

# Application of serum SNCA gene in the diagnosis of Parkinson's disease

Yi Quan

Beijing Tiantan Hospital <https://orcid.org/0000-0002-4929-7664>

Jia Wang

Beijing Tiantan Hospital

Shuo Wang

Beijing Tiantan Hospital

Jizong Zhao (✉ [zhaojz1205@outlook.com](mailto:zhaojz1205@outlook.com))

Beijing Tiantan Hospital, Capital Medical University <https://orcid.org/0000-0002-5906-6149>

---

## Research

**Keywords:** Bioinformatics, SNCA, multi-omics analysis, Parkinson's disease

**Posted Date:** April 28th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-394588/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** SNCA plays an important role in the development of Parkinson's disease. However, the mechanism in PD is still unclear. We integrate multi-omics data for systematic analysis to explore the potential mechanism and therapeutic target of SNCA in PD.

**Materials and methods:** We download the gene expression, miRNA expression, and methylation data sets from GEO. After integration and standardization, analyze the diagnostic value and potential function of SNCA. Furthermore, according to the expression of SNCA, the important role of the genetic and epigenetic mechanism of SNCA co-expression in PD was analyzed by RNA-Seq, miRNA-Seq, and methylation data.

**Results:** Six expression profiling data sets showed that SNCA is lowly expressed in PD and could be used as a potential PD marker. SNCA co-expressed genes in serum and substantia nigra samples have significant differences. The lncRNAs which most relevant to SNCA expression are related to a variety of tumors and nerve-related diseases meanwhile the miRNA which most relevant to SNCA expression is related to multiple neurodegenerative disease pathways. There are multiple CpG methylation sites in the SNCA promoter region, of which cg15133208 is significantly hypomethylated in PD.

**Conclusion:** This study proposes low expression of SNCA as a diagnostic biomarker for PD, explores the unique genome and apparent landscape associated with expression of SNCA, provides ideas for a better understanding of the molecular basis of PD, and serves as a diagnostic biomarker for expression of SNCA or as a potential therapeutic target for PD patients which provides more direct evidence.

## 1 Background

Parkinson's disease (PD) is the second most common neurodegenerative disease after Alzheimer's disease. It has a 1% incidence [1] in the 65-year-old population and 4% in the 80-year-old population [2]. PD mainly affects the patient's motor system and its symptoms include tremor, stiffness, slow movement while early PD patients only show non-motor symptoms (NMS), such as constipation, olfactory dysfunction, depression, and sleep disorders [3]. Clinical motor symptoms include bradykinesia, stiffness, tremor and postural instability, and asymmetric seizures. Other motor dysfunctions include changes in gait and posture (presented by panic gait or chaotic walking with rapid bending), spelling or swallowing difficulties, and masked expressions [4]. PD is caused by a variety of factors, including age, genetics, oxidative stress, calcium overload, mitochondrial dysfunction, and environmental factors [5]. But the mechanism of occurrence in PD is still unclear. Furthermore, the detailed process of controlling the occurrence and development of PD is still unknown lead to remain a key obstacle on the path of PD treatment. The development of robust and accurate biomarkers will greatly facilitate the early detection and identification of PD biological characteristics or potential biomarkers.

With the development of high-throughput technologies, we have been able to benefit from large sample cohort studies. Recently, there have been many omics-based biomarker studies. For example, Liu et al. Used miRNA expression spectrum data to present a variety of important miRNA in PD as a migraine

serum biomarker [6]. Wang et al. used gene expression and methylation data to identify 53 genetic markers that can be used as PD serum markers [7], George et al. used a protein interaction network and successfully screened multiple PD-related drug targets [8] prove that using omics data to study the mechanism of the development of complex diseases has become an effective method.

The internal protein  $\alpha$ -synuclein is mainly encoded by the SNCA gene in the nucleus [9] which is composed of Lewy bodies, is related to the neuropathological features of PD [10] and plays an important role in neurodegenerative diseases. The overexpression of  $\alpha$ -synuclein has a close relationship with the pathogenesis of PD [11]. The abnormal accumulation of  $\alpha$ -synuclein and the formation of Lewy bodies may trigger the body's inflammatory response. It has recently been confirmed that overexpression is one of the reasons [12], and it may also be related to epigenetic regulation [13]. Therefore, it is necessary to analyze the pathogenesis of PD correlation with  $\alpha$ -synuclein expression based on multi-omics.

In this study, we propose a low expression of SNCA as a serum diagnostic biomarker for PD. We also explore the unique genomic and apparent landscape associated with the expression of SNCA. Compared with current basic experiments of SNCA, our work is to use SNCA expression provides more direct evidence as a diagnostic biomarker or as a potential therapeutic target for PD patients.

## 2 Materials And Methods

### 2.1 Data collection and processing

We downloaded the PD sample's data sets from the Gene Expression Omnibus (GEO) database then screened three sets of serum expression profile data sets (GSE6613[14, 15], GSE57475[16], GSE99039[17]) and three sets of substantia nigra expression profile datasets(GSE7621[18], GSE20141[19], GSE49036[20]), one set of brain tissue data that matching with gene expression and miRNA expression(GSE110716[21], GSE110717[21]), and two sets of serum sample's methylation data sets (GSE111629 [22, 23], GSE72774[24, 23]), each data statistics are showed as(Table 1). We download these standardized data sets on 2019-03-01. For the microarray expression profile data, the probe corresponds to the gene for setting that if the probe is used to correspond to multiple genes, then it will be removed, or multiple probes correspond to one gene then getting median. For methylated data, the  $\beta$  value is evaluated for the CpG quantitative of methylation. Remove the methylation sites with NA samples greater than 30 and remove the CpG sites with cross-reactive in the genome, according to the cross-reactive sites provided by Chen et al [25].

Table 1  
All sample statistics for each data set

Accession	Platform	PD	Control
GSE6613	Affymetrix Human Genome U133A Array	50	22
GSE57475	Illumina HumanHT-12 V3.0 expression beadchip	93	49
GSE99039	Affymetrix Human Genome U133 Plus 2.0 Array	205	233
GSE7621	Affymetrix Human Genome U133 Plus 2.0 Array	16	9
GSE20141	Affymetrix Human Genome U133 Plus 2.0 Array	10	8
GSE49036	Affymetrix Human Genome U133 Plus 2.0 Array	15	8
GSE110716	Illumina HiSeq 1000	8	8
GSE110717	Illumina HiSeq 1000	8	8
GSE111629	Illumina HumanMethylation450 BeadChip	335	237
GSE72774	Illumina HumanMethylation450 BeadChip	289	219

## 2.2 Co-expression network construction

Firstly, we separately calculate the MAD of each gene in the data set and select the top 50% of the MAD gene (TMADG). Then, we used the WGCNA [26] package in R to construct a scale-free co-expression network for the TMADGs. At first, the Pearson's correlation matrices and average linkage method were both performed for all pair-wise genes. Then, a weighted adjacency matrix was constructed using a power function  $A_{mn} = |C_{mn}|^\beta$  ( $C_{mn}$  = Pearson's correlation between gene m and gene n;  $A_{mn}$  = adjacency between gene m and gene n).  $\beta$  was a soft-thresholding parameter that could emphasize strong correlations between genes and penalize weak correlations. After choosing the power of  $\beta$ , the adjacency was transformed into a topological overlap matrix (TOM), which could measure the network connectivity of a Gene defined as the sum of its adjacency with all other Genes for network generation, and the corresponding dissimilarity (1-TOM) was calculated. To classify Genes with similar expression profiles into genes modules, average linkage hierarchical clustering was conducted according to the TOM-based dissimilarity measure with a minimum size (Gene group) of 30 for the Genes dendrogram. To further analyze the module, we calculated the dissimilarity of module eigen Genes, chose a cut line for module dendrogram and merged some modules. Finally, select the module where SNCA is located as the SNCA co-expression module.

## 2.3 Chip re-annotation

To get the latest lncRNAs expression profile, we re-annotate the lncRNAs of the HG-U133 2.0 array. Firstly, we download the long non-coding RNA transcript sequences of the GRCh38.p12 version from GENCODE. After that we use seqmap [27] to compare the probe clusters sequences to the lncRNA sequence, set the mismatch to 0, and select at least 11 clusters to be matched to the same lncRNA probe means that the re-

annotation was successful. Finally, the total 2448 Affymetrix probe sets were included in our following analysis.

## 2.4 miRNA target gene prediction

miRTarBase[28] is a curated database of MicroRNA-Target Interactions, which provides the largest available prediction and experimentally verified collection of miRNA-target interactions, with a variety of novel and unique features. We plan to use the miRTarBase online platform for predicting miRNA target genes meanwhile using default threshold filters target genes.

## 2.5 Functional enrichment analyses

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed using the R package clusterprofiler[29] for genes, to identify over-represented KEGG pathway. For this analysis, an FDR <0 .05 was considered to denote statistical significance.

ssGSEA[30] was performed by the GSVA[31] using the MSigDB[32] C2 Canonical pathways gene set collection, which contains 1320 gene sets. Gene sets with a p-value less than 0.05 after performing 1000 permutations were considered to be significantly enriched.

## 2.6 Statistical analysis

All analyses that do not specify parameters used the default parameters of the software. Significance was set at  $P < 0.05$ . All statistical analyses were performed in R 3.4.3. The correlation uses the Pearson correlation coefficient, and the network visualization uses Cytoscape[33] (<http://www.cytoscape.org/>). ROC analysis using R package plotROC and drawing heat map with R package pheatmap.

# 3 Results

## 3.1 SNCA can be used as a serum marker in PD.

The expression of SNCA is related to the progress of PD. For the relationship between the transcription level of SNCA in serum and PD, we extracted the expression level of SNCA in each dataset from serum and substantia nigra samples. In serum samples, the transcriptional levels of SNCA in PD were significantly lower than that of the healthy group( Fig. 1A: GSE6613, GSE57475, GSE99039). Similarly, the PD in substantia nigra samples is also significantly lower than that of healthy group samples such as (Fig. 1A: GSE7621, GSE20141, GSE49036). To predict the performance of transcription level on PD, we further analyzed the diagnostic curve where the average AUC = 0.727 in the substantia nigra sample and the average AUC = 0.642 in the whole blood sample as (Fig. 1B), which means that the low expression of SNCA is related to the occurrence of PD and can be used as a diagnostic marker in PD.

## 3.2 SNCA has different regulatory mechanisms in different types of samples.

In order to observe the relationship between SNCA expression and pathways, we used ssGSEA to analyze the enrichment scores of KEGG Pathway in each dataset and selected the top 20 KEGG pathways which most related. In serum samples, the SNCA gene was mainly enriched to GLYCOSPHINGOLIPID BIOSYNTHESIS LACTO AND NEOLACTO SERIES, NITROGEN METABOLISM, PORPHYRIN AND CHLOROPHYLL METABOLISM, CITRATE CYCLE TCACYCLE, AMINOACYL TRNA BIOSYNTHESIS and other energy metabolism-related pathways shown in (Fig. 2A). In nigra samples, the SNCA gene was mainly enriched in PARKINSONS DISEASE, PATHWAYS IN CANCER, PROSTATE CANCER, LEUKEMIA, and other complex diseases related pathways shown in (Fig. 2B). There are obvious differences in the pathways of the SNCA expression in the two types of samples (Fig. 2C). It is worth mentioning that the most relevant AMINOACYL TRNA BIOSYNTHESIS pathways in two types of samples showing the opposite results (Fig. 2D). These results suggest that SNCA may have different regulatory mechanisms in different types of samples.

### **3.3 Relationship between genome-wide gene expression and SNCA expression**

In order to further study the biological role of SNCA in the occurrence of PD, gene expression profiles closely related to SCNA were obtained based on whole-genome sequencing analysis. The R software package WGCNA was used to analyze the SCNA co-expression modules in different data sets, which were obtained from the data set GSE6613. A total of 19 modules were included, and the module related to SNCA expression was brown, as shown in (Fig. 3A), with a total of 485 co-expressed genes. A total of 11 modules were obtained from the dataset GSE99039, and the module related to SNCA expression was a red module, such as (Fig. 3B), which included 184 genes, a total of 9 modules were obtained from the data set GSE57475, of which the module related to SNCA expression is black as in (Fig. 3C), which contains a total of 139 genes, and a total of 17 modules were obtained from the data set GSE49036, of which SNCA expression related. The module is turquoise such as (Fig. 3D), which contains 4,916 genes, and a total of 22 modules are obtained from the data set GSE20141. Among them, the module related to SNCA expression is turquoise, such as (Fig. 3E), which contains 2745 genes, and a total of 28 genes from the data set GSE7621. Module, the module related to SNCA expression is turquoise as (Fig. 3F), which contains 4125 genes, we analyze the relationship between the genes in the SNCA co-expressed module in the data set, as expected, there is less intersection between genes in the serum sample data set, there is more intersection between the substantia nigra sample data set, and between the two types of sample data sets, the intersection is less like (Fig. 3G). Further, use the intersection gene of each data set in the two types of samples for KEGG functional enrichment, select  $FDR < 0.05$ , the serum sample contains a total of 61 genes, enriched into 1 KEGG Pathway, such as with Porphyrin and chlorophyll metabolism, the substantia nigra sample is enriched in 22 pathways, containing a total of 826 genes such as (WGCNA-Cross-gene). The most significant pathways are mainly PD, AD, and other pathways such as (WGCNA-Tp-KEGG). Among the six datasets, 6 genes are co-expressed with SNCA, such as GLRX5, HAGH, MPP1, MKRN1, XK, and BABAM1. These genes may be potential regulators in the occurrence and development of PD.

## 3.4 Relationship between SNCA and lncRNA expression

lncRNAs play an important role in the occurrence and development of complex diseases. We re-annotated the gene chip, re-annotated 2448 lncRNAs from the GPL570 chip such as (lnc\_probe), and used the Pearson correlation coefficient to calculate the relationship between SNCA and lncRNA as in (Fig. 4A), Select  $FDR < 0.01$  and  $|R| > 0.35$  to get the 15 most relevant lncRNAs, all of which are positively correlated. These lncRNAs are closely related to SNCA expression. (Fig. 4B), where ST70T2 is related to autism [34], MEG3 is a tumor suppressor[35], involved in the invasion and metastasis of various tumors[36–39], HILS1 is associated with the prognosis of gliomas[40], FLJ22536 is associated with the aggressiveness of neuroblastoma[41], in short, these lncRNAs are involved in many complex diseases which the development process occurs.

## 3.5 SNCA miRNA expression relationship

In order to analyze the relationship between SNCA expression and miRNA, we screened the GEO database for data sets that simultaneously detected the expression of miRNA and SNCA (GSE110716 and GSE110717), and extracted miRNA expression profiles and SNCA expression profiles, which contained a total of 2565 miRNAs such as (mir .exp), use the Pearson correlation coefficient to calculate the relationship between SNCA and miRNA as in (Fig. 5A), choose  $p < 0.01$  and  $|R| > 0.6$  to get a total of 40 miRNA, and their expression profile heat map with SNCA is as in (Fig. 5B), of which 31 are negative Correlation, 9 positive correlations such as (mir.exp.cor), the correlation heat map between these miRNAs is as in (Fig. 5C), miRNAs mainly inhibit mRNA transcription by targeting mRNA, and further select 31 negatively related miRNAs to use the mirTarBase database to predict the target genes of these miRNAs, select at least 70% of the target genes targeted by negatively related miRNAs to contain a total of 102. These genes are mainly enriched in the Calcium signaling pathway, Dopaminergic synapse is equal to nerve conduction related pathways as in (Fig. 5D). The results indicate that these miRNAs may participate in the pathogenesis of neurodegenerative diseases by regulating SNCA expression.

## 3.6 Relationship between genome-wide methylation and SNCA expression

Epigenetic disorders are essential for the development of various neurodegenerative diseases such as PD [42]. In order to observe the SNCA-related epigenetic landscape in PD samples, we first analyzed all CpG sites in the SNCA promoter region. In the two sets of serum methylation data sets, we identified that the site numbered cg15133208 was significantly hypomethylated in PD as in (Fig. 6A, B), which is consistent with previous reports. Hypomethylation may affect the expression of SNCA and thus participate in the occurrence of PD and the progress process, we calculated the CPG sites that were significantly related to cp15133208 and selected the sites with  $FDR < 0.01$  and  $|R| > 0.6$ . After taking the intersection of the two sets of methylation data sets, a total of 735 CpG sites were obtained. All of them are positively correlated such as (methyl.cor). The heat map of these most relevant methylation sites is as shown in (Fig. 6C, D). Their low methylation is related to PD. These differential methylation sites are mainly distributed in the open-sea region is as shown in (Fig. 6E). Gene expression is closely related to the methylation of the

promoter region [43]. The promoter region of 670 genes annotated by these CpG sites is (1st.cpgs.anno), these genes are enriched to Rap1 signaling pathway, transcriptional misregulation in cancer, and other important biological pathways such as (Fig. 6F).

## 4 Discussion

The pathogenesis of PD is unclear, but previous studies have reported many abnormal molecular characteristics, such as an abnormal expression or altered epigenetic modification of PARK1-15LRRK2, MAPT, and GBA to be related to PD pathology [44, 45]. These characteristics can be used as diagnostic factors and promote the understanding of the pathogenesis of PD. However, these molecules are only suitable for some specific conditions and cannot meet the needs of clinical applications, which means that identifying robust molecular characteristics is still a challenge and requires more research to clarify the mechanism of the occurrence and development of PD. SNCA-encoded  $\alpha$ -synuclein is the main component of Lewy bodies. Over-expressed  $\alpha$ -synuclein relates to the pathogenesis of PD [11]. The abnormal accumulation of synuclein and the formation of Lewy bodies may trigger the body's inflammatory response. It may be subject to multiple omics molecular regulation in PD. We systematically analyzed the diagnostic value of SNCA in PD. SNCA is low in PD in serum and substantia nigra sample types Up, which can serve as an effective marker serum.

Genetics plays an important role in the etiology of many complex human diseases. Genes are the functional unit of genetic material. It is believed whether the expression of genes will affect the synthesis of downstream proteins, which are a basic part of the human body. However, recent studies have shown that individual genes cannot function alone. Instead, genes interact and affect human health together. Therefore, we identify genes that are co-expressed with SNCA based on gene expression profiles. First, we use WGCNA to separately calculate SNCA co-expression in different data sets. Module, screening co-expression module, analysis of the intersection of genes in the SNCA co-expression module in different data sets for mutual verification, of which the co-expressed genes in the three data sets of the substantia nigra sample have high reproducibility. The three datasets have poor reproducibility, which may be related to blood samples being easily affected by the environment. The genes co-expressed in the substantia nigra samples are mainly enriched in neurodegenerative disease pathways such as Parkinson disease, Huntington disease, Alzheimer disease, etc., Which is consistent with the expected result, and finally, we sift In the six data sets, six genes co-expressed with SNCA, such as GLRX5, HAGH, MPP1, MKRN1, XK, and BABAM1, are reported. Among them, GLRX5 is associated with tumor [46], anemia [47], schizophrenia [48]diseases, HAGH is associated with schizophrenia[49], autism[50], and other neurological diseases, MPP1 is associated with tumor drug resistance [51, 52], MKRN1 is involved in the development of gastric cancer [53], these genes are in a variety of complex. They play an important role in the occurrence and development of diseases. In PD, their co-expression with SNCA may be a new candidate marker for PD.

Recent studies have shown that genetics and epigenetic mechanisms play an important role in tumorigenesis. The former mainly focuses on abnormal gene expression, and the latter mainly include post-transcriptional regulation of microRNA and DNA methylation. Therefore, we further the possible

mechanisms by which SNCA highly influences the results of these three aspects were studied. First, the SNCA table was analyzed using miRNA expression profiling for related miRNAs, the 40 most relevant miRNAs were identified, including 31 negatively related miRNAs. Through functional enrichment, these miRNAs were found to be mainly enriched in the Calcium signaling pathway, Dopaminergic synapse is equal to nerve conduction related pathways, which means The abnormal expression of these miRNAs may play an important role in the occurrence of PD. Eryilmaz et al. found that the SNCA promoter region is hypomethylated in early-onset Parkinson's disease[54]. Based on the DNA methylation profile, we identified SNCA. There are significant hypomethylation sites in PD in the promoter region, and 735 CpG sites that are significantly related to this site were further identified. Cluster analysis revealed that the methylation of these sites was highly consistent. Sexuality is generally hypomethylated in PD. These CpG sites are mainly distributed in the open sea region. A total of 670 genes was identified by functional annotation of these CpG sites, and these genes were enriched in many important biological pathways. This suggests that these abnormal methylation sites may play an important role in the epigenetic regulation of PD.

Although we have comprehensively analyzed the role of SNCA expression in the development of PD through bioinformatics technology from multiple omics levels, these results can help us better understand the occurrence and development of PD. However, our research also has some Limitations. First, the source of our sample cohort is not uniform. Second, this work is a study in the field of bioinformatics, and its purpose is often to verify the accuracy and reliability of the analysis results or biomarkers through statistical significance and scientific literature verification of the diagnostic value of SNCA. In future research, if conditions permit, we will increase the experimental design and conduct more detailed mechanism research to further verify our conclusions.

In summary, in this study, we used a large sample research cohort to provide direct evidence that the low expression of SNCA is related to the occurrence of PD. The data analysis based on multi-omics clarifies the potential mechanism of SNCA in the occurrence and development of PD.

## 5 Conclusions

This study proposes low expression of SNCA as a diagnostic biomarker for PD, explores the unique genome and apparent landscape associated with expression of SNCA, provides ideas for a better understanding of the molecular basis of PD, and serves as a diagnostic biomarker for expression of SNCA or as a potential therapeutic target for PD patients which provides more direct evidence.

## Declarations

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

## Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

## Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Author Contributions

Yi Quan had full access to all of the study data and takes responsibility for the integrity and accuracy of the data analysis; Study concept and design: Yi Quan, Jia Wang; Critical revision of the manuscript for important intellectual content: All authors; Statistical analysis: Yi Quan, Jia Wang; Administrative, technical, and, study supervision: Shuo Wang, Jizong Zhao

## Funding

Funding This study is supported by the “13th Five-Year Plan” National Science and Technology Supporting Plan (2015BAI12B04), and the Beijing Municipal Administration of Hospitals’ Mission Plan (SML20150501).

## Acknowledgments

We thank the Central laboratory of China National Clinical Research Center for Neurological Diseases for technical support.

## References

1. Kalia LV, Lang AE. Parkinson's disease. *Lancet*. 2015;386(9996):896–912. doi:10.1016/S0140-6736(14)61393-3.
2. Wirdefeldt K, Adami HO, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. *Eur J Epidemiol*. 2011;26(Suppl 1):1–58. doi:10.1007/s10654-011-9581-6.
3. Dexter DT, Jenner P. Parkinson disease: from pathology to molecular disease mechanisms. *Free Radic Biol Med*. 2013;62:132–44. doi:10.1016/j.freeradbiomed.2013.01.018.
4. Jankovic J. Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry*. 2008;79(4):368–76. doi:10.1136/jnnp.2007.131045.
5. Warner TT, Schapira AH. (2003) Genetic and environmental factors in the cause of Parkinson's disease. *Ann Neurol* 53 Suppl 3:16–23; discussion S23-15. doi:10.1002/ana.10487.

6. Liu X, Chen J, Guan T, Yao H, Zhang W, Guan Z, Wang Y. miRNAs and target genes in the blood as biomarkers for the early diagnosis of Parkinson's disease. *BMC Syst Biol.* 2019;13(1):10. doi:10.1186/s12918-019-0680-4.
7. Wang C, Chen L, Yang Y, Zhang M, Wong G. Identification of potential blood biomarkers for Parkinson's disease by gene expression and DNA methylation data integration analysis. *Clin Epigenetics.* 2019;11(1):24. doi:10.1186/s13148-019-0621-5.
8. George G, Valiya Parambath S, Lokappa SB, Varkey J. Construction of Parkinson's disease marker-based weighted protein-protein interaction network for prioritization of co-expressed genes. *Gene.* 2019;697:67–77. doi:10.1016/j.gene.2019.02.026.
9. Goers J, Manning-Bog AB, McCormack AL, Millett IS, Doniach S, Di Monte DA, Uversky VN, Fink AL. Nuclear localization of alpha-synuclein and its interaction with histones. *Biochemistry.* 2003;42(28):8465–71. doi:10.1021/bi0341152.
10. Siddiqui IJ, Pervaiz N, Abbasi AA. The Parkinson Disease gene SNCA: Evolutionary and structural insights with pathological implication. *Sci Rep.* 2016;6:24475. doi:10.1038/srep24475.
11. Takahashi M, Suzuki M, Fukuoka M, Fujikake N, Watanabe S, Murata M, Wada K, Nagai Y, Hohjoh H. Normalization of Overexpressed alpha-Synuclein Causing Parkinson's Disease By a Moderate Gene Silencing With RNA Interference. *Mol Ther Nucleic Acids.* 2015;4:e241. doi:10.1038/mtna.2015.14.
12. Allen Reish HE, Standaert DG. Role of alpha-synuclein in inducing innate and adaptive immunity in Parkinson disease. *J Parkinsons Dis.* 2015;5(1):1–19. doi:10.3233/JPD-140491.
13. Schmitt I, Kaut O, Khazneh H, deBoni L, Ahmad A, Berg D, Klein C, Frohlich H, Wullner U. L-dopa increases alpha-synuclein DNA methylation in Parkinson's disease patients in vivo and in vitro. *Mov Disord.* 2015;30(13):1794–801. doi:10.1002/mds.26319.
14. Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D, Schwarzschild MA, Schlossmacher MG, Hauser MA, Vance JM, Sudarsky LR, Standaert DG, Growdon JH, Jensen RV, Gullans SR. Molecular markers of early Parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci U S A.* 2007;104(3):955–60. doi:10.1073/pnas.0610204104.
15. Scherzer CR, Grass JA, Liao Z, Pepivani I, Zheng B, Eklund AC, Ney PA, Ng J, McGoldrick M, Mollenhauer B, Bresnick EH, Schlossmacher MG. GATA transcription factors directly regulate the Parkinson's disease-linked gene alpha-synuclein. *Proc Natl Acad Sci U S A.* 2008;105(31):10907–12. doi:10.1073/pnas.0802437105.
16. Locascio JJ, Eberly S, Liao Z, Liu G, Hoelsing AN, Duong K, Trisini-Lipsanopoulos A, Dhima K, Hung AY, Flaherty AW, Schwarzschild MA, Hayes MT, Wills AM, Shivraj Sohur U, Mejia NI, Selkoe DJ, Oakes D, Shoulson I, Dong X, Marek K, Zheng B, Ivinson A, Hyman BT, Growdon JH, Sudarsky LR, Schlossmacher MG, Ravina B, Scherzer CR. Association between alpha-synuclein blood transcripts and early, neuroimaging-supported Parkinson's disease. *Brain.* 2015;138(Pt 9):2659–71. doi:10.1093/brain/awv202.
17. Shamir R, Klein C, Amar D, Vollstedt EJ, Bonin M, Usenovic M, Wong YC, Maver A, Poths S, Safer H, Corvol JC, Lesage S, Lavi O, Deuschl G, Kuhlenbaeumer G, Pawlack H, Ulitsky I, Kasten M, Riess O,

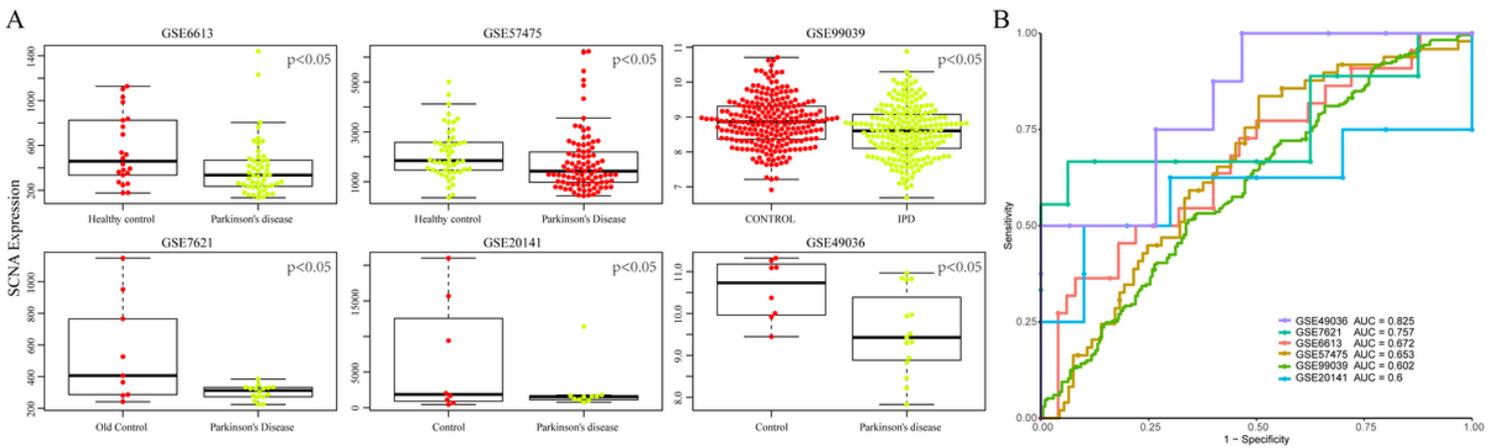
- Brice A, Peterlin B, Krainc D. Analysis of blood-based gene expression in idiopathic Parkinson disease. *Neurology*. 2017;89(16):1676–83. doi:10.1212/WNL.0000000000004516.
18. Lesnick TG, Papapetropoulos S, Mash DC, Ffrench-Mullen J, Shehadeh L, de Andrade M, Henley JR, Rocca WA, Ahlskog JE, Maraganore DM. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet*. 2007;3(6):e98. doi:10.1371/journal.pgen.0030098.
19. Zheng B, Liao Z, Locascio JJ, Lesniak KA, Roderick SS, Watt ML, Eklund AC, Zhang-James Y, Kim PD, Hauser MA, Grunblatt E, Moran LB, Mandel SA, Riederer P, Miller RM, Federoff HJ, Wullner U, Papapetropoulos S, Youdim MB, Cantuti-Castelvetri I, Young AB, Vance JM, Davis RL, Hedreen JC, Adler CH, Beach TG, Graeber MB, Middleton FA, Rochet JC, Scherzer CR, Global PDGEC. PGC-1alpha, a potential therapeutic target for early intervention in Parkinson's disease. *Sci Transl Med*. 2010;2(52):52ra73. doi:10.1126/scitranslmed.3001059.
20. Dijkstra AA, Ingrassia A, de Menezes RX, van Kesteren RE, Rozemuller AJ, Heutink P, van de Berg WD. Evidence for Immune Response, Axonal Dysfunction and Reduced Endocytosis in the Substantia Nigra in Early Stage Parkinson's Disease. *PLoS One*. 2015;10(6):e0128651. doi:10.1371/journal.pone.0128651.
21. Schulze M, Sommer A, Plotz S, Farrell M, Winner B, Grosch J, Winkler J, Riemenschneider MJ. Sporadic Parkinson's disease derived neuronal cells show disease-specific mRNA and small RNA signatures with abundant deregulation of piRNAs. *Acta Neuropathol Commun*. 2018;6(1):58. doi:10.1186/s40478-018-0561-x.
22. Chuang YH, Paul KC, Bronstein JM, Bordelon Y, Horvath S, Ritz B. Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Med*. 2017;9(1):76. doi:10.1186/s13073-017-0466-5.
23. Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging*. 2015;7(12):1130–42. doi:10.18632/aging.100859.
24. Czarnecki CM, Olivero DK, McVey AS, Pessin MF. Myocardial levels of calcium, glycogen and triglycerides in ethanol-fed turkey poults treated with allopurinol. *Comp Biochem Physiol C*. 1987;87(1):79–82.
25. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203–9. doi:10.4161/epi.23470.
26. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. doi:10.1186/1471-2105-9-559.
27. Hawkins TB, Dantzer J, Peters B, Dinauer M, Mockaitis K, Mooney S, Cornetta K. Identifying viral integration sites using SeqMap 2.0. *Bioinformatics*. 2011;27(5):720–2. doi:10.1093/bioinformatics/btq722.
28. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, Huang WC, Sun TH, Tu SJ, Lee WH, Chiew MY, Tai CS, Wei TY, Tsai TR, Huang HT, Wang CY, Wu HY, Ho SY, Chen PR, Chuang CH, Hsieh PJ, Wu YS, Chen WL, Li MJ, Wu YC, Huang XY, Ng FL, Buddhakosai W, Huang PC, Lan KC, Huang CY,

- Weng SL, Cheng YN, Liang C, Hsu WL, Huang HD. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2018;46(D1):D296–302. doi:10.1093/nar/gkx1067.
29. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–7. doi:10.1089/omi.2011.0118.
  30. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics.* 2007;23(23):3251–3. doi:10.1093/bioinformatics/btm369.
  31. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7. doi:10.1186/1471-2105-14-7.
  32. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* 2011;27(12):1739–40. doi:10.1093/bioinformatics/btr260.
  33. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498–504. doi:10.1101/gr.1239303.
  34. Vincent JB, Petek E, Thevarkunnel S, Kolozsvari D, Cheung J, Patel M, Scherer SW. The RAY1/ST7 tumor-suppressor locus on chromosome 7q31 represents a complex multi-transcript system. *Genomics.* 2002;80(3):283–94.
  35. Zhou Y, Zhang X, Klibanski A. MEG3 noncoding RNA: a tumor suppressor. *J Mol Endocrinol.* 2012;48(3):R45–53. doi:10.1530/JME-12-0008.
  36. Lyu Y, Lou J, Yang Y, Feng J, Hao Y, Huang S, Yin L, Xu J, Huang D, Ma B, Zou D, Wang Y, Zhang Y, Zhang B, Chen P, Yu K, Lam EW, Wang X, Liu Q, Yan J, Jin B. Dysfunction of the WT1-MEG3 signaling promotes AML leukemogenesis via p53-dependent and -independent pathways. *Leukemia.* 2017;31(12):2543–51. doi:10.1038/leu.2017.116.
  37. Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, Mitra S, Mohammed A, James AR, Hoberg E, Moustakas A, Gyllensten U, Jones SJ, Gustafsson CM, Sims AH, Westerlund F, Gorab E, Kanduri C. MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. *Nat Commun.* 2015;6:7743. doi:10.1038/ncomms8743.
  38. Wang P, Chen D, Ma H, Li Y. LncRNA MEG3 enhances cisplatin sensitivity in non-small cell lung cancer by regulating miR-21-5p/SOX7 axis. *Onco Targets Ther.* 2017;10:5137–49. doi:10.2147/OTT.S146423.
  39. Wei GH, Wang X. lncRNA MEG3 inhibit proliferation and metastasis of gastric cancer via p53 signaling pathway. *Eur Rev Med Pharmacol Sci.* 2017;21(17):3850–6.
  40. Liu B, Liu J, Liu K, Huang H, Li Y, Hu X, Wang K, Cao H, Cheng Q. A prognostic signature of five pseudogenes for predicting lower-grade gliomas. *Biomed Pharmacother.* 2019;117:109116. doi:10.1016/j.biopha.2019.109116.
  41. Maris JM, Mosse YP, Bradfield JP, Hou C, Monni S, Scott RH, Asgharzadeh S, Attiyeh EF, Diskin SJ, Laudenslager M, Winter C, Cole KA, Glessner JT, Kim C, Frackelton EC, Casalunovo T, Eckert AW,

- Capasso M, Rappaport EF, McConville C, London WB, Seeger RC, Rahman N, Devoto M, Grant SF, Li H, Hakonarson H. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *N Engl J Med*. 2008;358(24):2585–93. doi:10.1056/NEJMoa0708698.
42. Ammal Kaidery N, Tarannum S, Thomas B. Epigenetic landscape of Parkinson's disease: emerging role in disease mechanisms and therapeutic modalities. *Neurotherapeutics*. 2013;10(4):698–708. doi:10.1007/s13311-013-0211-8.
43. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484–92. doi:10.1038/nrg3230.
44. Coppede F. (2012) Genetics and epigenetics of Parkinson's disease. *ScientificWorldJournal* 2012:489830. doi:10.1100/2012/489830.
45. Miranda-Morales E, Meier K, Sandoval-Carrillo A, Salas-Pacheco J, Vazquez-Cardenas P, Arias-Carrion O. Implications of DNA Methylation in Parkinson's Disease. *Front Mol Neurosci*. 2017;10:225. doi:10.3389/fnmol.2017.00225.
46. Ishikawa R, Amano Y, Kawakami M, Sunohara M, Watanabe K, Kage H, Ohishi N, Yatomi Y, Nakajima J, Fukayama M, Nagase T, Takai D. The chimeric transcript RUNX1-GLRX5: a biomarker for good postoperative prognosis in Stage IA non-small-cell lung cancer. *Jpn J Clin Oncol*. 2016;46(2):185–9. doi:10.1093/jjco/hyv187.
47. Liu G, Guo S, Anderson GJ, Camaschella C, Han B, Nie G. Heterozygous missense mutations in the GLRX5 gene cause sideroblastic anemia in a Chinese patient. *Blood*. 2014;124(17):2750–1. doi:10.1182/blood-2014-08-598508.
48. Yang B, Niu W, Chen S, Xu F, Li X, Wu X, Cao Y, Zhang R, Yang F, Wang L, Li W, Xu Y, He L, He G. Association study of the GLRX5 rs1007814 polymorphism with schizophrenia in the Han Chinese population. *Psychiatr Genet*. 2017;27(2):76–7. doi:10.1097/YPG.000000000000160.
49. Bangel FN, Yamada K, Arai M, Iwayama Y, Balan S, Toyota T, Iwata Y, Suzuki K, Kikuchi M, Hashimoto T, Kanahara N, Mori N, Itokawa M, Stork O, Yoshikawa T. Genetic analysis of the glyoxalase system in schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry*. 2015;59:105–10. doi:10.1016/j.pnpbp.2015.01.014.
50. Kovac J, Podkrajsek KT, Luksic MM, Battelino T. Weak association of glyoxalase 1 (GLO1) variants with autism spectrum disorder. *Eur Child Adolesc Psychiatry*. 2015;24(1):75–82. doi:10.1007/s00787-014-0537-8.
51. Pitre A, Ge Y, Lin W, Wang Y, Fukuda Y, Temirov J, Phillips AH, Peters JL, Fan Y, Ma J, Nourse A, Sinha C, Lin H, Kriwacki R, Downing JR, Gruber TA, Centonze VE, Naren AP, Chen T, Schuetz JD. An unexpected protein interaction promotes drug resistance in leukemia. *Nat Commun*. 2017;8(1):1547. doi:10.1038/s41467-017-01678-y.
52. Wittig R, Nessling M, Will RD, Mollenhauer J, Salowsky R, Munstermann E, Schick M, Helmbach H, Gschwendt B, Korn B, Kioschis P, Lichter P, Schadendorf D, Poustka A. Candidate genes for cross-resistance against DNA-damaging drugs. *Cancer Res*. 2002;62(22):6698–705.

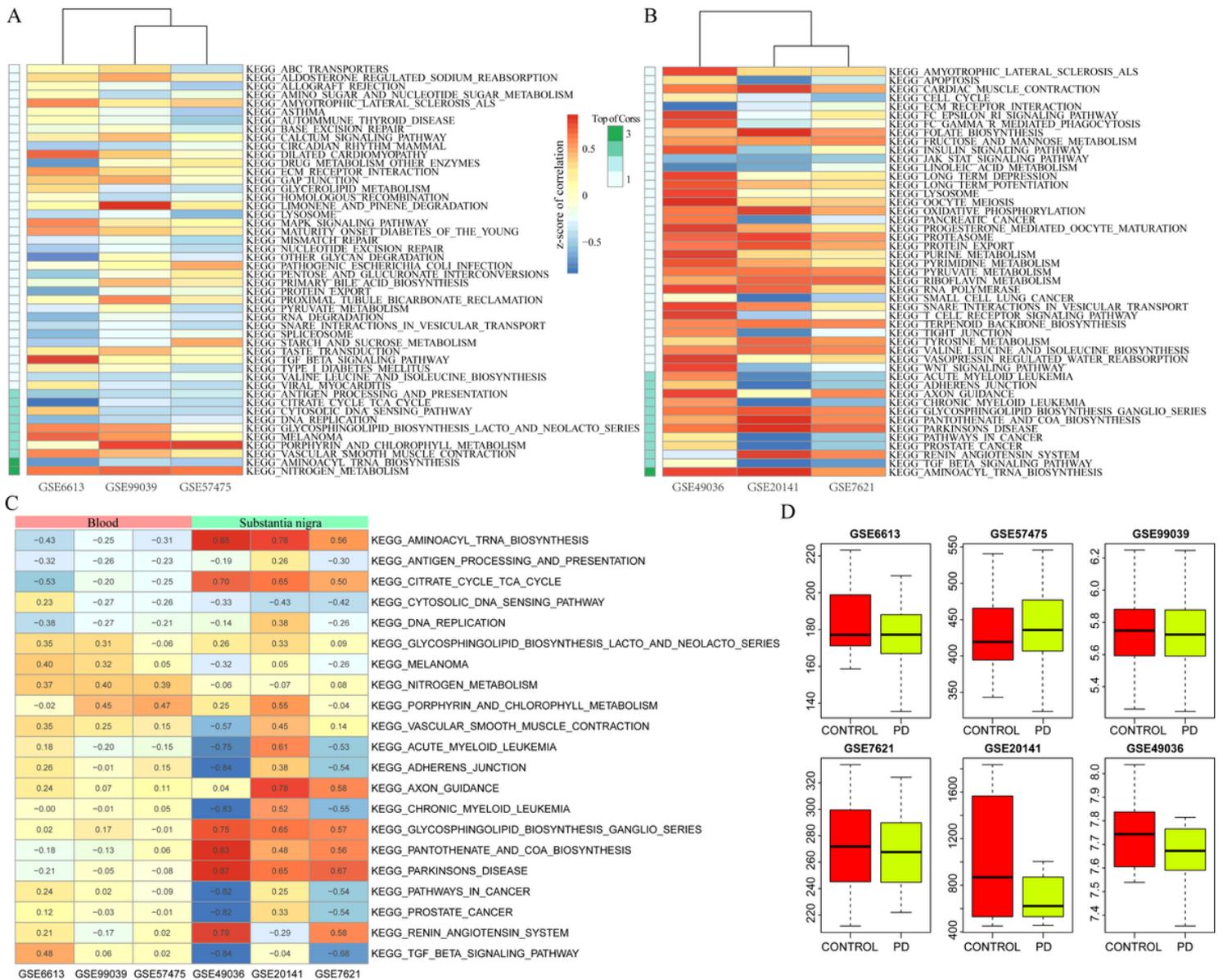
53. Ko A, Shin JY, Seo J, Lee KD, Lee EW, Lee MS, Lee HW, Choi IJ, Jeong JS, Chun KH, Song J. Acceleration of gastric tumorigenesis through MKRN1-mediated posttranslational regulation of p14ARF. *J Natl Cancer Inst.* 2012;104(21):1660–72. doi:10.1093/jnci/djs424.
54. Eryilmaz IE, Cecener G, Erer S, Egeli U, Tunca B, Zarifoglu M, Elibol B, Bora Tokcaer A, Saka E, Demirkiran M, Akbostanci C, Dogu O, Colakoglu B, Kenangil G, Kaleagasi H. Epigenetic approach to early-onset Parkinson's disease: low methylation status of SNCA and PARK2 promoter regions. *Neurol Res.* 2017;39(11):965–72. doi:10.1080/01616412.2017.1368141.

## Figures



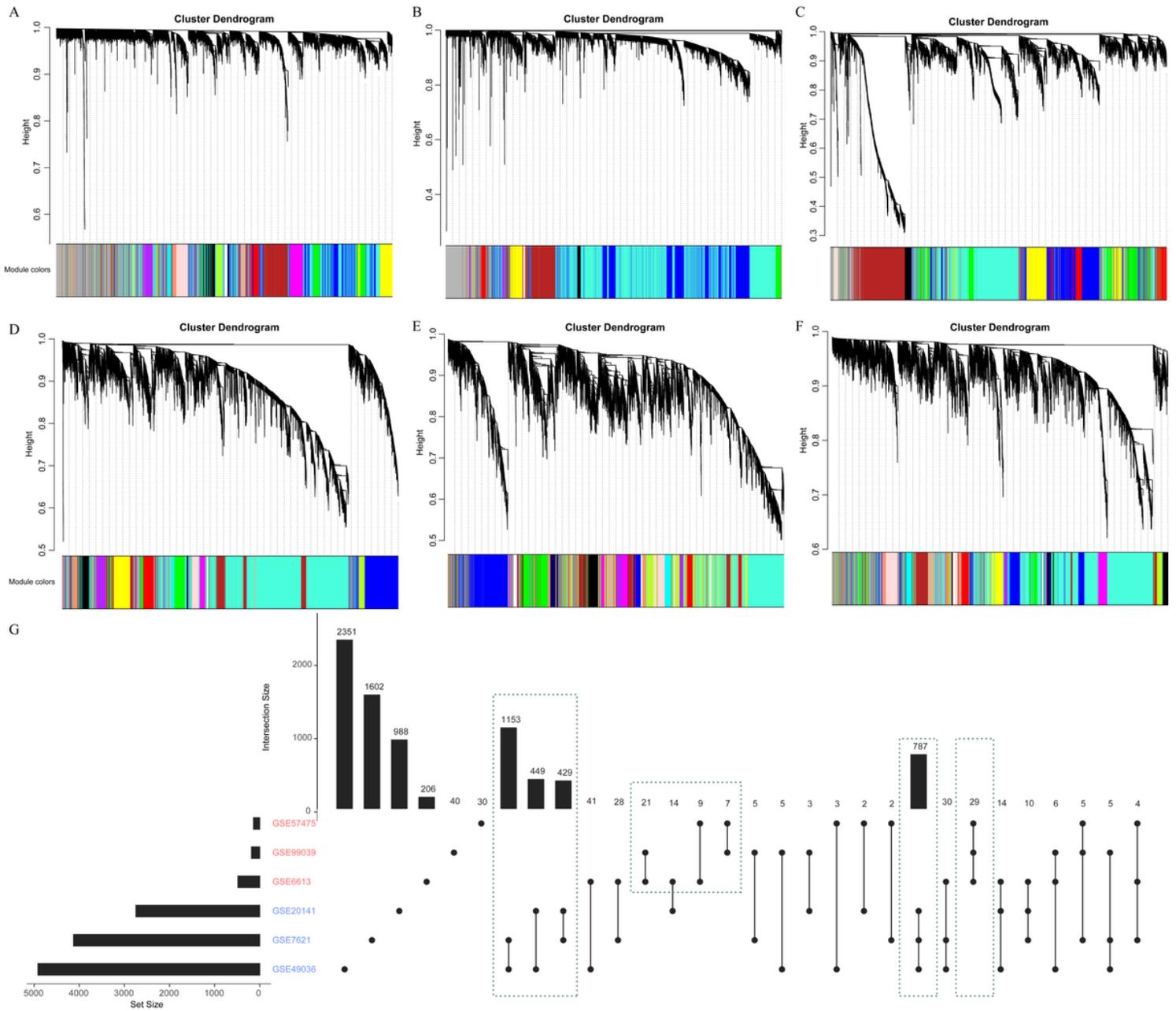
**Figure 1**

A: The expression difference of SNCA in PD and control group in 6 data sets; B: ROC curve of SNCA expression in 6 data sets to predict PD



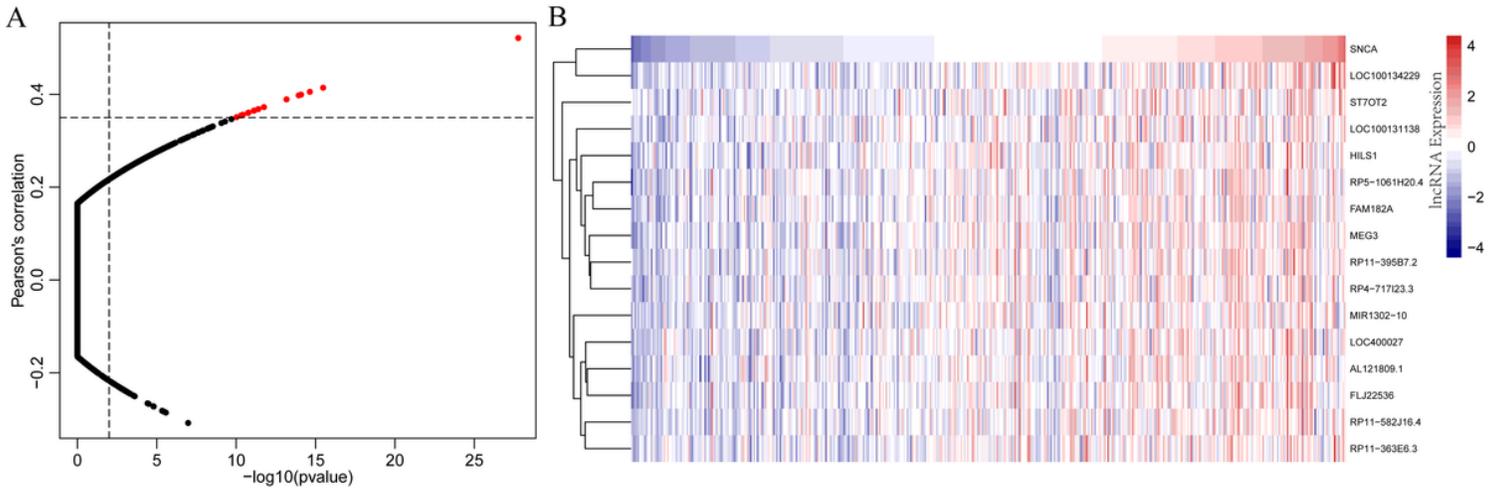
**Figure 2**

A: Pathway correlation heat map of the top20 most relevant to SNCA expression in the three serum sample data sets; B: Pathway correlation heat map of the top20 most relevant to SNCA expression in the three substantia nigra sample data sets; C: Serum sample and substantia nigra Correlation heat maps of the top20 paths in the samples in at least two data sets in each data set. D: Box plot of the expression distribution of AMINOACYL TRNA BIOSYNTHESIS pathway genes in Control and PD samples in each data set.



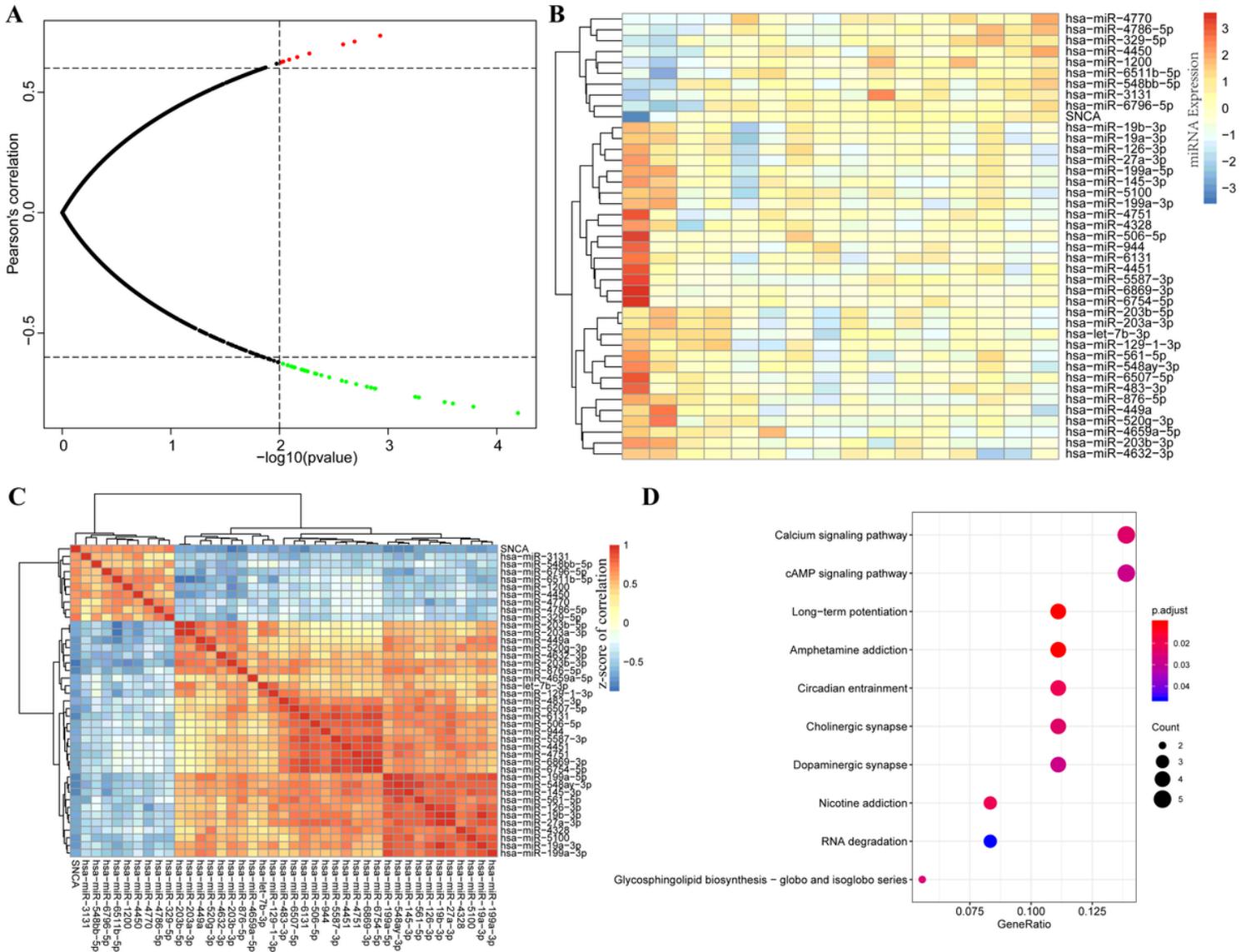
**Figure 3**

A-F The six data sets use the co-expression modules identified by WGCNA; G: the intersection relationship between the genes of the SNCA co-expressed modules in the six data sets.



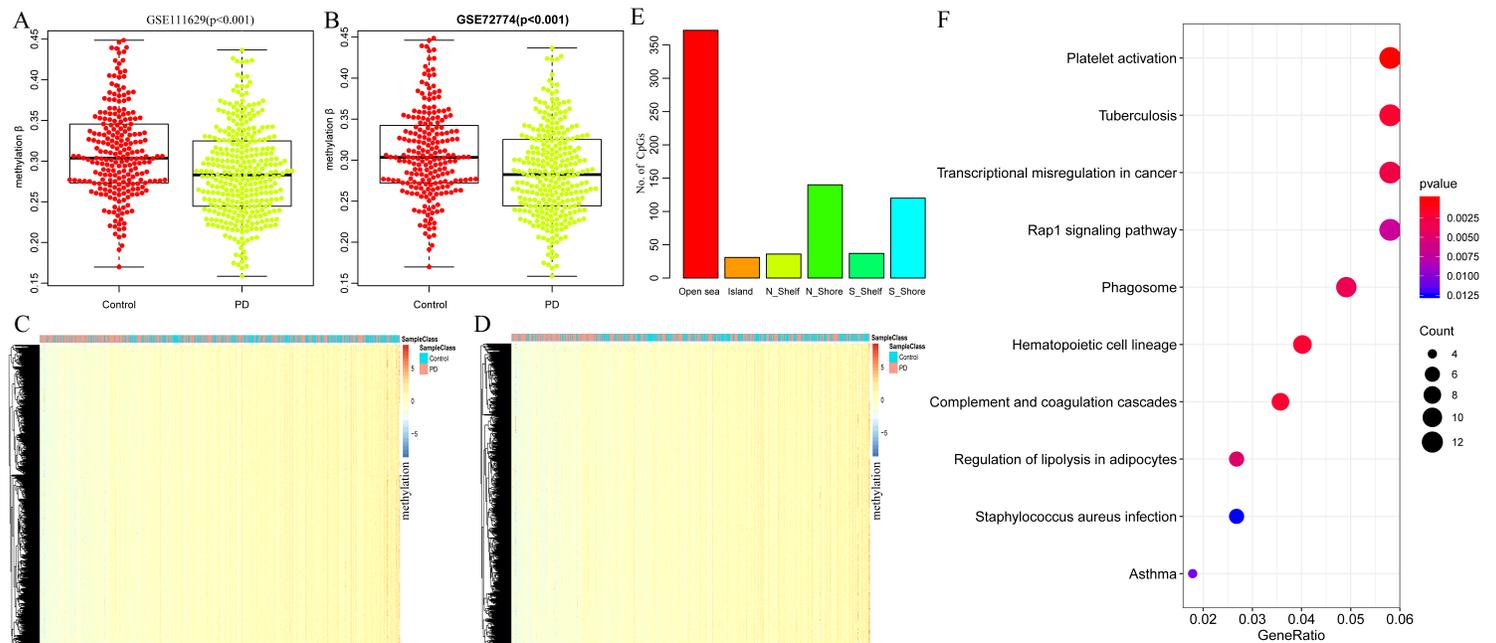
**Figure 4**

A: IncRNA and SNCA expression correlation distribution, x-axis is FDR, y-axis is correlation. B: Heat map of the expression profiles of significantly related IncRNA and SNCA.



## Figure 5

A: The distribution of miRNA and SNCA expression correlation, x-axis is p value, y-axis is correlation. B: Heat map of the expression profile of significantly correlated miRNA and SNCA; C: Heat map of the correlation coefficient between significantly correlated miRNA and SNCA; D: KEGG pathway enriched by negatively correlated miRNA target genes, the color indicates significance, the size of the dot indicates the number of genes.



## Figure 6

A-B: The difference distribution of cg15133208 sites in the two methylation data sets; C-D: the methylation spectra of the most relevant CpG sites in the two data sets in each sample. E: the distribution of the most relevant CpG loci of cg15133208 on the genome; F: the KEGG enrichment results of the genes annotated to the most relevant CpG loci of cg15133208, the dot size in the figure indicates the number of genes, and the color indicates the p value.

## Supplementary Files

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

- [Additionalfile1.xlsx](#)