

## LETTER TO EDITOR

# MicroRNAs from extracellular vesicles as a signature for Parkinson's disease

Dear Editor,

In the present study, we have demonstrated that extracellular vesicles (EVs) derived from cerebrospinal fluid (CSF) represent a promising source for the identification of a novel miRNA signatures in Parkinson's disease (PD). Using next-generation small-RNA sequencing, we present for the first time the complete and quantitative microRNAome of EVs isolated from human CSF of PD and age-correlated controls (CTR). In parallel, we performed CSF proteomic profiling of overlapping patient cohorts, which revealed the deregulation of disease-relevant pathways similar to the ones obtained with the parallel miRNA analyses, supporting the results for the identified signature.

Novel molecular signatures and disease biomarkers are urgently needed for PD, not only to improve diagnostic precision, but also to enable monitoring of treatment responses, as well as stratification of patients according to the molecular background, rather than solely on clinical phenotypes.<sup>1</sup> Circulating miRNAs are auspicious targets for biomarker studies because their expression reflects the functional state of cells and is directly influenced by pathological stimuli.<sup>2</sup> CSF is in direct contact with the brain parenchyma, and molecular alterations in its composition may reflect specific changes related to PD pathology in the brain.

MiRNA species circulating in CSF seem to overlap with miRNAs expressed in brain tissue.<sup>3</sup> Furthermore, miRNAs and other small-RNAs are enriched in the vesicular fraction of human CSF.<sup>3,4</sup> In order to characterize the size and particle distribution in our CSF EV preparations, we used Nanoparticle Tracking Analyses and observed a similar enrichment as previously reported (Figure 1A). Small-RNA sequencing ratified the miRNA abundance in CSF EVs—they represented, on average, 97.4% of all mapped small-RNAs in the discovery cohort (Figure 1C). In total, we detected 688 miRNAs. A total of 208 of these had a base mean higher than 5 reads and were analyzed further. Differential expression analyses revealed differences in the

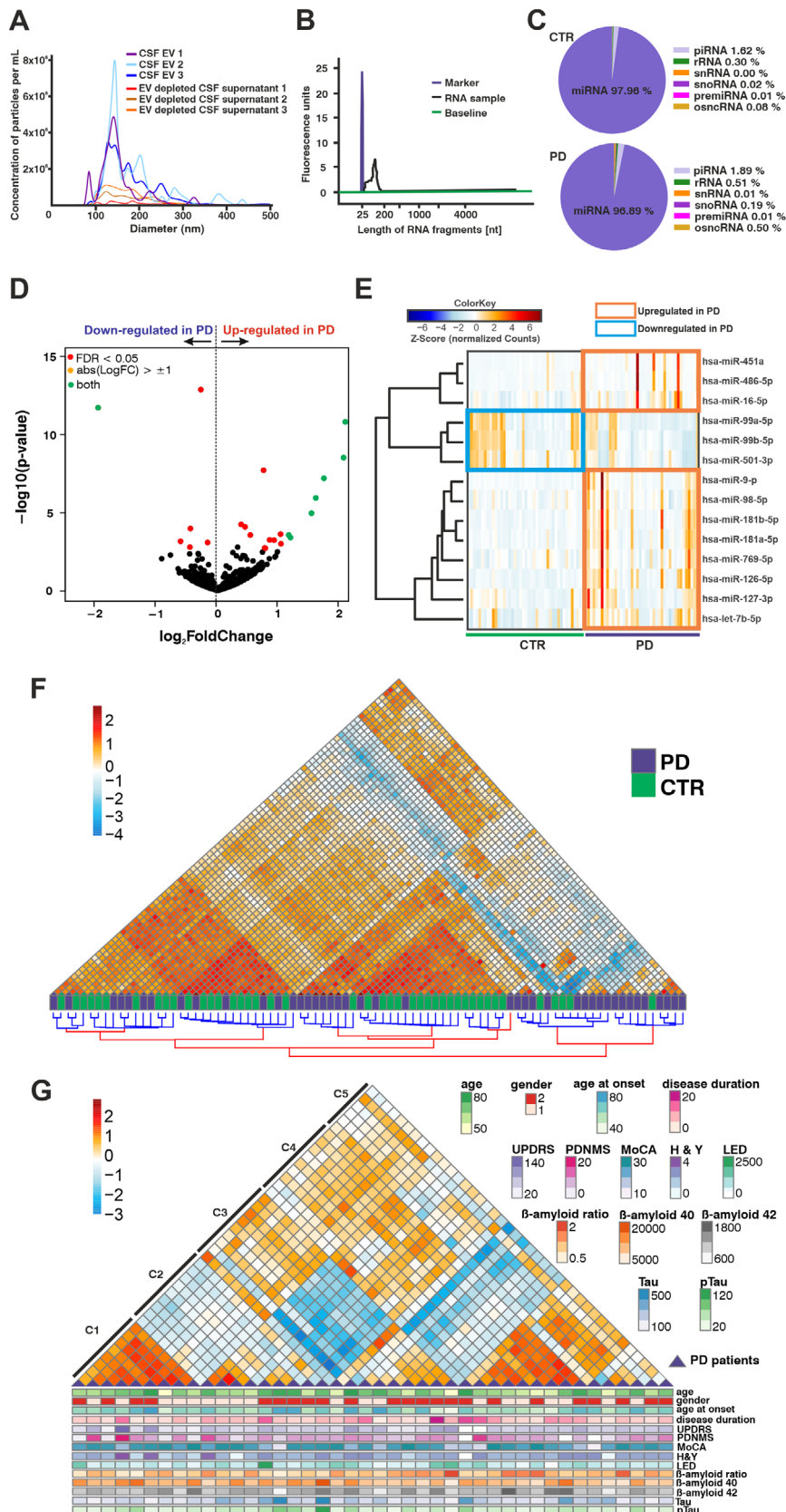
levels of 22 miRNAs in the PD versus CTR comparison (Figure 1D). The majority of the differentially expressed miRNAs were upregulated in PD subjects, whereas downregulated miRNAs showed only subtle levels of deregulation ( $-0.60 \leq \log_2FC \leq -0.16$ ). Among upregulated species figured brain-enriched miRNAs miR-9-5p, let-7b, miR-181a-5p, and miR-181b-5p (Figures 1D and 1E). To reduce bias by a potential erythrocyte contamination, we strictly selected CSF samples with a low number of red blood cells ( $<100/\mu\text{l}$  CSF). Furthermore, because miR-451a is highly enriched in red blood cells, it was excluded from feature-selection analyses.

To explore the overall miRNA expression differences in the cohorts, hierarchical clustering analyses were performed (Figure 1F). Grouping samples based on miRNA expression levels revealed differences in the overall miRNA abundance between PD and CTR samples. PD samples showed expression heterogeneity, as some of these subjects clustered close to/among CTRs. Repeating the analysis with PD samples only revealed five different sub-clusters (Figure 1G) that did not correlate with the distribution of clinical parameters (e.g., disease duration; age of death; Levodopa-equivalent dose; scores for disease severity [PDNMS; MDS-UPDRS; MoCA; mH&Y]). This suggests a molecular diversity in PD cases that is reflected by miRNA expression.

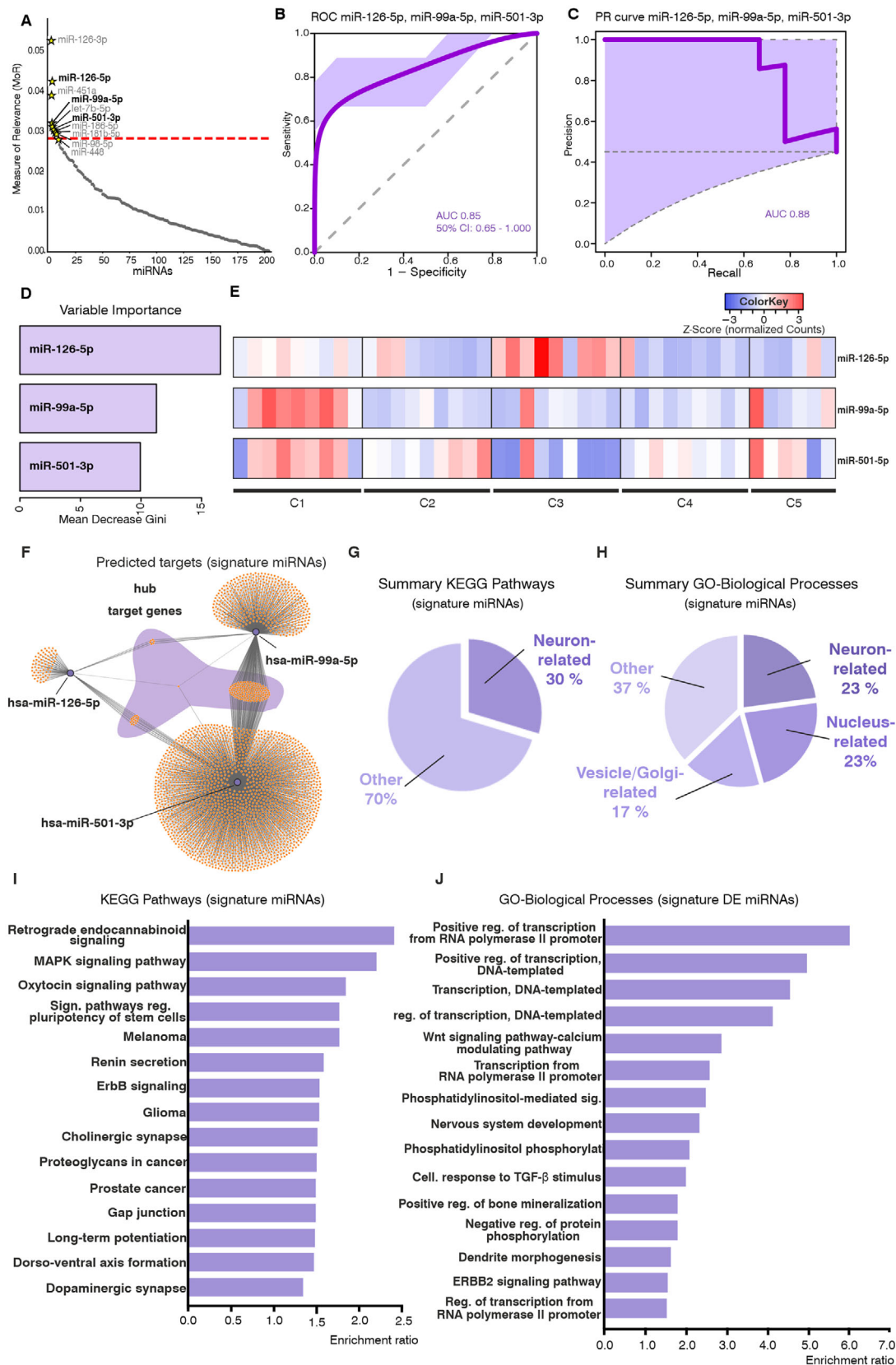
Using machine learning approaches (measure of relevance [MoR]; reliability analysis [RiA]; random forest) with the small-RNA sequencing data, we found an iterative signature comprising miR-126-5p, miR-99a-5p, and miR-501-3p, which could differentiate PD and CTR samples in our discovery cohorts (42 PD; 43 CTR) (Figures 2A–2D). Sample numbers for the discovery cohort were similar to other studies in the field<sup>4</sup> and were shown to be adequate for algorithm training. The panel was able to classify PD/CTR samples in an independent validation cohort (nine PD; 11 CTR) with an area under the curve (AUC) value of 0.85 (Precision–recall AUC = 0.88; sensitivity = 0.78; specificity = 0.95). Mean decrease in Gini

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics



**FIGURE 1** Small RNA sequencing from CSF EVs. (A) Size and particle distribution in EV-enriched CSF and EV-depleted supernatant fractions obtained with nanoparticle tracking analysis (NTA) after CSF ultracentrifugation. (B) Electropherogram showing the profile of RNA isolated from CSF extracellular vesicles. (C) Pie charts showing the average proportions of the different small RNA species in the small RNA libraries of CTR and PD subjects. (D) Volcano plot showing all detected miRNAs in the CSF of PD patients and CTR. Significantly different expressed miRNAs between CTR and PD are indicated in red (FDR < 0.05) and green (FDR < 0.05;  $\log_2FC > 1$ ). (E) Heatmap with individual expression levels for significantly regulated miRNAs in the discovery cohort (PD,  $n = 42$ ; CTR,  $n = 43$ ). (F) Hierarchical clustering of PD patients and CTR subjects in the discovery cohort. Samples with similar miRNA expression profiles are located close to each other. (G) Unbiased hierarchical clustering of PD patients from the discovery cohort according to similar miRNA expression profiles, showing also the corresponding clinical parameters. CTR, control; PD, Parkinson's disease; CSF, cerebrospinal fluid; EVs, extracellular vesicles; NTA, nanoparticle tracking analysis; C1–C5, distinct PD sample clusters; sex: 1 = female, 2 = male; age, age at onset and disease duration given in years; UPDRS, Movement Disorders Society-Unified Parkinson's Disease Rating Scale; PDNMS, Parkinson's Disease Non-Motor Scale; MoCA, Montreal Cognitive Assessment; H&Y, modified Hoehn & Yahr stage; LED, levodopa equivalent dose in mg; beta-amyloid-(1-40) and beta-amyloid-(1-42) in pg/mL; beta-amyloid ratio: beta-amyloid-(1-42)/beta-amyloid-(1-40)  $\times 10$ ; Tau in pg/mL; pTau: phosphorylated (181) Tau in pg/mL



**FIGURE 2** A miRNA signature identified in PD CSF. (A) MoR analysis for identification of relevant miRNAs that discriminate between PD and CTR in the discovery cohort (PD,  $n = 42$ ; CTR,  $n = 43$ ). The red line indicates the critical MoR value cutoff. miRNAs located above this line are considered relevant. The light-gray miRNAs were excluded after mean score filtering in the feature selection procedure. The combination of miRNAs that was tested to discriminate PD and CTR subjects in an independent validation cohort is indicated in bold black



(Figure 2D) indicated miR-126-5p as the most discriminative variable, followed by miR-99a-5p and miR-501-3p. A third independent cohort (25 PD; 25 CTR) was used for validation purposes. Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) experiments confirmed the differential expression of miR-126-5p and miR-99a-5p when comparing PD and CTR cohorts (Figure S2). The individual expression of each signature miRNA in PD subjects of the discovery cohort (Figure 2E) delineated a similar heterogeneity to the one observed in the global miRNA analysis (Figure 1G), confirming the molecular diversity within PD cases. Subclusters 1 and 3 present opposing expression for the signature miRNAs, whereas subclusters 2 and 5 present similar levels for these candidates. These findings suggest that the identified signature would be a useful tool for distinguishing disease subgroups based on miRNA expression. On the other hand, the inclusion of additional patients/cohorts with variate compositions might explain the lack of reproducibility of studies in the field,<sup>5</sup> as well as the discrepant results for some candidates during the additional validation studies we presented here. Using samples from patients with different molecular backgrounds, which cannot be distinguished by clinical phenotype alone, as well as the smaller size of the validation cohort might explain the differences in the results observed for miR-501-3p with RNA sequencing and qRT-PCR experiments.

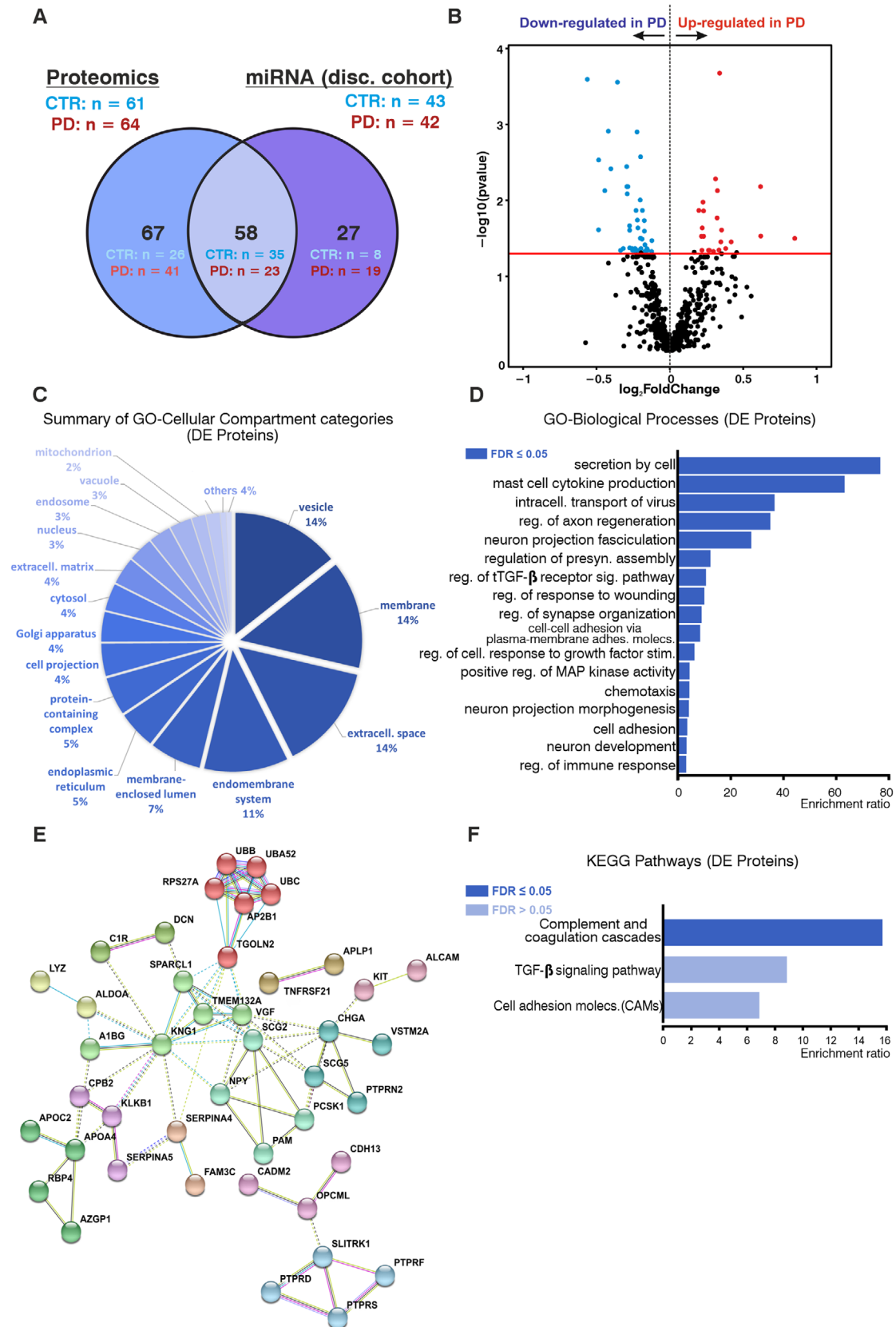
Regarding the biological role of the three signature miRNAs, functional annotation analyses with their predicted targets indicated that these candidates likely originate in neurons. Neuron-related terms comprised the most frequent enriched categories for *Gene Ontology-Biological Processes (GO-BP)* results (8/35 enriched GO-BP terms), indicating their neuronal origin. Terms including *neuron death*, *vesicle-mediated transport*, and *proteasomal-protein catabolic process* indicate the participation of these miRNAs in processes directly related to PD pathogenesis<sup>6</sup> (Figures 2H and 2J). These findings are corroborated by KEGG

pathway (*Kyoto Encyclopedia of Genes and Genomes*) enrichment results: 19 out of 64 annotated KEGG pathways were neuron related (Figures 2G and 2I). Among the top 15 pathways figure *retrograde endocannabinoid signaling* and *cholinergic-dopaminergic synapse*, categories with important involvement in PD pathology.<sup>6,7</sup> Furthermore, each candidate of our panel has been linked to neurodegenerative mechanisms previously: miR-126 has been linked to insulin/IGF-1/PI3K signaling and found in increased levels in PD substantia nigra<sup>8</sup>; miR-99a-5p has been associated with neuroinflammation/neurodegeneration processes by regulating microglial functions<sup>9</sup>; miR-501-3p is a regulator of dendritic spine remodeling, and was also found upregulated in Alzheimer's disease brains.<sup>10</sup>

Aiming to identify differentially expressed proteins and to explore disease-relevant pathways further, an overlapping cohort (64 PD; 61 CTR) was analyzed using mass spectrometry using total CSF (Figure 3A). In total, 67 proteins were found differentially expressed between conditions (45 downregulated in PD/22 upregulated in PD) (Figure 3B). Functional annotation showed an important enrichment for inflammatory/immune-related terms, as well as neuronal-related terms (e.g., *axon regeneration*; *neuronal development*; *synapse organization* for GO-BP terms; *complement/coagulation cascades* for KEGG pathways) (Figure 3D). Remarkably, these results overlap with the pathways annotated for the signature miRNAs, especially for the regulation of *neuron development/morphogenesis* and *synapse- and secretion-related* terms. PPI networks with deregulated proteins revealed important hub-proteins (TGOLN2; SCG2; KNG1; APOA4) (Figure 3E). Proteins that have been previously postulated as PD biomarkers (VGF and EPHA4) were also identified in our studies (Table S4). Overall, although the parallel studies differed regarding the analyzed CSF compartments and the cohorts did not overlap completely, several disease-relevant pathways were coincidental, further supporting the results of the miRNA study.

---

lettering (miR-126-5p, miR-99a-5p, and miR-501-3p). (B) ROC curve showing the performance of the three signature miRNAs for the discrimination of PD and CTR subjects in an independent validation cohort (PD,  $n = 9$ ; CTR,  $n = 11$ ). Training of the model was performed on the discovery cohort with a 10-fold cross-validation. The lilac area indicates the 50% confidence interval; an AUC of 0.85 was obtained. (C) Precision–recall curve for the performance of the three signature miRNAs in the validation cohort with an AUC of 0.88. (D) Variable importance indicated by the mean decrease in Gini showing the relevance of the individual miRNAs in the signature. (E) Heatmap showing individual expression levels for the three signature miRNAs in the distinct PD sub-clusters from the discovery cohort, identified by unbiased hierarchical clustering (PD,  $n = 42$ ). (F–J) Differential expression and target-gene functional annotation results for CSF small RNA sequencing. (F) Predicted targets for signature miRNAs. Hub target genes common to the three miRNAs are highlighted by the shadowing in purple. (G) Summary of enriched KEGG pathways and (H) *GO-cellular compartment* categories in the functional annotation for predicted targets of the signature miRNAs. Neuron-related pathways are enriched in the dataset. (I) Top-15 KEGG pathways and (J) *GO-biological processes* terms enriched for the predicted targets of the signature miRNAs. Bars represent the enrichment ratio results from the WEBGESTALT algorithm. CTR, control; PD, Parkinson's disease; MoR, measure of relevance; ROC, receiver operatin characteristic; AUC, area under the curve; C1–C5, distinct PD sample clusters; CSF, cerebrospinal fluid; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate



**FIGURE 3** Proteomics analysis with CSF of PD and CTR subjects. (A) Venn diagram indicating the intersection between the discovery cohort from the small RNA sequencing (miRNA) studies and the subgroup of subjects analyzed by proteomics. Stratification of subjects: discovery cohort small RNA sequencing: PD,  $n = 42$ ; CTR,  $n = 43$ . Proteomics: PD,  $n = 64$ ; CTR,  $n = 61$ . (B) Volcano plot showing all detected proteins in the total CSF of PD patients and CTR. Differentially expressed proteins between CTR and PD are indicated in blue (downregulated

Limitations for the identification of molecular signatures in CSF EVs must be critically considered: the starting volume of CSF for isolation of sequencing-quality RNA (~4.5 mL) is relatively high, limiting the number of available samples/additional analyses that can be performed. More efficient EV/RNA isolation protocols will significantly improve further CSF multiomics studies. It is important to highlight that the identification of such a miRNA signature in PD CSF must be taken as a starting point, and both the individual expression of each miRNA candidate as well as the combinatorial diagnostic value of the proposed panel must be validated in subsequent multicentric studies. Furthermore, we aimed to strictly select PD patients with a clear clinical phenotype to evaluate miRNA changes in a more advanced stage of the disease. A subsequent study recruiting patients shortly after onset of motor symptoms would be an important follow-up for this work to assess the value of the signature for the identification of early PD patients.

In summary, we identified a novel miRNA signature in PD CSF composed of miR-126-5p, miR-99a-5p, and miR-501-3p. This signature could potentially contribute to an improved PD diagnosis, as well as to delineate future drugable targets for the disease by revealing important pathophysiological mechanisms. The validity of this signature as a diagnostic biomarker panel should be subsequently validated in larger multicentric studies. Our small-RNA data also indicate that profiling miRNA expression in CSF EVs might identify clinically inapparent subgroups of PD patients, which could be ultimately used for personalized diagnostic and therapeutic strategies for the disease.

## ACKNOWLEDGMENTS

The authors appreciate the participation of patients in this study. We thank Woori Koh, Anna Fischbach, and Matthias Börger for their participation in patient recruitment. We also thank Barbara Müller for technical assistance and the Transcriptome Analysis Laboratory Göttingen for the performance of the small RNA sequencing.

## CONFLICT OF INTEREST


The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT


RNA sequencing and proteomics datasets were deposited in public repositories — the European Genome-phenome


Archive (EGA) and the Proteomics Identifications Database (PRIDE), respectively — following the applicable guidelines. RNA sequencing data are stored under the accession number EGAD00001006629. Proteomics data are available via ProteomeXchange with identifier PXD022234.

Lucas Caldi Gomes<sup>1,2</sup> 

Anna-Elisa Roser<sup>2</sup> 


Gaurav Jain<sup>3</sup> 

Tonatiuh Pena Centeno<sup>3</sup> 

Fabian Maass<sup>2</sup> 


Lukas Schilde<sup>4,5</sup>


Caroline May<sup>4,5</sup> 

Anja Schneider<sup>6</sup> 

Mathias Bähr<sup>2</sup> 

Katrin Marcus<sup>4,5</sup> 

André Fischer<sup>3,7</sup> 

Paul Lingor<sup>1,8</sup> 

<sup>1</sup> Department of Neurology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, München, Germany

<sup>2</sup> Department of Neurology, University Medical Center Göttingen, Göttingen, Germany

<sup>3</sup> Department for Epigenetics and Systems Medicine in Neurodegenerative Diseases, German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany

<sup>4</sup> Medical Faculty, Medizinisches Proteom-Center, Ruhr University Bochum, Bochum, Germany

<sup>5</sup> Medical Proteome Analysis, Center for Protein Diagnostics (PRODI), Ruhr-University Bochum, Bochum, Germany

<sup>6</sup> Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Clinic Bonn and German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

<sup>7</sup> Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany

<sup>8</sup> German Center for Neurodegenerative Diseases (DZNE), site München, München, Germany

## Correspondence

Paul Lingor, Department of Neurology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Ismaninger Str. 22, 81675 München, Bavaria, Germany.

Email: [paul.lingor@tum.de](mailto:paul.lingor@tum.de)

in PD) and red (upregulated in PD) (FDR < 0.05). (C) Summary of enriched GO-cellular compartment categories in the functional annotation for differentially expressed proteins. (D) Enriched GO-biological process categories and (E) KEGG pathways in the functional annotation of differentially expressed proteins. Bars represent enrichment ratio results from the WEBGESTALT algorithm. (F) STRING analysis for the differentially expressed proteins. Clusters defined by the Markov Cluster Algorithm in STRING using default parameters. CSF, cerebrospinal fluid; PD, Parkinson disease; CTR, control subjects; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes


Lucas Caldi Gomes and Anna-Elisa Roser contributed equally to this work.

#### Funding information

This study was supported by the Cluster of Excellence and DFG Research Center Nanoscale Microscopy and Molecular Physiology of the Brain (CNMPB), Göttingen, Germany, by Bridging Funds from the Göttingen Graduate Center for Neurosciences, Biophysics, and Molecular Biosciences (GGNB) (LCG), by the German Network for Bioinformatics Infrastructure (de.NBI - FKZ 031 A 534A) (Federal Ministry of Education and Research), P.U.R.E. (Protein Research Unit Ruhr within Europe), and Ruhr-University Bochum, Bochum, Germany ProDi (Center for protein diagnostics), Ministry of Innovation, Science and Research of North-Rhine Westphalia, Germany (KM).

#### ORCID

Lucas Caldi Gomes  <https://orcid.org/0000-0003-4959-2169>

Anna-Elisa Roser  <https://orcid.org/0000-0001-6233-1567>

Gaurav Jain  <https://orcid.org/0000-0003-0987-1323>  
Tonatiuh Pena Centeno  <https://orcid.org/0000-0002-9329-8912>

Fabian Maass  <https://orcid.org/0000-0002-4077-2852>

Caroline May  <https://orcid.org/0000-0002-6532-5979>

Anja Schneider  <https://orcid.org/0000-0001-9540-8700>

Mathias Bähr  <https://orcid.org/0000-0002-9570-1592>

Katrin Marcus  <https://orcid.org/0000-0002-3313-8845>

André Fischer  <https://orcid.org/0000-0001-8546-1161>

Paul Lingor  <https://orcid.org/0000-0001-9362-7096>

#### REFERENCES

- Schlossmacher MG, Mollenhauer B. Biomarker research in Parkinson's disease: objective measures needed for patient stratification in future cause-directed trials. *Biomarkers Med.* 2010;4:647-650.
- Saba R, Schrott GM. MicroRNAs in neuronal development, function and dysfunction. *Brain Res.* 2010;1338:3-13.
- Yagi Y, Ohkubo T, Kawaji H, et al. Next-generation sequencing-based small RNA profiling of cerebrospinal fluid exosomes. *Neurosci Lett.* 2017;636:48-57.
- Jain G, Stuenkel A, Rao P, et al. A combined miRNA-piRNA signature to detect Alzheimer's disease. *Transl Psychiatry.* 2019;9:250.
- Roser AE, Caldi Gomes L, Schünemann J, et al. Circulating miRNAs as diagnostic biomarkers for Parkinson's disease. *Front Neurosci.* 2018;12:625.
- Dauer W, Przedborski S. Parkinson's disease: mechanisms and models. *Neuron.* 2003;39:889-909.
- Pisani V, Madeo G, Tassone A, et al. Homeostatic changes of the endocannabinoid system in Parkinson's disease: ECB changes in PD. *Mov Disord.* 2011;26:216-222.
- Kim W, Lee Y, McKenna ND, et al. miR-126 contributes to Parkinson's disease by dysregulating the insulin-like growth factor/phosphoinositide 3-kinase signaling. *Neurobiol Aging.* 2014;35:1712-1721.
- Butovsky O, Jedrychowski MP, Moore CS, et al. Identification of a unique TGF- $\beta$  dependent molecular and functional signature in microglia. *Nat Neurosci.* 2014;17:131-143.
- Hara N, Kikuchi M, Miyashita A, et al. Serum microRNA miR-501-3p as a potential biomarker related to the progression of Alzheimer's disease. *Acta Neuropathol Commun.* 2017;5:10.

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.