scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Chromatin accessibility and transcriptome in human neuronal model exposed to Parkinson's environmental toxins

Jingchao Hong[1,2,3] & Juanjuan Huang[1] ✉

Parkinson's disease (PD) is a complex neurodegenerative disorder where environmental factors play a predominant role in sporadic cases. While environmental toxins are implicated in PD pathogenesis via epigenetic pathways, the specific alterations in chromatin accessibility induced by these toxins remain poorly characterized. To address this gap, we established *in vitro* models using the human neuroblastoma cell line SH-SY5Y, a well-established neuronal model, exposed to the PD-associated environmental toxins rotenone and MPP⁺. We performed RNA sequencing (RNA-seq) to profile the transcriptome and Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) to map genome-wide chromatin accessibility landscapes under toxin-exposed conditions. This integrated dataset provides a comprehensive resource detailing both gene expression and chromatin accessibility dynamics in response to PD-relevant environmental neurotoxins and will facilitate investigations into transcriptional regulation, biomarker discovery, and the epigenetic basis of environmentally linked sporadic PD.

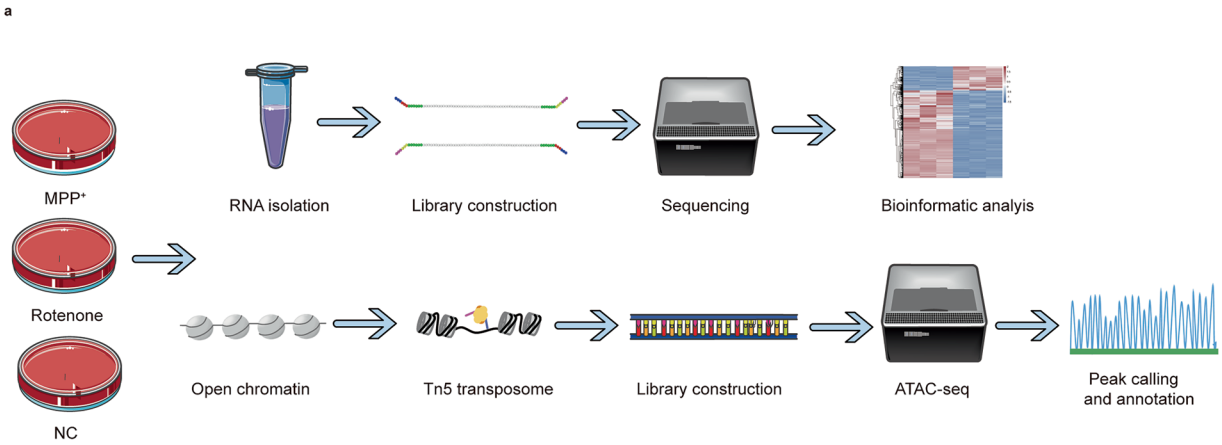
Background & Summary

Parkinson's disease (PD), the second most common neurodegenerative disorder globally, is clinically defined by motor symptoms including bradykinesia, rigidity, and postural instability[1–3]. Its neuropathological hallmarks are the accumulation of α-synuclein-positive Lewy bodies and the selective loss of dopaminergic neurons in the substantia nigra pars compacta[4,5]. While the precise etiology is unknown, a complex interplay of genetic susceptibility and environmental exposures is recognized[6]. Critically, sporadic PD represents 80–90% of cases, underscoring the major contribution of non-genetic factors[7].

Epidemiological studies strongly associate exposure to specific environmental toxins, such as the pesticides rotenone and paraquat, and the synthetic compound 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) or its metabolite MPP^+, with increased PD risk[8,9]. These toxins, particularly MPP^+ and rotenone (widely used to model PD), induce neuronal damage partly through mitochondrial complex I inhibition, leading to energy failure, oxidative stress, and apoptosis[10–12]. However, emerging research indicates that their impact extends beyond direct cytotoxicity, implicating epigenetic dysregulation as a critical mediator of environmentally driven PD pathology. This concept is supported by alterations in histone modifications (e.g., acetylation) observed in PD patient brains and toxin-exposed cellular models[13,14].

Despite this, a fundamental layer of epigenetic regulation—genome-wide chromatin accessibility, which governs the exposure of regulatory DNA elements to transcription factors and other macromolecules—remains largely unexplored in the context of environmental toxin-induced PD pathology[15]. The lack of integrated datasets capturing both chromatin accessibility and gene expression changes in relevant models hinders a systems-level understanding of the transcriptional dysregulation underlying environmental toxin-induced neurotoxicity.

[1]Department of Neurology, Xiangya Hospital, Central South University, Changsha, 410008, Hunan, China. [2]Centre for Medical Genetics & Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, 410008, Hunan, China. [3]Institute of Molecular Precision Medicine, Xiangya Hospital, Key Laboratory of Molecular Precision Medicine of Hunan Province, Central South University, Changsha, 410008, Hunan, China. ✉e-mail: jjhuang_2413@csu.edu.cn



b

Step	Description
RNA Sample Collection	Collect samples for RNA-seq.
RNA Extraction and Quality Control	Extract RNA and perform quality control checks.
Library Construction, QC, Sequencing	Construct sequencing libraries, perform quality control, and sequence.
Raw Data	Obtain raw sequencing data (in FastQ format).
Data Quality Assessment	Assess the quality of sequencing data.
Data Filtering and Cleaning	Filter out low-quality reads and adapter sequences.
Alignment to Reference Genome/Transcriptome	Align filtered reads to the reference genome or transcriptome.
Alignment Result Evaluation	Evaluate the quality and alignment rate of the results.
Gene Expression Quantification	Quantify gene expression levels using tools like HTSeq or featureCounts.
Normalization	Normalize gene expression data (e.g., TPM, FPKM, RPKM, or DESeq2).
Differential Expression Analysis	Analyze differentially expressed genes between conditions or groups using tools like DESeq2 or edgeR.
Functional Annotation	Annotate and categorize differentially expressed genes.

c

Step	Description
ATAC-seq Sample Collection	Collect samples for ATAC-seq.
Library Construction, QC, Sequencing	Perform library construction, quality control, and sequencing.
Raw Data	Obtain raw data from sequencing.
Data Quality Assessment	Assess the quality of sequencing data.
Data Filtering	Filter the raw data to remove low-quality data.
Alignment to Reference Genome	Align the data to the reference genome.
Alignment Result Evaluation	Evaluate the quality of the alignment results.
Differential Peak Analysis	Analyze differential peaks under different conditions.

Fig. 1 Overview of study design and data analysis workflow. (**a**) Collection and preparation of experimental samples. (**b**) The data analysis workflow for mRNA sequencing data. (**c**) The data analysis workflow for ATAC-seq data.

To fill this critical data gap and provide a foundational resource for the PD research community, we employed an integrated multi-omics approach. We utilized the human dopaminergic neuroblastoma cell line SH-SY5Y, a well-established and extensively characterized *in vitro* model for PD research, exposed to two prototypical PD-linked environmental toxins: MPP^+ and rotenone. Our study generated two primary datasets: 1. Transcriptome Profiles: Comprehensive mRNA expression data obtained via RNA-seq in MPP^+ and rotenone 24 h treatments. 2. Chromatin Accessibility Maps: Genome-wide profiles of open chromatin regions generated using ATAC-seq methodology (Fig. 1). This dataset thus provides a comprehensive analysis of environmental toxin effects on the neuronal chromatin accessibility, offering novel insights into PD pathogenesis.

Methods

Sample preparation and collection. SH-SY5Y cells (ATCC) were cultured in DMEM/F12 (1:1) medium (Gibco, MA, USA) supplemented with 10% fetal bovine serum (Gibco, MA, USA). Cells were plated in 6-well plates at a density of 5×10^5 cells/mL and incubated overnight at 37 °C with 5% CO_2. Experimental groups were treated with 5 mM MPP^+ (Selleck, Houston, USA) and 10 μM rotenone (Sigma, Saint Louis, USA). These concentrations were selected based on established literature demonstrating their ability to induce significant PD-relevant neurotoxicity in SH-SY5Y cells[16,17]. Negative control (NC) wells received equivalent volumes of dimethyl sulfoxide (DMSO) vehicle (Sigma, Saint Louis, USA). After 24-hour exposure, cells were harvested for RNA sequencing and ATAC-seq analyses. Three biological replicates per group (NC, MPP^+, Rotenone) were processed for RNA-seq, while two replicates per group were analyzed by ATAC-seq (Table 1).

RNA sequencing library preparation. Total RNA was extracted from experimental samples using TRIzol reagent (Thermo Fisher Scientific, San Diego, USA) according to the manufacturer's protocol. RNA integrity was verified by 1% agarose gel electrophoresis, and genomic DNA contamination was assessed. RNA purity (OD260/280 and OD260/230 ratios) was determined using a NanoDrop spectrophotometer, concentration quantified with a Qubit 2.0 Fluorometer, and integrity evaluated via an Agilent 2100 Bioanalyzer.

Group	Sample ID	Sequencing strategy
NC	NC1, NC2, NC3	mRNA sequencing
	N1, N2	ATAC sequencing
MPP^+	MPP1, MPP2, MPP3	mRNA sequencing
	M1, M2	ATAC sequencing
Rotenone	Rot1, Rot2, Rot3	mRNA sequencing
	R1, R2	ATAC sequencing

Table 1. Overview of experimental samples and sequencing strategy.

mRNA was enriched using poly(A) capture beads and purified. The enriched mRNA was fragmented by incubation at elevated temperature. First-strand cDNA synthesis was performed using fragmented mRNA as template in a reverse transcription system (Thermo Fisher Scientific, San Diego, USA). During second-strand synthesis, end repair and 3′-adenylation (A-tailing) were simultaneously completed. Illumina adapters were ligated to the cDNA fragments, followed by purification with Hieff NGS® DNA Selection Beads (Yeasen, Shanghai, China) for size selection. Libraries were amplified by PCR and quality-controlled. Final sequencing was conducted on an Illumina NovaSeq X Plus platform with paired-end 150 bp reads.

Transcripts expression quantification. Raw sequencing reads containing adapter contaminants or low-quality bases were processed to ensure data integrity for downstream analyses. Quality filtering was performed using fastp (v0.18.0) with the following parameters[18]:

(1) Removal of adapter-containing reads;
(2) Exclusion of reads with >10% undefined nucleotides (N);
(3) Discarding reads where >50% of bases exhibited Q-scores ≤ 20.

Subsequently, Bowtie2 (v2.2.8) aligned reads to a ribosomal RNA (rRNA) reference database, with rRNA-mapped reads eliminated[19]. The resulting high-quality clean reads were utilized for transcriptome assembly and quantification.

A HISAT2 (v2.1.0) index was constructed from the reference genome Ensembl_release111, and filtered paired-end reads were mapped using default parameters. Transcript assembly was conducted with StringTie (v1.3.1) in reference-guided mode[20,21]. Expression abundance of each transcript was quantified as TPM (Transcripts Per Kilobase Million) values through RSEM (v1.2.19)[22], reflecting normalized read counts relative to gene length and sequencing depth. It is critical to note that TPM values were generated for visualization and within-sample comparison only.

Principal component analysis (PCA). PCA was implemented using the R package gmodels. This dimensionality reduction technique transforms high-dimensional correlated variables (e.g., gene expression profiles) into orthogonal principal components that capture maximal variance. PCA visualization elucidated intrinsic sample relationships and dataset structure.

Differential expression analysis. Differential expression analysis of mRNAs between the MPP^+-treated and control groups, as well as the rotenone-treated and control groups, was performed using DESeq2 (v1.20.0)[23]. The analysis utilized the raw read count matrix as input. DESeq2 employs a negative binomial Wald test to assess statistical significance. To account for multiple testing, P-values were adjusted using the Benjamini-Hochberg method. Differentially expressed mRNAs (DEmRNAs) were identified with the thresholds of an adjusted P-value < 0.05 and $|\log_2(\text{fold change})| \geq 1$ (Fig. S1a,b). The complete DEmRNA datasets have been deposited in Figshare and are available in the file titled "RNA-seq differential expression". Subsequently, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted using the OmicShare tools for the following gene sets: DEmRNAs specific to MPP^+ treatment, those specific to rotenone treatment, and the overlapping DEmRNAs between the two treatment groups. The background gene set for all enrichment analyses consisted of all protein-coding genes detected in our SH-SY5Y cell transcriptome data. The tool employs a hypergeometric test for enrichment analysis, and P-values were adjusted for multiple testing using the Benjamini-Hochberg method. Significantly enriched pathways were defined as those with an adjusted P-value < 0.05 (Fig. S1c–f).

Library construction for ATAC-seq. ATAC-seq was performed according to established protocols[24]. Briefly, 50,000 cells per sample were lysed to isolate nuclei. The nuclei were then extracted and incubated with Tn5 transposase (37 °C, 30 min) for tagmentation. Illumina TruSeq adapters (N5 and N7) were ligated at a final concentration of 0.5 μM, followed by library amplification via PCR. Amplified libraries were purified using Novoprotein DNA Clean Beads (Novoprotein, Soochow, China). DNA concentration was quantified with a Qubit 2.0 Fluorometer (Thermo Fisher, MA, USA), and fragment size distribution was assessed using an Agilent 2100 Bioanalyzer. After quality control, 150-bp paired-end sequencing was conducted on an Illumina NovaSeq 6000 platform (PersonalBio, Shanghai, China).

ATAC-seq data analysis. Raw reads containing adapter contaminants and low-quality bases were filtered using fastp (v0.22.0) to obtain high-quality clean data[18]. Paired-end clean reads were aligned to the Ensembl_release_111 reference genome using Bowtie2 (v2.4.1) with default parameters[25]. Peak calling was performed with

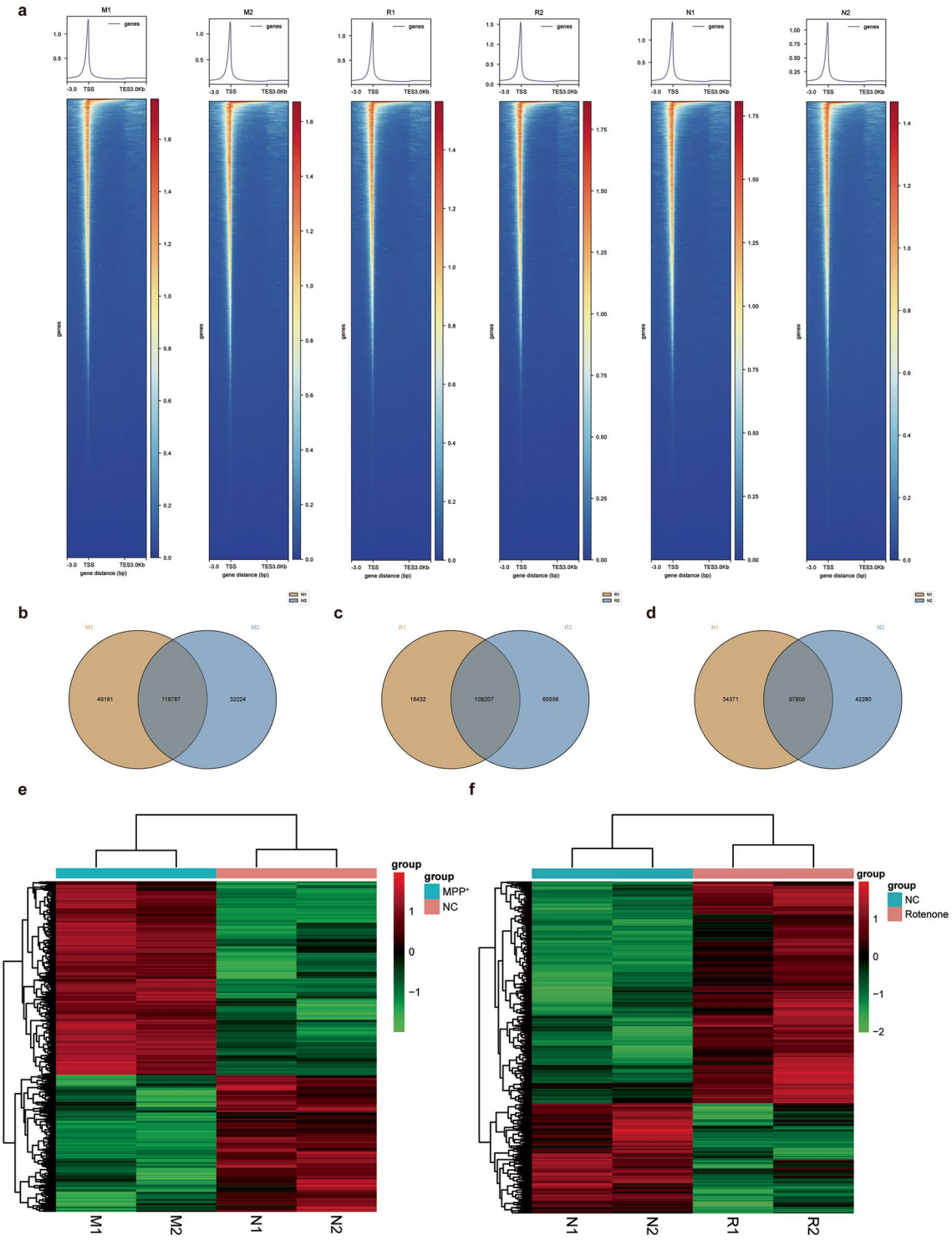


Fig. 2 Reads distribution analysis and peak calling of ATAC-seq data. (**a**) Read density within 5 kb upstream of the transcription start site (TSS) and 3 kb downstream of the transcription end site (TES). (**b**) Overlapping relationship between M1 and M2 within the MPP^+-treated group. (**c**) Overlapping relationship between R1 and R2 within the rotenone-treated group. (**d**) Overlapping relationship between N1 and N2 within the NC group. (**e**) Hierarchical clustering analysis of MPP^+-treated group and NC group. (**f**) Hierarchical clustering analysis of rotenone-treated group and NC group.

MACS3 (v3.0.0a7) in narrowPeak mode (adjusted p-value < 0.05)[26]. Signal distributions across genomic regions were visualized using deepTools (v3.5.1) (Fig. 2a)[27].

The reproducibility of biological replicates was assessed by identifying peaks common to and unique among replicates within each treatment group using BEDTools intersect (v2.30.0), with a requirement of a minimum

Sample ID	Raw reads	Clean reads	Q20 (%)	Q30 (%)	GC content (%)
NC1	44294584	44172280	98.05	94.62	48.84
NC2	37582116	37487114	97.96	94.39	48.90
NC3	42347934	42234850	97.93	94.28	49.01
MPP1	47251854	47119914	97.93	94.38	49.24
MPP2	44101138	43993890	98.00	94.46	49.20
MPP3	39339706	39234612	98.03	94.58	49.44
Rot1	41864082	41770868	98.15	94.78	48.64
Rot2	40262316	40176668	98.05	94.58	48.80
Rot3	39089090	38977398	98.05	94.63	48.92

Table 2. Summary statistics for mRNA sequencing data.

Sample	Clean reads	Mapped Reads(%)	Unmapped Reads(%)
NC1	44172280	434598 (0.98%)	43737682 (99.02%)
NC2	37487114	378354 (1.01%)	37108760 (98.99%)
NC3	42234850	433660 (1.03%)	41801190 (98.97%)
MPP1	47119914	174710 (0.37%)	46945204 (99.63%)
MPP2	43993890	166238 (0.38%)	43827652 (99.62%)
MPP3	39234612	142764 (0.36%)	39091848 (99.64%)
Rot1	41770868	1058342 (2.53%)	40712526 (97.47%)
Rot2	40176668	1001042 (2.49%)	39175626 (97.51%)
Rot3	38977398	967388 (2.48%)	38010010 (97.52%)

Table 3. Summary statistics of ribosome mapping for mRNA sequencing data.

Sample ID	Total	Unique Mapped(%)	Multiple Mapped(%)	Total Mapped(%)
NC1	43737682	32247855 (73.73%)	892434 (2.04%)	33140289 (75.77%)
NC2	37108760	27341985 (73.68%)	749698 (2.02%)	28091683 (75.70%)
NC3	41801190	30595633 (73.19%)	838137 (2.01%)	31433770 (75.20%)
MPP1	46945204	40070001 (85.35%)	1183887 (2.52%)	41253888 (87.88%)
MPP2	43827652	37691620 (86.00%)	1110466 (2.53%)	38802086 (88.53%)
MPP3	39091848	33488724 (85.67%)	990339 (2.53%)	34479063 (88.20%)
Rot1	40712526	33884730 (83.23%)	1108402 (2.72%)	34993132 (85.95%)
Rot2	39175626	32659449 (83.37%)	1088228 (2.78%)	33747677 (86.14%)
Rot3	38010010	31673785 (83.33%)	1049895 (2.76%)	32723680 (86.09%)

Table 4. Summary of mapping results from RNA-seq.

of 50% reciprocal overlap. These overlaps were quantified and visualized using Venn diagrams (Fig. 2b–d). For downstream quantitative comparison, a unified set of consensus peaks was generated by merging peak sets from all samples using BEDTools merge (v2.30.0)[28]. Read counts for each consensus peak region were quantified using FeatureCounts (v2.0.3).

Differential peak analysis was conducted using DESeq2 (v1.38.3) with a significance threshold of absolute $\log_2$ fold change > 1 and P-value < 0.05, identifying chromatin accessibility changes in both MPP^+-treated versus control and rotenone-treated versus control comparisons[23]. Hierarchical clustering heatmaps visualizing these differential peaks were generated with the pheatmap package (v1.0.12) in R, utilizing z-score normalized read densities across all significant peaks (Fig. 2e,f)[29].

All consensus peaks were annotated with ChIPseeker (v1.34.1) using hierarchical priority: Promoter > 5′ UTR > 3′ UTR > Exon > Intron > Downstream > Distal Intergenic[30]. Nearest genes were defined as the closest transcriptional units within 100 kb of peak summits. All differential peak genomic coordinates and quantitative matrices have been deposited in Figshare. The genomic annotations for all peaks, together with the lists of peaks exhibiting significantly increased or decreased accessibility in both the MPP^+- versus control and rotenone- versus control comparisons, are provided in the file "ATAC peak differential expression". The gene lists associated with these differential peaks for each comparison are available in the separate file "ATAC MPP ROT Diff Gene". The corresponding numbers of up- and down-regulated genes are presented in Fig. S2a,b, respectively, while the overlapping genes are depicted in Fig. S2c.

The joint analysis of ATAC-sequencing and RNA-sequencing. An integrated analysis of ATAC-seq and RNA-seq data was performed to identify high-confidence candidate genes that exhibited concurrent chromatin remodeling and transcriptional changes in response to MPP^+ and rotenone. Genes associated with differential

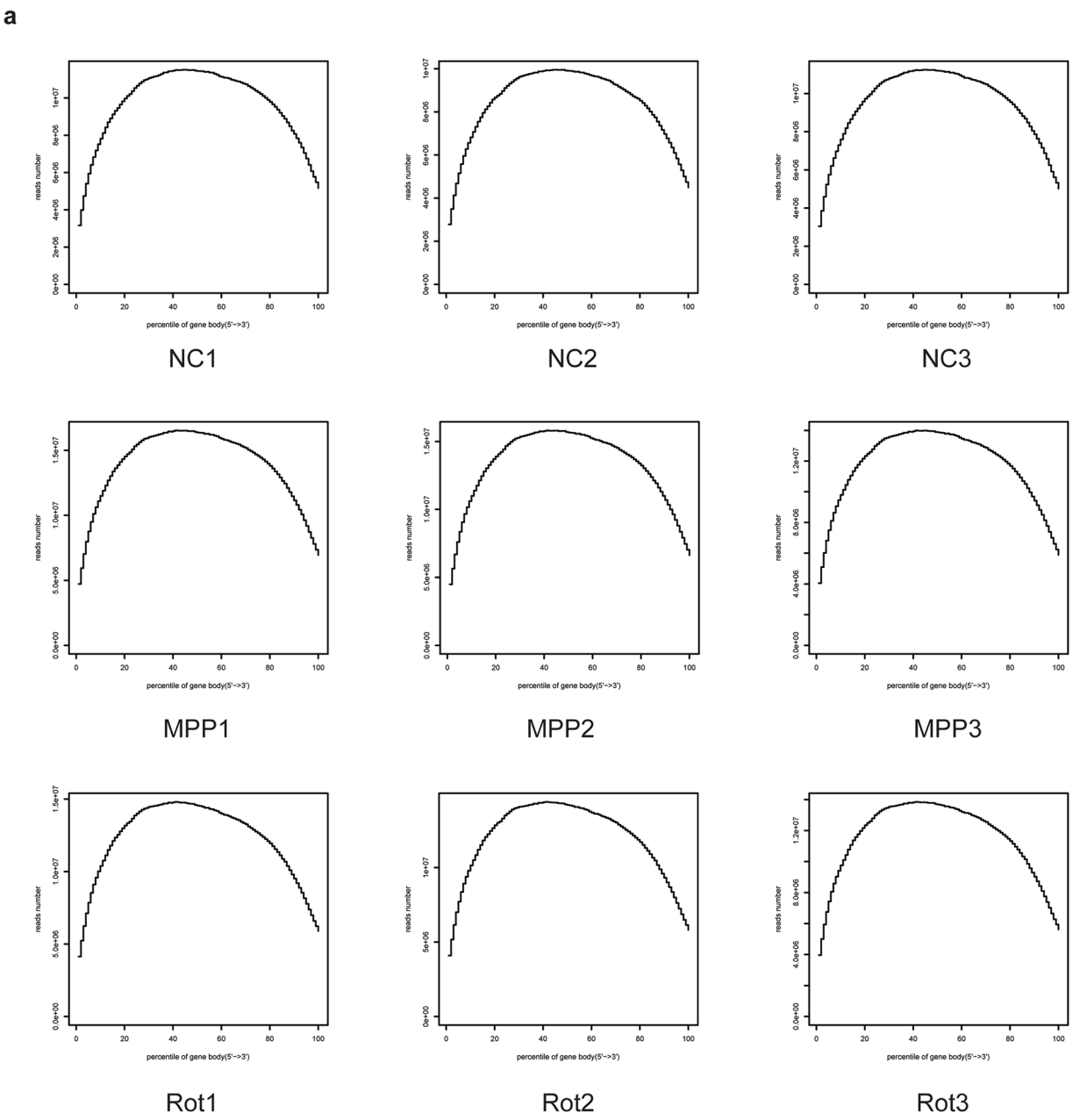


Fig. 3 Gene body coverage analysis of RNA-sequencing data. (**a**) Read coverage across the gene body (from 5′ to 3′ end) is shown for MPP^+-treated, rotenone-treated, and NC samples.

chromatin accessibility peaks were intersected with DEmRNAs (Fig. S2d). The resulting overlapping gene set was then subjected to KEGG pathway enrichment analysis using the OmicShare tools with P-value < 0.05 set as the significance threshold to identify associated signaling pathways (Fig. S2e). The gene sets comprising overlapping differentially accessible genes from ATAC-seq and overlapping DEmRNAs from RNA-seq that were commonly altered by both MPP^+ and rotenone treatments have been deposited in Figshare under the file "MPP ROT ATAC RNASEQ Diff Gene". The final high-confidence candidate gene set, derived from the intersection of ATAC-seq and RNA-seq data, is available in the separate file "MPP ROT ATAC RNASEQ overlap Diff Gene". Additionally, the complete results of the KEGG pathway enrichment analysis have been deposited in Figshare under the file "ATAC and RNA seq diff overlap functional analysis".

Data Records

The datasets generated in this study have been deposited in publicly accessible repositories as follows:

Raw sequencing data. All raw FASTQ files for RNA-seq and ATAC-seq are available in the China National GeneBank Database (CNGBdb) under SuperSeries accession number CNP0008342 or in the Gene Expression Omnibus (GEO) repository under the accession number GSE315111 and GSE315113[31–33].

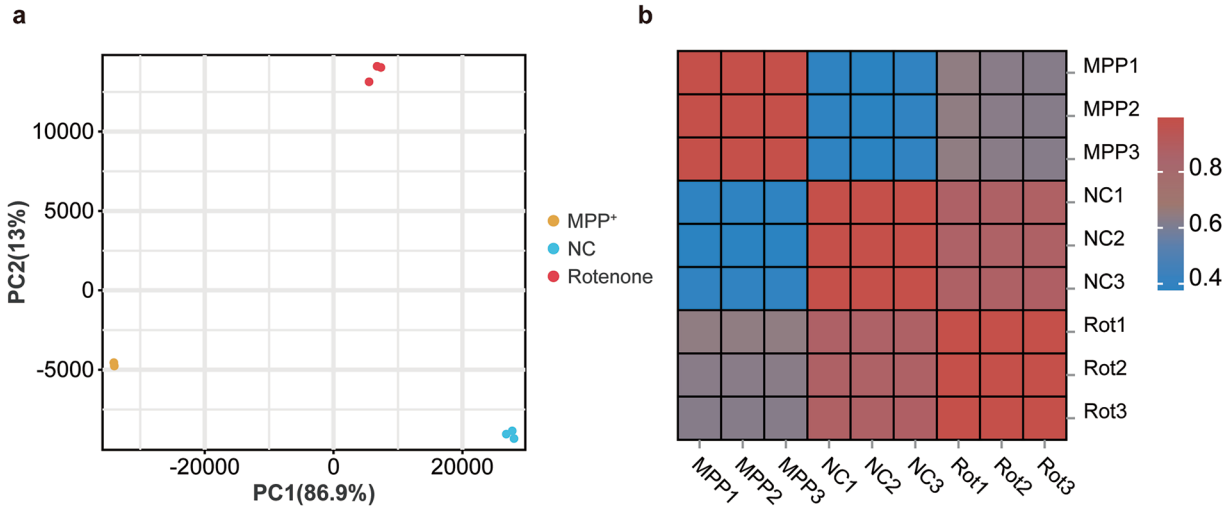


Fig. 4 RNA-sequencing reproducibility analysis. (**a**) PCA plot showing clustering between MPP$^+$-treated samples, rotenone-treated samples and NC samples based on their similarity. (**b**) Heatmap of Pearson correlation analysis of experimental samples within MPP$^+$-treated, rotenone-treated and control groups.

Processed data (Figshare). The following processed data files are available via Figshare[34]:

RNA-seq differential expression.zip. Contains gene expression data and differential expression results comparing MPP$^+$ with the control groups and rotenone with the control groups.

ATAC peak differential expression.zip. Contains annotated chromatin accessibility peaks for all samples (control, MPP$^+$ and rotenone), along with differential peak annotations.

ATAC MPP ROT Diff Gene.xlsx. Lists genes annotated from differential peaks in the MPP$^+$ vs. control and rotenone vs. control comparisons.

MPP ROT ATAC RNASEQ Diff Gene.xlsx. Provides the overlapping genes identified from separate RNA-seq and ATAC-seq analyses for each treatment comparison.

MPP ROT ATAC RNASEQ overlap Diff Gene.xlsx. Contains the overlapping differentially expressed genes from both RNA-seq and ATAC-seq analyses in the MPP$^+$ and rotenone treatments.

ATAC and RNA seq diff overlap functional analysis.zip. Includes KEGG pathway enrichment results for the overlapping differentially expressed genes from both omics analyses.

Technical Validation

RNA-sequencing quality validation. Raw sequencing data underwent filtration via Fastp (v0.18.0) to eliminate adapter-contaminated reads, sequences lacking 3′ adapters or insert tags, reads containing >10% undetermined bases (N), homopolymeric tracts (poly A/T/G/C), and low-quality reads with >50% bases exhibiting Q-scores ≤ 20. Post-filtering metrics confirmed retention of ≥97% high-quality clean reads across all samples, with Q20 (>97%), Q30 (>94%), and GC content (45–55%) documented in Table 2. The remaining high-quality sequencing reads were aligned to the species-specific ribosomal RNA (rRNA) reference database using Bowtie2, with rRNA-mapped reads constituting <3% of total clean reads across all samples (Table 3)[25]. Subsequent alignment to the human reference genome achieved an average mapping rate of 83.33% across samples (Table 4). Gene body coverage analysis further revealed highly uniform read distribution along transcripts from the 5′ to 3′ end, with a smooth and nearly symmetrical profile. The absence of apparent 3′-end bias indicated high RNA integrity and consistent library preparation quality among all samples (Fig. 3).

RNA-sequencing reproducibility control. To validate RNA-seq reproducibility, PCA was performed on sample-level expression profiles, assessing inter-sample variance and identifying potential outliers. Distinct clustering patterns emerged for MPP$^+$-treated, rotenone-treated, and control groups in the PCA plot (Fig. 4a), confirming group-specific transcriptional signatures. Inter-replicate consistency was further quantified by pairwise Pearson correlation coefficients, with all biological replicates exhibiting $R > 0.99$ (Fig. 4b), demonstrating exceptional biological reproducibility.

Differential gene expression analysis verification. To independently validate the differential gene expression results from RNA-seq, we performed RT-qPCR on the same RNA samples. The expression pattern of a selected gene, Trib3, as determined by RT-qPCR, was highly concordant with the RNA-seq data (Fig. 5a,b).

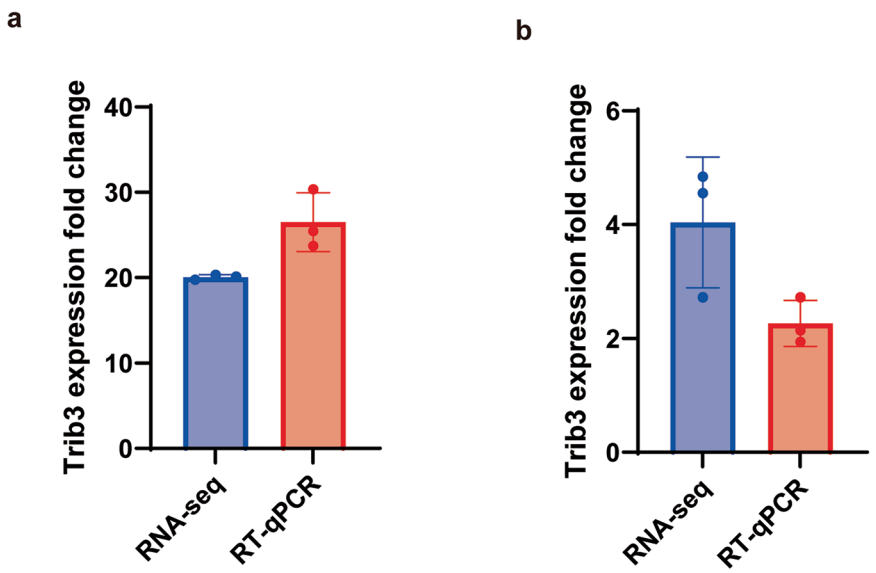


Fig. 5 Differential gene expression analysis verification. (**a**) Comparison of Trib3 mRNA fold change following MPP^+ treatment, as measured by RNA-seq and RT-qPCR, N = 3. (**b**) Comparison of Trib3 mRNA fold change following rotenone treatment, as measured by RNA-seq and RT-qPCR, N = 3.

Sample ID	Raw reads	Clean reads	Q20 (%)	Q30 (%)	Non-mitochondrial mapped (%)
N1	64777880	54809296	95.21	89.09	54009324 (98.54)
N2	67755814	57248942	95.01	88.76	56439089 (98.58)
M1	67615168	56523844	95.22	89.02	55781966 (98.68)
M2	63046482	53746566	95.02	89.12	53029986 (98.66)
R1	65510306	53728482	93.64	85.45	52983129 (98.61)
R2	85277564	70312270	94.86	88.45	69418323 (98.72)

Table 5. Summary statistics for ATAC-seq data.

Sample ID	All reads	Mapped reads	Mapped rate(%)
N1	54809296	54009324	98.54
N2	57248942	56439089	98.58
M1	56523844	55781966	98.68
M2	53746566	53029986	98.66
R1	53728482	52983129	98.61
R2	70312270	69418323	98.72

Table 6. Summary of mapping results from ATAC-seq.

Sample ID	Unique reads	Peak reads	FRiP score
N1	40293812	20269215	0.5030
N2	45586371	18928400	0.4152
M1	46588651	27308808	0.5861
M2	44143875	22080774	0.5002
R1	39699359	18830236	0.4743
R2	52757275	32896197	0.6235

Table 7. Summary of FRiP scores for each sample from ATAC-seq.

ATAC-sequencing quality control. ATAC-seq data processing was initiated by adapter trimming and quality filtering with fastp (v0.18.0) to remove reads containing >50% bases with a Q-score ≤ 20[18]. Quality assessment verified the retention of ≥82% high-quality clean reads across all samples, with Q20 (>93%) and Q30 (>85%) rates summarized in Table 5. Cleaned reads were subsequently aligned to the reference genome using Bowtie2 (v2.4.1)[25], achieving a mapping rate exceeding 98% (Table 6). Following alignment, mitochondrial-mapped

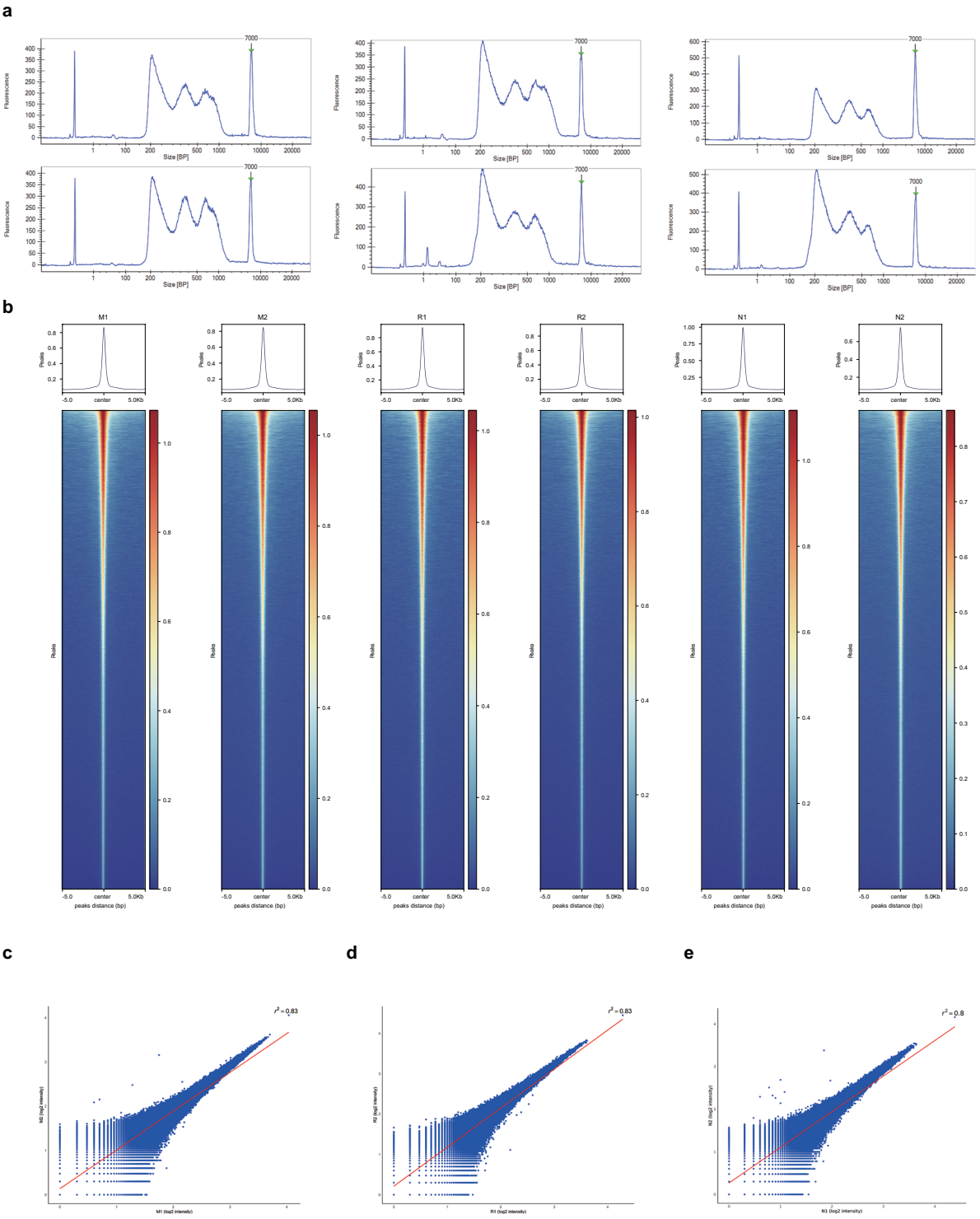


Fig. 6 ATAC-sequencing quality assessments. (**a**) ATAC-seq data quality control metrics of fragment size distribution. (**b**) ATAC-seq sequencing read enrichment around transcription start sites (TSS). (**c**) Pearson correlation between MPP^{+}-treatment samples. (**d**) Pearson correlation between rotenone-treatment samples. (**e**) Pearson correlation between NC samples.

sequences and PCR duplicates were removed. The FRiP (Fraction of Reads in Peaks) scores were all above 0.4, indicating good signal-to-noise ratio (Table 7). Fragment size distribution displayed characteristic nucleosome periodicity, coupled with significant read density enrichment near transcription start sites (Fig. 6a,b), confirming robust chromatin accessibility profiling. Inter-replicate reproducibility was assessed through Pearson correlation analysis of normalized chromatin accessibility profiles. High correlation coefficients were maintained within treatment groups: 0.83 (MPP^{+}-treated), 0.83 (rotenone-treated), and 0.80 (control) (Fig. 6c–e), demonstrating

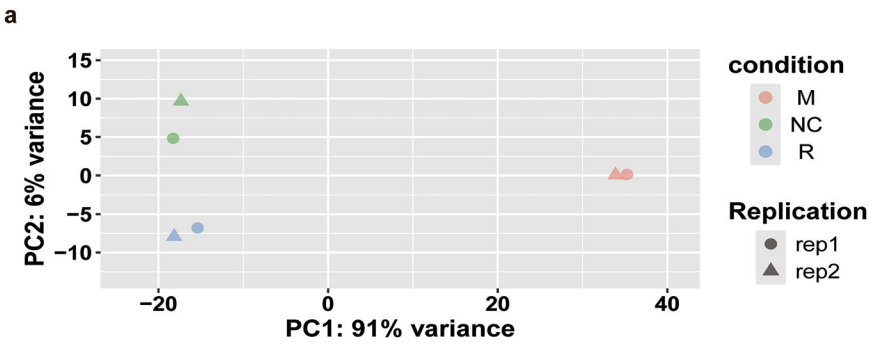


Fig. 7 ATAC-sequencing reproducibility analysis. (**a**) PCA plot showing clustering between MPP$^+$(M)-treated samples, rotenone(R)-treated samples and NC(N) samples based on their similarity.

consistent chromatin landscape capture across biological replicates. Furthermore, principal component analysis (PCA) revealed group-specific signatures of chromatin accessibility, reinforcing the differential accessibility patterns among experimental conditions (Fig. 7).

Data availability

Raw sequencing data from this study are publicly available in the CNGBdb under the accession number CNP0008342 or in the GEO repository under the accession number GSE315111 and GSE315113[31–33]. The processed data files are available at the Figshare repository[34].

Code availability

The parameters of bioinformatic tools applied in this study have been described in the Method section. If no parameter is provided, the default is set. No custom code was employed.

Received: 28 July 2025; Accepted: 12 January 2026;
Published online: 05 February 2026

References

1. Berardelli, A., Sabra, A. & Hallett, M. Physiological mechanisms of rigidity in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* **46**, 45–53, https://doi.org/10.1136/jnnp.46.1.45 (1983).
2. Berardelli, A., Rothwell, J. C., Thompson, P. D. & Hallett, M. Pathophysiology of bradykinesia in Parkinson's disease. *Brain* **124**, 2131–2146, https://doi.org/10.1093/brain/124.11.2131 (2001).
3. Mirelman, A. *et al*. Gait impairments in Parkinson's disease. *The Lancet Neurology* **18**, 697–708, https://doi.org/10.1016/S1474-4422(19)30044-4 (2019).
4. Gibb, W. & Lees, A. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* **51**, 745–752, https://doi.org/10.1136/jnnp.51.6.745 (1988).
5. Surmeier, D. J. Determinants of dopaminergic neuron loss in Parkinson's disease. *The FEBS Journal* **285**, 3657–3668, https://doi.org/10.1111/febs.14607 (2018).
6. Warner, T. T. & Schapira, A. H. Genetic and environmental factors in the cause of Parkinson's disease. *Ann. Neurol.* **53**, S16–S25, https://doi.org/10.1002/ana.10487 (2003).
7. Thomas, B. & Beal, M. F. Parkinson's disease. *Hum. Mol. Genet.* **16**, R183–R194, https://doi.org/10.1093/hmg/ddm159 (2007).
8. Tanner, C. M. *et al*. Rotenone, paraquat, and Parkinson's disease. *Environ. Health Perspect.* **119**, 866–872, https://doi.org/10.1289/ehp.100283 (2011).
9. Langston, J. W. The MPTP story. *Journal of Parkinson's disease* **7**, S11–S19, https://doi.org/10.3233/JPD-179006 (2017).
10. Perier, C., Bové, J., Vila, M. & Przedborski, S. The rotenone model of Parkinson's disease. *Trends Neurosci.* **26**, 345–346, https://doi.org/10.1016/S0166-2236(03)00144-9 (2003).
11. Smeyne, R. J. & Jackson-Lewis, V. The MPTP model of Parkinson's disease. *Mol. Brain. Res.* **134**, 57–66, https://doi.org/10.1016/j.molbrainres.2004.09.017 (2005).
12. Sherer, T. B., Betarbet, R. & Greenamyre, J. T. Environment, mitochondria, and Parkinson's disease. *The Neuroscientist* **8**, 192–197, https://doi.org/10.1177/1073858402008003 (2002).
13. Toker, L. *et al*. Genome-wide histone acetylation analysis reveals altered transcriptional regulation in the Parkinson's disease brain. *Mol. Neurodegener.* **16**, 1–20, https://doi.org/10.1186/s13024-021-00450-7 (2021).
14. Park, G. *et al*. Regulation of histone acetylation by autophagy in Parkinson disease. *J. Biol. Chem.* **291**, 3531–3540, https://doi.org/10.1074/jbc.M115.675488 (2016).
15. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**, 207–220, https://doi.org/10.1038/s41576-018-0089-8 (2019).
16. Aymerich, M. S. *et al*. Neuroprotective effect of JZL184 in MPP+-treated SH-SY5Y cells through CB2 receptors. *Mol. Neurobiol.* **53**, 2312–2319, https://doi.org/10.1007/s12035-015-9213-3 (2016).
17. Wu, X. *et al*. Rifampicin prevents SH-SY5Y cells from rotenone-induced apoptosis via the PI3K/Akt/GSK-3β/CREB signaling pathway. *Neurochem. Res.* **43**, 886–893, https://doi.org/10.1007/s11064-018-2494-y (2018).
18. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890, https://doi.org/10.1093/bioinformatics/bty560 (2018).
19. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359, https://doi.org/10.1038/nmeth.1923 (2012).
20. Pertea, M. *et al*. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295, https://doi.org/10.1038/nbt.3122 (2015).
21. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667, https://doi.org/10.1038/nprot.2016.095 (2016).

22. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* **12**, 1–16, https://doi.org/10.1186/1471-2105-12-323 (2011).
23. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 *Genome Biol.* **15**, 1–21, https://doi.org/10.1186/s13059-014-0550-8 (2014).
24. Zhang, X. *et al.* Cold-induced FOXO1 nuclear transport aids cold survival and tissue storage. *Nat. Commun.* **15**, 2859, https://doi.org/10.1038/s41467-024-47095-w (2024).
25. Langdon, W. B. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *Biodata Min.* **8**, 1–7, https://doi.org/10.1186/s13040-014-0034-0 (2015).
26. Zeng, L. & Zhang, B. Exploring the Genomic Landscape: An In-depth ChIP-seq Analysis Protocol for Uncovering Protein-DNA Interactions. *Curr. Protocol.* **3**, e909, https://doi.org/10.1002/cpz1.909 (2023).
27. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191, https://doi.org/10.1093/nar/gku365 (2014).
28. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842, https://doi.org/10.1093/bioinformatics/btq033 (2010).
29. Kolde, R. & Kolde, M. R. Package 'pheatmap'. *R package* **1**, 790 (2015).
30. Wang, Q. *et al.* Exploring epigenomic datasets by ChIPseeker. *Curr. Protocol.* **2**, e585, https://doi.org/10.1002/cpz1.585 (2022).
31. Hong, J. & Huang, J. *CNGBdb: CNP0008342.* https://db.cngb.org/data_resources/project/CNP0008342 (2025).
32. Hong, J. & Huang, J. *GEO.* https://identifiers.org/geo/GSE315111 (2025).
33. Hong, J. & Huang, J. *GEO.* https://identifiers.org/geo/GSE315113 (2025).
34. Hong, J. & Huang, J. Chromatin accessibility and transcriptome in human neuronal model exposed to Parkinson's environmental toxins. *Figshare* https://doi.org/10.6084/m9.figshare.29649077 (2025).

Acknowledgements

We thank the Personalbio Technologies Co. Ltd. and Gene Denovo Biotechnology Co. Ltd. for assisting in sequencing.

Author contributions

J.J.H. conceived and supervised the study. J.C.H. and J.J.H. performed the experiments. J.J.H. and J.C.H. performed data analyses. J.C.H. and J.J.H. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-026-06626-4.

Correspondence and requests for materials should be addressed to J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.