Original article

Screening and bioinformatics analysis of differentially expressed genes reveals DAPL1 as a critical regulatory gene in the substantia nigra of patients with Parkinson's disease based on GEO database

Xi Cheng*, Qingyan Su, Jinghui Lai

[1] *Rehabilitation Hospital affiliated to Fujian University of Traditional Chinese Medicine, No. 13, Hudong Branch Road, Gulou District, Fuzhou City, Fujian Province, China*

Received: 16/08/2022 – **Accepted:** 10/10/2023

DOI: 10.15761/0101-60830000000718

ABSTRACT

Objective: To investigate the screening of differential genes in the substantia nigra of Patients with Parkinson's disease and their related signaling pathways and biological processes, to provide clues for subsequent experimental verification studies. **Methods**: (1) In the GEO (Gene Expression Omnibus) database, the data set of substantia nigra and Parkinson's disease were retrieved, and the R language was used to analyze the chip data. The differential genes between different chip datasets were intersected to obtain the common difference genes. GO and KEGG enrichment analysis were applied by Cytoscape. The String database and GeneMANIA online database was used for protein-protein interaction (PPI) network analysis. The LOSSS database were used to analyze the most promising hub genes. Treefam and Aminode was used to analyze their gene family conservatism. The human protein atlas database analyzed their expression distribution in normal brain tissue. **Results** In this study, 45 co-differentiated genes were obtained from the two datasets. GO and KEGG analysis revealed that these genes were enriched in the regulation of synapse organizations, and Serotonergic synapse, et al. PPI showed 20 nodes and 22 protein interactions. Finally, 2 co-differential genes, DAPL1 and SLC2A13, were obtained through intersection with the DISEASES database. Treefam and Aminode analysis further showed that DAPL1 was genetically and proteinally conserved across species and was primarily expressed in the normal hypothalamus, midbrain, and cortex. DAPL1 was also significantly low expression in the validation set GSE49036 in the substantia substantia of patients with Parkinson's disease, and the PPI network shows that its role may be closely related to the binding of the death domain. **Conclusion** Based on geo database, DAPL1, a key gene involved in the occurrence and development of Parkinson's disease, was screened by bioinformatics, and data support was needed for further research.

Cheng X et al / Arch Clin Psychiatry. 2023;50(6):127-137.

Keywords: Parkinson's disease, substantia nigra, DAPL1, bioinformatics, GEO

Introduction

Parkinson's disease (PD) is a common neurodegenerative disease in middle-aged and elderly people, with an incidence of up to 1% in the elderly over 65 years of age, second only to Alzheimer's disease, which has the highest incidence[1]. Parkinson's disease is characterized by symptom groups of progressive slow movements, muscle rigidity, resting tremor, and postural disorders, resulting from the loss of dopaminergic neurons in the substantia nigra[2]. The latest research results showed that the substantia nigra was the main site of the presence of catecholamine neurons, while neuronal cells containing neuromelanin were more susceptible to Parkinson's disease-like pathological changes[3].

Therefore, in this study, we used the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) of the National Center for Biotechnology Information to discover abnormally expressed genes and associated biological processes in the substantia nigra of patients with Parkinson's disease[4]. In addition, the emergence of the key regulatory gene DAPL1 was evaluated, providing new ideas for the treatment of patients with Parkinson's disease.

Materials and methods

GEO datasets

In the GEO (https://www.ncbi.nlm.nih.gov/ geo/) database, the relevant chips were retrieved by keywords such as "substantia nigra" and "Parkinson's disease", and the dataset containing the gene expression information of substantia nigra and control group in Parkinson's disease patients was screened, and two sets of datasets were extracted from them: GSE7621 and GSE42966[5, 6].

The GSE7621 dataset used the Human Genome U133 Plus 2.0 Array gene chip provided by Affymetrix to detect RNA expression in 16 postmortem Parkinson's patients with Parkinson's disease and in nine normal human substantia nigra groups; the GSE42966 dataset used Agilent-014850 Whole Human Genome Microarray 4x44K from Agilent G4112F Microarray gene

Address for correspondence: Xi Cheng, Rehabilitation Hospital affiliated to Fujian University of Traditional Chinese Medicine, No. 13, Hudong Branch Road, Gulou District, Fuzhou City, Fujian Province, China

E-mail: chengxiemail1974@163.com

chip, RNA expression detection of postmortem substantia nigra tissue and 6 normal human substantia group of 9 patients with Parkinson's disease.The GSE49036 dataset used the Human Genome U133 Plus 2.0 Array gene chip provided by Affymetrix to detect RNA expression in 20 post-mortem substantia and 8 normal human substantia nigra groups with Parkinson's disease[7]. The dataset would be used to verify the expression level of the hub gene.

List of Parkinson's disease genes

A list of human Parkinson's disease genes is available from the DISEASES portal (https://maayanlab.cloud/Harmonizome/resource/DISEASES), which contains a total of 98 genes[8].

Screening for differential co-expression genes

Differentially expressed genes (DEGs) were screened between the GSE7621 and GSE42966 experimental and control groups using the GEO2R online analysis tool (nlm.nih.gov/geo/geo2r/ http://www.ncbi. P) < 0.05, the logarithmic absolute value of the fold change (FC) | logFC|>1. Co-expressing genes for the DEGs of the two datasets plotted online to obtain the intersection of the Venny diagram (https://bioinformatics.psb.ugent.be/webtools/Venn/).

Volcano mapping and genetic heat mapping

The series matrix table of GSE7621 and GSE42966 were downloaded from the GEO database, and then based on the ggplot2 [version 3.3.3] of R software (version 4.0.2) and the Complex Heatmap package [version 2.2.0], the volcano map and gene expression heat map were drawn[9, 10].

GO/KEGG pathway enrichment analysis of co-expression DEGs

To identify co-expression of DEGs-related pathways and functional annotations, the R (version 3.6.3)-based cluster Profiler package [version 3.14.3] was employed for enrichment analysis of gene ontologies (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG)[11].

PPI network construction

The data of DEGs interactions between GSE7621 and GSE42966were obtained by using STRING online database (https://cn.string-db.org/), and the PPI network was visualized with Cytoscape v3.8.2[12, 13]. The PPI network of hub Gene DAPL1 was explored by using Gene MANIA prediction server (https://genemania.org)[14].

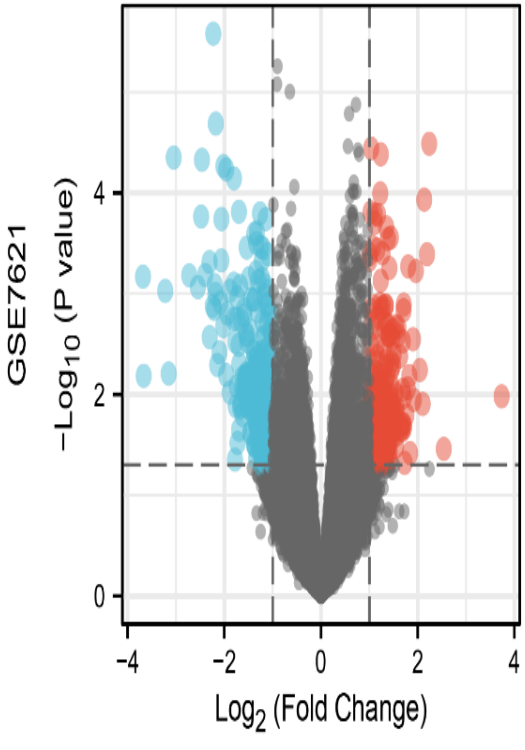
Other online databases

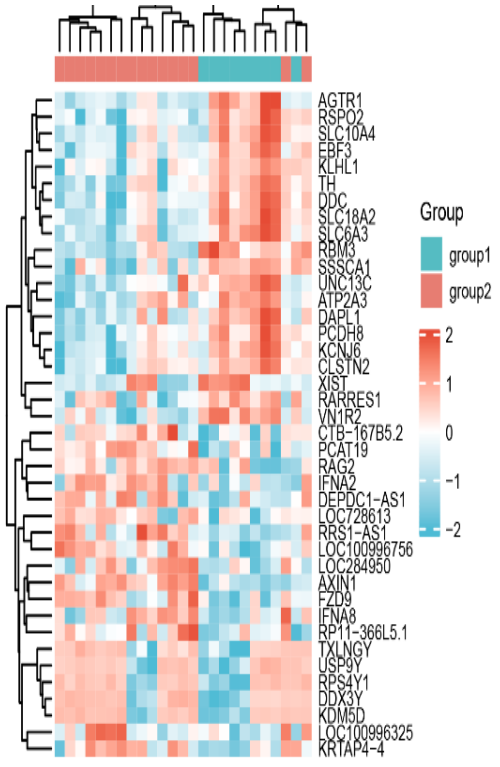
The TreeFam database was applied to obtain DAPL1 gene family information[15]. Aminode was employed to analyze the sequence conservatory of amino acids across families[16]. The human protein atlas database analyzed the expression distribution of DAPL1 in normal brain tissues[17]; and shinygeo was used to verify the expression of DAPL1 in the dataset of the substantia nigra samples in Parkinson's disease patients[18].

Results

Screening of Parkinson's disease genes with expression differences

Based on adj. P< 0.05 and | the screening criteria of logFC |>1, the gene expression profile of GSE7621 identified 527 DEGs, of which 246 were upregulated and 281 were down-regulated; 986 DEGs were screened out in GSE42966, of which 487 were upregulated and 499 were down-regulated; and the gene volcano maps and sample clustering results of the two groups of DEGs in the experimental group and the control group were shown in **Figure 1A-B**. Venn plots were plotted on the acquired differential gene datasets, and 45 co-expressed DEGs (**Figure 1C, Table 1**) were obtained, of which 5 were upregulated genes, namely PLA2G1B, RELN, IL7R, SLC9A3R2, and PLCXD1. There are 40 downregulated genes, namely PAPPA, TUBGCP5, TRPV2, CNIH3, ZAR1, RERG, KLHL13, SYNGR3, CGREF1, GFRA1, GPX3, PTPRN, TUB, RGS6, SLIT1, LXN, CPLX2, SLC2A13, DOK6, CADPS2, TPBG, ROBO2, CNTN6, EN1, TNMD, CBLN1, CEACAM21, HTR1F, DLK1, PCSK1, LPO, ALDH1A1, CLSTN2, AGTR1, PCDH8, DAPL1, SLC10A4, SLC18A2, DDC, TH.





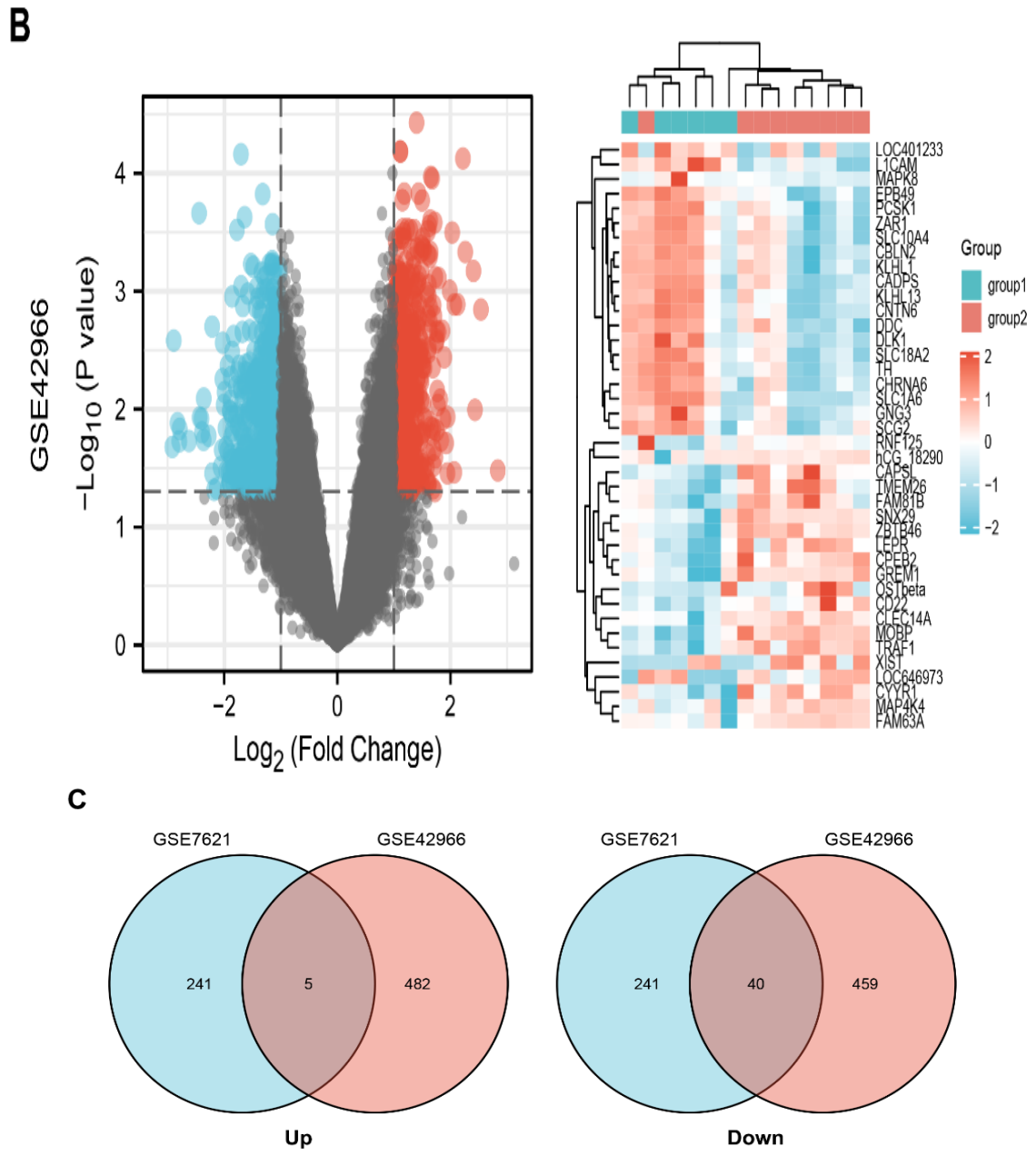


Figure 1: Screening of differentially expressed genes in substantia nigra of Parkinson's disease. Volcano plot representing the DEGs and heat map of the 20 most upregulated and downregulated genes in of GSE7621 (A) and GSE42966 (B), satisfying the criteria of log2 (fold-change) value >1 or <-1 and P<0.05. Significantly expressed genes are represented as red or blue dots. (C) Venn diagram based on the co-up-regulated and down-regulated DEGs between GSE7621 and GSE42966.

Table 1(a) **Common DEGs of GSE7621 and GSE42966**

Up	Down
PLA2G1B	PAPPA
RELN	TUBGCP5
IL7R	TRPV2
SLC9A3R2	CNIH3
PLCXD1	ZAR1
	RERG
	KLHL13
	SYNGR3
	CGREF1
	GFRA1
	GPX3
	PTPRN
	TUB

Table 1(b) **Common DEGs of GSE7621 and GSE42966**

Up	Down
	RGS6
	SLIT1
	LXN
	CPLX2
	SLC2A13
	DOK6
	CADPS2
	TPBG
	ROBO2
	CNTN6
	EN1
	TNMD
	CBLN1
	CEACAM21
	HTR1F
	DLK1
	PCSK1
	LPO
	ALDH1A1
	CLSTN2
	AGTR1
	PCDH8
	DAPL1
	SLC10A4
	SLC18A2
	DDC
	TH

GO/KEGG enrichment and protein interaction Analysis of DEGs in Parkinson's disease

GO/KEGG enrichment analysis showed that in BP, common DEGs were mainly enriched in the regulation of synapse organization, regulation of synapse structure or activity, regulation of synapse assembly, etc. In CC, DEGs were mainly involved in PR synapse, parallel fiber to Purkinje cell synapse, transport vesicle, etc. In MF, common DEGs were primarily involved in the organic hydroxy compound transmembrane transporter activity. The results of KEGG pathway enrichment analysis further showed that the co-differential genes were enriched in serotonergic synapse, Cocaine addiction, Amphetamine addiction, etc. (**Figure 2A-B, Table 2**).

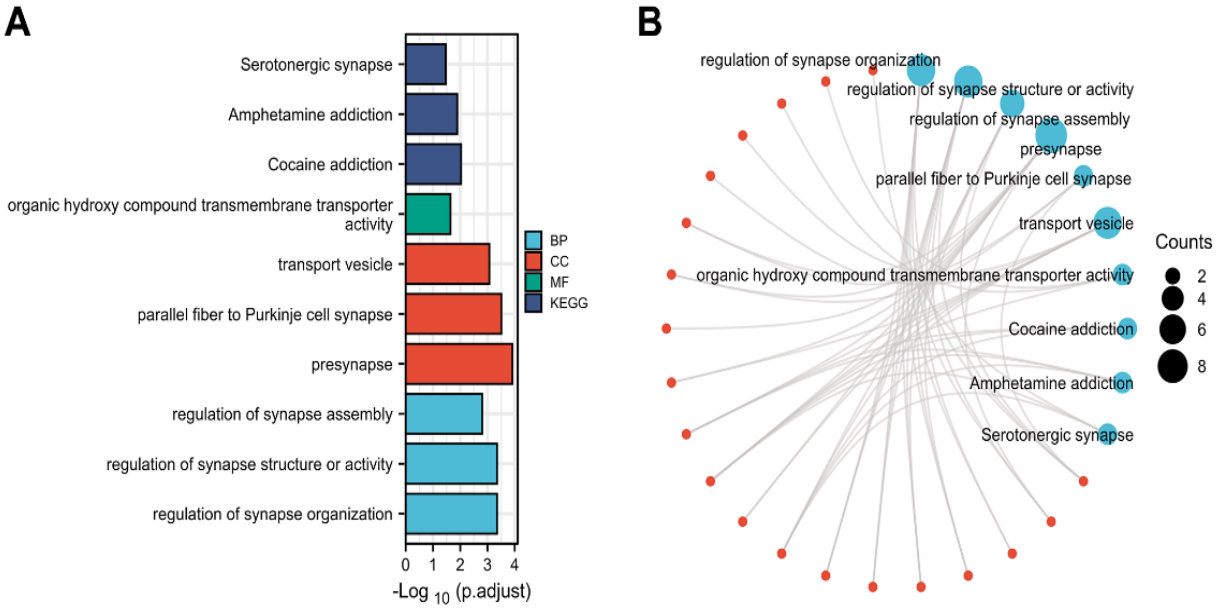


Figure 2: GO/KEGG enrichment and protein interaction analysis of DEGs between GSE7621 and GSE42966. **(A)** Histogram for GO and KEGG analysis of 45 DEGs between GSE7621 and GSE42966. **(B)** Circular map for GO and KEGG analysis of 45 DEGs between GSE7621 and GSE42966

Table 2 **GOKEGG analysis of common DEGs**

Ontology	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0050807	regulation of synapse organization	7/44	218/18670	7.14e-07	4.47e-04	3.64e-04
BP	GO:0050803	regulation of synapse structure or activity	7/44	227/18670	9.36e-07	4.47e-04	3.64e-04
BP	GO:0051963	regulation of synapse assembly	5/44	106/18670	4.89e-06	0.002	0.001
CC	GO:0098793	presynapse	9/44	491/19717	1.12e-06	1.22e-04	8.12e-05
CC	GO:0098688	parallel fiber to Purkinje cell synapse	3/44	16/19717	5.69e-06	3.10e-04	2.07e-04
CC	GO:0030133	transport vesicle	7/44	392/19717	2.36e-05	8.59e-04	5.73e-04
MF	GO:1901618	organic hydroxy compound transmembrane transporter activity	3/42	44/17697	1.54e-04	0.023	0.019
KEGG	hsa05030	Cocaine addiction	3/18	49/8076	1.61e-04	0.009	0.008
KEGG	hsa05031	Amphetamine addiction	3/18	69/8076	4.44e-04	0.013	0.011
KEGG	hsa04726	Serotonergic synapse	3/18	115/8076	0.002	0.034	0.029

C

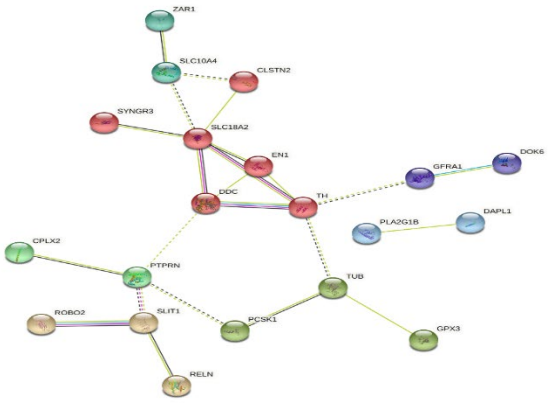


Figure 2C: STRING protein network map of the 45 DEGs. nodes and lines indicate the individual protein identified and the direct or indirect association, respectively

Table 3 **STRING PPI network analysis of common DEGs**

Node1	Node2	coexpression	Experimentally_determined_interaction	Database annotated	Automated_textmining	Combined_score
SLIT1	ROBO2	0.096	0.572	0.8	0.97	0.997
TH	DDC	0.063	0.126	0.9	0.922	0.992
SLC18A2	TH	0.066	0.176	0	0.925	0.937
SLC18A2	DDC	0.062	0.176	0	0.914	0.928
GFRA1	DOK6	0	0	0.9	0.2	0.916
SYNGR3	SLC18A2	0.065	0	0	0.641	0.651
EN1	SLC18A2	0.062	0	0	0.604	0.613
SLC10A4	ZAR1	0.078	0	0	0.594	0.609
EN1	TH	0	0	0	0.605	0.605
EN1	DDC	0	0	0	0.573	0.573
SLC18A2	CLSTN2	0	0	0	0.542	0.542
SLIT1	RELN	0.098	0	0	0.511	0.54
TUB	GPX3	0	0	0	0.5	0.499
SLC10A4	CLSTN2	0.062	0	0	0.487	0.498
TUB	PCSK1	0.096	0	0	0.444	0.476
PTPRN	PCSK1	0.232	0	0	0.339	0.471
GFRA1	TH	0.062	0	0	0.455	0.466
SLC10A4	SLC18A2	0.049	0	0	0.459	0.464
DAPL1	PLA2G1B	0	0	0	0.448	0.448
PTPRN	CPLX2	0.361	0	0	0.16	0.44
PTPRN	DDC	0	0	0	0.415	0.414
TUB	TH	0.063	0	0	0.397	0.411
SLIT1	PTPRN	0.129	0.148	0	0.257	0.4

The protein interaction network results showed 20 nodes, 22 inter-protein interactions, of which only 1 gene with DAPL1 having a locus of action (**Figure 2C, Table 3**).

Conservative analysis and brain tissue expression distribution analysis of potentially pathogenic gene DAPL1 in Parkinson's disease

In the DISEASES database, there were 98 genes closely associated with Parkinson's disease (**Table 4**), which were intersected with the common difference genes DEGs described above by Venny and finally 2 co-differentially expressed genes, DAPL1 and SLC2A13 were obtained (**Figure 3A**). Treefam analysis showed that DAPL1 presented interspecies gene sequence conservatism (**Figure 3B**).

Aminode amino acid sequence analysis further revealed that amino acid conserved sites were in the 13-33, 46-58, and 82-90 regions (**Figure 3C**).

Analysis of the Human protein atlas database showed that DAPL1 was primarily expressed in the normal hypothalamus, midbrain, and cortex (**Figure 3D**).

Table 4(a) **DISEASES for PD**

Symbol	Name	Standardized Value
CRHR1-IT1	CRHR1 intronic transcript 1	1.70133
LRRK2	leucine-rich repeat kinase 2	1.68372
WNT3	wingless-type MMTV integration site family, member 3	1.2206
BST1	bone marrow stromal cell antigen 1	1.17894
SNCA	synuclein, alpha (non A4 component of amyloid precursor)	1.14929
CCDC62	coiled-coil domain containing 62	0.932337
RAB25	RAB25, member RAS oncogene family	0.850113
MCCC1	methylcrotonoyl-CoA carboxylase 1 (alpha)	0.682635
CTIF	CBP80/20-dependent translation initiation factor	0.602429
GAK	cyclin G associated kinase	0.58494
ADAMTSL1	ADAMTS-like 1	0.52461
WIPF3	WAS/WASL interacting protein family, member 3	0.464771
TMEM175	transmembrane protein 175	0.449873
SLC50A1	solute carrier family 50 (sugar efflux transporter), member 1	0.438544
RREB1	ras responsive element binding protein 1	0.438544
TMEM55A	transmembrane protein 55A	0.438544
SREBF1	sterol regulatory element binding transcription factor 1	0.429338
DNAJC1	DnaJ (Hsp40) homolog, subfamily C, member 1	0.429338
ARL15	ADP-ribosylation factor-like 15	0.4215
CYP17A1	cytochrome P450, family 17, subfamily A, polypeptide 1	0.4215
NUCKS1	nuclear casein kinase and cyclin-dependent kinase substrate 1	0.4215
CTNNA3	catenin (cadherin-associated protein), alpha 3	0.408752
HYI	hydroxypyruvate isomerase (putative)	0.403461
MMRN1	multimerin 1	0.403461
SLC41A1	solute carrier family 41 (magnesium transporter), member 1	0.403461
RFX4	regulatory factor X, 4 (influences HLA class II expression)	0.368757
TYW1B	tRNA-yW synthesizing protein 1 homolog B (S. cerevisiae)	0.368757
RIT2	Ras-like without CAAX 2	0.368757
ACMSD	aminocarboxymuconate semialdehyde decarboxylase	0.368757
CCDC82	coiled-coil domain containing 82	0.368757
TMC3	transmembrane channel-like 3	0.368757
KANSL1	KAT8 regulatory NSL complex subunit 1	0.368757
SPPL2C	signal peptide peptidase like 2C	0.368757
MAPT	microtubule-associated protein tau	0.368757
SH3BGRL	SH3 domain binding glutamate-rich protein like	0.345945
UNC13C	unc-13 homolog C (C. elegans)	0.345945
PRSS53	protease, serine, 53	0.327313
FAM47E	family with sequence similarity 47, member E	0.327313
TAS1R2	taste receptor, type 1, member 2	0.311576
SH3GL2	SH3-domain GRB2-like 2	0.311576
QSER1	glutamine and serine rich 1	0.311576
AXIN1	axin 1	0.311576
COL13A1	collagen, type XIII, alpha 1	0.297866
DAPL1	death associated protein-like 1	0.285417
CNKSR3	CNKSR family member 3	0.264872
STAP1	signal transducing adaptor family member 1	0.256228
HACE1	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	0.256228
CTSB	cathepsin B	0.256228

Table 4(b) **DISEASES for PD**

Symbol	Name	Standardized Value
SLC2A13	solute carrier family 2 (facilitated glucose transporter), member 13	0.256228
NSF	N-ethylmaleimide-sensitive factor	0.256228
PHKA2	phosphorylase kinase, alpha 2 (liver)	0.256228
DLG2	discs, large homolog 2 (Drosophila)	0.197766
PARD3	par-3 family cell polarity regulator	0.197766
DGKQ	diacylglycerol kinase, theta 110kDa	0.197766
DIAPH3	diaphanous-related formin 3	0.197766
CSMD1	CUB and Sushi multiple domains 1	0.197766
TMEM72	transmembrane protein 72	0.197766
STK39	serine threonine kinase 39	0.15771
RBMS3	RNA binding motif, single stranded interacting protein 3	0.15771
UNC13B	unc-13 homolog B (C. elegans)	0.15771
SLC45A3	solute carrier family 45, member 3	0.15771
PAK1	p21 protein (Cdc42/Rac)-activated kinase 1	0.15771
MIR4315-2	microRNA 4315-2	0.15771
USP15	ubiquitin specific peptidase 15	0.15771
GFPT2	glutamine-fructose-6-phosphate transaminase 2	0.15771
ITGAL	integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)	0.15771
OCA2	oculocutaneous albinism II	0.15771
IGSF5	immunoglobulin superfamily, member 5	0.15771
RAP1A	RAP1A, member of RAS oncogene family	0.15771
C9	complement component 9	0.126165
ULK2	unc-51 like autophagy activating kinase 2	0.126165
USP36	ubiquitin specific peptidase 36	0.126165
GPM6B	glycoprotein M6B	0.126165
DSG3	desmoglein 3	0.126165
CLRN3	clarin 3	0.126165
TASP1	taspase, threonine aspartase, 1	0.100789
DKK2	dickkopf WNT signaling pathway inhibitor 2	0.100789
C4ORF26	chromosome 4 open reading frame 26	0.100789
LMNB1	lamin B1	0.100789
PRDM15	PR domain containing 15	0.100789
ITGA8	integrin, alpha 8	0.100789
HIVEP2	human immunodeficiency virus type I enhancer binding protein 2	0.080593
SPANXN3	SPANX family, member N3	0.080593
OTOG	otogelin	0.080593
COLGALT2	collagen beta(1-O)galactosyltransferase 2	0.080593
MAP2K6	mitogen-activated protein kinase kinase 6	0.080593
LMO2	LIM domain only 2 (rhombotin-like 1)	0.052255
ATF6	activating transcription factor 6	0.052255
SEMA5A	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	0.052255
SCN2A	sodium channel, voltage gated, type II alpha subunit	0.052255
ETV6	ets variant 6	0.052255
CTC1	CTS telomere maintenance complex component 1	0.043246
CYB5RL	cytochrome b5 reductase-like	0.043246
RNF114	ring finger protein 114	0.043246
AAK1	AP2 associated kinase 1	0.043246
TCF12	transcription factor 12	0.043246
SURF6	surfeit 6	0.043246
OR9G1	olfactory receptor, family 9, subfamily G, member 1	0.035343

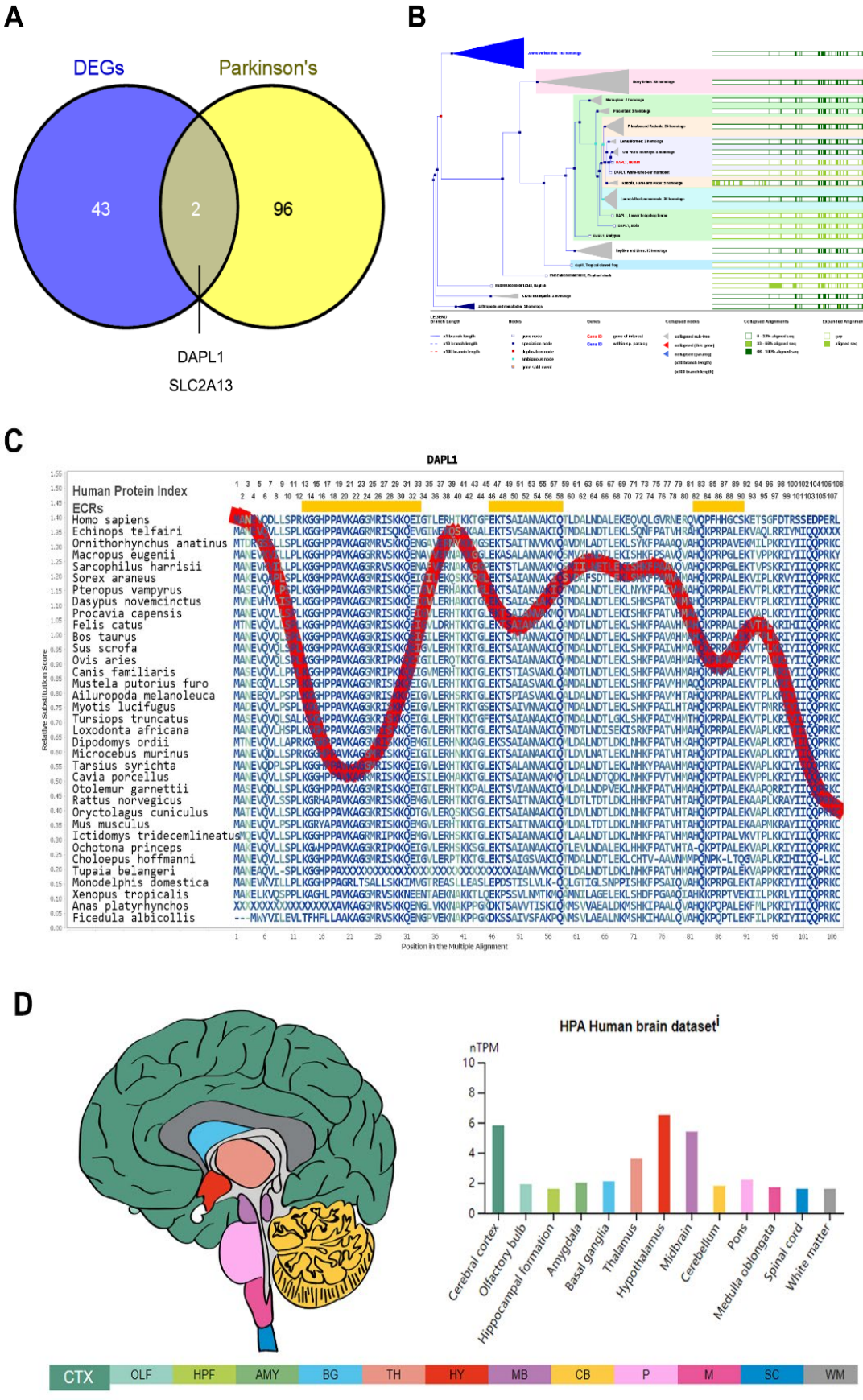


Figure 3: Bioinformatics analysis of DAPL1. (A) Venn diagram of the DEGs and risk genes in Parkinson's disease. (B) the pruned tree for Tree Fam family TF329716 (DAPL1) showing model species only. (C) Aminode graphical analysis of DAPL1. (D) DAPL1 expression in human brain tissues

Expression verification in authentication set GSE49036 and PPI network construction of DAPL 1

Based on the expression data of Parkinson's disease substantia (15 samples) and normal tissue (8 samples) collected in the GSE49036 database, whether DAPL1 was expressed significantly was explored. The results of the analysis showed (**Figure 4A**) that the expression of DAPL1 in the substantia tissues of patients with Parkinson's disease was decreased compared with that in normal substantia nigra tissues, with a statistical difference ($p<0.05$). Further PPI network construction found that proteins that interact closely with DAPL1 including DAP, NRG2, BTBD11, etc., which were significantly related to the death domain binding signal pathway (**Figure 4B**).

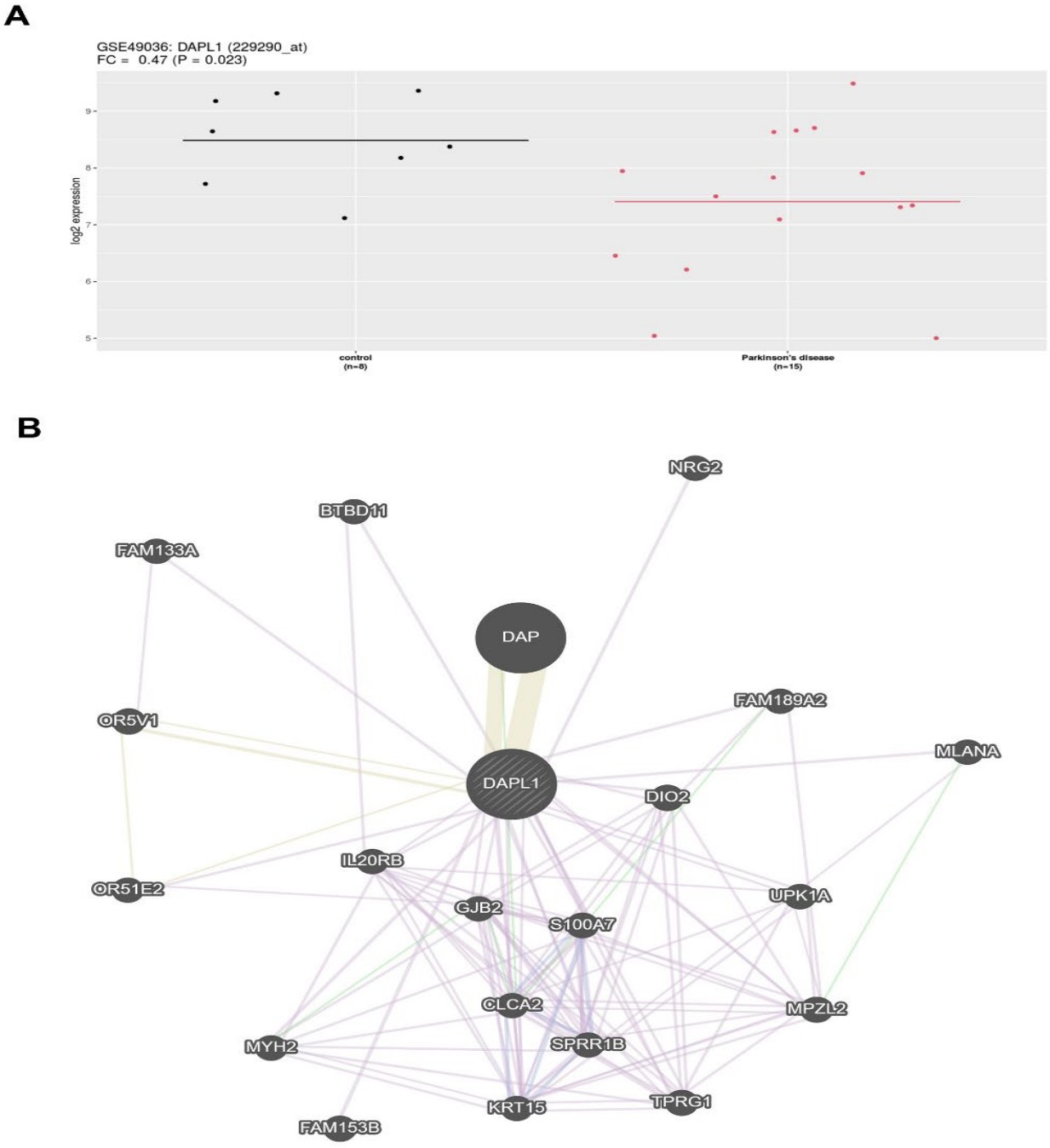


Figure 4: Expression validation and PPI network of DAPL1. (A) DAPL1 expression in GSE49036 validation dataset. (B) PPI networks of DAPL1. Edge thickness shows the interaction strength.

Discussion

According to the epidemiological data of Parkinson's disease in recent years, the prevalence of Parkinson's disease in the elderly population over 65 years old in China was about 1.7%[19]. According to this data, there are currently more than 2.7 million Parkinson's patients in China. With the development of the aging population, the disease burden of Parkinson's disease was also gradually increasing, and it was expected that the number of patients with Parkinson's disease in China may reach 5 million, accounting for 50% of Parkinson's disease patients worldwide by 2030.

This data should not be underestimated, behind it was the grim situation of the task of Parkinson's disease prevention and treatment in China[20]. The onset of Parkinson's disease may be

closely related to a variety of influencing factors, but the more important are genetic and environmental factors. In the later stages of Parkinson's disease, the drug levodopa becomes less effective at treating symptoms due to the unstoppable loss of neurons that released dopamine[21]. But in a new preclinical study, researchers from Northwestern University's Feinberg School of Medicine have discovered a gene therapy that targets the small brain region where these neurons were located, the gene therapy restored the ability of neurons in the substantia nigra to convert levodopa into dopamine[22]. Therefore, it was particularly necessary to explore the abnormally expressed genes and biological functions in the substantia nigra of patients with Parkinson's disease. Based on this background, we screened and obtained Parkinson's disease chips GSE7621, GSE42966 and GSE49036 from the GEO database, mined and analyzed the gene expression data of the substantia nigra they contained, and finally found that the low expression of DAPL1 may play a key role in the progression of Parkinson's disease. First, we go through adj. P< 0.05 and | the screening criteria of FC |>1, and 45 differential co-expressed genes were screened in GSE7621 and GSE42966 (Figure 1). Further ANALYSIS OF GO/KEGG enrichment and protein interactions shows that co-DEGs are primarily closely related to the regulation of synaptic activity (Figure 2).

Diseases-gene-related evidence from the DISEASES database through the integration of experimental datasets (GWAS), hand-searched literature, or automated text mining. To further screen out potential key genes in the substantia nigra in the development of Parkinson's disease, we searched the DISEASES database and found that there were 98 genes closely associated with Parkinson's disease, and further intersected by Venny and the previously obtained common difference genes DEGs, and finally obtained 2 co-differentially expressed genes, DAPL1 and SLC2A13 (Figure 3A). Several research groups, including ChunYu Li, have found that SLC2A13 is a risk gene in patients with Parkinson's disease[23-25]. The above data further confirms the reliability and forward-looking nature of our analytical results. DAPL1 is a prognostic indicator of tumors such as lung cancer and an important regulator of testosterone production by Leydig cells in mouse testicles[26, 27]. However, there are no literature reports on DAPL1 and Parkinson's disease. Treefam analysis showed that DAPL1 was highly conserved in interspecies gene sequences (Figure 3B); Aminode amino acid sequence analysis results further showed that amino acid conserved sites were in 13-33, 46-58, and 82-903 regions (Figure 3C); Human protein atlas database analysis showed that DAPL1 was mainly expressed in the normal hypothalamus, midbrain, and cortex (Figure 3D). The above results show that DAPL1 is likely to play a role in brain diseases.

To further investigate the role of DAPL1 in Parkinson's disease, the expression of DAPL1 in Parkinson's disease substantia tissue and normal tissue was verified by using the GSE49036 validation dataset, and the results found that the expression of DAPL1 in the substantia tissue of patients with GSE49036 Parkinson's disease was also significantly reduced similar to the screening results of GSE7621 and GSE42966. The DAPL1 PPI network found that DAPL1 interacts mainly with death domain binding-related proteins such as DAP, NRG2, BTBD11, etc., suggesting that DAPL1 may play a role in inhibiting the process of Parkinson's disease through the death domain binding signal.

Conclusion

Our research found that DAPL1 may be a risk gene of PD, but the specific mechanism still needs to be further explored.

Data Availability Statement

All data generated or analyzed during this study are included in this published article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Reference

[1] G. Li *et al.*, "Parkinson's disease in China: a forty-year growing track of bedside work," *Transl Neurodegener,* vol. 8, p. 22, 2019, doi: 10.1186/s40035-019-0162-z.

[2] M. J. Armstrong and M. S. Okun, "Diagnosis and Treatment of Parkinson Disease: A Review," *JAMA,* vol. 323, no. 6, pp. 548-560, Feb 11 2020, doi: 10.1001/jama.2019.22360.

[3] B. G. Trist, D. J. Hare, and K. L. Double, "Oxidative stress in the aging substantia nigra and the etiology of Parkinson's disease," *Aging Cell,* vol. 18, no. 6, p. e13031, Dec 2019, doi: 10.1111/acel.13031.

[4] D. Chicco, "geneExpressionFromGEO: An R Package to Facilitate Data Reading from Gene Expression Omnibus (GEO)," *Methods Mol Biol,* vol. 2401, pp. 187-194, 2022, doi: 10.1007/978-1-0716-1839-4_12.

[5] S. H. Liu, Y. L. Wang, S. M. Jiang, X. J. Wan, J. H. Yan, and C. F. Liu, "Identifying the hub gene and immune infiltration of Parkinson's disease using bioinformatical methods," *Brain Res,* vol. 1785, p. 147879, Jun 15 2022, doi: 10.1016/j.brainres.2022.147879.

[6] F. Ferraro, C. Fevga, V. Bonifati, W. Mandemakers, A. Mahfouz, and M. Reinders, "Correcting Differential Gene Expression Analysis for Cyto-Architectural Alterations in Substantia Nigra of Parkinson's Disease Patients Reveals Known and Potential Novel Disease-Associated Genes and Pathways," *Cells,* vol. 11, no. 2, Jan 7 2022, doi: 10.3390/cells11020198.

[7] E. McCarthy *et al.*, "Gene Co-expression Analysis of the Human Substantia Nigra Identifies ZNHIT1 as an SNCA Co-expressed Gene that Protects Against alpha-Synuclein-Induced Impairments in Neurite Growth and Mitochondrial Dysfunction in SH-SY5Y Cells," *Mol Neurobiol,* vol. 59, no. 5, pp. 2745-2757, May 2022, doi: 10.1007/s12035-022-02768-9.

[8] S. Pletscher-Frankild, A. Palleja, K. Tsafou, J. X. Binder, and L. J. Jensen, "DISEASES: text mining and data integration of disease-gene associations," *Methods,* vol. 74, pp. 83-9, Mar 2015, doi: 10.1016/j.ymeth.2014.11.020.

[9] J. L. Steenwyk and A. Rokas, "ggpubfigs: Colorblind-Friendly Color Palettes and ggplot2 Graphic System Extensions for Publication-Quality Scientific Figures," *Microbiol Resour Announc,* vol. 10, no. 44, p. e0087121, Nov 4 2021, doi: 10.1128/MRA.00871-21.

[10] Z. Gu and D. Hubschmann, "Make Interactive Complex Heatmaps in R," *Bioinformatics,* Dec 2 2021, doi: 10.1093/bioinformatics/btab806.

[11] T. Wu *et al.*, "clusterProfiler 4.0: A universal enrichment tool for interpreting omics data," *Innovation (N Y),* vol. 2, no. 3, p. 100141, Aug 28 2021, doi: 10.1016/j.xinn.2021.100141.

[12] D. Szklarczyk *et al.*, "The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Res,* vol. 49, no. D1, pp. D605-D612, Jan 8 2021, doi: 10.1093/nar/gkaa1074.

[13] N. T. Doncheva, J. H. Morris, J. Gorodkin, and L. J. Jensen, "Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data," *J Proteome Res,* vol. 18, no. 2, pp. 623-632, Feb 1 2019, doi: 10.1021/acs.jproteome.8b00702.

[14] M. Franz *et al.*, "GeneMANIA update 2018," *Nucleic Acids*

Res, vol. 46, no. W1, pp. W60-W64, Jul 2 2018, doi: 10.1093/nar/gky311.

[15] F. Schreiber, M. Patricio, M. Muffato, M. Pignatelli, and A. Bateman, "TreeFam v9: a new website, more species and orthology-on-the-fly," *Nucleic Acids Res,* vol. 42, no. Database issue, pp. D922-5, Jan 2014, doi: 10.1093/nar/gkt1055.

[16] K. T. Chang, J. Guo, A. di Ronza, and M. Sardiello, "Aminode: Identification of Evolutionary Constraints in the Human Proteome," *Sci Rep,* vol. 8, no. 1, p. 1357, Jan 22 2018, doi: 10.1038/s41598-018-19744-w.

[17] A. Digre and C. Lindskog, "The Human Protein Atlas-Spatial localization of the human proteome in health and disease," *Protein Sci,* vol. 30, no. 1, pp. 218-233, Jan 2021, doi: 10.1002/pro.3987.

[18] J. Dumas, M. A. Gargano, and G. M. Dancik, "shinyGEO: a web-based application for analyzing gene expression omnibus datasets," *Bioinformatics,* vol. 32, no. 23, pp. 3679-3681, Dec 1 2016, doi: 10.1093/bioinformatics/btw519.

[19] Z. Song *et al.*, "Prevalence of Parkinson's Disease in Adults Aged 65 Years and Older in China: A Multicenter Population-Based Survey," *Neuroepidemiology,* vol. 56, no. 1, pp. 50-58, 2022, doi: 10.1159/000520726.

[20] Y. Zhang *et al.*, "Mortality from Parkinson's disease in China: Findings from a ten-year follow up study in Shanghai," *Parkinsonism Relat Disord,* vol. 55, pp. 75-80, Oct 2018, doi: 10.1016/j.parkreldis.2018.05.012.

[21] H. Bogetofte, A. Alamyar, M. Blaabjerg, and M. Meyer, "Levodopa Therapy for Parkinson's Disease: History, Current Status and Perspectives," *CNS Neurol Disord Drug Targets,* vol. 19, no. 8, pp. 572-583, 2020, doi: 10.2174/1871527319666200722153156.

[22] P. Gonzalez-Rodriguez *et al.*, "Disruption of mitochondrial complex I induces progressive parkinsonism," *Nature,* vol. 599, no. 7886, pp. 650-656, Nov 2021, doi: 10.1038/s41586-021-04059-0.

[23] M. Kim, J. H. Won, J. Youn, and H. Park, "Joint-Connectivity-Based Sparse Canonical Correlation Analysis of Imaging Genetics for Detecting Biomarkers of Parkinson's Disease," *IEEE Trans Med Imaging,* vol. 39, no. 1, pp. 23-34, Jan 2020, doi: 10.1109/TMI.2019.2918839.

[24] C. Li *et al.*, "Mutation analysis of seven SLC family transporters for early-onset Parkinson's disease in Chinese population," *Neurobiol Aging,* vol. 103, pp. 152 e1-152 e6, Jul 2021, doi: 10.1016/j.neurobiolaging.2021.02.022.

[25] W. Satake *et al.*, "Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease," *Nat Genet,* vol. 41, no. 12, pp. 1303-7, Dec 2009, doi: 10.1038/ng.485.

[26] Y. Y. Feng, Y. Yang, Y. Y. Wang, X. X. Bai, P. Hai, and R. Zhao, "[Hypomethylation of DAPL1 associated with prognosis of lung cancer patients with EGFR Del19 mutation]," *Zhonghua Zhong Liu Za Zhi,* vol. 43, no. 12, pp. 1264-1268, Dec 23 2021, doi: 10.3760/cma.j.cn112152-20190923-00618.

[27] H. B. Chen *et al.*, "DAPL1 is a novel regulator of testosterone production in Leydig cells of mouse testis," *Sci Rep,* vol. 11, no. 1, p. 18532, Sep 17 2021, doi: 10.1038/s41598-021-97961-6.