MicroRNA Candidate Biomarkers for Parkinson's Disease and Idiopathic REM Sleep Behavior Disorder

Ligang Wu (lgwu@sibcb.ac.cn)
Institute of Biochemistry and Cell Biology State Key Laboratory of Molecular Biology: Chinese Academy of Sciences Center for Excellence in Molecular Cell Science https://orcid.org/0000-0003-4010-9118

Jun Liu
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital https://orcid.org/0000-0001-8300-8646

Yuanyuan Li
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Ying Cao
Chinese Academy of Sciences State Key Laboratory of Molecular Biology

Wei Liu
Chinese Academy of Sciences State Key Laboratory of Molecular Biology

Hongdao Zhang
Chinese Academy of Sciences State Key Laboratory of Molecular Biology

Aonan Zhao
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Ningdi Luo
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Research article

Keywords: Parkinson's disease, iRBD, plasma, extracellular vesicle, miRNA, biomarker

Posted Date: December 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1163742/v1

Abstract

Background: Parkinson's disease (PD), a severe neurodegenerative disorder, and idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD), a parasomnia recognized as the prodromal stage of synucleinopathies (including PD), both lack reliable, non-invasive biomarker tests for early intervention and management.

Objectives: To investigate whether plasma extracellular vesicle (EV)-associated sncRNAs could discriminate PD and/or iRBD patients from healthy individuals.

Methods: We optimized a cDNA library construction method, EVsmall-seq, for high throughput sequencing of sncRNAs associated with plasma EVs. We profiled EV-sncRNAs from the plasma of 60 normal controls, 56 iRBD patients, and 53 PD patients, and constructed a support vector machine (SVM) classifier to identify the informative miRNA features to distinguish PD and/or iRBD patients from healthy individuals.

Results: First, a sixteen-miRNA signature was found to distinguish PD patients from healthy individuals with 88% sensitivity, 90.43% specificity, and 89.13% accuracy. Second, a three-miRNA signature was found to distinguish iRBD patients from healthy individuals with 96% sensitivity, 86.36% specificity, and 91.49% accuracy. Third, tweenty 20 miRNAs were found consistently increased or decreased in expression from healthy subjects to iRBD to PD patients, which might be linked to PD development through iRBD.

Conclusions: Current study provides a valuable and highly informative dataset of EV-associated sncRNAs from plasma of iRBD and PD patients. We identified miRNA signature features that could serve as minimally-invasive, blood-based surveillance biomarkers for distinguishing iRBD or PD from healthy individuals with high sensitivity, specificity, and accuracy.

Background

Parkinson's disease (PD) is the second most common neurodegenerative disease, affecting approximately 1% of people over the age of 60 globally, and imposes increasingly heavy social and economic burdens on aging societies[1,2]. Idiopathic rapid eye movement (REM) sleep behavior disorder (iRBD) is a parasomnia characterized by the loss of normal atonia during the REM stage of sleep which results in overt motor behaviors that frequently represent the enactment of dreams[3]. In the past decade, iRBD has been established as one of the earliest and most specific prodromal signs of α-synucleinopathies, including PD, dementia with Lewy bodies, and multiple system atrophy[3]. Thus, accessible and reliable biomarkers for early diagnosis of PD and iRBD are urgently needed to identify candidate therapeutic targets and to monitor disease progression during therapeutic interventions[4,5].

Currently, the diagnosis of PD mostly relies on clinical symptoms, which hampers the detection of the earliest phases of the disease, the time at which treatment may have the greatest therapeutic effect. A

variety of biomarkers for diagnosing PD are under investigation, including factors based on pathological, imaging, biochemical, and genetic data[4]. Biofluid biomarkers have markedly expanded over past 5 years, including α-synuclein, lysosomal enzymes and neurofilament light chain in CSF[6,7]. Blood biomarkers, such as α-synuclein, are also under investigation; whereas its quantities are strongly influenced by red blood cell (RBC) contamination and haemolysis, which limit the utility for diagnostic purposes[8]. Besides, although non-invasive neuroimaging techniques can provide in-depth information about brain structure and function, these methods require major investments in infrastructure which limit their wide clinical deployment[9]. Therefore, a minimally invasive test for early detection of iRBD and PD and for monitoring disease progression still poses a major challenge.

In consideration of potential biomarkers, microRNAs (miRNAs) are small non-coding RNAs (sncRNAs) that are generally expressed in all eukaryotic cells and which perform critical regulatory functions at the posttranscriptional level. Cellular miRNAs released into body fluids can be readily detected, making them ideal biomarkers for diagnosis of various diseases[10]. Previous studies have reported that salivary miR-153 and miR-223 can be used as biomarkers for idiopathic PD[11], while serum miR-221 is a potential predictor of PD[12]. In addition, circulating brain-enriched miRNAs have been used to distinguish idiopathic and genetic PD[13]. However, all of these studies used reverse transcription followed by real-time quantitative PCR (RT-qPCR) method to measure the relative expression of a few miRNAs in bodily fluids. Since this method detects a limited number of known miRNAs and is also inefficient in distinguishing miRNAs with similar sequences[14], the most diagnostically informative miRNAs could be potentially overlooked.

Extracellular vesicles (EV) are nano-scale, membrane-enclosed particles released from (possibly) all eukaryotic cells to transport proteins, lipids, RNAs, and DNA fragments[15], a process shown to have important biological functions[16,17]. EVs have been found in serum, plasma, urine, saliva, cerebrospinal fluid, and breast milk[15]. Moreover, the EV membrane effectively prevents degradation of the enclosed miRNAs by ribonucleases that are abundant in biofluids[18,19]. These features of EV-associated miRNAs cumulatively enhance miRNA biomarker reliability compared with unprotected, cell-free, miRNAs[20–23]. However, the full spectrum of EV-associated miRNAs present in the plasma of iRBD and PD patients remains unknown, and characterizing the miRNA population in PD or iRBD patients may unveil several effective diagnostic biomarkers. In this study, we used high throughput sequencing to profile the EV-associated sncRNA population in plasma of iRBD and PD patients. Here, we report several such candidate miRNA biomarkers for diagnosis of iRBD and PD patients.

Methods

Study population

A total of 169 participants were included in this study and divided into three groups: 60 healthy individuals, 56 patients with iRBD, and 53 patients with PD. The healthy individuals were community

volunteers without neurodegenerative disorders. The diagnosis of iRBD was made based on video-polysomnography evidence according to standard International Classification of Sleep Disorders-II criteria. All iRBD patients were examined by neurologists to exclude those with motor signs of parkinsonism or secondary causes. The diagnosis of PD was performed according to the International Parkinson and Movement Disorder Society (MDS) diagnostic criteria by at least two neurologists skilled in movement disorders[24], and all of the patients were diagnosed as idiopathic PD.

This study was approved by the ethics committee of Ruijin Hospital affiliated with the Shanghai JiaoTong University School of Medicine and was carried out at the Department of Neurology and Institute of Neurology of Ruijin Hospital. All participants or their guardians provided written informed consent.

Clinical assessment

For iRBD and PD patients, essential demographic and clinical information, including a study questionnaire for motor and nonmotor manifestations of their disease, was collected and documented. The motor subscale of the Unified PD Rating Scale (UPDRS) was used to evaluate motor symptoms and ON medications. iRBD symptoms and their severity were evaluated by the REM Sleep Behavior Disorder Screening Questionnaire (RBDSQ). Nonmotor symptoms and autonomic dysfunction were evaluated by the Non-Motor Symptom Questionnaire (NMSQ) and Scale for Outcomes in PD-Autonomic (SCOPA-AUT), respectively. Depressive state was measured using the 17-item Hamilton Depression Rating Scale (HAMD); the Sniffin' Sticks 16-item test (SS-16) was performed to assess olfactory function.

Plasma sample collection

Plasma samples (0.5-1 ml) were collected from subjects in the fasting state. For each blood collection, 2 ml of venous blood was collected into 10-ml BD K_2EDTA Vacutainer tubes (BD; Cat# 367525) to prevent coagulation. The anticoagulant-treated blood samples were immediately inverted several times and transferred to 2-ml conical tubes (Eppendorf; Cat# 0030120094). To obtain plasma, blood samples were centrifuged at 1,300 × *g* for 10 min at room temperature (RT). The upper layer containing plasma was transferred to new tubes and centrifuged twice at 2,500 × *g* for 15 min at RT to obtain platelet-poor plasma (PPP). Each plasma supernatant was carefully transferred to fresh 1-ml tubes and then stored at -80 °C until further use.

EV isolation and RNA extraction

Frozen plasma samples were thawed in a thermostat water bath for 2 min at 37 °C and then centrifuged at 2,500 × *g* for 15 min at 4 °C to remove precipitated proteins and lipids. Approximately 0.5-1 ml of plasma was used as the starting material for EV isolation with an exoRNeasy Midi kit (Qiagen; Cat# 77144) according to the manufacturer's instructions. For EV-RNA extraction, 1 ml of TRIzol reagent (Invitrogen; Cat# 15596026) was added to the column, which was centrifuged at 3,000 × *g* for 1 min. Total EV-RNA was isolated by TRIzol reagent according to the manufacturer's recommendations. The RNA pellet was dissolved in 11 μl of deionized water treated with diethylpyrocarbonate (DEPC); 1 μl was

used for yield measurement of EV-RNA by Quant-iT RiboGreen RNA Assay Kit (Thermo Fisher Scientific; Cat# R11490); 2 μl was used for quality and size distribution examination by a 2200 Bioanalyzer (Agilent Technologies, CA, USA). The remaining 8 μl was used for sncRNA library construction and sequencing. For EV characterization, 1 ml of XE buffer was added to the column and centrifuged at 3,000 × *g* for 1 min. Eluates were concentrated by ultrafiltration using Amicon Ultra-0.5 ml Centrifugal Filters with a molecular weight cutoff of 100 kDa (Millipore; Cat# UFC5003). EV was resuspended in 40 μl of phosphate-buffered saline (PBS) and was ready for downstream applications, including transmission electron microscopy (TEM), nanoparticle tracking analysis (NTA), and western blotting.

Transmission electron microscopy (TEM) analysis

Purified concentrated EV in PBS were mixed with an equal volume of 4% paraformaldehyde (PFA). Five microliters of resuspended pellets were added to a formvar-carbon-coated EM grid and absorbed for 20 min in a dry environment. The grids were washed once with PBS for 1 min, followed by once with 1% glutaraldehyde for 5 min and seven times with Milli-Q water for 2 min per wash. The grids were then negatively stained with 2.5% uranyl-oxalate solution for 10 min and air-dried for 5 min under incandescent light. Images were acquired on a Tecnai G2 Spirit TEM (FEI, The Netherlands) with a wide-angle AMT 2k CCD camera operating at 120 kV at the Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences.

Nanoparticle tracking analysis (NTA)

The particle size and concentration of purified EV in PBS were analyzed using a ZetaView PMX 110 (Particle Metrix, Germany) equipped with a 405-nm laser and a high-sensitivity sCMOS camera. The ZetaView system was calibrated using 110 nm polystyrene particles. All measurements were performed at RT. Furthermore, NTA measurement was recorded and analyzed at 11 positions for 60 seconds at a frame rate of 30 frames/second. The result for each sample is presented as an average of the 3 measurements. Particle movement was analyzed using NTA software (ZetaView 8.04.02 SP2).

Western blotting

For protein analysis, the protein concentrations of purified EV in PBS were quantified using a bicinchoninic acid (BCA) assay (Beyotime; Cat# P0010S). Ten micrograms of total protein of each sample were denatured in 4× sodium dodecyl sulfonate (SDS) buffer at 95 °C for 5 min before separated through 8% polyacrylamide SDS gels and transferred to nitrocellulose membranes. The membranes were blocked with 5% (wt/vol) fat-free milk in PBS with Tween (PBST) for 1 hour at RT prior to incubation with anti-Alix antibody (1:1000 dilution; Abcam; Cat# 186429), anti-CD63 antibody (1:200 dilution; Abcam; Cat# 193349), anti-Calnexin antibody (1:1000 dilution; Abcam; Cat# 22595), and anti-GRP94 antibody (1:500 dilution; Abcam; Cat# 108606). All primary antibodies were diluted in 5% fat-free milk, and blots were probed overnight at 4 °C. The membranes were washed four times for 10 min and incubated with HRP-conjugated secondary antibodies for 1 hour at RT. The secondary antibodies were goat anti-rabbit IgG for Alix, GRP94, or Calnexin (1:5000 dilution; MultiSciences; Cat# GAR007) and goat anti-mouse IgG

for CD63 (1:5000 dilution; MultiSciences; Cat# GAM007). The protein bands were detected with a Pierce ECL western blotting substrate kit (Thermo Fisher Scientific; Cat# R32209).

Construction of the cDNA libraries from plasma EV-RNA

EV-RNA obtained from the plasma of healthy controls, iRBD patients, and PD patients was profiled by EVsmall-seq. For each sample, 0.5 ng RNA per sample was used as input material for generating sncRNA libraries. Ligation reaction for 5' and 3' adapters was performed according to the procedure described previously with some modifications (manuscript in preparing). In brief, adapter-ligated RNA was mixed with ProtoScript II Reverse Transcriptase, RNase Inhibitor, and reaction Buffer (NEB; Cat# M0358), and incubated for 1 hour at 44 °C. After PCR amplification of RT products, the sequence-specific sgRNAs were assembled with Cas9 to cleave self-ligating adapters and RNA contaminants in the libraries. The oligo sequences of 3'adapter, biotin-RT primer, 5'adapter, adapter-sgRNA, ysRNA-sgRNA (Y RNA-derived small non-coding RNA sgRNA), P5 primer, and P7 primer used for sncRNA library construction were listed in **Additional file 1**. Size selection of PCR products was performed by high-resolution polyacrylamide gel electrophoresis through 6% gels. For miRNAs, the bands corresponding to 135-145 bp were purified for sequencing.

High throughput sequencing and data analyses of sncRNAs

All the cDNA libraries of sncRNAs were sequenced on NovaSeq (Illumina). We used cutadapt to clip adapters and filter out low-quality reads[25]. Reads failing to match the adaptor or reads with lengths shorter than 17 nt were discarded. Redundant sequences were collapsed as useful reads for further analysis[26]. Then, we aligned the reads to reference sequences by bowtie[27]. The reads that matched the 5' start site of annotated miRNAs and matched the 3' ends with at most 3 nt deletions and/or 3 nt additional sequences derived from pre-miRNAs were counted in the abundance of miRNAs. The miRNA expression level was normalized by size factor[28]. SncRNAs with more than one annotation were characterized in the following order: miRNA, ysRNA, tsRNA (tRNA-derived small non-coding RNA), rsRNA (rRNA-derived small non-coding RNA), snRNA, snoRNA, lncRNA, and mRNA. Sequences that were not annotated with any of the RNA categories above were classified as others. The miRNA expression profiles in human tissues were downloaded from TissueAtlas[29]. The gene ontology enrichment analysis of miRNA' target genes was performed by miRPathDB[30]. miRNA's target sites for corresponding genes were predicted by TargetScan[31].

Sources for sequences and genome assemblies

The genomic sequences of human (hg38) were downloaded from the University of California Santa Cruz (UCSC) Genome Browser [32] (http://genome.ucsc.edu/). Known RNA sequences were retrieved from the following databases: miRNA, miRbase (version 22.0)[33]; tRNAs, Genomic tRNA Database (http://lowelab.ucsc.edu/GtRNAdb); Y RNAs, rRNAs, 5.8S, 18S, 28S and 45S from NCBI GenBank

(http://www.ncbi.nlm.nih.gov/), 5S from Ensembl (http://www.ensembl.org/index.html); and snoRNAs, lncRNAs and mRNAs from Ensembl (http://www.ensembl.org/index.html).

Statistical analysis

We used the R software package for statistical analysis. Batch correction was conducted by limma[34]. Differentially expressed miRNAs were calculated by limma[34] (p value<0.05, fold change>1.4). We used the prcomp package for principal component analysis (PCA), which was visualized by ggplot2[35]. Unsupervised hierarchical clustering analysis was conducted and visualized from pheatmap[36]. The comparisons between two groups were performed with a Wilcoxon rank sum test, whereas comparisons among three or more groups were performed with a Kruskal-Wallis rank sum test. The statistical tests were performed and visualized by ggplot2 and ggpurb[37]. For the miRNAs with increasing expression in the healthy-iRBD-PD hierarchy, the mean value of miRNA expression in each group has to meet the condition of healthy < iRBD < PD and *P* value <0.05 based on the Kruskal-Wallis sum test; for the miRNAs with decreasing expression in the healthy-iRBD-PD hierarchy, the mean value of miRNA expression in each group has to meet the condition of healthy > iRBD > PD and *P* value <0.05 based on the Kruskal-Wallis rank sum test.

Features were selected with the R package boruta to find all relevant variables for machine learning[38]. Classification of different groups was carried out with the svm function in the R package e1071[39]. The log_2-transformed normalized read counts of every miRNA were defined as the input data. Models were trained on a training set (60% of data) and evaluated on a validation set (remaining 40% of the data), with an equal distribution of individuals in the disease groups **(Supplementary Table S1)**. The diagnostic efficacy was evaluated by receiver operating characteristic (ROC) curve analysis for the training and validation cohorts[40]. The comparison between areas under the curve (AUCs) of different classifiers was evaluated by the bootstrap method with 100 iterations.

Results

Plasma sample collection and EV-RNA isolation

A total of 169 participants were enrolled in this study and subsequently divided into three groups consisting of 56 iRBD patients, 53 PD patients, and 60 healthy individuals. The demographics and clinical characteristics of the enrolled participants are listed in Table 1. No significant differences were found in age distribution among the three groups, although the PD group contained a significantly higher proportion of female subjects. Average disease duration for iRBD patients were 7 years, and 5 years for PD patients. In addition, the iRBD patient group exhibited a higher average RBDSQ score than that of the PD group ($P < 0.001$), while PD patients showed lower SS-16 ($P = 0.03$) and higher SCOPA-AUT ($P = 0.001$) scores, on average, than iRBD patients. The HAMD scores were higher in PD patients than in healthy controls ($P < 0.001$). No significant differences in NMSQ and HAMD scores were observed between iRBD and PD groups.

Table 1
Demographic and clinical profiles of healthy, iRBD, and PD participants

Clinical parameters	Healthy	iRBD	PD	*P* value
No. of participants	60	56	53	/
Age	63.5±9.0	64.0±7.3	63.0±9.0	0.62
Sex (M/F)	35/25	34/22	25/28	<0.001
Disease duration (y)	/	7.0±5.1	6.9±12.4	/
UPDRS III	/	/	30.6±18.0	/
RBDSQ	/	8.1±3.2	4.4±3.2	<0.001
NMSQ	/	7.0±3.9	8.2±3.4	0.09
SS-16	/	8.8±4.4	7.1±3.5	0.03
HAMD	/	3.6±4.0	5.1±5.0	0.11
SCOPA-AUT	/	7.8±4.7	14.4±12.9	0.001

UPDRS III: the Unified PD Rating Scale III; RBDSQ: the REM Sleep Behavior Disorder Screening Questionnaire; NMSQ: Non-Motor Symptom Questionnaire; SS-16: the Sniffin' Sticks 16-item test; HAMD: the 17-item Hamilton Depression Rating Scale; SCOPA-AUT: Scale for Outcomes in PD-Autonomic.

To ensure obtain high-quality plasma EVs from patients, we established a procedure for plasma sample collection and protocols for EV isolation **(Supplementary Fig. S1A**, also see Materials and Methods**)**. EVs were isolated from patient plasma using an exoRNeasy Midi kit (Qiagen). The morphology and size distribution of EVs were evaluated by transmission electron microscope (TEM) and nanoparticle tracking analysis (NTA). The size of isolated EVs was ranged from 30 - 500 nm with a median diameter of 105 nm **(Supplementary Fig. S1B, C)**. Furthermore, enrichment for the EV markers Alix and CD63 was observed in the EVs isolated from the plasma samples **(Supplementary Fig. S1D)**. In contrast, Calnexin and GRP94 were undetectable in isolated EVs, which indicated the absence of contamination by endoplasmic reticulum proteins **(Supplementary Fig. S1D)**. Total EV RNA was then extracted using TRIzol reagent instead of QIAzol, which is the gold standard for RNA extraction and highly efficient for isolating small RNAs. Aliquots of purified EV RNAs were further analyzed by a Quant-iT RiboGreen RNA Assay Kit and 2200 Bioanalyzer. The EV RNA obtained from 0.5 milliliter (ml) of plasma ranged between 0.61 and 2.93 ng, and showed normal size distribution **(Supplementary Fig. S1E)** consistent with the previous study[20]. Approximately 0.5 ng of EV-RNA that passed the quantity and quality control criteria was used for small-RNA library construction and sequencing analyses.

Optimize sncRNA sequencing library construction method for low RNA inputs

Profiling the relative abundance of specific sncRNAs by high throughput sequencing typically requires tens of nanograms of total RNA[41,42] and our initial attempts to construct a highly quality EV-associated sncRNA library using commercially available kits were unsuccessful, mainly due to the effects of low RNA inputs that result in 5’ and 3’ adapters generating the predominant ligation byproducts which severely compromise subsequent PCR amplification and sequencing quality. We therefore optimized the conditions for 5’ and 3’ adapter ligation and introduced an extra 5’ exonuclease treatment step that effectively reduced adapter dimer formation. Previous studies have shown that Y RNA-derived small non-coding RNAs (ysRNAs) can comprise the most abundant sncRNA type in plasma[43,44], and the prevalence of ysRNAs can significantly increase sequencing cost and interfere with the amplification of other, low abundance sncRNAs during cDNA library construction. Since the sgRNA-guided Cas9 nuclease is capable of cleaving double-stranded DNA bearing a protospacer adjacent motif (PAM) sequence both *in vitro* and *in vivo*, we designed sgRNAs and introduced a Cas9/sgRNA *in vitro* cleavage step to reduce both the ysRNAs **(Supplementary Fig. S1F)** and the adapter dimers **(Additional file 1)**. The EV RNAs (0.5 to 2 ng) treated with Cas9/sgRNA showed comparable and clear miRNA bands, indicating that we were able to reduce the input of EV RNA to 1 ng or less (Fig. 1A). The Cas9/sgRNA *in vitro* cleavage step led to an effective reduction of both ysRNAs and adapter dimer byproducts, enabled a straightforward PAGE-based method for detection and recovery of miRNA products (~140 bp) (Fig. 1B, C). Although the Cas9/sgRNA treatment couldn’t remove ysRNAs completely, the cleavage efficiency could be further increased by raising the Cas9/sgRNA concentration in the reaction. The comparison of sequencing results between the pairwise samples with or without Cas9/sgRNA treatment showed high correlations among miRNA profiles (Spearman correlation coefficient ≥ 0.99) (Fig. 1D, E), which indicated that Cas9/sgRNA treatment had no obvious effect on the individual miRNA abundance. Through these modifications, we successfully reduced the total required EV-RNA inputs to 0.5 ng. We named this optimized library construction method for sncRNA sequencing with low EV-RNA inputs as EVsmall-seq.

Sequencing EV-associated sncRNAs in plasma samples

We then used the EVsmall-seq for high throughput sequencing-based profiling of EV-associated sncRNA expression in 169 human plasma samples. Each library was sequenced to a mean depth of 19 million reads (median depth: 15 million) **(**Fig. 1F, **Additional file 2)**, which is sufficient for the detection of low-abundance sncRNAs with 0.5 ng input EV-RNAs **(Supplementary Fig. S2)**. After low-quality and short reads being discarded, the average proportion of high quality reads in these samples was 82.3% (Fig. 1G), and 65.2% of these high-quality reads could be mapped to the reference sequence for sncRNAs (Fig. 1H), indicating the overall high quality of the sncRNA libraries. We found that the miRNAs (17-24 nt), ysRNAs (25-33 nt), tRNA-derived small non-coding (tsRNAs, 30-33 nt), and rRNA-derived small non-coding (rsRNAs, 17-20 nt) were the most prevalent types of sncRNAs associated with EVs (Fig. 1I, Fig. 1J). On average, 390 different miRNA species were detected in each sample (average normalized read count > 1, expressed frequency of all samples > 25%) **(**Fig. 1K, **Additional file 2-3)**.

Identification of PD-specific miRNA biomarkers

We next sought to determine whether plasma EV-associated miRNA profiles could be used to distinguish PD patients from healthy individuals. We identified a total of 89 upregulated and 38 downregulated miRNAs unique to the PD patient group (Fig. 2A, B, **Additional file 4**). Unsupervised hierarchical clustering and principal component analysis (PCA) of the differentially expressed miRNAs separated PD patient samples from healthy controls with minor overlap (Fig. 2C). We then constructed a support vector machine (SVM) classifier to distinguish PD patients. Feature selection from the training set (n=67) identified 16 informative miRNAs (**Additional file 5**), including 10 upregulated miRNAs (miR-27b-3p, miR-199a-5p, miR-151a-3p, miR-584-5p, miR-889-3p, miR-619-5p, miR-130b-5p, miR-197-3p, miR-4433b-5p, and miR-4433a-3p) and 6 downregulated miRNAs (miR-182-5p, miR-96-5p, miR-155-5p, miR-150-5p, miR-150-3p, and miR-3615). We then used a 5-fold cross-validation SVM algorithm to classify PD patients from healthy individuals, which showed 100% sensitivity, 87.5% specificity, and 94.03% accuracy in the training set (35 healthy controls and 32 patients with PD) and 88% sensitivity, 90.43% specificity and 89.13% accuracy in the validation data set (25 healthy controls and 21 patients with PD) (Fig. 2D). The classifier could also identify patients with PD from healthy individuals in both the training (AUC=0.918; 95% CI, 0.913-0.922) and validation (AUC=0.913; 95% CI, 0.907-0.92) sets (Fig. 2E). These results suggested that the 16 miRNA features from the sncRNA profile of PD patients could be used to distinguish PD samples from healthy controls.

Identification of iRBD-specific miRNA biomarkers

Given the high risk of developing α-synucleinopathies by iRBD patients, such as PD, PD dementia, dementia with Lewy bodies or multiple system atrophy, in addition to the high specificity and long interval between iRBD onset and clinical manifestations of α-synucleinopathies, the prodromal phase of this disorder represents a unique opportunity for potential disease interventions[45]. To identify patients in this early stage of pathogenesis, we further explored whether EV miRNAs could distinguish iRBD patients from healthy individuals. A total of 75 differentially upregulated and 46 downregulated miRNAs were identified among the total plasma EV-associated sncRNAs expressed in iRBD patients compared with healthy subjects (Fig. 3A, B, **Additional file 6**). Unsupervised hierarchical clustering and PCA based on the differentially detected miRNAs separated iRBD patient samples from control samples with minor overlap (Fig. 3C).

We performed feature selection (n=69) and resulted in a miRNA signature comprised of three informative features. Specifically, miR-27b-3p was upregulated in the iRBD group, while miR-182-5p and miR-7-5p were downregulated (**Additional file 7**). The diagnostic performance of this training set showed 97.14% sensitivity, 88.24% specificity, and 92.75% accuracy (Fig. 3D). Independent validation of the training model with a separate validation set (n=47) revealed 96% sensitivity, 86.36% specificity, and 91.49% accuracy (Fig. 3D). Ultimately, this classifier could effectively distinguish iRBD patients from control subjects in both the training (AUC 0.97; 95% CI, 0.967-0.972) and validation (AUC 0.969; 95% CI, 0.966-0.973) datasets (Fig. 3E), which indicated that the three miRNA signature features could sufficiently distinguish iRBD patients.

Notably, the predicted targets of miR-7-5p, miR-27b-3p, and miR-182-5p, the signature miRNAs for iRBD, were enriched in pathways such as ‘neuron projection’, ‘neurotransmitter secretion’, ‘generation of neurons’, and ‘nervous system development’ (**Supplementary Fig. S3**), including α-Synuclein, NLRP3, PINK1, and GDNF that have been previously shown to contribute to PD pathogenesis[46–49]. Moreover, we found that some miRNAs in iRBD 3-miRNA or PD 16-miRNA signature were highly expressed in the human brain, including miR-7-5p and miR-27b-3p (Fig. 3F), which suggested that those EV-associated miRNAs may be secreted by neuronal cells in the brain. In addition, miR-199a-5p and miR-182-5p were predicted to target the 3’UTR of several genes previously reported to contribute to PD pathogenesis, such as ATP13A2 and SCARB2[4] (**Supplementary Fig. S4**).

Consistently altered EV miRNAs in a healthy-iRBD-PD hierarchy

To next investigate whether (and which of) these miRNAs might be linked to PD development through iRBD, we compared the miRNA profiles in the plasma EVs derived from healthy, iRBD, and PD samples. We found 20 miRNAs that consistently increased or decreased in expression from control to iRBD to PD patients (Fig. 4A, B). In particular, let-7i-5p, miR-4433b-5p, miR-335-3p, miR-130b-5p, miR-766-3p, miR-744-5p, miR-1301-3p, miR-4433a-3p, miR-889-3p, and miR-4433a-5p all significantly increased in transcript abundance following the control<iRBD<PD hierarchy (Fig. 4A). By contrast, miR-10b-5p, miR-150-5p, miR-342-3p, miR-186-5p, miR-192-5p, miR-361-3p, miR-155-5p, miR-3615, miR-150-3p, and miR-500a-3p all decreased in expression following the order of control>iRBD>PD (Fig. 4B), which suggested that these miRNAs could possibly participate in both iRBD and PD. Consistent with this hypothesis, miR-150-5p expression was previously reported to be dysregulated during PD progression[50]. Extensive research conducted on let-7 revealed that let-7 could be involved in regulating neuronal degeneration in PD[51] and recognized by TLR7 to promote inflammation and neuronal death[52]. In addition, the accumulation of let-7 could lead to increased α-synuclein expression and decreased autophagy[53]. Intriguingly, miR-10b-5p, miR-342-3p, miR-186-5p, miR-361-3p, miR-155-3p, miR-130b-5p, and miR-744-5p also showed aberrant expression levels in neurodegenerative diseases such as Alzheimer’s disease, multiple system atrophy, and prion diseases[54]. However, other miRNAs, including miR-4433a-3p, miR-4433a-5p, miR-4433b-5p, and miR-192-5p, have not been previously linked to PD and therefore warrant closer investigation.

Discussion

Accessible and reliable biomarkers for early diagnosis of PD and iRBD are urgently needed to identify candidate therapeutic targets and to monitor disease progression during therapeutic interventions[5]. In this study, we investigated the expression of sncRNAs associated with plasma-derived EVs from 169 individuals using an improved library preparation protocol for high throughput sequencing. Our analysis revealed 16 EV-associated miRNA features that were diagnostically informative for PD, and three miRNA signature features that could serve as biomarkers for iRBD. These results demonstrated that high throughput sequencing-based detection and quantification of sncRNAs expression can provide more informative profiles that include all types and relative abundance of sncRNAs and their sequence

variants, thereby significantly improving the overall performance of machine learning diagnostic classifiers and ultimately the AUC values of miRNA biomarkers[11,13].

However, the current study also has several limitations. First, as a single-center study, it is uncertain whether these predictive markers are applicable to other populations with different exposures to environmental or genetic factors (i.e., ethnicity). Future work with larger cohorts from multiple centers is required to confirm these results, ideally through a prospective validation study. Second, only the miRNA signature was used in this study. A combination of miRNAs with v-PSG, neuroimaging, and neuropathological assessment may improve the sensitivity and specificity of diagnosis for prodromal PD (i.e., iRBD) and PD.

Conclusion

In conclusion, we established a cDNA library construction method, named EVsmall-seq, for high throughput sequencing of sncRNAs in plasma-EV using a low as 0.5 ng of total RNA input and identified miRNA signature features that could serve as biomarkers for distinguish iRBD and/or PD from healthy individuals with high sensitivity and accuracy. Moreover, our study provides a valuable resource for sncRNA profiles in plasma EVs from the plasma of iRBD and PD patients and reveals an effective and non-invasive diagnostic strategy, with relevant biomarkers, for neurodegenerative diseases.

Abbreviations

AUC: Areas under the curve; BCA: Bicinchoninic acid; DEPC: Diethylpyrocarbonate; EV: Extracellular vesicle; HAMD: Hamilton Depression Rating Scale; iRBD: Idiopathic rapid eye movement sleep behavior disorder; miRNAs: microRNAs; MDS: Movement Disorder Society; NTA: Nanoparticle tracking analysis; NMSQ: Non-Motor Symptom Questionnaire; PD: Parkinson's disease; PCA: Principal component analysis; PAM: Protospacer adjacent motif; PPP: Platelet-poor plasma; PAGE: Polyacrylamide gel electrophoresis; PBMC: Peripheral blood mononuclear cell; PBS: Phosphate-buffered saline; PFA: Paraformaldehyde; REM: Rapid eye movement; RBC: Red blood cell; RT-qPCR: Real-time quantitative PCR; TEM: Transmission electron microscope; RT: Room temperature; RBDSQ: REM Sleep Behavior Disorder Screening Questionnaire; ROC: Receiver operating characteristic; rsRNAs: rRNA-derived small non-coding; SVM: Support vector machine; sncRNAs: Small non-coding RNAs; SCOPA-AUT: Scale for Outcomes in PD-Autonomic; SS-16: Sniffin' Sticks 16-item test; tsRNAs: tRNA-derived small non-coding RNAs; UPDRS: Unified PD Rating Scale; ysRNAs: Y RNA-derived small non-coding RNAs

Declarations

Acknowledgments

We are grateful to the patients and relatives for their participation in this project. We thank the healthy control subjects who voluntarily contributed samples for this study. We also thank members of our

laboratory for their helpful discussions. We would like to acknowledge the HPC Storage & Network Service Platform from Shanghai Institute of Biochemistry and Cell Biology for providing IT support and supplying the computing resources.

Author contributions

Ligang Wu and Jun Liu designed the study. Yuanyuan Li collected clinical samples and processed the blood samples. Ying Cao processed the plasma EV, characterized the EV, extracted EV-RNA, and constructed the sncRNA libraries. Aonan Zhao and Ningdi Luo contribute to the sample collection of iRBD patients. Hongdao Zhang optimized the method for EV-associated sncRNAs library construction. Haisheng Zhou contributed to the extraction of EV and RNAs. Wei Liu performed the bioinformatics analysis and interpreted the data. Wei Liu, Ying Cao, Yuanyuan Li, and Ligang Wu wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB19040102 to Ligang Wu), the National Key R&D Program of China (2017YFA0504401 to Ligang Wu), the National Natural Science Foundation of China (31970607 and 31470781 to Ligang Wu; 81471287, 81501097, 61901250 and 91949129 to Jun Liu; 82001341 to Yuanyuan Li), and the Shanghai Sailing Program (19YF1429800 to Yuanyuan Li).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The deep sequencing data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus under accession number GSE166070 (secure token evijmiektjqhngx).

Ethics approval and consent to participate

The study was reviewed and approved by all countries' respective Ethics committees and all participants signed an informed consent to take part in the research.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding agency

Strategic Priority Research Program of the Chinese Academy of Sciences (Ligang Wu), National Key R&D Program of China (Ligang Wu), National Natural Science Foundation of China (Ligang Wu, Jun Liu and Yuanyuan Li), and Shanghai Sailing Program (Yuanyuan Li).

Relevant conflicts of interests/financial disclosures

The authors have declared no conflicts of interest.

References

1. de Lau, L. M. L. & Breteler, M. M. B. Epidemiology of Parkinson's disease. *Lancet Neurol.* **5**, 525–535 (2006).
2. Ascherio, A. & Schwarzschild, M. A. The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol.* **15**, 1257–1272 (2016).
3. Matar, E. & Lewis, S. J. REM sleep behaviour disorder: not just a bad dream. *Med. J. Aust.* **207**, 262–268 (2017).
4. Kalia, L. V. & Lang, A. E. Parkinson's disease. *The Lancet* **386**, 896–912 (2015).
5. Miller, D. B. & O'Callaghan, J. P. Biomarkers of Parkinson's disease: present and future. *Metabolism.* **64**, S40-46 (2015).
6. Eusebi, P. *et al.* Diagnostic utility of cerebrospinal fluid α-synuclein in Parkinson's disease: A systematic review and meta-analysis. *Mov. Disord. Off. J. Mov. Disord. Soc.* **32**, 1389–1400 (2017).
7. Bäckström, D. C. *et al.* Cerebrospinal Fluid Patterns and the Risk of Future Dementia in Early, Incident Parkinson Disease. *JAMA Neurol.* **72**, 1175–1182 (2015).
8. Barbour, R. *et al.* Red blood cells are the major source of alpha-synuclein in blood. *Neurodegener. Dis.* **5**, 55–59 (2008).
9. Teixeira dos Santos, M. C., Bell, R. & da Costa, A. N. Recent developments in circulating biomarkers in Parkinson's disease: the potential use of miRNAs in a clinical setting. *Bioanalysis* **8**, 2497–2518 (2016).
10. Cardo, L. F. *et al.* Profile of microRNAs in the plasma of Parkinson's disease patients and healthy controls. *J. Neurol.* **260**, 1420–1422 (2013).
11. Cressatti, M. *et al.* Salivary microR-153 and microR-223 Levels as Potential Diagnostic Biomarkers of Idiopathic Parkinson's Disease. *Mov. Disord.* mds.27935 (2019) doi:10.1002/mds.27935.
12. Ma, W. *et al.* Serum miR-221 serves as a biomarker for Parkinson's disease. *Cell Biochem. Funct.* **34**, 511–515 (2016).
13. Ravanidis, S. *et al.* Circulating Brain-enriched MicroRNAs for detection and discrimination of idiopathic and genetic Parkinson's disease. *Mov. Disord.* mds.27928 (2019) doi:10.1002/mds.27928.
14. Pillman, K. A., Goodall, G. J., Bracken, C. P. & Gantier, M. P. miRNA length variation during macrophage stimulation confounds the interpretation of results: implications for miRNA quantification by RT-qPCR. *RNA N. Y. N* **25**, 232–238 (2019).

15. O'Brien, K., Breyne, K., Ughetto, S., Laurent, L. C. & Breakefield, X. O. RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat. Rev. Mol. Cell Biol.* (2020) doi:10.1038/s41580-020-0251-y.
16. An, T. *et al.* Exosomes serve as tumour markers for personalized diagnostics owing to their important role in cancer metastasis. *J. Extracell. Vesicles* **4**, 27522 (2015).
17. Thind, A. & Wilson, C. Exosomal miRNAs as cancer biomarkers and therapeutic targets. *J. Extracell. Vesicles* **5**, 31292 (2016).
18. Théry, C. Cancer: Diagnosis by extracellular vesicles. *Nature* **523**, 161–162 (2015).
19. Vlaeminck-Guillem, V. Extracellular Vesicles in Prostate Cancer Carcinogenesis, Diagnosis, and Management. *Front. Oncol.* **8**, 222 (2018).
20. Sundar, I. K., Li, D. & Rahman, I. Small RNA-sequence analysis of plasma-derived extracellular vesicle miRNAs in smokers and patients with chronic obstructive pulmonary disease as circulating biomarkers. *J. Extracell. Vesicles* **8**, 1684816 (2019).
21. Zhang, J.-T. *et al.* Plasma extracellular vesicle microRNAs for pulmonary ground-glass nodules. *J. Extracell. Vesicles* **8**, 1663666 (2019).
22. Cheng, L. *et al.* Small RNA fingerprinting of Alzheimer's disease frontal cortex extracellular vesicles and their comparison with peripheral extracellular vesicles. *J. Extracell. Vesicles* **9**, 1766822 (2020).
23. Burgos, K. *et al.* Profiles of Extracellular miRNA in Cerebrospinal Fluid and Serum from Patients with Alzheimer's and Parkinson's Diseases Correlate with Disease Status and Features of Pathology. *PLoS ONE* **9**, e94839 (2014).
24. Postuma, R. B. *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord. Off. J. Mov. Disord. Soc.* **30**, 1591–1601 (2015).
25. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
26. Gordon, A. FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab.cshl.edu/fastx_toolkit. (2010).
27. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
28. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
29. Ludwig, N. *et al.* Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* **44**, 3865–3877 (2016).
30. Kehl, T. *et al.* miRPathDB 2.0: a novel release of the miRNA Pathway Dictionary Database. *Nucleic Acids Res.* **48**, D142–D147 (2020).
31. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, (2015).

32. Haeussler, M. *et al.* The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* **47**, D853–D858 (2019).
33. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
34. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
35. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
36. Kolde, R. *pheatmap: Pretty Heatmaps*. (2019).
37. Kassambara, A. *ggpubr: 'ggplot2' Based Publication Ready Plots*. (2018).
38. Kursa, M. B. & Rudnicki, W. R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **36**, (2010).
39. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. (2019).
40. Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. ROCR: visualizing classifier performance in R. *Bioinformatics* **21**, 7881 (2005).
41. Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539–543 (2008).
42. Ohnishi, Y. *et al.* Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. *Nucleic Acids Res.* **38**, 5141–5151 (2010).
43. Yeri, A. *et al.* Total Extracellular Small RNA Profiles from Plasma, Saliva, and Urine of Healthy Subjects. *Sci. Rep.* **7**, 44061 (2017).
44. Godoy, P. M. *et al.* Large Differences in Small RNA Composition Between Human Biofluids. *Cell Rep.* **25**, 1346–1358 (2018).
45. Högl, B., Stefani, A. & Videnovic, A. Idiopathic REM sleep behaviour disorder and neurodegeneration - an update. *Nat. Rev. Neurol.* **14**, 40–55 (2018).
46. Junn, E. *et al.* Repression of α-synuclein expression and toxicity by microRNA-7. *Proc. Natl. Acad. Sci.* **106**, 13052–13057 (2009).
47. Zhou, Y. *et al.* MicroRNA-7 targets Nod-like receptor protein 3 inflammasome to modulate neuroinflammation in the pathogenesis of Parkinson's disease. *Mol. Neurodegener.* **11**, 28 (2016).
48. Kim, J. *et al.* miR-27a and miR-27b regulate autophagic clearance of damaged mitochondria by targeting PTEN-induced putative kinase 1 (PINK1). *Mol. Neurodegener.* **11**, 55 (2016).
49. Roser, A.-E. *et al.* miR-182-5p and miR-183-5p Act as GDNF Mimics in Dopaminergic Midbrain Neurons. *Mol. Ther. Nucleic Acids* **11**, 9–22 (2018).
50. Li, H. *et al.* MicroRNA-150 serves as a diagnostic biomarker and is involved in the inflammatory pathogenesis of Parkinson's disease. *Mol. Genet. Genomic Med.* **8**, e1189 (2020).
51. Goh, S. Y., Chao, Y. X., Dheen, S. T., Tan, E.-K. & Tay, S. S.-W. Role of MicroRNAs in Parkinson's Disease. *Int. J. Mol. Sci.* **20**, 5649 (2019).

52. Lehmann, S. M. *et al.* An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration. *Nat. Neurosci.* **15**, 827–835 (2012).
53. Shamsuzzama, null, Kumar, L. & Nazir, A. Modulation of Alpha-synuclein Expression and Associated Effects by MicroRNA Let-7 in Transgenic C. elegans. *Front. Mol. Neurosci.* **10**, 328 (2017).
54. Basak, I., Patil, K. S., Alves, G., Larsen, J. P. & Møller, S. G. microRNAs as neuroregulators, biomarkers and therapeutic agents in neurodegenerative diseases. *Cell. Mol. Life Sci. CMLS* **73**, 811–827 (2016).

Figures

Figure 1

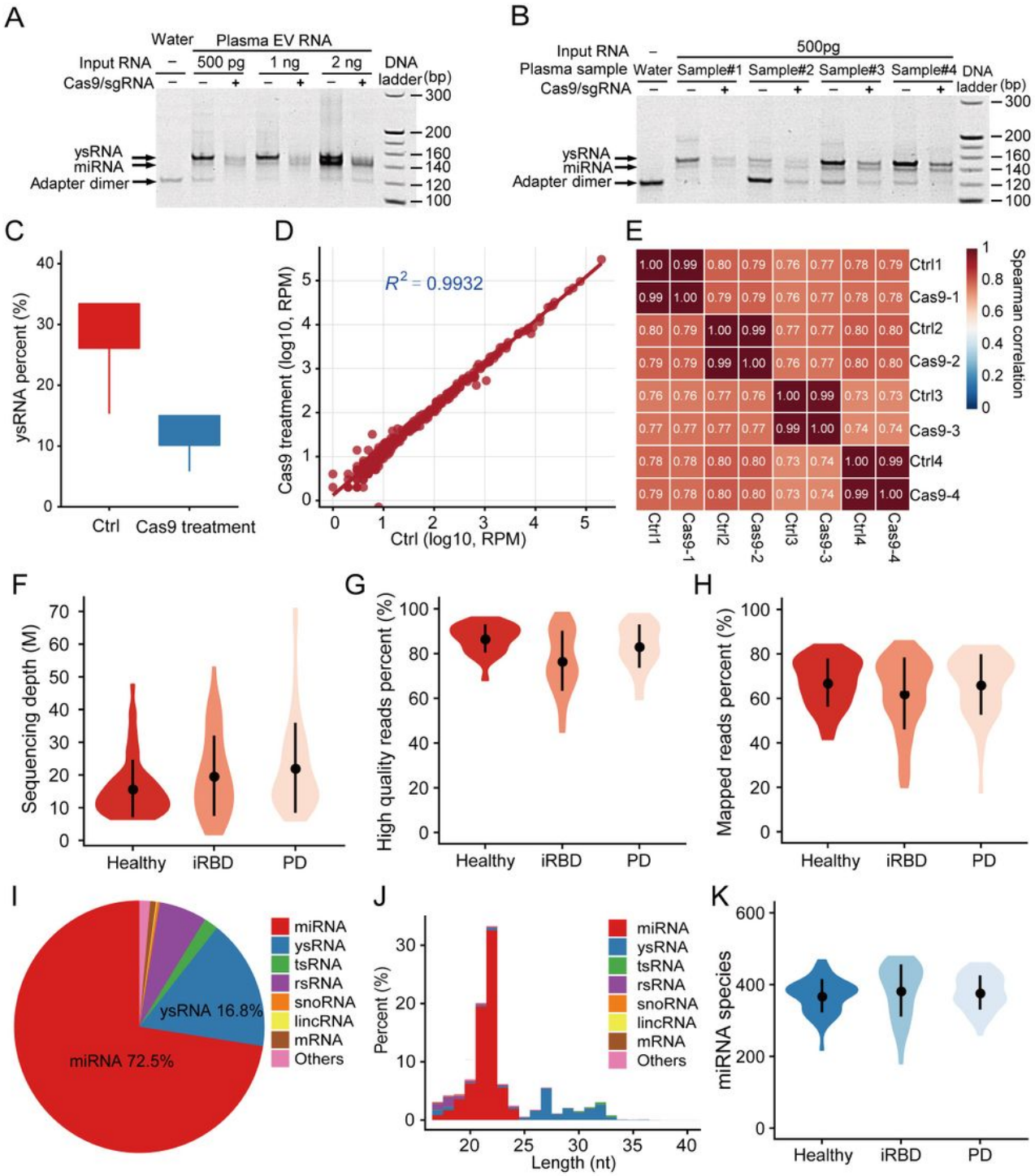

Figure 1

Plasma EV-associated sncRNAs profiling by high throughput sequencing

A. Comparison of the miRNA bands generated from 0.5 to 2 nanograms (ng) of EV RNA from human plasma. The miRNA band was approximately 140 bp and detected by 6% polyacrylamide gel electrophoresis (PAGE). **B.** Comparison of cDNA library construction with or without Cas9/sgRNA

treatment. For each plasma sample, 0.5 ng of total EV-RNA was used as the input, and the PCR amplification products were analyzed by 6% PAGE. **C.** A boxplot of the percentage of ysRNAs among total mapped reads in the cDNA library with (blue) or without (red) Cas9/sgRNA treatment. Each group has four biological replicates. **D.** Representative scatter plot of miRNA expression in the cDNA library with (y-axis) or without (x-axis) Cas9/sgRNA treatment. Each dot represents one miRNA species. The adjusted R-squared value of the linear regression is shown in blue. **E.** Spearman correlation heatmap of miRNA expression in the cDNA libraries treated with or without Cas9/sgRNA. The correlation coefficient of four groups of paired samples is shown in the plot. Ctrl: without Cas9/sgRNA treatment, Cas9: Cas9/sgRNA treatment. **F**. The sequencing depth of sncRNAs in healthy, iRBD, and PD samples. M: one million reads. The mean and standard deviation of sequencing depth is shown in black. **G, H**. The percentage of high-quality reads (G) or mapped reads (H) in healthy, iRBD, and PD samples. **I**, **J**. The average proportion (I) and length distribution (J) of various types of sncRNAs in plasma-derived EVs. **K**. The number of detected miRNA species in plasma EVs.

Figure 2

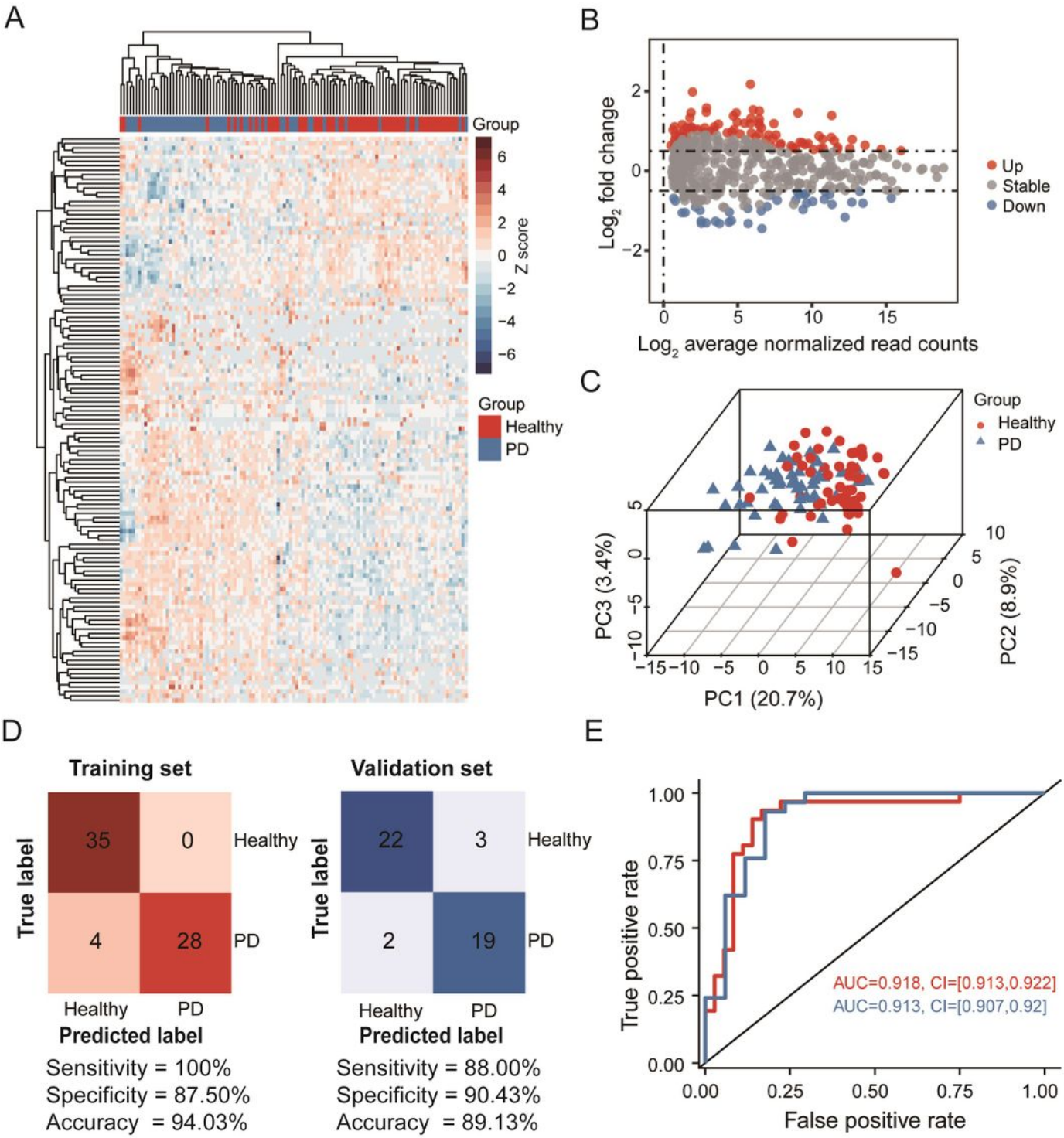

Figure 2

EV-associated miRNA biomarkers for detecting PD

A. Unsupervised clustering heatmap of differentially expressed miRNAs for PD patients (blue) and healthy individuals (red). Each row represents a differentially expressed miRNA. **B**. MA-plot of differentially expressed miRNAs between PD patients and healthy individuals by transforming read

abundance onto M (log_2 fold change) and A (log_2 average normalized read counts) scales. Up- or down-regulated miRNAs are shown in red or blue, respectively. **C**. Three-dimensional scatter plot of principal component analysis (PCA) for the differentially expressed miRNAs between PD patients and healthy individuals. The percentage of variance explained by PC1, PC2, and PC3 are shown in the label. **D**. Confusion matrix of the support vector machine (SVM) classifier for PD patients and healthy individuals in the training set (left) and validation set (right). The sensitivity, specificity, and accuracy are shown at the bottom. **E**. The receiver operating characteristic (ROC) curve of the SVM classifier in the training set (Red, n = 67) and validation set (blue, n = 46). The values of area under the curve (AUC) and 95% confidence interval (CI) are shown in the plot.

Figure 3

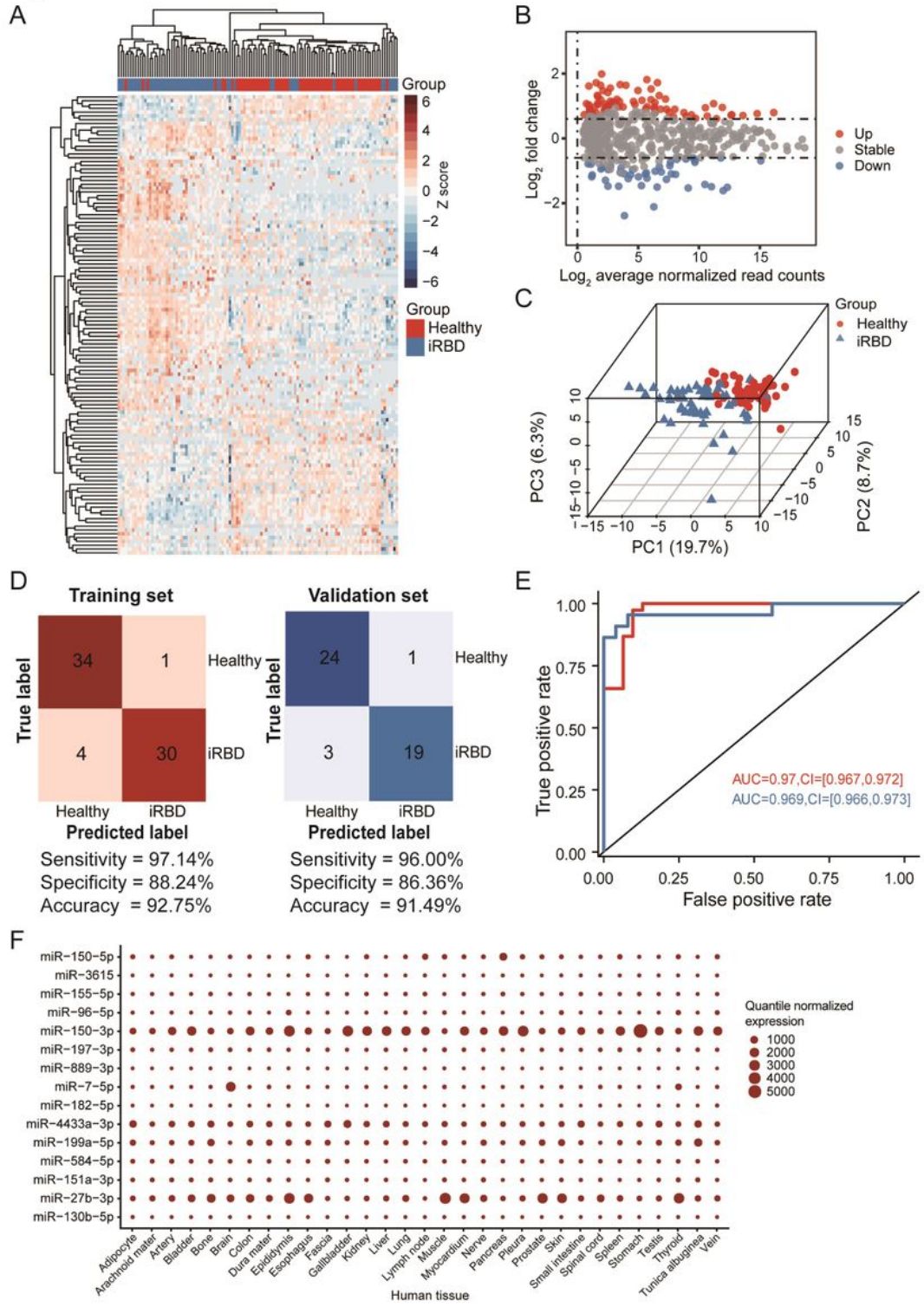

Figure 3

EV-associated miRNA biomarkers for detecting iRBD

A. Unsupervised clustering heatmap of differentially expressed miRNAs of iRBD patients and healthy individuals (red). Each row represents a differentially expressed miRNA. **B.** MA-plot of the differentially expressed miRNAs between iRBD patients and healthy individuals. The up- or down-regulated miRNAs are

shown in red or blue, respectively. **C.** Three-dimensional scatter plot of PCA for the differentially expressed miRNAs between iRBD patients and healthy individuals. **D**. Confusion matrix of the SVM classifier for iRBD patients and healthy individuals in the training set (left) and validation set (right). The sensitivity, specificity, and accuracy are shown at the bottom. **E.** The ROC curve of the SVM classifier in training set (red, n = 69) and validation set (blue, n = 47). The values of AUC and 95% confidence interval are shown in the plot. **F**. The expression levels of PD- and iRBD-specific miRNA biomarkers in human tissues.

Figure 4

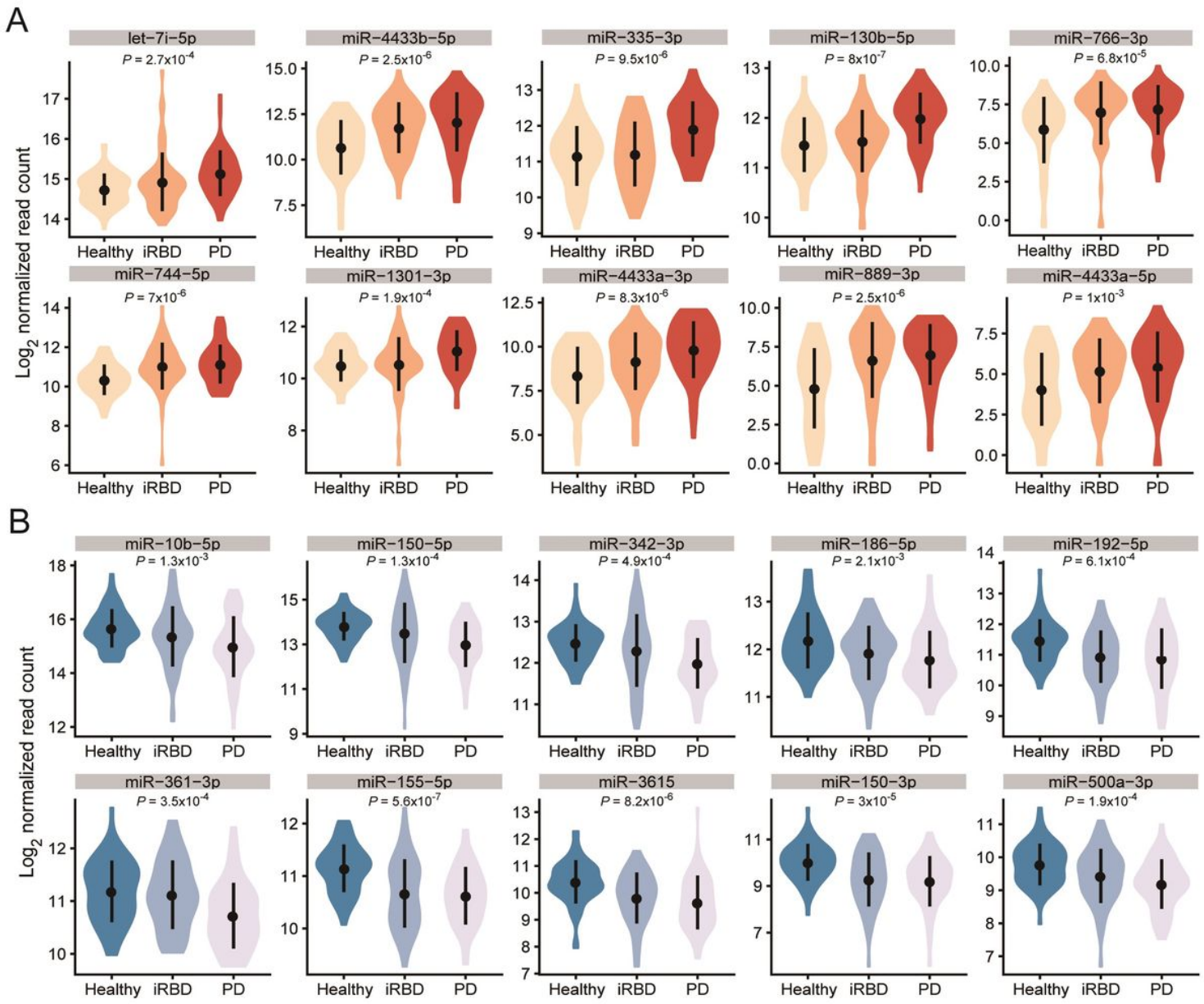

Figure 4

miRNAs with consistently increases or decreases in expression from healthy to iRBD to PD samples

A, B. Violin plot of miRNAs with an expression that progressively increases (A) or decreases (B) from healthy control to iRBD to PD samples. The *P* values of a Kruskal-Wallis rank sum test for comparisons among three groups are shown in the boxplot for each miRNA. The mean and standard deviation of miRNA expression levels are shown in black.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx
- Additionalfile4.xlsx
- Additionalfile5.xlsx
- Additionalfile6.xlsx
- Additionalfile7.xlsx
- SupplementaryFigureS1.tif
- SupplementaryFigureS2.tif
- SupplementaryFigureS3.tif
- SupplementaryFigureS4.tif
- SupplementaryTableS1.xlsx