Characterizing the Expression Patterns of Parkinson's Disease Associated Genes

Bin Li[1,2,3], Guihu Zhao[1,2], Kuokuo Li[4], Zheng Wang[1], Zhenghuan Fang[4], Xiaomeng Wang[4], Tengfei Luo[4], Yi Zhang[1], Yijing Wang[4], Qian Chen[1], Yuanfeng Huang[1], Lijie Dong[4], Jifeng Guo[1,4], Beisha Tang[1,4]* and Jinchen Li[1,2,4]*

[1] Department of Geriatrics, National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China, [2] Department of Neurology, Xiangya Hospital, Central South University, Changsha, China, [3] Mobile Health Ministry of Education–China Mobile Joint Laboratory, Xiangya Hospital, Central South University, Changsha, China, [4] Center for Medical Genetics and Hunan Key Laboratory, School of Life Sciences, Central South University, Changsha, China

Background: The expression pattern represents a quantitative phenotype that provides an in-depth view of the molecular mechanism in Parkinson's disease (PD); however, the expression patterns of PD-associated genes (PAGs) and their relation to age at onset (AAO) remain unclear.

Methods: The known PD-causing genes and PD-risk genes, which were collected from latest published authoritative meta-analysis, were integrated as PAGs. The expression data from Genotype-Tissue Expression database, Allen Brian Map database, and BrainSpan database, were extracted to characterize the tissue specificity, inhibitory-excitatory neuron expression profile, and spatio-temporal expression pattern of PAGs, respectively. The AAO information of PD-causing gene was download from Gene4PD and MDSgene database.

Results: We prioritized 107 PAGs and found that the PAGs were more likely to be expressed in brain-related tissues than non-brain tissues and that more PAGs had higher expression levels in excitatory neurons than inhibitory neurons. In addition, we identified two spatio-temporal expression modules of PAGs in human brain: the first module showed a higher expression level in the adult period than in the prenatal period, and the second module showed the opposite features. It showed that more PAGs belong to the first module that the second module. Furthermore, we found that the median AAO of patients with mutations in PD-causing genes of the first module was lower than that of the second module.

Conclusion: In conclusion, this study provided comprehensive landscape of expression patterns, AAO features and their relationship for the first time, improving the understanding of pathogenesis, and precision medicine in PD.

Keywords: Parkinson's disease, PD-associated genes, expression pattern, gene expression, age at onset

BACKGROUND

Parkinson's disease (PD) is the second most common neurodegenerative disease (Dorsey et al., 2007; Alzheimer's Association, 2014), characterized by motor syndromes, such as resting tremor, bradykinesia, rigidity, and postural instability, or non-motor symptoms such as dementia, depression, and sleep disorders (Kalia and Lang, 2015). The prevalence of PD is approximately 1% in the population over the age of 60 years, but reaches up to 5% in individuals older than age 85 years, highlighting the impact of advancing age on the risk of this disease (Ascherio and Schwarzschild, 2016; Lim et al., 2019). The economic burden of PD is up to $23 billion USD per year (Findley, 2007). As life expectancy and aging populations increase worldwide, the number of people with PD is expected to increase by more than 1.5-fold by 2030, making this disease an urgent global health issue (Dorsey et al., 2007).

Although both genetic and environmental factors likely contribute to the development of PD, it shows a high heritability (Blauwendraat et al., 2019a; Li et al., 2019). With the development of high throughput technologies in the last two decades (Jansen et al., 2017; Sandor et al., 2017; Siitonen et al., 2017), thousands of studies have focused on the genetic architectures of PD to understand its pathogenesis (Kalia and Lang, 2015; Labbe et al., 2016; Blauwendraat et al., 2019b). Up to now, genome-wide association studies (GWAS) on the contribution of genetics in a common form of PD (sporadic PD) identified more than 100 loci that were associated with disease risk at a genome-wide significance level, which could sever as potential biomarkers and risk factors for PD (International Parkinson Disease Genomics Consortium, Nalls et al., 2011; Nalls et al., 2014; Chang et al., 2017; Blauwendraat et al., 2019a; Nalls et al., 2019). Additionally, over 20 genes were regarded as PD-causing genes (Blauwendraat et al., 2019b). Meanwhile, we have replicated *PRKN* (Guo et al., 2008), *PINK1* (Tang et al., 2006; Guo et al., 2008), *PARK7* (Tang et al., 2006; Guo et al., 2008), *ATP13A2* (Guo et al., 2008), *PLA2G6* (Shi et al., 2011), *CHCHD2* (Liu et al., 2015), *RAB39B* (Kang et al., 2016), *TMEM230* (Yan et al., 2017), *GCH1* (Xu et al., 2017), and other genes (Tian et al., 2012; Yang et al., 2019) in patients with PD in China. Furthermore, we performed multiplex ligation-dependent probe amplification assays and whole-exome sequencing for 1,676 unrelated patients with PD in a mainland Chinese population. Our data indicate that PD patients with an age at onset (AAO) of 40 years may benefit from genetic counseling, especially those from families with a recessive inheritance pattern (Zhao et al., 2020).

Gene expression studies have been a ubiquitous methodology in biomedical research to deeper understand pathological mechanisms driving diseases (Amar et al., 2015). Meanwhile, gene expression in the patient-personalized medicine may predict the fate of disease, treatment effectiveness, as well as may improve diagnostics and epidemiological disease at a population level (Gire et al., 2014; Fischer et al., 2015; Roychowdhury and Chinnaiyan, 2016; Judge et al., 2020). It

is to be noted that gene expression in human brain plays an important role in the development of PD. However, whether PD-associated genes (PAGs) present with a specific expression pattern is unclear. To investigate the specific genetic components of PAGs expression, we systematically characterized the expression pattern of PAGs. This study is expected to provide a comprehensive genetic landscape of PAGs for improving the understanding of the pathogenesis of PD.

MATERIALS AND METHODS

The PAGs were integrated with PD-causing genes from a systematic review and PD-risk genes (Blauwendraat et al., 2019b) from iPDGC Locus Browser[1] (Grenn et al., 2020), which is collecting three GWAS researches (Iwaki et al., 2019; Nalls et al., 2019; Foo et al., 2020). This gene set was further used to investigate the tissue specificity, excitatory-inhibitory neuron expression profile, as well as the spatiotemporal expression pattern.

To explore the tissue specificity of PAGs in different tissues, we sourced the expression table of 54 tissues across 948 donors from the Genotype-Tissue Expression (GTEx)[2] database (GTEx Consortium et al., 2017). The average expression level of each gene in each tissue was then calculated. For each gene, tissues with the top 50% expression levels in the 53 tissues were defined as preferential expression tissue. In addition, a Reads Per Kilobase per Million mapped reads (RPKM) value of 1 is generally used for defining the threshold of expressed genes. Fisher's exact test was used to calculate the P-value for each tissue.

We analyzed the inhibitory-excitatory neuron expression profile of PAGs based on single-nucleus RNA sequencing data which was downloaded from the Allen Brian Map database[3]. This data set contains 75 transcriptionally distinct cell types from 15,928 nuclei, including 45 GABAergic (inhibitory) neuron types, 24 glutamatergic (excitatory) neuron types, and six non-neuronal types. For each gene in each nucleus, the RPKM value, i.e., the mRNA expression level, was calculated based on the counts per million. Then, the average expression level of each neuronal type for each gene was counted. For the sake of comparison and analysis, we normalized the mean expression values by applying a conversion formula, as follows:

$$Normalized\ mean\ expression\ value$$
$$= log_2(expression\ mean\ value) \qquad (1)$$

We then used normalized mean expression values to assess the transcriptional profiles in inhibitory and excitatory neurons, and differences in transcription levels were statistically evaluated by the Wilcoxon rank test.

Developmental human brain RNA sequencing data were extracted from the BrainSpan database[4], which contained 524 tissue-samples of expression data across different developmental stages (from the fetal period to the adult period) and 16

---

[1] https://pdgenetics.shinyapps.io/GWASBrowser/
[2] https://gtexportal.org/
[3] https://portal.brain-map.org/atlases-and-data/rnaseq
[4] http://www.brainspan.org/

brain regions. To characterize the spatio-temporal expression pattern of PAGs, we performed a weighted gene co-expression network analysis (WGCNA) crossing all tissue samples from BrainSpan to cluster the spatio-temporal expression patterns of PAGs using the standard method with a power of seven. The expression values were normalized by a conversion formula, as follows:

$$Normalized\ mean\ expression\ value$$
$$= log_2 \frac{expression\ value}{mean\ expression\ value} \qquad (2)$$

In each co-expression module, the temporal-pattern was divided into 12 time-nodes (six before birth time-nodes and six after birth time-nodes); whereas the spatio-pattern was divided into 16 different brain districts, including the hippocampus, amygdala, striatum, mediodorsal nucleus of thalamus, cerebellar cortex, and 11 cerebral cortical districts.

Association Between Age at Onset and Expression Pattern

The AAO information of each PD-causing genes were collected from MDSgene database (Lill et al., 2016)[5]. We characterized the

[5] https://www.mdsgene.org/g4d

TABLE 1 | PD associated genes.

PD-causing genes[#]	PD-risk genes[##]
SNCA, *PRKN*,	*ADRA2A, ASXL3, BAG3, BIN3, BRIP1, BST1, C5orf24, CAB39L,*
UCHL1, PARK7	*CAMK2D, CASC16, CD19, CHD9, CHRNB1, CLCN3, CRHR1, CRLS1,*
LRRK2, *PINK1*,	*CTSB, DLG2, DNAH17, DYRK1A, ELOVL7, FAM171A2, FAM47E,*
POLG, HTRA2,	*FAM49B, FBRSL1, FCGR2A, FGF20, FYN, GAK, GALC, GBAP1, GBF1,*
ATP13A2, FBXO7,	*GCH1, GPNMB, GS1-124K5.11, HIP1R, HLA-DRB5, IGSF9B, INPP5F,*
GIGYF2, GBA,	*IP6K2, ITGA8, ITPKB, KCNIP3, KCNS3, KPNA1, KRTCAP2, LCORL,*
PLA2G6, EIF4G1,	*LINC00693, LOC100131289,* **LRRK2**, *MAP4K4, MBNL2, MCCC1,*
VPS35, DNAJC6,	*MED12L, MEX3C, MIPOL1, NOD2, NUCKS1, PAM, PMVK, RAB29,*
SYNJ1, DNAJC13,	*RETREG3, RIMS1, RIT2, RNF141, RPS12, RPS6KL1, SATB1, SCAF11,*
TMEM230, **VPS13C**,	*SCARB2, SETD1A, SH3GL2, SIPA1L2, SLC44A1,* **SNCA**, *SPPL2B,*
LRP10	*SPTSSB, STK39, SV2C, SYT17, TMEM163, TMEM175, TRIM40, UBAP2, UBTF, VAMP4,* **VPS13C**, *WBSCR17, WNT3*

[#]*Known PD-causing genes [Blauwendraat et al. (2019b). The genetic architecture of Parkinson's disease. Lancet Neurol].*
[##]*Known PD-risk genes [Grenn et al. (2020). The Parkinson's Disease Genome-Wide Association Study Locus Browser. Mov Disord].*
Three genes in bold terms, such as SNCA, LRRK2, VPS13C, are regarded as both PD-causing genes and PD-risk genes.

association of spatio-temporal expression patterns and AAO by comparing the median AAO of genes in a different module, which was achieved by performing a WGCNA. All genes were classified into one of three sections: juvenile-onset ($\leq$30 years), early onset (30–50 years), or late-onset ($>$50 years), according to the median of the AAO. Then we compared genes with more than five AAO terms to assess the association between AAO and PD-causing genes. After that, we check the association of spatio-temporal expression patterns and AAO by analyzing the other AAO data from Gene4PD[6].

RESULTS

In total, we sourced 21 known PD-casuing genes (Blauwendraat et al., 2019b) and 89 PD risk genes (**Table 1**) from iPDGC Locus

[6] http://genemed.tech/gene4pd
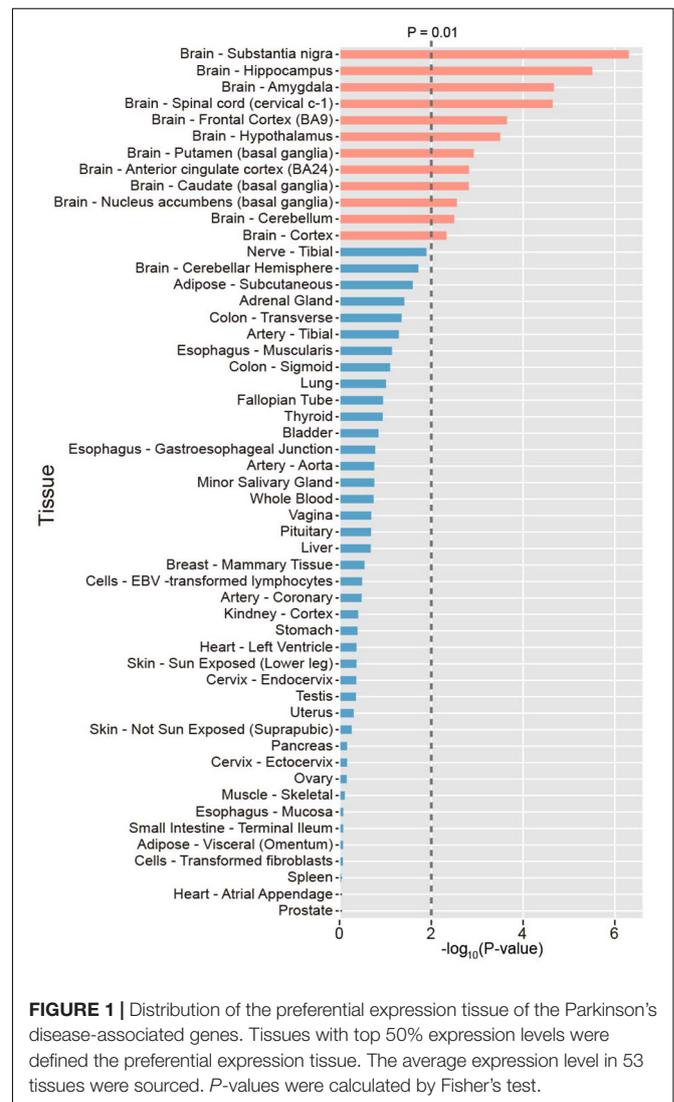
FIGURE 1 | Distribution of the preferential expression tissue of the Parkinson's disease-associated genes. Tissues with top 50% expression levels were defined the preferential expression tissue. The average expression level in 53 tissues were sourced. *P*-values were calculated by Fisher's test.

Browser[7] (Grenn et al., 2020), which is collecting three GWAS researches (Iwaki et al., 2019; Nalls et al., 2019; Foo et al., 2020). We noted that three PD-causing genes, such as *LRRK2, SNCA,* and *VPS13C*, were also regarded as PD-risk genes, the other genes in which have suggestive genetic evidence that confer risk to PD from GWAS. Totally, 107 PAGs were collected in this study.
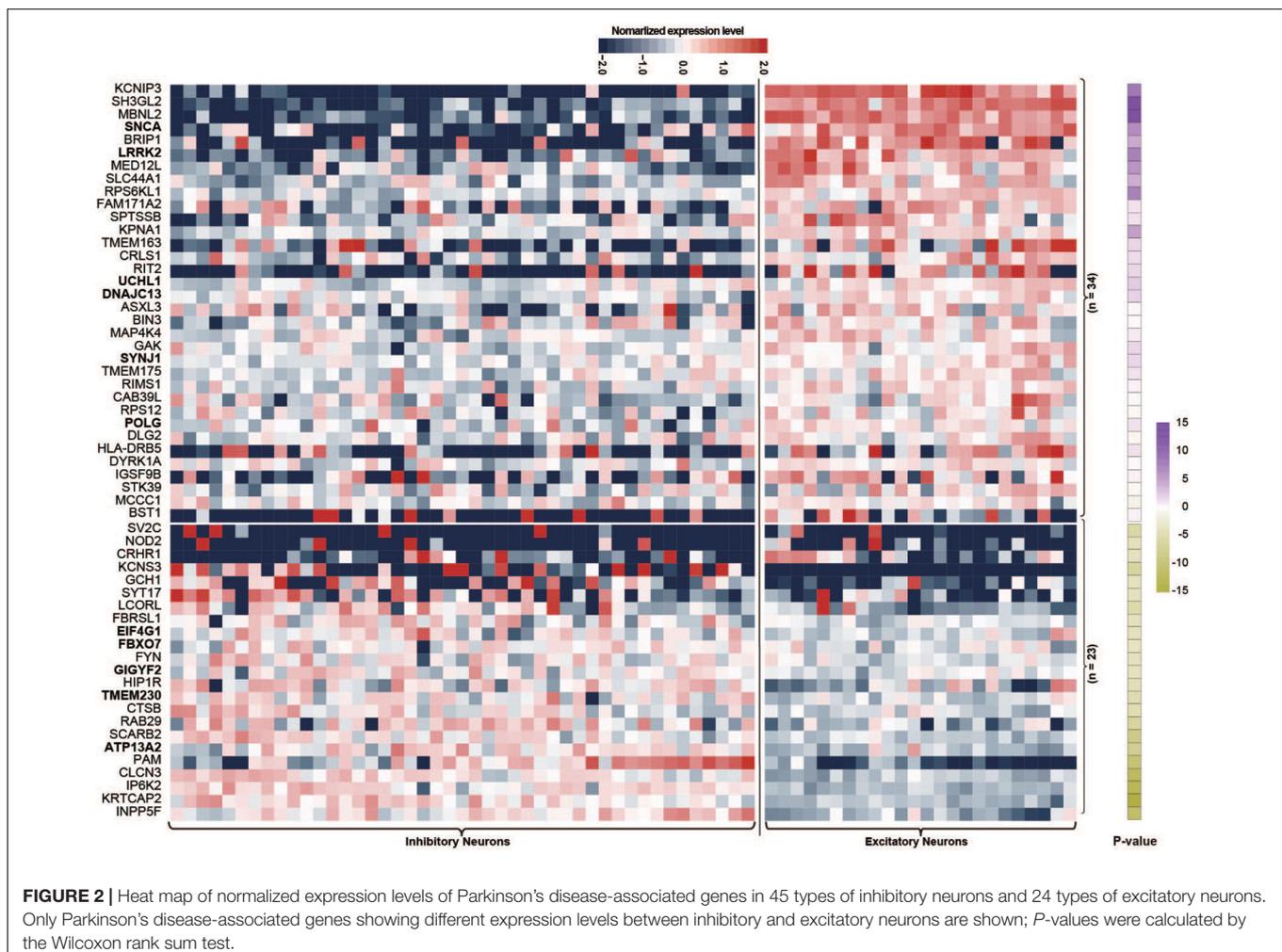
PAGs Presented Specific Expression Patterns

To investigate the tissue specificity of PAGs, we calculated the preferential expression of tissues of each gene and systematically tested the enrichment of preferential expression of tissues of the 107 PAGs using expression data derived from the GTEx project. As a result, we found that the 107 PAGs were significantly expressed in 12 tissues ($P < 0.01$), all of which belong to the brain, including the substantia nigra, hippocampus, amygdala, spinal cord, frontal cortex, hypothalamus, putamen, anterior cingulate cortex, caudate, nucleus accumbens, and cerebellum (**Figure 1**). This result provides novel evidence that PAGs are likely involved

---

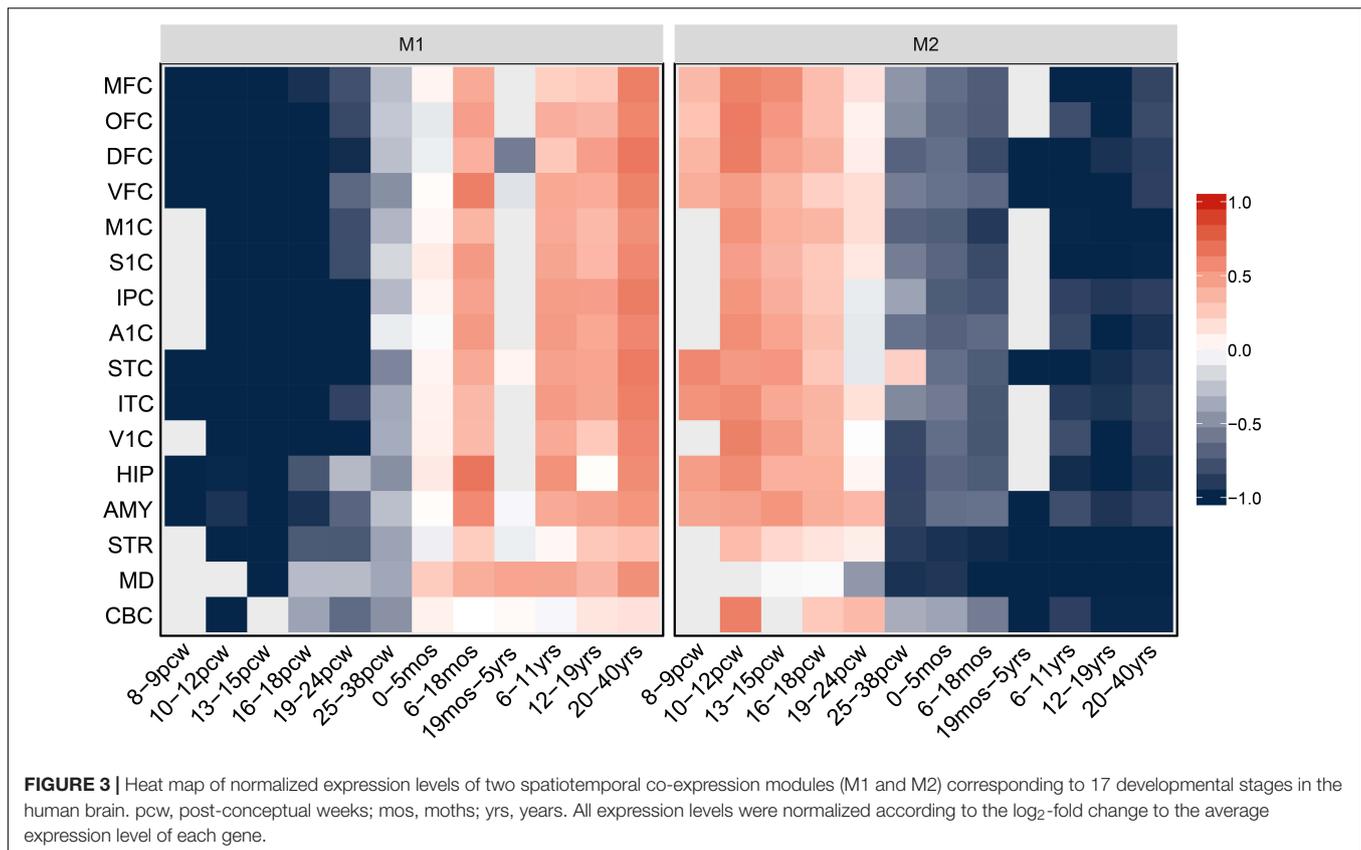[7]https://pdgenetics.shinyapps.io/GWASBrowser/

in the other brain specific gene regulation, as well as the known brain-specific gene regulation.

Single-nucleus RNA sequencing is a good method for evaluating the biological features of genes in neuronal subtypes. In this study, we investigated the expression levels of PAGs in 45 inhibitory neuron types and 24 excitatory neuron types based on The Allen Human Brain Reference Atlas database. For the 107 PAGs, 50 genes did not exhibit significantly different expression levels between excitatory and inhibitory neurons, whereas the other 57 genes presented significant differences in the expression level in the two type of neurons (**Supplementary Table 1**). Specifically, 34 of the 57 genes (59.65%) including six PD-causing genes (*LRRK2, SNCA, UCH1, DNAJC13, SYNJ1, POLG*) had higher expression levels in excitatory neurons than in inhibitory neurons (**Figure 2**), whereas the other 23 genes (40.35%) including five PD-causing genes (*ATP13A2, TMEM230, EIF4G1, GIGYF2, FBXO7*) exhibited the opposite features (**Figure 2**). This result highlights the essential role of excitatory neurons in the cortical circuit in PD; more functional validation is needed in future studies.

To further characterize the spatiotemporal expression patterns of the 107 PAGs during brain development from the fetal

FIGURE 2 | Heat map of normalized expression levels of Parkinson's disease-associated genes in 45 types of inhibitory neurons and 24 types of excitatory neurons. Only Parkinson's disease-associated genes showing different expression levels between inhibitory and excitatory neurons are shown; *P*-values were calculated by the Wilcoxon rank sum test.

FIGURE 3 | Heat map of normalized expression levels of two spatiotemporal co-expression modules (M1 and M2) corresponding to 17 developmental stages in the human brain. pcw, post-conceptual weeks; mos, moths; yrs, years. All expression levels were normalized according to the $\log_2$-fold change to the average expression level of each gene.

period to the adult period, we performed WGCNA in brain tissue samples from the BrainSpan database and detected two independent co-expression modules comprising 69 PAGs (**Supplementary Table 1**). The first module (M1) integrated genes showing low expression in the human embryonic phase which gradually increased in expression starting in the prenatal period (16–18 post-conceptual weeks), and reached a highly stable expression level after birth (both in the adolescent and adult period) (**Figure 3**). The M1 contained 47 genes, 10 of which are known PD-causing genes (*LRRK2, SNCA, SYNJ1, TMEM230, FBXO7, PINK1, PLA2G6, DNAJC6, GBA, PRKN*) (**Supplementary Figure 1**). In addition, the second module, which integrated genes highly expressed during the embryonic and early to-middle fetal periods (8–24 post-conceptual weeks), but showed decreased expression starting in the prenatal period and adult period (**Figure 3**), contained 22 genes including two known PD-causing genes (*VPS35, GIGYF2*) (**Supplementary Figure 2**). Since PAGs were preferentially expressed in the postnatal period compared to the prenatal period, this suggests that genes with enriched postnatal expression are more likely to be associated with PD.
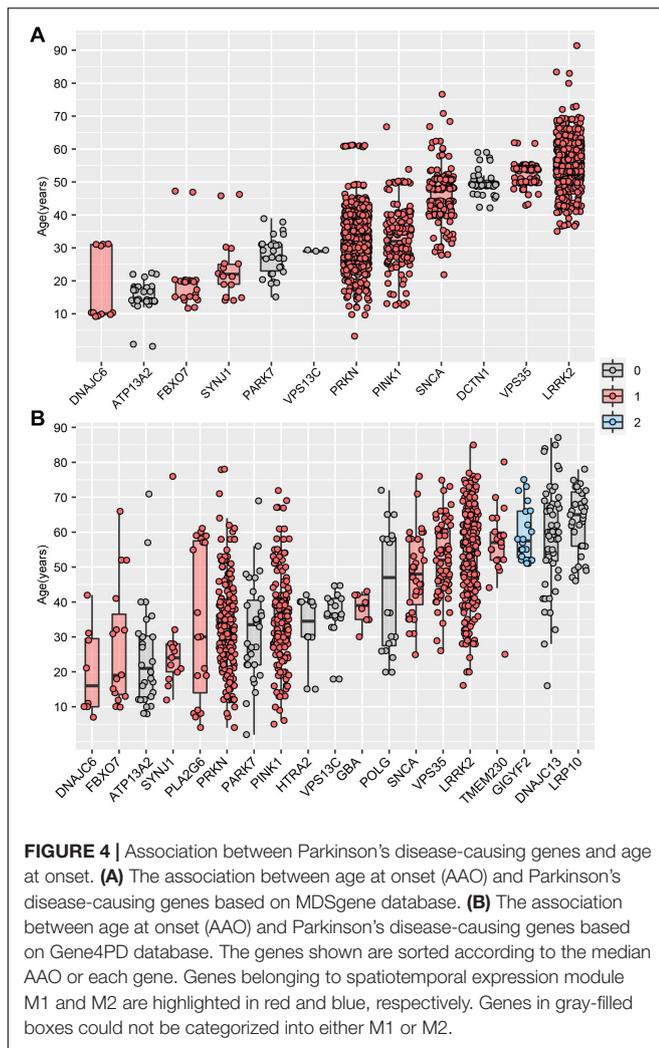
AAO Is Associated With the Expression Patterns

MDSgene database (Lill et al., 2016) (see text footnote 5) integrated AAO data and genetic data of 12 PD-causing genes, providing an unprecedented opportunity to identify the vital

association between the AAO and PD-causing genes on a large scale. By analyzing the PD-causing genes with more than five AAO items, we found that five genes (*DNAJC6, ATP13A2, FBXO7, SYNJ1, PARK7*) were associated with a juvenile-onset, four genes (*PARKN, PINK1, SNCA, DCTN1*) were associated with an early onset, while *VPS35* and *LRRK2* were associated with a late-onset (**Figure 4A**). Interestingly, we found that the PD-causing genes in MDSgene were associated with specific spatiotemporal expression pattern (**Figure 4A**). Specifically, nine of the 12 reported genes in MDSgene belonged to M1, while none of them belonged to M2, and the remaining three genes did not belong to either modules.

We then replicate above results by analyzing the AAO data and genetic data from Gene4PD database[8]. Based on the 21 PD-causing genes reported in the latest review (Blauwendraat et al., 2019b), we analyzed 19 PD-causing genes with more than 5 AAO terns in Gene4PD database. As a result, we found that five of the 19 PD-causing genes were associated with a juvenile-onset, eight genes were associated with an early onset, whereas another six genes were associated with a late-onset (**Figure 4B**). Meanwhile, 11 of the 19 genes belonged to M1, only *GIGYF2* one gene belonged to M2, and seven genes did not belong to either module. Furthermore, the median AAO of patients with rare variants in *GIGYF2* genes, which is belong to M2, is higher than other genes belong to M1 (**Figure 4B**). This results above

---

[8]http://www.genemed.tech/gene4pd

FIGURE 4 | Association between Parkinson's disease-causing genes and age at onset. **(A)** The association between age at onset (AAO) and Parkinson's disease-causing genes based on MDSgene database. **(B)** The association between age at onset (AAO) and Parkinson's disease-causing genes based on Gene4PD database. The genes shown are sorted according to the median AAO or each gene. Genes belonging to spatiotemporal expression module M1 and M2 are highlighted in red and blue, respectively. Genes in gray-filled boxes could not be categorized into either M1 or M2.

provided novel insight the association among AAO, PD-causing genes, and spatiotemporal expression patterns.

DISCUSSION

With the development of biotechnology, numerous candidate genes associated with PD have been reported from thousands of published studies. The ability to comprehensively investigate these PAGs from expression patterns in human, would be helpful for studying the pathogenesis of PD (Blauwendraat et al., 2019b).

Analysis of the spatiotemporal gene expression pattern is a widely accepted method for evaluating cellular function at the transcription level in a specific tissue (Li et al., 2017). In this study, we comprehensively analyzed the expression pattern of PAGs with a novel landscape, and many meaningful results were achieved. First, we demonstrated that PAGs were more likely to have higher expression levels in brain tissues than in other tissues. Second, we found that most PAGs had broad expression profiles in excitatory neurons rather than in inhibitory neurons, suggesting that excitatory neurons are more likely associated

with PD (Zhang et al., 2016). Furthermore, an interesting study, focused on analyzing the specific gene expression pattern in different cell types associated with PD risk and described a single-cell atlas in human substantia nigra and cortex (Agarwal et al., 2020). According to this atlas, we investigated whether the PAGs were enriched in any cell types. For cortex, it showed that the PAGs were enriched in excitatory neurons ($P = 1.18e-05$) and other four cell types, including astrocyte ($P = 7.64-e03$), microglia ($P = 0.048$), oligodendrocytes ($P = 0.01$), oligo-precursor cells ($P = 1.78e-03$). We noted that the PAGs were also enriched in inhibitory neurons ($P = 0.07$), which is close to be significant (**Supplementary Table 2**). These results were consisted with our conclusion in this study. For substantia nigra, we also find that the PAGs enriched in Astrocytes ($P = 4.62e-03$), Dopaminergic neurons ($P = 2.60e-05$), GABAergic neurons ($P = 1.92e-03$), Oligodendrocytes ($P = 2.39e-03$), Oligo-precursor cells ($P = 0.02$), but not Endothelial ($P = 0.15$) in substantia nigra (**Supplementary Table 2**). Notably, the PAGs were most enriched in dopaminergic neurons with the most significant p-value. Third, we identified two spatiotemporal expression patterns (M1 and M2) associated with PD. More PAGs belonged to M1 than to M2, highlighting that the high expression level in the adult period is associated with neurodegeneration. However, M2 genes including two PD-causing genes (VPS35, GIGYF2), were highly expressed during the embryonic and early to-middle fetal period, suggesting that they may associated with neurodevelopment. For example, GIGYF2 was identified to be associated with macrocephaly (Guo H. et al., 2018). This finding underlines the importance of further research on the mechanisms contributing to the correlation between PD and neurodevelopmental disorders (Hu et al., 2017). Notably, NUS1, which belongs to M2, is a novel PD risk gene, which was found to be associated with the early onset of PD in a Han Chinese population in our previous study (Guo J.F. et al., 2018). However, the innovative results in this study were based on the analysis of bioinformatics; thus, the pathogenic mechanisms will need to be verified with cell or animal experiments. Moreover, 47 PAGs including 10 known PAGs in M1, were steadily highly expressed in the adult period (i.e., M1), which is consisted with the feature of adult onset of PD. It can be inferred that other 37 PAGs (**Supplementary Table 1**), co-expressed with these 10 known PAGs, acting on the same signaling pathway. Our study adds to the evidence that these 37 PAGs contribute to PD. Moreover, another spatiotemporal expression patterns (i.e., M2) which highly expressed during the embryonic and early to-middle fetal period, suggest that not all the PAGs are highly expressed in adult periods, and the pathogenesis of these genes is worth further investigation, which may be helpful to the molecular typing of PD. These hint that the genes, enriched in the postmitotic cells and progenitors or developmental cell states, might involve into different unknown signal ways, respectively. We highly encourage further study to perform more comprehensive analysis with more data.

PD is an age-dependent neurodegenerative condition. The onset age of PD has been widely reported to be related to the variation in phenotypes and genotypes (Wickremaratchi et al., 2011; Kilarski et al., 2012; Nalls et al., 2015). PD-causing genes, such as ATP13A2 (Ramirez et al., 2006), PLA2G6

(Paisan-Ruiz et al., 2009), *FBXO7* (Di Fonzo et al., 2009), *PRKN* (Lucking et al., 2000; Djarmati et al., 2004), *PINK1* (Bonifati et al., 2005; Kumazawa et al., 2008), and *PARK7* (Djarmati et al., 2004) have been suggested be strongly associated with AAO. Based on two integrated studies on analyzing the AAO data and genetic data from MDSgene database and Gene4PD database, respectively, we confirmed the association between commonly reported PD-causing genes and AAO. Interestingly, we found that most of the PD-causing genes were belonged to M1, suggesting that the PD-causing genes may be highly expressed after birth. Data from both databases showed very similar trends of each gene, which further examining the relationship between PD-causing genes, AAO and expression pattern.

CONCLUSION

In conclusion, we prioritized 107 functionally related PAGs and characterized the genetic landscape of the PAGs by analyzing the expression patterns, associations between AAO and expression pattern. This provided novel insight among the AAO, PAGs, and spatiotemporal expression patterns, as well as provided novel clues for understanding the pathology of PD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

BL, BT, and JL: study design. BL, GZ, KL, ZW, ZF, XW, TL, YZ, YW, QC, YH, and JL: literature search and data collection. BL, GZ, KL, ZW, BT, and JL: figures. BL, GZ, ZF, XW, TL, YZ, YW, QC, LD, and JL: data analysis. BL, GZ, YW, QC, JG, BT, and JL: data interpretation. BL, GZ, KL, ZW, ZF, XW, TL, YZ, YW, QC, YH, LD, JG, BT, and JL: writing of the manuscript draft and all revision stages. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.629156/full#supplementary-material

REFERENCES

Agarwal, D., Sandor, C., Volpato, V., Caffrey, T. M., Monzon-Sandoval, J., Bowden, R., et al. (2020). A single-cell atlas of the human substantia nigra reveals cell-specific pathways associated with neurological disorders. *Nat. Commun.* 11:4183.

Alzheimer's Association (2014). 2014 Alzheimer's disease facts and figures. *Alzheimers Dement* 10, e47–e92.

Amar, D., Hait, T., Izraeli, S., and Shamir, R. (2015). Integrated analysis of numerous heterogeneous gene expression profiles for detecting robust disease-specific biomarkers and proposing drug targets. *Nucleic Acids Res.* 43, 7779–7789. doi: 10.1093/nar/gkv810

Ascherio, A., and Schwarzschild, M. A. (2016). The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol.* 15, 1257–1272.

Blauwendraat, C., Heilbron, K., Vallerga, C. L., Bandres-Ciga, S., von Coelln, R., Pihlstrom, L., et al. (2019a). Parkinson's disease age at onset genome-wide association study: defining heritability, genetic loci, and alpha-synuclein mechanisms. *Mov. Disord.* 34, 866–875.

Blauwendraat, C., Nalls, M. A., and Singleton, A. B. (2019b). The genetic architecture of Parkinson's disease. *Lancet Neurol.* 19, 170–178.

Bonifati, V., Rohe, C. F., Breedveld, G. J., Fabrizio, E., De Mari, M., Tassorelli, C., et al. (2005). Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes. *Neurology* 65, 87–95. doi: 10.1212/01.wnl.0000167546.39375.82

Chang, D., Nalls, M. A., Hallgrimsdottir, I. B., Hunkapiller, J., van der Brug, M., and Cai, F., et al. (2017). A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* 49, 1511–1516.

Di Fonzo, A., Dekker, M. C., Montagna, P., Baruzzi, A., Yonova, E. H., Correia Guedes, L., et al. (2009). FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. *Neurology* 72, 240–245. doi: 10.1212/01.wnl.0000338144.10967.2b

Djarmati, A., Hedrich, K., Svetel, M., Schafer, N., Juric, V., Vukosavic, S., et al. (2004). Detection of Parkin (PARK2) and DJ1 (PARK7) mutations in early-onset Parkinson disease: Parkin mutation frequency depends on ethnic origin of patients. *Hum. Mutat.* 23:525. doi: 10.1002/humu.9240

Dorsey, E. R., Constantinescu, R., Thompson, J. P., Biglan, K. M., Holloway, R. G., Kieburtz, K., et al. (2007). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology* 68, 384–386. doi: 10.1212/01.wnl.0000247740.47667.03

Findley, L. J. (2007). The economic impact of Parkinson's disease. *Parkinsonism Relat. Disord.* 13(Suppl.), S8–S12.

Fischer, N., Indenbirken, D., Meyer, T., Lutgehetmann, M., Lellek, H., Spohn, M., et al. (2015). Evaluation of unbiased next-generation sequencing of RNA (RNA-seq) as a diagnostic method in influenza virus-positive respiratory samples. *J. Clin. Microbiol.* 53, 2238–2250. doi: 10.1128/jcm.02495-14

Foo, J. N., Chew, E. G. Y., Chung, S. J., Peng, R., Blauwendraat, C., Nalls, M. A., et al. (2020). Identification of risk loci for parkinson disease in asians and comparison of risk between asians and europeans: a genome-wide association study. *JAMA Neurol.* 77, 746–754. doi: 10.1001/jamaneurol.2020.0428

Gire, S. K., Goba, A., Andersen, K. G., Sealfon, R. S., Park, D. J., Kanneh, L., et al. (2014). Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* 345, 1369–1372.

Grenn, F. P., Kim, J. J., Makarious, M. B., Iwaki, H., Illarionova, A., Brolin, K., et al. (2020). The Parkinson's disease genome-wide association study locus browser. *Mov. Disord.* 35, 2056–2067.

GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)— Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NHGRI, et al. (2017). Genetic effects on gene expression across human tissues. *Nature* 550, 204–213. doi: 10.1038/nature24277

Guo, H., Wang, T., Wu, H., Long, M., Coe, B. P., Li, H., et al. (2018). Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol. Autism.* 9:64.

Guo, J. F., Xiao, B., Liao, B., Zhang, X. W., Nie, L. L., Zhang, Y. H., et al. (2008). Mutation analysis of Parkin, PINK1, DJ-1 and ATP13A2 genes in Chinese patients with autosomal recessive early-onset Parkinsonism. *Mov. Disord.* 23, 2074–2079. doi: 10.1002/mds.22156

Guo, J. F., Zhang, L., Li, K., Mei, J. P., Xue, J., Chen, J., et al. (2018). Coding mutations in NUS1 contribute to Parkinson's disease. *Proc. Natl. Acad. Sci. U.S.A.* 115, 11567–11572.

Hu, Y., Pan, Z., Hu, Y., Zhang, L., and Wang, J. (2017). Network and pathway-based analyses of genes associated with Parkinson's disease. *Mol. Neurobiol.* 54, 4452–4465. doi: 10.1007/s12035-016-9998-8

International Parkinson Disease Genomics Consortium, Nalls, M. A., Plagnol, V., Hernandez, D. G., Sharma, M., Sheerin, U. M., et al. (2011). Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 377, 641–649. doi: 10.1016/s0140-6736(10)62345-8

Iwaki, H., Blauwendraat, C., Leonard, H. L., Kim, J. J., Liu, G., Maple-Grodem, J., et al. (2019). Genomewide association study of Parkinson's disease clinical biomarkers in 12 longitudinal patients' cohorts. *Mov. Disord.* 34, 1839–1850.

Jansen, I. E., Ye, H., Heetveld, S., Lechler, M. C., Michels, H., Seinstra, R. I., et al. (2017). Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing. *Genome Biol.* 18:22.

Judge, M., Parker, E., Naniche, D., and Le Souef, P. (2020). Gene expression: the key to understanding HIV-1 infection? *Microbiol. Mol. Biol. Rev.* 2:84.

Kalia, L. V., and Lang, A. E. (2015). Parkinson's disease. *Lancet* 386, 896–912.

Kang, J. F., Luo, Y., Tang, B. S., Wan, C. M., Yang, Y., Li, K., et al. (2016). RAB39B gene mutations are not linked to familial Parkinson's disease in China. *Sci. Rep.* 6:34502.

Kilarski, L. L., Pearson, J. P., Newsway, V., Majounie, E., Knipe, M. D., Misbahuddin, A., et al. (2012). Systematic review and UK-based study of PARK2 (parkin), PINK1, PARK7 (DJ-1) and LRRK2 in early-onset Parkinson's disease. *Mov. Disord.* 27, 1522–1529. doi: 10.1002/mds.25132

Kumazawa, R., Tomiyama, H., Li, Y., Imamichi, Y., Funayama, M., Yoshino, H., et al. (2008). Mutation analysis of the PINK1 gene in 391 patients with Parkinson disease. *Arch. Neurol.* 65, 802–808.

Labbe, C., Lorenzo-Betancor, O., and Ross, O. A. (2016). Epigenetic regulation in Parkinson's disease. *Acta Neuropathol.* 132, 515–530.

Li, G., Ma, J., Cui, S., He, Y., Xiao, Q., Liu, J., et al. (2019). Parkinson's disease in China: a forty-year growing track of bedside work. *Transl. Neurodegener.* 8:22.

Li, J., Wang, L., Guo, H., Shi, L., Zhang, K., Tang, M., et al. (2017). Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. *Mol. Psychiatry* 22, 1282–1290. doi: 10.1038/mp.2017.140

Lill, C. M., Mashychev, A., Hartmann, C., Lohmann, K., Marras, C., Lang, A. E., et al. (2016). Launching the movement disorders society genetic mutation database (MDSGene). *Mov. Disord.* 31, 607–609. doi: 10.1002/mds.26651

Lim, S. Y., Tan, A. H., Ahmad-Annuar, A., Klein, C., Tan, L. C. S., Rosales, R. L., et al. (2019). Parkinson's disease in the western pacific region. *Lancet Neurol.* 18, 865–879.

Liu, Z., Guo, J., Li, K., Qin, L., Kang, J., Shu, L., et al. (2015). Mutation analysis of CHCHD2 gene in Chinese familial Parkinson's disease. *Neurobiol. Aging* 36, e3117–3117.

Lucking, C. B., Durr, A., Bonifati, V., Vaughan, J., De Michele, G., Gasser, T., et al. (2000). Association between early-onset Parkinson's disease and mutations in the parkin gene. *N. Engl. J. Med.* 342, 1560–1567.

Nalls, M. A., Blauwendraat, C., Vallerga, C. L., Heilbron, K., Bandres-Ciga, S., Chang, D., et al. (2019). Identification of novel risk loci, causal insights,

and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* 18, 1091–1102.

Nalls, M. A., Escott-Price, V., Williams, N. M., Lubbe, S., Keller, M. F., Morris, H. R., et al. (2015). International Parkinson's Disease Genomics C: genetic risk and age in Parkinson's disease: continuum not stratum. *Mov. Disord.* 30, 850–854. doi: 10.1002/mds.26192

Nalls, M. A., Pankratz, N., Lill, C. M., Do, C. B., Hernandez, D. G., Saad, M., et al. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* 46, 989–993.

Paisan-Ruiz, C., Bhatia, K. P., Li, A., Hernandez, D., Davis, M., Wood, N. W., et al. (2009). Characterization of PLA2G6 as a locus for dystonia-parkinsonism. *Ann. Neurol.* 65, 19–23. doi: 10.1002/ana.21415

Ramirez, A., Heimbach, A., Grundemann, J., Stiller, B., Hampshire, D., Cid, L. P., et al. (2006). Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat. Genet.* 38, 1184–1191. doi: 10.1038/ng1884

Roychowdhury, S., and Chinnaiyan, A. M. (2016). Translating cancer genomes and transcriptomes for precision oncology. *CA Cancer J. Clin.* 66, 75–88. doi: 10.3322/caac.21329

Sandor, C., Honti, F., Haerty, W., Szewczyk-Krolikowski, K., Tomlinson, P., Evetts, S., et al. (2017). Whole-exome sequencing of 228 patients with sporadic Parkinson's disease. *Sci. Rep.* 7:41188.

Shi, C. H., Tang, B. S., Wang, L., Lv, Z. Y., Wang, J., Luo, L. Z., et al. (2011). PLA2G6 gene mutation in autosomal recessive early-onset parkinsonism in a Chinese cohort. *Neurology* 77, 75–81. doi: 10.1212/wnl.0b013e318221acd3

Siitonen, A., Nalls, M. A., Hernandez, D., Gibbs, J. R., Ding, J., Ylikotila, P., et al. (2017). Genetics of early-onset Parkinson's disease in Finland: exome sequencing and genome-wide association study. *Neurobiol. Aging* 53, 195.e197–95.e110.

Tang, B., Xiong, H., Sun, P., Zhang, Y., Wang, D., Hu, Z., et al. (2006). Association of PINK1 and DJ-1 confers digenic inheritance of early-onset Parkinson's disease. *Hum. Mol. Genet.* 15, 1816–1825. doi: 10.1093/hmg/ddl104

Tian, J. Y., Guo, J. F., Wang, L., Sun, Q. Y., Yao, L. Y., Luo, L. Z., et al. (2012). Mutation analysis of LRRK2, SCNA, UCHL1, HtrA2 and GIGYF2 genes in Chinese patients with autosomal dorminant Parkinson's disease. *Neurosci. Lett.* 516, 207–211. doi: 10.1016/j.neulet.2012.03.086

Wickremaratchi, M. M., Knipe, M. D., Sastry, B. S., Morgan, E., Jones, A., Salmon, R., et al. (2011). The motor phenotype of Parkinson's disease in relation to age at onset. *Mov. Disord.* 26, 457–463. doi: 10.1002/mds.23469

Xu, Q., Li, K., Sun, Q., Ding, D., Zhao, Y., Yang, N., et al. (2017). Rare GCH1 heterozygous variants contributing to Parkinson's disease. *Brain* 140:e41. doi: 10.1093/brain/awx110

Yan, W., Tang, B., Zhou, X., Lei, L., Li, K., Sun, Q., et al. (2017). TMEM230 mutation analysis in Parkinson's disease in a Chinese population. *Neurobiol. Aging* 49, 219.e211–219.e213.

Yang, N., Zhao, Y., Liu, Z., Zhang, R., He, Y., Zhou, Y., et al. (2019). Systematically analyzing rare variants of autosomal-dominant genes for sporadic Parkinson's disease in a Chinese cohort. *Neurobiol. Aging* 76, 215.e211–215.e217.

Zhang, Y., Tan, F., Xu, P., and Qu, S. (2016). Recent advance in the relationship between excitatory amino acid transporters and Parkinson's disease. *Neural. Plast.* 2016:8941327.

Zhao, Y., Qin, L., Pan, H., Liu, Z., Jiang, L., He, Y., et al. (2020). The role of genetics in Parkinson's disease: a large cohort study in Chinese mainland population. *Brain* 143, 2220–2234.