct classification of $94.2 \pm 3.1\%$ participants with PD and 100% controls. The biomarker profile of older adults with PD was defined by higher circulating levels of interleukin (IL) 8, macrophage inflammatory protein (MIP)-1$\beta$, phosphoethanolamine, and proline, and by lower concentrations of citrulline, IL9, and MIP-1$\alpha$. Our innovative approach allowed identifying and evaluating the classification performance of a set of potential biomarkers for PD in older adults. Future studies are warranted to establish whether these biomolecules could serve as biomarkers for PD as well as unveil new targets for interventions.

## Introduction

Over the last years, neurological disorders have become the leading cause of disability globally (GBD 2015 Neurological Disorders Collaborator Group 2017). Among neurodegenerative conditions, Parkinson's disease (PD) has shown the fastest growth in prevalence, due to the demographic transition occurring worldwide, the longer disease duration, and the greater exposure to environmental risk factors (GBD 2016 Parkinson's Disease Collaborators 2018; Hou et al. 2019).

From a neuropathological point of view, the cardinal elements of PD are progressive dopaminergic denervation affecting the substantia nigra (SN) and widespread $\alpha$-synuclein intracellular accumulation (Dickson 2018). Clinical manifestations of the disease include paradigmatic motor symptoms (i.e., bradykinesia, rigidity, and tremor at rest) as well as a constellation of non-motor signs and symptoms (e.g., cognitive dysfunction, constipation, depression, dysautonomia, hyposmia, sleep problems) that have a major impact on overall health and quality of life of affected people (Poewe et al. 2017).

The pathophysiology of PD has yet to be clearly defined, but emerging evidence points towards a relevant role played by biological mechanisms underlying the aging process, the so-called hallmarks of aging (Hou et al. 2019). Genomic instability, telomere attrition,

epigenetic alterations, loss of proteostasis, mitochondrial dysfunction, cellular senescence, deregulated nutrient sensing, stem cell exhaustion, and altered intercellular communication contribute to increasing susceptibility to PD and may be involved in both its development and progression (Hou et al. 2019). The heterogeneous nature of PD and its multifaceted pathophysiology have hampered the definition of effective treatments for the disease that currently rely on symptomatic therapies, mainly based on dopamine replacement (Kulisevsky et al. 2018).

The development of reliable biomarkers for early diagnosis, tracking of disease progression, and monitoring of intervention effects is a critical issue to be addressed for optimal management of PD (He et al. 2018). Several clinical, biochemical, neuroimaging, and genetic parameters have been proposed as biomarkers for PD (Emamzadeh and Surguchov 2018; He et al. 2018). However, due to the intrinsic complexity of PD pathophysiology, a single biomarker may not be able to capture all possible contributors to the development of the disease and their changes over time (Chen-Plotkin et al. 2018). Recently, innovative multi-marker approaches have been proposed in order to identify the best descriptors and predictors of complex conditions, including PD (LeWitt et al. 2017; Ahmadi Rastegar et al. 2019; Picca et al. 2019b; Posavi et al. 2019). These strategies share the implementation of high-throughput analytical platforms to simultaneously measure panels of circulating mediators pertaining to different processes, coupled with multivariate statistical modeling (LeWitt et al. 2017; Ahmadi Rastegar et al. 2019; Picca et al. 2019a, 2019c; Posavi et al. 2019). As recently shown by our group, the application of this strategy allowed the identification of an amino acid signature in serum (Picca et al. 2019a) as well as a mitochondrial fingerprint in circulating small extracellular vesicles in older adults with PD (Picca et al. 2020). However, the relationship among mediators of inflammation, neurogenesis, neural plasticity, and amino acid metabolism that represent critical processes that may precede and contribute to PD (Hatano et al. 2016; Qin et al. 2016; Kim et al. 2018; Rahmani et al. 2019) remains unexplored. Therefore, we complemented the findings from previous reports of the EXosomes in PArkiNson Disease (EXPAND) study (Picca et al. 2019a, 2020) with measures of markers of inflammation, neurogenesis, and neural plasticity. Biomarker selection was accomplished through a novel multi-platform

regression method called sequential and orthogonalized covariance selection (SO-CovSel) (Biancolillo et al. 2019). SO-CovSel is a variable selection method especially suited for the investigation of biomarkers for complex conditions, such as PD, due to its ability to handle highly correlated variables organized in multi-block datasets (Biancolillo et al. 2019).

The implementation of such an innovative strategy may lead to the identification of new biomarkers for PD and help characterize older people with specific PD subtypes, predict the disease course, and devise possible personalized treatments.

## Methods

### Study population

Analyses were conducted in a convenience sample of 50 older adults enrolled as part of the EXosomes in PArkiNson Disease (EXPAND) study (Picca et al. 2019c). Prior to enrolment, all participants provided written informed consent. The study was conducted in agreement with legal requirements and international norms (Declaration of Helsinki 1964). The study protocol was approved by the Ethics Committee of the Università Cattolica del Sacro Cuore, Rome, Italy (protocol no. 0045298/17).

Participants were recruited at the Fondazione Policlinico Universitario "Agostino Gemelli" IRCCS (Rome, Italy), as described elsewhere (Picca et al. 2019a). Briefly, cases were men and women aged 70+ diagnosed with PD according to the Queen Square Brain Bank criteria (Hughes et al. 1992), under stable dopaminergic therapy for at least 1 month prior to enrollment. Controls were age-matched people without any signs of Parkinsonism, potential premotor symptoms, or family history of PD. For both cases and controls, progressive neurological diseases and cognitive impairment (i.e., Mini Mental State Examination score < 24/30) were considered exclusion criteria.

### Participant assessment

Age, sex, comorbid conditions, and medications were recorded for all participants. The Unified Parkinson's Disease Rating Scale (UPDRS) (Antonini et al. 2013) and the Hoehn and Yahr staging scale (Martinez-Martin 2010) during the "on" state were used to evaluate

clinical characteristics of PD participants. The levodopa equivalent daily dose (LEDD, mg) was calculated according to published conversion factors for individual antiparkinsonian drugs (Tomlinson et al. 2010).

### Blood sampling and processing

Blood samples were obtained in the morning after overnight fasting. Following 30 min clotting at room temperature, samples were centrifuged at $1000 \times g$ for 15 min at 4 °C. The serum fraction was collected, aliquoted, and stored at −80 °C until analysis.

### Measurement of candidate serum biomarkers

A biomarker panel was designed based on previous studies by our group in older adult populations (Marzetti et al. 2019; Picca et al. 2019a, 2019d, 2020). Ultraperformance liquid chromatography/mass spectrometry (UPLC/MS) was applied to serum samples to measure the concentration of 37 amino acids and derivatives, as described previously (Calvani et al. 2018; Picca et al. 2019a).

Twenty-seven inflammatory markers, growth factors, and chemokines were measured simultaneously in serum through a magnetic bead–based immunoassay kit on a Bio-Plex® System with Luminex xMap Technology (BioRad Laboratories, Inc., Hercules, CA, USA), as detailed elsewhere (Marzetti et al. 2019). The panel included interleukin (IL) 1β, IL1 receptor agonist (IL1-ra), IL2, IL4, IL5, IL6, IL7, IL8, IL9, IL10, IL12, IL13, IL15, IL17, fibroblast growth factor (FGF) β, granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), interferon (IFN) γ, C-C motif chemokine ligand (CCL)2 (also known as monocyte chemoattractant protein 1 (MCP-1)), CCL3 (also known as macrophage inflammatory protein 1-α (MIP-1α)), CCL4 (also known as MIP-1β), CCL5 (also known as Regulated on Activation, Normal T Cell Expressed and Secreted (RANTES)), CCL11 (also known as eutaxin), C-X-C motif chemokine ligand (CXCL) 10 (also known as IFN-γ-induced protein 10 (IP10)), platelet-derived growth factor (PDGF) bb, and tumor necrosis factor (TNF) α. C-reactive protein (CRP), myeloperoxidase (MPO), FGF21, and brain-derived neurotrophic factor (BDNF) were assayed by commercially available kits on an ELLA™ automated immunoassay system (Bio-Techne, San Jose, CA, USA).

## Statistical analysis

### Descriptive statistics

The normal distribution of data was ascertained through the Kolmogorov-Smirnov test. Comparisons between PD and control participants for normally distributed continuous variables were performed by $t$ test statistics. The non-parametric Mann-Whitney $U$ test was applied to assess differences for non-normally distributed continuous data. Differences in categorical variables between groups were determined via $\chi^2$ statistics. All descriptive analyses were performed using the GraphPrism 5.03 software (GraphPad Software, Inc., San Diego, CA), with statistical significance set at $p < 0.05$.

### SO-CovSel analysis

SO-CovSel analysis was run under MATLAB R2015b environment (The Mathworks, Natick, MA, USA), by means of in-house written functions (freely downloadable at www.chem.uniroma1.it/romechemometrics/research/algorithms/). After import into MATLAB, serum concentrations of candidate biomarkers were organized into three matrices according to their biological pathways (Table 1).

The identification of putative biomarkers coupled with classification modeling of differences between participants with and without PD was accomplished through the SO-CovSel algorithm (Biancolillo and Næs 2019; Biancolillo et al. 2019). The latter is a multi-block regression algorithm which couples a very parsimonious variable selection step with sequential modeling of the contribution from different blocks of data. If a dummy response variable coding for category belonging is used, the SO-CovSel can also be used for classification, as it was done in the present study. To discriminate between participants with and without PD, a binary-coded response vector $\mathbf{y}$ was used (i.e., "1" for PD and "0" for non-PD).

When multiple blocks of data $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_b$ are collected, the SO-CovSel algorithm operates as follows:

1. A reduced set of predictors is extracted from the matrix $\mathbf{X}_1$ by applying the CovSel procedure. The CovSel selects the first variable as the one showing maximum covariance with the response $\mathbf{y}$; then, it orthogonalizes both $\mathbf{X}_1$ and $\mathbf{y}$ relative to the selected variable and iterates the procedure on the orthogonalized $\mathbf{X}_1$ and $\mathbf{y}$ until the desired number of variables is extracted.

2. The first regression model between the selected variables in the first block $\mathbf{X}_{1\mathrm{sel}}$ and the responses is built as:

$$\mathbf{y} = \mathbf{X}_{1\mathrm{sel}}\mathbf{b}_1 + \mathbf{e}_1 \tag{1}$$

$\mathbf{b}_1$ and $\mathbf{e}_1$ being the regression vector and the residual of the first regression, respectively.

3. The second block is then orthogonalized relative to the variables selected from the first block

$$\mathbf{X}_{2\mathrm{Orlh}} = \left(\mathbf{I} - \mathbf{X}_{1\mathrm{sel}}\left(\mathbf{X}_{1\mathrm{sel}}^{\mathrm{T}}\mathbf{X}_{1\mathrm{sel}}\right)^{-1}\mathbf{X}_{1\mathrm{sel}}^{\mathrm{T}}\right)\mathbf{X}_2 \tag{2}$$

4. A reduced set of predictors is extracted from the orthogonalized second block by applying the CovSel algorithm to $\mathbf{X}_{2\mathrm{Orlh}}$ and $\mathbf{e}_1$.

5. The second regression model between the selected variables in the second orthogonalized block $\mathbf{X}_{2\mathrm{Orth,sel}}$ and the $\mathbf{y}$-residual responses is built as:

$$\mathbf{e}_1 = \mathbf{X}_{2\mathrm{Orth,sel}}\mathbf{b}_2 + \mathbf{e}_2 \tag{3}$$

6. The procedure continues until all of the blocks have been used to build the model.

If classification is the final aim, a linear discriminant analysis (LDA) model can be built either on the selected variables from the different blocks or on the matrix of predicted responses.

Operationally, the optimal combination of variables from the different blocks, which could as well involve the possibility of a block not contributing any predictor (i.e., being completely redundant), was selected as the one leading to the lowest error in internal validation. In particular, to allow unbiased estimate of the predictive ability of the models and, at the same time, selection on an internal validation set, a repeated double cross-validation (rDCV) procedure was adopted. In the rDCV, the dataset is first split into a certain number of cancelation groups which, in turn, are left out as external validation set (outer loop). The remaining individuals are further split according to a second (internal) cross-validation loop, which is used for model selection. For the procedure not to rely on a particular splitting scheme, it was repeated 50 times and the results were averaged. Moreover, to further rule out the possibility of good results stemming from chance correlations, the

**Table 1** Composition of the multi-block dataset used for sequential and orthogonalized covariance selection analysis

| Data block | Biological pathway(s) | Variables |
|---|---|---|
| Matrix 1 | Inflammation | CCL5, CCL11, CRP, FGF-β, GM-CSF, IFN-γ, IL1β, IL1-ra, IL4, IL6, IL8, IL9, IL17, IP-10, MCP-1, MIP-1α, MIP-1β, MPO, PDGF bb, TNF-α |
| Matrix 2 | Protein/amino acid metabolism | 1-methylhistidine, 3-methylhistidine, 4-hydroxyproline, α-aminobutyric acid, β-alanine, β-aminobutyric acid, γ-aminobutyric acid, alanine, aminoadipic acid, arginine, asparagine, aspartic acid, citrulline, cystine, ethanolamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, ornithine, phenylalanine, phosphoethanolamine, proline, sarcosine, serine, taurine, threonine, tryptophan, tyrosine, valine |
| Matrix 3 | Neurogenesis/neural plasticity | BDNF, FGF21 |

*BDNF*, brain-derived neurotrophic factor; *CCL*, C-C motif chemokine ligand; *CRP*, C-reactive protein; *FGF*, fibroblast growth factor; *GM-CSF*, granulocyte-macrophage colony-stimulating factor; *IFN*, interferon; *IL*, interleukin; *IL1-ra*, IL1 receptor agonist; *IP*, interferon γ-induced protein; *MCP-1*, monocyte chemoattractant protein 1; *MIP*, macrophage inflammatory protein; *MPO*, myeloperoxidase; *PDGF*, platelet-derived growth factor; *TNF-α*, tumor necrosis factor α

significance of the classification accuracy was estimated by comparing the values of the number of misclassifications and of the area under the receiver operating characteristic (AUROC) curve with their respective distributions under the null curve estimated by means of permutation tests with 10,000 randomizations.

## Results

### Characteristics of the study population

Analyses were conducted in 20 older adults with PD (mean age $73.1 \pm 10.2$ years) and 30 age- and sex-matched control participants (mean age $74.6 \pm 4.3$ years). Demographic and clinical characteristics of study participants are listed in Table 2. The two groups were comparable for age, sex distribution, and number of comorbid conditions and medications.

As for the clinical characteristics of participants with PD, eight were stage 1, two stage 2, five stage 3, and five stage 4 according to the Hoehn and Yahr staging scale. The mean LEDD was $564.4 \pm 270.8$ mg, while the average disease duration was $98.9 \pm 69.5$ months.

### Selection of candidate biomarkers by SO-CovSel analysis

Circulating levels of 68 inflammatory cytokines, growth factors, neurogenesis and neural plasticity mediators, and amino acids and derivatives were measured through multiple analytical platforms. Of the assayed biomolecules,

the concentrations of anserine, carnosine, cystathionine, phosphoserine, IL2, IL5, IL7, IL10, IL12, IL13, IL15, IL17, and G-CSF were below the detection limit. Several SO-CovSel models were built using a multi-matrix dataset containing 55 candidate biomarkers. Among the various models tested, the one that allowed the best classification of PD and control participants with the smallest number of variables was selected. This criterion was selected to fulfill requirements in biomarker discovery process for implementation in the clinical arena (Sprott 2010; Justice et al. 2018). The most "parsimonious" model was built using only seven biomarkers (Table 3).

The rate of correct classification was $94.2 \pm 3.1\%$ for participants with PD and 100% for controls ($97.7 \pm 1.2\%$ in the whole study population), with an average AUROC curve very close to 1. When compared with

**Table 2** Main characteristics of study participants

| | PD (n = 20) | Controls (n = 30) | p |
|---|---|---|---|
| Age, years (mean ± SD) | 73.1 ± 10.2 | 74.6 ± 4.3 | 0.1067 |
| Sex (female), n (%) | 9 (45) | 16 (53) | 0.4280 |
| Number of disease conditions (mean ± SD)[1] | 3.1 ± 1.6 | 2.9 ± 2.0 | 0.8046 |
| Number of medications (mean ± SD)[2] | 3.5 ± 1.3 | 2.8 ± 1.7 | 0.1034 |

[1] Includes hypertension, coronary artery disease, prior stroke, peripheral vascular disease, diabetes, chronic obstructive pulmonary disease, and osteoarthritis

[2] Includes prescription and over-the-counter medications

*PD*, Parkinson's disease; *SD*, standard deviation

**Table 3** Serum concentrations of discriminant biomolecules as resulted from sequential and orthogonalized covariance selection analysis. Italicized inputs indicate amino acids. Bold inputs correspond to inflammatory mediators

|  | PD ($n = 20$) | Controls ($n = 30$) |
|---|---|---|
| *Citrulline (µmol/L)* | *28.5 (9.5)* | *37.6 (17.7)* |
| *Phosphoethanolamine (µmol/L)* | *2.9 (1.8)* | *1.4 (1.0)* |
| *Proline (µmol/L)* | *295.7 (142.5)* | *194.8 (62.2)* |
| **IL8 (pg/mL)** | **14.0 (8.1)** | **6.6 (1.7)** |
| **IL9 (pg/mL)** | **96.6 (3.5)** | **113.3 (44.4)** |
| **MIP-1α (pg/mL)** | **2.5 (3.1)** | **13.5 (1.9)** |
| **MIP-1β (pg/mL)** | **176.8 (17.9)** | **87.4 (52.0)** |

Data are shown as median (interquartile range)

*IL*, interleukin; *MIP*, macrophage inflammatory protein; *PD*, Parkinson's disease

their distributions under the null hypothesis, all of the classification figures of merit were statistically significant ($p < 0.0001$). The SO-CovSel model chosen had higher predictive value than any selected biomarkers alone or in combination, although MIP-1α and MIP-1β alone showed remarkable classification performance (Table 4).

Among the candidate PD biomarkers selected by the SO-CovSel model, circulating levels of the amino acids phosphoethanolamine and proline were higher in participants with PD, while those of citrulline were higher in controls. The inflammatory profile of participants with PD was characterized by higher serum levels of IL8 and MIP-1β, and lower concentrations of IL9 and MIP-1α (Fig. 1 and Table 3). Unexpectedly, neurotrophic factors were not selected by the CovSel analysis. A preliminary analysis was conducted to explore possible relationships between discriminant biomarkers and age or disease stage in participants with PD (Figs. 2 and 3). Circulating levels of biomolecules selected by the SO-CovSel analysis were not significantly influenced by age. Serum concentrations of phosphoethanolamine and proline tended to be lower with increasing disease severity (Fig. 3). However, the small study sample does not allow assessing the actual validity of this finding, which merits further investigation.

## Discussion

The implementation of new high-throughput technologies and associated computational modeling, the so-

called omics revolution, has allowed investigating the pathophysiology of complex conditions in unprecedented detail (Loscalzo and Barabasi 2011; Chen and Snyder 2013). This marks a significant step towards the achievement of a true personalized medicine and the redefinition of healthcare pathways, from conventional symptom-oriented paradigms to individualized, multi-targeted approaches (Chen and Snyder 2013). In this scenario, the definition of reliable biomarkers to be used in clinical practice has become of utmost importance (Justice et al. 2018).

In this proof of concept study, we applied an innovative multi-marker approach to identify and evaluate the classification performance of a set of potential biomarkers for PD in older adults. The main finding of the present investigation was that the SO-CovSel-based analytical strategy was able to (almost) perfectly distinguish people with PD from controls using only seven circulating markers (Table 3). Participants with PD were characterized by a specific inflammatory profile, defined by higher levels of IL8 and MIP-1β, and lower levels of IL9 and MIP-1α. This peripheral inflammatory phenotype may represent a peculiar signature of inflamm-aging in PD. Low-grade, chronic, systemic inflammation is a hallmark of aging and age-related conditions, including neurodegeneration (Giunta 2006; Hou et al. 2019). A number of studies showed the existence of specific inflammatory perturbations in both sporadic and familial forms of PD (Collins et al. 2012; Deleidi and Gasser 2013; Qin et al. 2016). Moreover, alterations in peripheral immune cell repertoire and function have been described in PD, including a T helper (Th) 1-based adaptive immune response and dysregulation in circulating monocyte response and composition (Grozdanov et al. 2014; Kustrimovic et al. 2018).

Notably, MIP-1α and IL8 are among the cyto/chemokines aberrantly expressed by peripheral mononuclear cells at baseline and following stimulation by lipopolysaccharide in persons with PD (Reale et al. 2009). Further to this, circulating levels of MIP-1α, MIP-1β, and IL8 were associated with motor symptom severity, depression, and functional status, and predicted their changes over time in a longitudinal study of a well-characterized cohort of people with PD (Ahmadi Rastegar et al. 2019). MIP-1α and MIP-1β showed remarkable predictive capacity also when considered alone (Table 4), which supports their value as possible biomarkers for PD. Evidence from animal models

**Table 4** Discriminant power of individual biomarkers identified by sequential and orthogonalized covariance selection. Biomarker performance was estimated by linear discriminant analysis on single predictors coupled with 50 runs of repeated double cross-validation (rDCV). Data are shown as percentage of classification accuracy (mean ± standard deviation) on the outer (external validation) rDCV loop

|                     | PD ($n = 20$)   | Controls ($n = 30$) | Total          |
|---------------------|-----------------|---------------------|----------------|
| Citrulline          | $82.3 \pm 2.9$  | $66.3 \pm 1.0$      | $72.7 \pm 1.3$ |
| Phosphoethanolamine | $62.3 \pm 2.9$  | $82.8 \pm 1.5$      | $74.6 \pm 1.2$ |
| Proline             | $65.0 \pm 0.0$  | $80.0 \pm 2.0$      | $74.0 \pm 1.2$ |
| IL8                 | $40.7 \pm 4.3$  | $99.2 \pm 1.4$      | $75.8 \pm 1.8$ |
| IL9                 | $99.8 \pm 0.9$  | $53.4 \pm 2.7$      | $72.0 \pm 1.7$ |
| MIP-1$\alpha$       | $90.0 \pm 0.0$  | $100.0 \pm 0.0$     | $96.0 \pm 0.0$ |
| MIP-1$\beta$        | $100.0 \pm 0.0$ | $86.2 \pm 1.4$      | $91.7 \pm 0.9$ |

*IL*, interleukin; *MIP*, macrophage inflammatory protein; *PD*, Parkinson's disease

suggests that MIP-1$\alpha$ and MIP-1$\beta$ are involved in neurodegeneration, in particular in neuronal regenerative/repair activities following different insults (Perrin et al. 2005; Pattarini et al. 2007). IL9 is a pleiotropic cytokine with multiple activities on several cell types both in the immune compartment and in the central nervous system (CNS) (Elyaman and Khoury 2017). IL9-producing Th9 cells may be involved both in neurodegeneration and autoimmune CNS diseases (Saresella et al. 2011; Elyaman and Khoury 2017; Sommer et al. 2017). Interestingly, IL9 may be either pro-inflammatory or regulatory depending on the context and the source of the producing cells, and may serve neuroprotective and repair functions (Elyaman et al. 2009; Elyaman and Khoury 2017). Collectively, our findings suggest a dysregulation in peripheral immune cell activity in PD, thus corroborating previous evidence (Grozdanov et al. 2014; Kim et al. 2018). The reduced concentrations of IL9 and MIP-1$\alpha$ may also be indicative of impaired neuroprotection/repair capacity in older adults with PD.

Three amino acids concurred with the abovementioned inflammatory biomolecules to define the PD multi-marker profile. Serum levels of phosphoethanolamine and proline were higher in older participants with PD, while those of citrulline were higher in controls. Phosphaethanolamine is a primary amine involved in phospholipid turnover as a precursor of phosphatidylethanolamine and
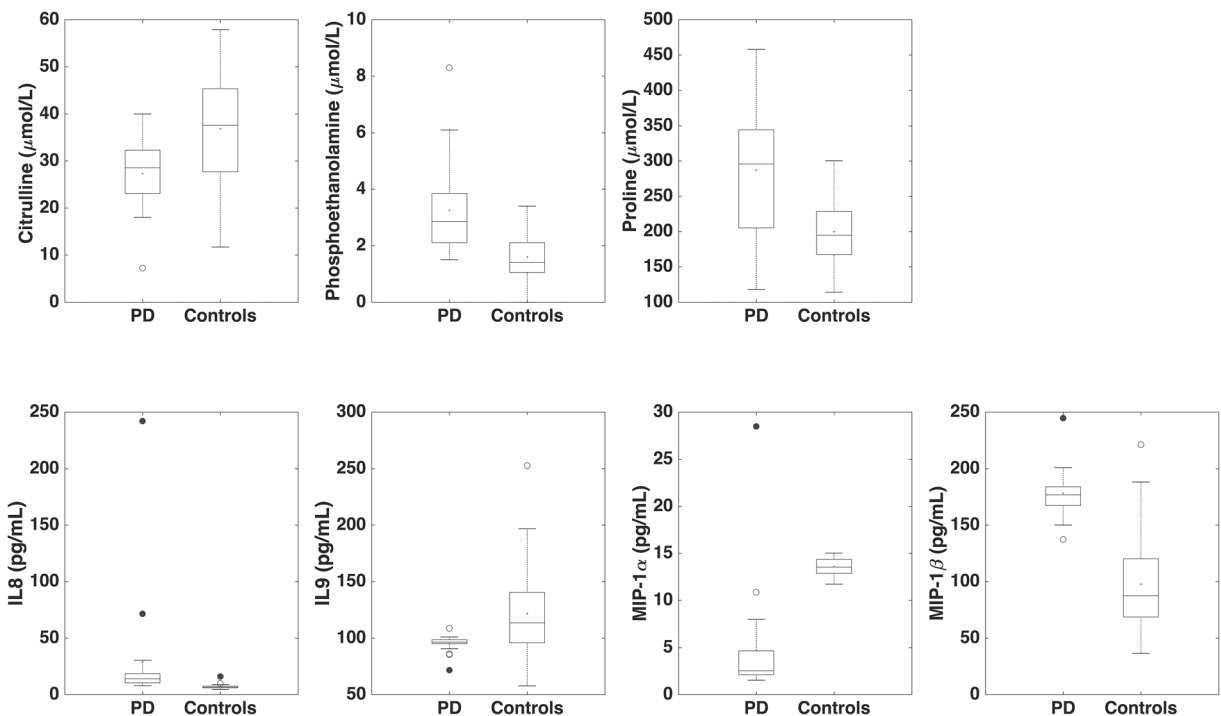


**Fig. 1** Comparison of serum levels of biomarkers identified by sequential and orthogonalized covariance selection analysis between participants with Parkinson's disease (PD) and controls. The box boundaries closest to and farthest from zero indicate the 25th and 75th percentiles, respectively. The black line within the box marks the median. Whiskers above and below the box indicate 95% confidence interval for the median. IL, interleukin; MIP, macrophage inflammatory protein
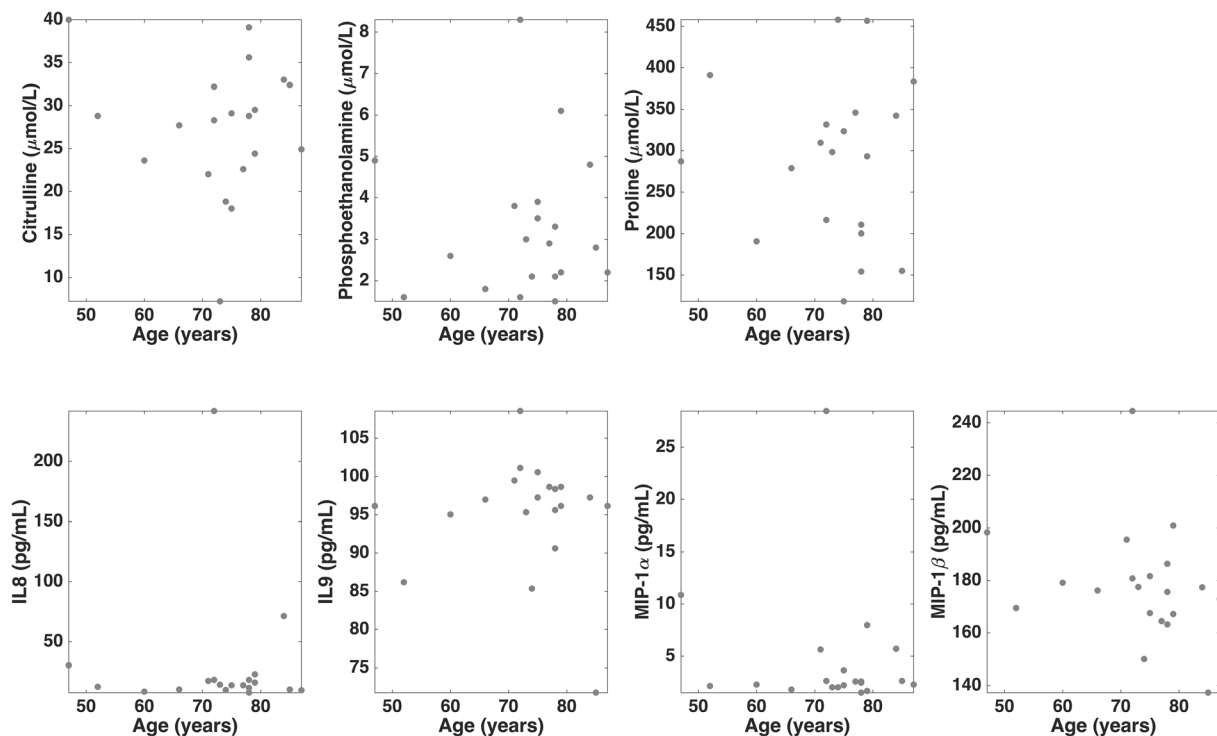
**Fig. 2** Distribution of serum levels of biomarkers identified by sequential and orthogonalized covariance selection analysis according to age in participants with Parkinson's disease. IL, interleukin; MIP, macrophage inflammatory protein

phosphatidylcholine (Patel and Witt 2017). As such, phosphoethanolamine has a remarkable role in influencing membrane composition and topology, promoting membrane fusion at the cellular and organellar level, modulating the activity of several respiratory complexes, and regulating mitochondrial quality control (Calzada et al. 2016). Interestingly, binding of $\alpha$-synuclein to lipids seems to be necessary for its pathophysiological roles (Davidson et al. 1998; Alecu and Bennett 2019). Hence, dysregulation in (phospho)lipid metabolism may promote aggregation and propagation of $\alpha$-synuclein (Lashuel et al. 2013; Alecu and Bennett 2019). Alterations in circulating glycerophospholipids, sphingolipids, and precursors (including the phosphoethanolamine companion ethanolamine) were recently found in a metabotyping study in early-disease-state PD patients (Stoessel et al. 2018). Our finding of increased circulating levels of phosphoethanolamine may be indicative of a perturbation in the CDP-ethanolamine pathway, the major route of phosphatidylethanolamine production (van der Veen et al. 2017).

Proline and citrulline are crucial intermediate at the crossroads of urea cycle, nitric oxide synthesis, glutamine, and polyamine metabolism (Caldwell et al. 2015).

A metabolomics-based pathway enrichment analysis highlighted the relevance of proline metabolism in all major neurodegenerative diseases (Kori et al. 2016). Moreover, citrulline may have neuroprotective effects through its action on arginine/nitric oxide metabolism (Yabuki et al. 2013). The presence of the dyad proline-citrulline among the discriminant predictors suggests a perturbation in nitrogen homeostasis in PD that may be associated with impaired neuroprotection (Breuillard et al. 2015; Caldwell et al. 2015).

Unexpectedly, serum BDNF levels did not differ between participants with PD and controls. BDNF is a neurotrophin that has pleiotropic activities, including promotion of survival and function of dopaminergic neurons in the SN (Mercado et al. 2017). A recent systematic review and meta-analysis showed lower circulating concentrations of BDNF in people with PD (Rahmani et al. 2019). However, factors such as antiparkinsonian medicaments and compensatory mechanisms may increase blood BDNF levels in PD patients, which could partly explain the lack of association in our sample (Okazawa et al. 1992; Scalzo et al. 2010; Rahmani et al. 2019).

Although reporting novel findings, the present study has some limitations that should be acknowledged.
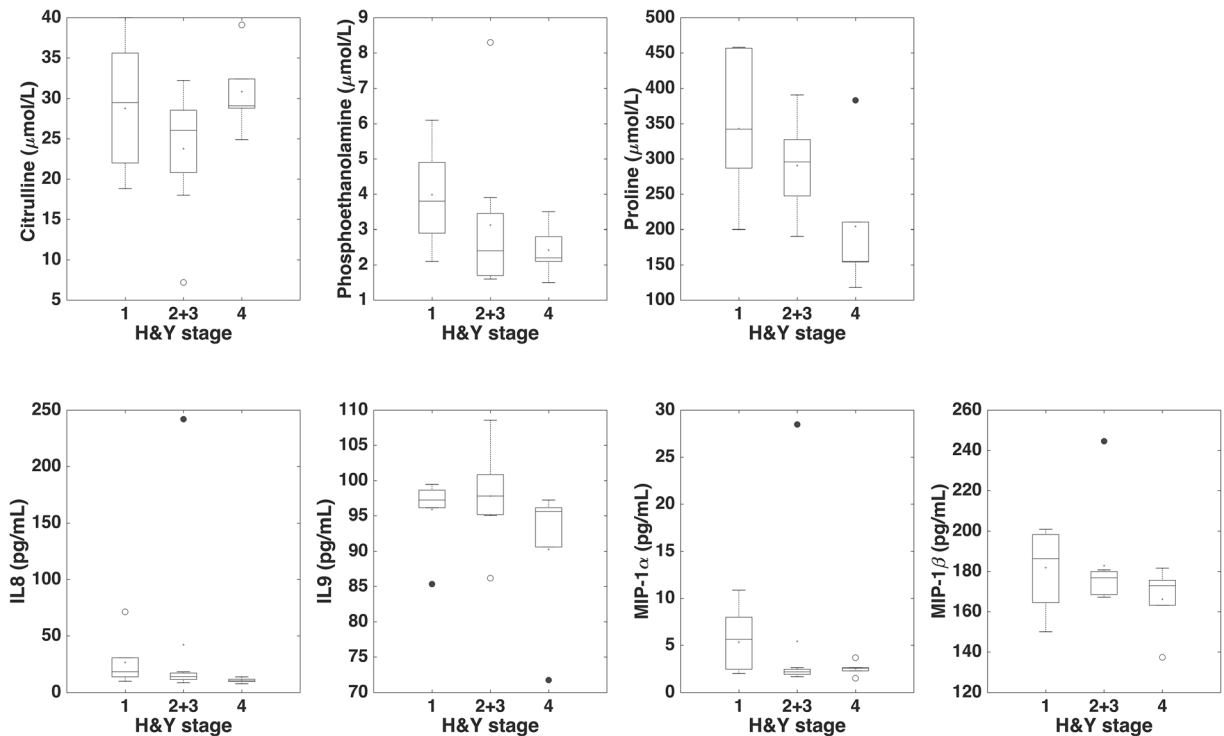
**Fig. 3** Comparison of serum levels of biomarkers identified by sequential and orthogonalized covariance selection analysis in participants with Parkinson's disease according to Hoehn and Yahr (H&Y) rating scale. The box boundaries closest to and farthest from zero indicate the 25th and 75th percentiles, respectively. The black line within the box marks the median. Whiskers above and below the box indicate 95% confidence interval for the median. IL, interleukin; MIP, macrophage inflammatory protein

First, the study population was accurately selected and well-characterized from a clinical point of view; yet, the sample size was quite limited and experimental variables numerous. Nonetheless, the SO-CovSel is especially suited for experimental settings in which variables are interdependent and organized in multiblock datasets (Biancolillo and Næs 2019; Biancolillo et al. 2019). Only Caucasian people were enrolled; thus, validation in other ethnic groups is needed to allow generalization of our findings. Likewise, future investigations are warranted to test the predictive value of the selected biomarkers in other cohorts. Dopaminergic medications and aromatic L-amino decarboxylase inhibitors may directly or indirectly influence amino acid concentrations in biofluids, as well as exert immunomodulatory actions (Beck et al. 2004; Havelund et al. 2017; Figura et al. 2018). To reduce the possibility that antiparkinsonian therapies might affect the concentrations of assayed biomolecules, only patients under stable dopaminergic treatment for at least 1 month were included.

Finally, the cross-sectional design of the present investigation did not allow inferring temporal relationships between selected mediators and PD.

In conclusion, our innovative SO-CovSel-based approach allowed selecting a panel of candidate biomarkers for PD and obtaining new hints into PD pathophysiology. The implementation of such an innovative multi-marker analytical strategy in longitudinal cohorts could allow validating novel sets of biomarkers for PD and identifying new targets for interventions.

**Code availability**  Not applicable.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethics approval** The study was conducted in agreement with legal requirements and international norms (Declaration of Helsinki 1964). The study protocol was approved by the Ethics Committee of the Università Cattolica del Sacro Cuore, Rome, Italy (protocol no. 0045298/17).

**Consent to participate** Prior to enrolment, all participants provided written informed consent.

**Consent for publication** Not applicable.

# References

Ahmadi Rastegar D, Ho N, Halliday GM, Dzamko N. Parkinson's progression prediction using machine learning and serum cytokines. NPJ Parkinsons Dis. 2019;5:14. https://doi.org/10.1038/s41531-019-0086-4.

Alecu I, Bennett SAL. Dysregulated lipid metabolism and its role in α-synucleinopathy in Parkinson's disease. Front Neurosci. 2019;13:328. https://doi.org/10.3389/fnins.2019.00328.

Antonini A, Abbruzzese G, Ferini-Strambi L, Tilley B, Huang J, Stebbins GT, et al. Validation of the Italian version of the Movement Disorder Society–unified Parkinson's disease rating scale. Neurol Sci. 2013;34:683–7. https://doi.org/10.1007/s10072-012-1112-z.

Beck GC, Brinkkoetter P, Hanusch C, et al. Clinical review: immunomodulatory effects of dopamine in general inflammation. Crit Care. 2004;8:485–91. https://doi.org/10.1186/cc2879.

Biancolillo A, Næs T. The sequential and orthogonalized PLS regression for multiblock regression. In: Cocchi M, editor. Data fusion methodology and applications, vol. 31. Amsterdam: Elsevier Inc.; 2019. p. 157–77. https://doi.org/10.1016/B978-0-444-63984-4.00006-5.

Biancolillo A, Marini F, Roger J. SO-CovSel: a novel method for variable selection in a multiblock framework. J Chemom. 2019:e3120. https://doi.org/10.1002/cem.3120.

Breuillard C, Cynober L, Moinard C. Citrulline and nitrogen homeostasis: an overview. Amino Acids. 2015;47:685–91. https://doi.org/10.1007/s00726-015-1932-2.

Caldwell RB, Toque HA, Narayanan SP, Caldwell RW. Arginase: an old enzyme with new tricks. Trends Pharmacol Sci. 2015;36:395–405. https://doi.org/10.1016/j.tips.2015.03.006.

Calvani R, Picca A, Marini F, et al. A distinct pattern of circulating amino acids characterizes older persons with physical frailty and sarcopenia: results from the BIOSPHERE study. Nutrients. 2018;10:1691. https://doi.org/10.3390/nu10111691.

Calzada E, Onguka O, Claypool SM. Phosphatidylethanolamine metabolism in health and disease. Int Rev Cell Mol Biol. 2016;321:29–88. https://doi.org/10.1016/bs.ircmb.2015.10.001.

Chen R, Snyder M. Promise of personalized omics to precision medicine. Wiley Interdiscip Rev Syst Biol Med. 2013;5:73–82. https://doi.org/10.1002/WSBM.1198.

Chen-Plotkin AS, Albin R, Alcalay R, et al. Finding useful biomarkers for Parkinson's disease. Sci Transl Med. 2018;10: eaam6003. https://doi.org/10.1126/scitranslmed.aam6003.

Collins LM, Toulouse A, Connor TJ, Nolan YM. Contributions of central and systemic inflammation to the pathophysiology of Parkinson's disease. Neuropharmacology. 2012;62:2154–68. https://doi.org/10.1016/j.neuropharm.2012.01.028.

Davidson WS, Jonas A, Clayton DF, George JM. Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. J Biol Chem. 1998;273:9443–9. https://doi.org/10.1074/jbc.273.16.9443.

Deleidi M, Gasser T. The role of inflammation in sporadic and familial Parkinson's disease. Cell Mol Life Sci. 2013;70: 4259–73. https://doi.org/10.1007/s00018-013-1352-y.

Dickson DW. Neuropathology of Parkinson disease. Parkinsonism Relat Disord. 2018;46:S30–3. https://doi.org/10.1016/j.parkreldis.2017.07.033.

Elyaman W, Khoury SJ. Th9 cells in the pathogenesis of EAE and multiple sclerosis. Semin Immunopathol. 2017;39:79–87. https://doi.org/10.1007/s00281-016-0604-y.

Elyaman W, Bradshaw EM, Uyttenhove C, Dardalhon V, Awasthi A, Imitola J, et al. IL-9 induces differentiation of TH17 cells and enhances function of FoxP3+ natural regulatory T cells. Proc Natl Acad Sci U S A. 2009;106:12885–90. https://doi.org/10.1073/pnas.0812530106.

Emamzadeh FN, Surguchov A. Parkinson's disease: biomarkers, treatment, and risk factors. Front Neurosci. 2018;12:612. https://doi.org/10.3389/fnins.2018.00612.

Figura M, Kuśmierska K, Bucior E, Szlufik S, Koziorowski D, Jamrozik Z, et al. Serum amino acid profile in patients with Parkinson's disease. PLoS One. 2018;13:e0191670. https://doi.org/10.1371/journal.pone.0191670.

GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol. 2017;16:877–97. https://doi.org/10.1016/S1474-4422(17)30299-5.

GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2018;17:939–53. https://doi.org/10.1016/S1474-4422(18)30295-3.

Giunta S. Is inflammaging an auto[innate]immunity subclinical syndrome? Immun Ageing. 2006;3:12. https://doi.org/10.1186/1742-4933-3-12.

Grozdanov V, Bliederhaeuser C, Ruf WP, Roth V, Fundel-Clemens K, Zondler L, et al. Inflammatory dysregulation of

blood monocytes in Parkinson's disease patients. Acta Neuropathol. 2014;128:651–63. https://doi.org/10.1007/s00401-014-1345-4.

Hatano T, Saiki S, Okuzumi A, et al. Identification of novel biomarkers for Parkinson's disease by metabolomic technologies. J Neurol Neurosurg Psychiatry. 2016;87:295–301. https://doi.org/10.1136/jnnp-2014-309676.

Havelund JF, Andersen AD, Binzer M, Blaabjerg M, Heegaard NHH, Stenager E, et al. Changes in kynurenine pathway metabolism in Parkinson patients with L-DOPA-induced dyskinesia. J Neurochem. 2017;142:756–66. https://doi.org/10.1111/jnc.14104.

He R, Yan X, Guo J, Xu Q, Tang B, Sun Q. Recent advances in biomarkers for Parkinson's disease. Front Aging Neurosci. 2018;10:305. https://doi.org/10.3389/fnagi.2018.00305.

Hou Y, Dan X, Babbar M, Wei Y, Hasselbalch SG, Croteau DL, et al. Ageing as a risk factor for neurodegenerative disease. Nat Rev Neurol. 2019;15:565–81. https://doi.org/10.1038/s41582-019-0244-7.

Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry. 1992;55:181–4. https://doi.org/10.1136/jnnp.55.3.181.

Justice JN, Ferrucci L, Newman AB, Aroda VR, Bahnson JL, Divers J, et al. A framework for selection of blood-based biomarkers for geroscience-guided clinical trials: report from the TAME Biomarkers Workgroup. GeroScience. 2018;40: 419–36. https://doi.org/10.1007/s11357-018-0042-y.

Kim R, Kim H-J, Kim A, Jang M, Kim A, Kim Y, et al. Peripheral blood inflammatory markers in early Parkinson's disease. J Clin Neurosci. 2018;58:30–3. https://doi.org/10.1016/j.jocn.2018.10.079.

Kori M, Aydln B, Unal S, et al. Metabolic biomarkers and neurodegeneration: a pathway enrichment analysis of Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis. OMICS. 2016;20:645–61. https://doi.org/10.1089/omi.2016.0106.

Kulisevsky J, Oliveira L, Fox SH. Update in therapeutic strategies for Parkinson s disease. Curr Opin Neurol. 2018;31:439–47. https://doi.org/10.1097/WCO.0000000000000579.

Kustrimovic N, Comi C, Magistrelli L, Rasini E, Legnaro M, Bombelli R, et al. Parkinson's disease patients have a complex phenotypic and functional Th1 bias: cross-sectional studies of CD4+ Th1/Th2/T17 and Treg in drug-naïve and drug-treated patients. J Neuroinflammation. 2018;15:205. https://doi.org/10.1186/s12974-018-1248-8.

Lashuel HA, Overk CR, Oueslati A, Masliah E. The many faces of α-synuclein: from structure and toxicity to therapeutic target. Nat Rev Neurosci. 2013;14:38–48. https://doi.org/10.1038/nrn3406.

LeWitt PA, Li J, Lu M, Guo L, Auinger P, Parkinson Study Group–DATATOP Investigators. Metabolomic biomarkers as strong correlates of Parkinson disease progression. Neurology. 2017;88:862–9. https://doi.org/10.1212/WNL.0000000000003663.

Loscalzo J, Barabasi A-L. Systems biology and the future of medicine. Wiley Interdisc Rev Syst Biol Med. 2011;3: 619–27. https://doi.org/10.1002/wsbm.144.

Martinez-Martin P. Hoehn and Yahr staging scale. In: Kompoliti K, Verhagen Metman L, editors. Encyclopedia of movement disorders. Amsterdam: Elsevier Inc.; 2010. p. 23–5. https://doi.org/10.1016/B978-0-12-374105-9.00034-4.

Marzetti E, Picca A, Marini F, Biancolillo A, Coelho-Junior HJ, Gervasoni J, et al. Inflammatory signatures in older persons with physical frailty and sarcopenia: the frailty "cytokinome" at its core. Exp Gerontol. 2019;122:129–38. https://doi.org/10.1016/j.exger.2019.04.019.

Mercado NM, Collier TJ, Sortwell CE, Steece-Collier K. BDNF in the aged brain: translational implications for Parkinson's disease. Austin Neurol Neurosci. 2017;2:1021.

Okazawa H, Murata M, Watanabe M, Kamei M, Kanazawa I. Dopaminergic stimulation up-regulates the in vivo expression of brain-derived neurotrophic factor (BDNF) in the striatum. FEBS Lett. 1992;313:138–42. https://doi.org/10.1016/0014-5793(92)81430-t.

Patel D, Witt SN. Ethanolamine and phosphatidylethanolamine: partners in health and disease. Oxidative Med Cell Longev. 2017;2017:4829180. https://doi.org/10.1155/2017/4829180.

Pattarini R, Smeyne RJ, Morgan JI. Temporal mRNA profiles of inflammatory mediators in the murine 1-methyl-4-phenyl-1, 2,3,6-tetrahydropyrimidine model of Parkinson's disease. Neuroscience. 2007;145:654–68. https://doi.org/10.1016/j.neuroscience.2006.12.030.

Perrin FE, Lacroix S, Avilés-Trigueros M, David S. Involvement of monocyte chemoattractant protein-1, macrophage inflammatory protein-1alpha and interleukin-1beta in Wallerian degeneration. Brain. 2005;128:854–66. https://doi.org/10.1093/brain/awh407.

Picca A, Calvani R, Landi G, Marini F, Biancolillo A, Gervasoni J, 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. 2019a;128:110766. https://doi.org/10.1016/j.exger.2019.110766.

Picca A, Coelho-Junior HJ, Cesari M, Marini F, Miccheli A, Gervasoni J, et al. The metabolomics side of frailty: toward personalized medicine for the aged. Exp Gerontol. 2019b;126: 110692. https://doi.org/10.1016/j.exger.2019.110692.

Picca A, Guerra F, Calvani R, et al. Mitochondrial-derived vesicles as candidate biomarkers in Parkinson's disease: rationale, design and methods of the EXosomes in PArkiNson Disease (EXPAND) study. Int J Mol Sci. 2019c;20:2373. https://doi.org/10.3390/ijms20102373.

Picca A, Ponziani FR, Calvani R, et al. Gut microbial, inflammatory and metabolic signatures in older people with physical frailty and sarcopenia: results from the BIOSPHERE study. Nutrients. 2019d;12:65. https://doi.org/10.3390/nu12010065.

Picca A, Guerra F, Calvani R, et al. Mitochondrial signatures in circulating extracellular vesicles of older adults with Parkinson's disease: results from the EXosomes in PArkiNson's Disease (EXPAND) study. J Clin Med. 2020;9:504. https://doi.org/10.3390/jcm9020504.

Poewe W, Seppi K, Tanner CM, et al. Parkinson disease. Nat Rev Dis Prim. 2017;3:17013. https://doi.org/10.1038/nrdp.2017.13.

Posavi M, Diaz-Ortiz M, Liu B, Swanson CR, Skrinak RT, Hernandez-Con P, et al. Characterization of Parkinson's disease using blood-based biomarkers: a multicohort proteomic analysis. PLoS Med. 2019;16:e1002931. https://doi.org/10.1371/journal.pmed.1002931.

Qin X-Y, Zhang S-P, Cao C, Loh YP, Cheng Y. Aberrations in peripheral inflammatory cytokine levels in Parkinson disease: a systematic review and meta-analysis. JAMA Neurol. 2016;73:1316–24. https://doi.org/10.1001/jamaneurol.2016.2742.

Rahmani F, Saghazadeh A, Rahmani M, Teixeira AL, Rezaei N, Aghamollaii V, et al. Plasma levels of brain-derived neurotrophic factor in patients with Parkinson disease: a systematic review and meta-analysis. Brain Res. 2019;1704:127–36. https://doi.org/10.1016/j.brainres.2018.10.006.

Reale M, Iarlori C, Thomas A, et al. Peripheral cytokines profile in Parkinson's disease. Brain Behav Immun. 2009;23:55–63. https://doi.org/10.1016/j.bbi.2008.07.003.

Saresella M, Calabrese E, Marventano I, et al. Increased activity of Th-17 and Th-9 lymphocytes and a skewing of the post-thymic differentiation pathway are seen in Alzheimer's disease. Brain Behav Immun. 2011;25:539–47. https://doi.org/10.1016/j.bbi.2010.12.004.

Scalzo P, Kümmer A, Bretas TL, Cardoso F, Teixeira AL. Serum levels of brain-derived neurotrophic factor correlate with motor impairment in Parkinson's disease. J Neurol. 2010;257:540–5. https://doi.org/10.1007/s00415-009-5357-2.

Sommer A, Winner B, Prots I. The Trojan horse-neuroinflammatory impact of T cells in neurodegenerative diseases. Mol Neurodegener. 2017;12:78. https://doi.org/10.1186/s13024-017-0222-8.

Sprott RL. Biomarkers of aging and disease: introduction and definitions. Exp Gerontol. 2010;45:2–4. https://doi.org/10.1016/j.exger.2009.07.008.

Stoessel D, Schulte C, Teixeira dos Santos MC, et al. Promising metabolite profiles in the plasma and CSF of early clinical Parkinson's disease. Front Aging Neurosci. 2018;10:51. https://doi.org/10.3389/fnagi.2018.00051.

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 2010;25:2649–53. https://doi.org/10.1002/mds.23429.

van der Veen JN, Kennelly JP, Wan S, Vance JE, Vance DE, Jacobs RL. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochim Biophys Acta Biomembr. 2017;1859:1558–72. https://doi.org/10.1016/j.bbamem.2017.04.006.

Yabuki Y, Shioda N, Yamamoto Y, Shigano M, Kumagai K, Morita M, et al. Oral l-citrulline administration improves memory deficits following transient brain ischemia through cerebrovascular protection. Brain Res. 2013;1520:157–67. https://doi.org/10.1016/j.brainres.2013.05.011.