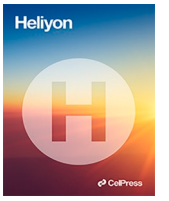
Heliyon 10 (2024) e40618

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Epigenomic and clinical analyses of striatal DAT binding in healthy individuals reveal well-known loci of Parkinson's disease

Arash Yaghoobi [a], Homa Seyedmirzaei [b,c], Marzie Jamaat [d], Moein Ala [e,*]

[a] School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, 19395-5746, Iran
[b] Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences, Tehran, Iran
[c] Students' Scientific Research Center, Tehran University of Medical Sciences, Tehran, Iran
[d] Islamic Azad University, Tehran North Branch, Faculty of Biological Sciences, Tehran, Iran
[e] Experimental Research Center, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
DAT scan
Parkinson's disease
Epigenome-wide association study
Cognitive domains
Healthy individuals

ABSTRACT

Background: Striatal dopamine transporter (DAT) binding is a sensitive and specific endophenotype for detecting dopaminergic deficits across Parkinson's disease (PD) spectrum. Molecular and clinical signatures of PD in asymptomatic phases help understand the earliest pathophysiological mechanisms underlying the disease. We aimed to investigate whether blood epigenetic markers are associated with inter-individual variation of striatal DAT binding among healthy elderly individuals. We also investigated whether this potential inter-individual variation can manifest as dysfunction of particular cognitive domains. Omics studies conducted on endophenotypes of PD among healthy asymptomatic individuals can provide invaluable insights into early detection, disease mechanisms, and potential therapeutic targets for PD.

Method: We conducted a blood epigenome-wide association study of striatal DAT binding on 96 healthy individuals using the Illumina EPIC array. For functional annotation of our top results, we employed the enhancer-gene mapping strategy using a midbrain single-nucleus multimodal dataset. Finally, we conducted several investigative regression analyses on several neuropsychological tests across five cognitive domains to assess their association with striatal DAT binding among 250 healthy subjects.

Results: We identified seven suggestive (P-value$<10^{-5}$) CpG probes. Specifically, three probes were colocalized with three risk loci previously identified in PD's largest Genome-Wide Association Study (GWAS). *UCN5A* and *APOE* loci were identified as suggestive DMRs associated with striatal DAT binding. Functional analyses prioritized the *FDFT1* gene as the potential target gene in the previously reported *CTSB* GWAS locus. We also showed that delayed recall memory impairment was correlated with reduced striatal DAT binding, irrespective of age.

Conclusion: Our study suggested epigenetic and cognitive signatures of striatal DAT binding among healthy individuals, providing valuable insights for future experimental and clinical studies of early PD.

* Corresponding author. School of Medicine, Tehran University of Medical Sciences (TUMS), Tehran, Iran.
E-mail address: moinala75@yahoo.com (M. Ala).

https://doi.org/10.1016/j.heliyon.2024.e40618
Received 16 May 2024; Received in revised form 20 November 2024; Accepted 20 November 2024
Available online 22 November 2024

1. Introduction

Dopamine transporter scan (DaTscan) is a nuclear imaging modality used to detect nigrostriatal dopaminergic deficits in Parkinson's disease (PD) and other Parkinsonian syndromes [1]. PD is the second most common neurodegenerative disorder, affecting approximately 1 % of people aged more than 60 years worldwide [2,3]. It is estimated that the first motor symptoms of PD do not appear until approximately 35%–45 % of striatal DAT activity is lost [4]. Therefore, PD is now seen as a continuum from asymptomatic individuals who are at high risk for PD to individuals with prodromal symptoms, all the way to PD patients. Considering this fundamental nature of PD, sensitive and specific biomarkers are valuable endophenotypes for quantifying the early stages of the disease [5]. Due to the high sensitivity (84.4 %) and specificity (96.2 %) of DaTscan [6], this imaging tool can serve as a robust endophenotype to quantify the amount of dopaminergic neuron loss in asymptomatic and prodromal phases of PD.

PD originates from pathological changes in various cells in the brain. These changes may be detected at the epigenomic level in the early stages of the disease [7]. One of the epigenetic mechanisms is DNA methylation, which controls gene expression. To date, few studies have proposed several genes whose methylations were altered in blood or brain tissues, such as *SNCA*, *DRD2*, *NPAS2*, *CRY1*, *MAPT*, and *PDE4D* [7]. However, these studies mostly compared the epigenomic profiles of patients with late-stage PD with healthy controls, which impedes our understanding of the biological pathways involved in early PD.

In addition to molecular changes in the brain, prodromal PD can manifest with clinical presentations, such as changes in cognition. It is proposed that approximately 40 % of patients with early-stage PD suffer from mild cognitive impairment. A large longitudinal study on healthy subjects of the Rotterdam cohort has shown that cognitive dysfunction can increase the risk of future Parkinsonism among healthy individuals, suggesting cognitive dysfunction as an early sign of PD [8]. Among numerous pathophysiological mechanisms proposed for cognitive dysfunction in PD, dopamine deficiency is one of the most known hypotheses [9]. Deficiency in nigrostriatal dopamine was associated with executive dysfunction among patients with early-stage drug naïve PD [10]. Studies on healthy people have also proposed correlations between striatal DAT binding and cognitive dysfunction, including verbal function and executive/working memory impairments [11,12]. However, studies on association between various domains of cognitive dysfunction and striatal DAT binding among healthy individuals are scarce and heterogeneous and remain to be elucidated.

Given the importance of understanding early biological alterations in PD, we hypothesized that inter-individual variation of striatal DAT binding among healthy elderly individuals could be a robust endophenotype for epigenomic studies. Therefore, we conducted an epigenome-wide association study (EWAS) and investigated the association between the whole blood DNA methylome profile and the inter-individual variation of striatal DAT binding measured by DaTscan among healthy individuals of the PPMI cohort. We also used functional omics approaches to prioritize the potential target genes in the identified loci. Given the importance of identifying the earliest clinical manifestations associated with the risk of PD, we also hypothesized that inter-individual variation of striatal DAT binding could manifest as mild cognitive impairment among healthy elderly individuals across specific domains. Thus, we explored the correlation of striatal DAT binding with several neuropsychological tests across five cognitive domains among healthy individuals. Our findings can bring a novel insight into the molecular and clinical aspects of early PD.

2. Methods

2.1. Study design and population

We used the PPMI cohort for our study. PPMI is a longitudinal, observational, and multi-center study that has collected various clinical, neuroimaging, and biological data from individuals across the PD spectrum. The consent forms have been obtained from all the participants. The details of this cohort can be found on its website (https://www.ppmi-info.org/).

Based on our research purpose, we included healthy individuals with no parkinsonism syndromes whose baseline DaTscan data were available. Then, we filtered subjects whose baseline cognitive tests or blood methylome data were available for subsequent analyses. This study was exempted from IRB approval as data from a pre-existing database was used.

2.2. DaTscan

DaTscan was performed in PPMI imaging centers. The imaging protocol of the cohort is available at http://www.ppmi-info.org/study-design/research-documents-and-sops/. All imaging preprocessing steps were performed according to the PPMI protocol. Count densities were extracted from the caudate nucleus and putamen. The occipital cortex was also considered the reference region. The regional SBR (striatal specific binding ratio) was calculated as follows: (SBR)= (striatal region)/(occipital)−1 [13]. We focused only on mean, right, and left striatal DAT bindings. The mean striatal SBR was calculated by the following formula: Mean striatal SBR= (left caudate SBR + right caudate SBR + left putamen SBR + right putamen SBR)/4.

2.3. Whole blood DNA methylation

Whole genome methylation of whole blood samples of the PPMI cohort was performed using the Illumina human methylation EPIC array at baseline visits. The methylation EPIC array covers >850,000 methylation sites across the whole genome (17 CpG sites per gene on average) at the single nucleotide resolution. We first downloaded the final normalized dataset of the PPMI cohort and separated the healthy individuals. Some quality control steps were performed using the Bioconductor package minfi to produce this dataset: Samples with a mean detection P-value >0.01 were excluded. Samples discordant between clinically reported sex and methylation-predicted

sex were also excluded. Specific sentrix arrays were excluded for subjects who were processed more than once. Probes with a detection P value > 0.01 in more than 20 % of the samples were identified. Blood cell composition was determined using the minfi "estimateCellCounts" function, which assesses the relative proportion of $CD8^+$ and $CD4^+$ T cells, natural killer cells, B cells, monocytes, and granulocytes using the Houseman algorithm. Multidimensional scaling was also performed. Finally, functional normalization was performed on data using "preprocessFunnorm," followed by removing 2769 failed probes. More details about the quality control steps can be found on the PPMI website. We conducted some additional quality control steps using the OSCA software (https://yanglab.westlake.edu.cn/software/osca/): We excluded CpG probes located on sex chromosomes and probes whose genomic positions were not available based on the "EPICv2.hg38.manifest" file provided by the Illumina company. Principal component analysis (PCA) was performed to detect outlier samples. We excluded samples whose first and second principal components were more than ± 3 standard deviation from the means. Compared to a similar previous study [14], we adopted a more stringent quality control strategy: we excluded probes with a standard deviation smaller than 0.025, a missing ratio of more than 0.01, and a mean beta value of more than 0.85 or less than 0.15.

2.4. EWAS analysis to identify differentially methylated positions (DMPs)

For EWAS analysis, we first adjusted the beta values of CpG probes for age, sex, blood cell composition, and BMI for each sample. The smoking data for healthy individuals were only available for approximately half of them in the PPMI dataset, who were non-smokers. Therefore, we did not include this covariate in our analysis. EWAS analysis was performed using the mixed linear model regression option ("moa") of the OSCA software (to correct for hidden confounding factors) with mean, right, and left striatal DAT binding as the dependent variables and all methylation probes as independent variables. We used age, gender, and handedness as potential covariates affecting striatal DAT binding. In our EWAS, based on the number of probes tested (379,158), the stringent P-value threshold was 1.31×10^{-7} (0.05/379,158). However, due to the modest sample size included in our study and the correlation between methylation probes, like many other EWASs with small sample sizes, we reported 'suggestively significant' findings with a p-value threshold of 1×10^{-5} [15–22]. Importantly, this approach provides a better overview of top findings in hypothesis-generating studies in which our primary goal is to provide hypotheses and insights into the biological implications of an EWAS. Since EWAS studies are prone to high inflation rates and bias, we first computed the inflation factor (λ) using the conventional approach. In the next step, we used the BACON package [23], specifically developed to control bias and inflation in EWAS studies, to obtain more accurate P-values. We finally reported the bias- and inflation-corrected test statistics computed by this package.

2.5. Differentially methylated region (DMR) analysis

We used the dmrff package [24] for DMR analysis. We used the default of the package. Briefly, this package identifies regions composed of methylation sites with EWAS P-values of less than 0.05 at a maximum 500 base-pair gap between CpG sites. For DMR analysis, based on the number of regions tested (19,345), our study's stringent threshold of P-value was 2.5×10^{-6} (0.05/19,345). However, we reported 'suggestively significant' findings with a p-value threshold of 1×10^{-5} based on a similar previous study [20]. Additionally, we reported only regions containing at least three probes.

2.6. Positional and functional mapping of DMPs

We first performed positional mapping of our significant DMPs using the "EPICv2.hg38.manifest" file provided by the Illumina website. We also searched for the overlap of our top DMPs with 90 PD GWAS loci previously reported in largest PD GWAS conducted by Nalls et al. [25]. We then searched for the overlap of significant DMPs with enhancers across B-cells, mononuclear phagocytes, and T-cells using the enhancer-gene map generated from 131 human cell types and tissues by the Activity By Contact (ABC) model [26]. To investigate the potential target genes of enhancer regions in the PD-related tissues, we used a single-nucleus multimodal dataset (snRNA-seq + snATAC-seq) obtained from the midbrain of 9 post-mortem healthy elderly individuals [27]. We first searched for the overlap of our EWAS DMPs identified in the pervious step with the chromatin-accessible regions. Due to the sparsity nature of single-cell data, we then used the Spearman correlation test to investigate the potential correlation between the expression of nearby genes at the EWAS top DMPs (± 200 kb) with the chromatin accessibilities across all cells of each sample. We adopted a correlation coefficient of more than 0.30 as a significant correlation.

2.7. Cognitive phenotypes

We extracted 12 test scores across five cognitive domains from healthy individuals of the PPMI cohort. These 12 clinical phenotypes were categorized as follows: Executive/working memory domain (letter number sequence test, semantic fluency animal sub score test, and trail making test B), language/executive domain (lexical fluency test and modified Boston naming test), attention/processing speed domain (symbol digit modalities test and trail-making test A), visuospatial processing domain (Benton judgment of line orientation test and clock drawing test), and memory domain (Hopkins verbal learning (HVL) immediate recall test, HVL delayed recall test, and HVL discrimination recognition test). Since data from healthy individuals of the PPMI cohort have been collected in various phases, the results of 5 tests (lexical fluency test, trail-making test A and B, modified Boston naming test, and clock drawing test) were not assessed among the healthy individuals of the original cohort which resulted in smaller sample sizes relative to other tests. Then, we transformed the scores of 12 cognitive tests into Z-scores (mean 0 and standard deviation = 1) by dividing the difference of scores

from the mean to the standard deviation of scores.

2.8. Investigating association of cognitive phenotypes with striatal DAT binding

To investigate subtle impairments of which cognitive domain may reflect striatal DAT binding variation among healthy individuals, we used multiple separate linear regression models for each of the 12 neuropsychological tests, with the Z-score of neuropsychological tests as the dependent variables and the mean, left, and right striatal DAT binding as the independent variables. The results were adjusted for age, gender, and education. Given the exploratory nature of the analysis, we deemed a P-value of less than 0.05 statistically significant. All the analyses were performed by the R version 4.3.

3. Results

3.1. EWAS results

The whole blood DNA methylome data of 865,859 CpG probes among 420 participants at their baseline visit were available in the PPMI cohort dataset. Among these subjects, 99 were healthy individuals. We included 96 healthy individuals whose DaTscan and BMI data were available. We then performed additional quality controls on these individuals. First, we excluded 19,663 probes due to their location on the sex chromosomes or missing genomic positions. No outlier was identified based on the top 2 principal components of the methylation profile. Then, 369,632 probes with standard deviations of less than 0.025 were excluded. 97,406 probes with mean beta values of less than 0.15 or more than 0.85 were also excluded. The missing ratio of no probes was more than 0.01. Finally, we adjusted the beta values of CpG probes for age, sex, blood cell compositions, and BMI for each sample. The mean age of the subjects was 61.75 (±11.04). Twenty-five subjects (26.04 %) were female, and 71 were male. The mean of mean, right, and left striatal DAT bindings were 2.48 (±0.51), 4.94 (±1.03), and 4.98 (±1.05), respectively. Finally, we performed EWAS using 379,158 CpG probes from 96 healthy individuals using age, gender, and handedness as covariates. We did not observe significant deflation or inflation in our three analyses (λ_{bacon} = 1.01, 1.04, and 1.05). No CpG probes reached our stringent P-value threshold. We identified seven suggestive DMPs associated with striatal DAT binding with P-values of less than 1×10^{-5}. The complete summary statistics of top seven DMPs are shown in Table 1. The Manhattan and Q-Q plots of our EWAS are depicted in Fig. 1(a and b). Among these seven DMPs, three (cg18316254, cg08118333, and cg14458991) were located in close vicinity of three previously reported GWAS hits in the largest PD GWAS to date. The genomic locations of two previously reported PD GWAS hits and our EWAS results are also shown in Fig. 2(a and b).

3.2. DMR analysis results

No DMRs reached the significant and suggestive threshold. Similar to a previous study, using the dmrff package, which reported top DMRs with a P-value less than 10^{-3} [28], we only report our two top regions with a more stringent P-value threshold of less than 10^{-4} to decrease the probability of false positives yet providing biological clues: In all three analyses, hypermethylation of two regions were suggestively (1×10^{-4}) associated with lower striatal DAT binding: A region on 5q35.2 in UNC5A gene (min P-value = 9.98×10^{-6} associated with right striatal DAT binding) containing three CpG probes and a region on 19q13.32 in CEACAM20 gene (APOE locus) (min P-value = 1.78×10^{-5} associated with mean striatal DAT binding) containing three CpG probes.

3.3. Functional annotation of DMPs

Among our top DMPs, only one DMP (cg14458991) was located in enhancer regions predicted by the ABC method across blood cells. Three target genes with the highest ABC score in B-cells were identified for this region: The *FDFT1*, *CTSB*, and *NEIL2* genes. The genomic locations of the enhancers for this DMP and the list of genes in the ABC model are shown in Supplementary Fig. 1(a and b). We then found that this DMP overlapped with chromatin-accessible regions in the single-nucleus ATAC-seq of the midbrain regions of 9

Table 1
EWAS Summary statistics of top 7 DMPs associated with striatal DAT binding.

CpG probe	Chr	Gene annotation	Previously reported PD GWAS loci	Effect size	SE	P-value
Mean striatal DAT binding						
cg18316254	12	*GALNT9-FBRSL1*	*	−0.39	0.08	8.53×10^{-6}
cg06311934	20	Promoter of *KCNQ2*		−0.36	0.07	5.95×10^{-6}
Right striatal DAT binding						
cg08118333	7	*KCTD7-RABGEF1*	*	0.58	0.13	9.52×10^{-6}
cg03367356	12	*C12orf42*		−0.84	0.18	7.28×10^{-6}
cg09697874	19	*MARK4*		−0.60	0.13	7.66×10^{-6}
cg14458991	8	*FDFT1*	*	0.83	0.18	6.27×10^{-6}
cg06311934	20	Promoter of *KCNQ2*		−0.72	0.15	6.74×10^{-6}
Left striatal DAT binding						
cg20494062	2	*OTOS-GPC1*		−0.68	0.14	1.93×10^{-6}

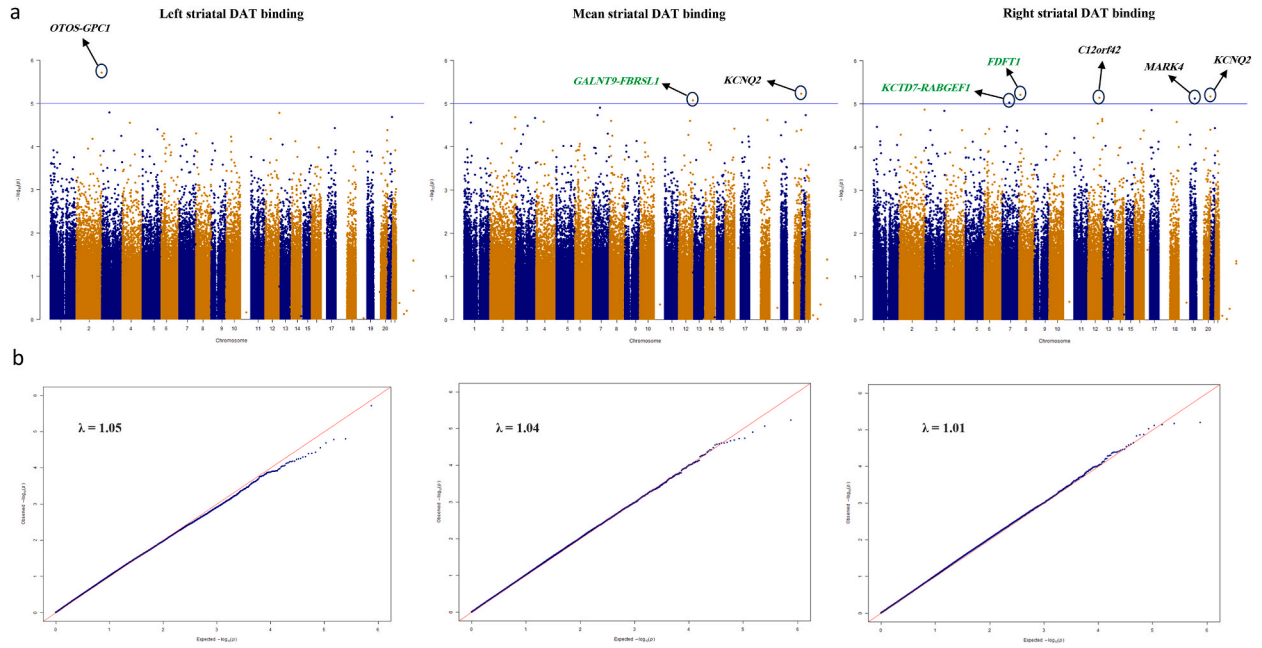


Fig. 1. The figures depict the Manhattan and Q-Q plots of mean, right, and left striatal DAT binding EWAS among 96 healthy individuals. Each figure has two panels:
a) This panel shows the Manhattan plot with the positional gene mapping of the top DMPs. The blue line shows the suggestive threshold (P-value$<10^{-5}$) of our EWAS. Loci previously reported as top PD GWAS hits are highlighted in green.
b) This panel illustrates the Q-Q plot of the P-values of our EWAS results with the bias- and inflation-corrected lambda of our analysis computed by the BACON package.

post-mortem healthy elderly individuals. Multiple correlational analyses between chromatin accessibility and gene expression of four nearby genes in this locus (*FDFT1*, *CTSB*, *NEIL2*, and *GATA4*) prioritized the *FDFT1* gene as the potential target gene of the enhancer. The *FDFT1* gene was the top gene across all 9 samples, and its expression was significantly correlated with the ATAC peak region (max correlation coefficient = 0.32). The genomic position of this DMP in the context of the GWAS hit in this locus, EWAS result of this DMP, and the complete results of our functional analyses are shown in Fig. 3(a–c).

3.4. Results of cognitive function analysis

The results of seven neuropsychological tests from 252 healthy individuals with DaTscan data were extracted. We also extracted the results of five neuropsychological tests from 62 healthy individuals from the other phase of the cohort whose DaTscan data were also available. The association of mean, right, and left striatal DAT binding with these neuropsychological tests were separately investigated. A few missing data were found in a few tests and excluded from the analysis. The mean age of 252 subjects was 61.53 (±11.64). Ninety-six (38.09 %) were female, and 156 were male. The mean of mean, right, and left striatal DAT binding were 2.62 (±0.55), 5.22 (±1.10), and 5.26 (±1.12), respectively. The mean of educational years was 16.11 (±3.08). Among the 62 healthy individuals, the mean age was 64.22 (±12.39). Thirty-four were male (54.83 %), and 28 were female. The mean of the mean, right, and left striatal DTA bindings were 2.77 (±0.49), 5.50 (±1.01), and 5.59 (±1.01), respectively. The mean of the educational years was 16.37 (±3.59). The number of subjects for each analysis is shown in Table 1. Only two of 12 neuropsychological tests were significantly correlated with mean and right striatal DAT bindings: HVLT delayed recall test (spearman correlation coefficient = 0.22, P-value = 4×10^{-4}) and HVLT discrimination recognition test (spearman correlation coefficient = 0.20, P-value = 10^{-3}). Among these two tests, only the HVLT delayed recall test remained statistically significant after adjusting for potential confounders, including age, gender, and education (P-value = 0.04). The complete results of various cognitive tests associations with the mean striatal DAT binding are shown in Table 2.

4. Discussion

To the best of our knowledge, this is the first study investigating the epigenetic signatures of striatal DAT binding among healthy individuals. Using whole blood methylome data from healthy individuals, we found that the altered methylations of seven loci, including three previously reported PD GWAS loci, were significantly associated with lower striatal DAT binding. Using a single-nucleus multimodal dataset obtained from the midbrain of post-mortem healthy elderly individuals, we prioritized the *FDFT1* gene as the candidate gene in the previously reported *CTSB* GWAS locus. Finally, multiple investigative linear regression analyses unveiled that reduced striatal DAT binding was markedly correlated with subtle impairments in delayed memory function among healthy

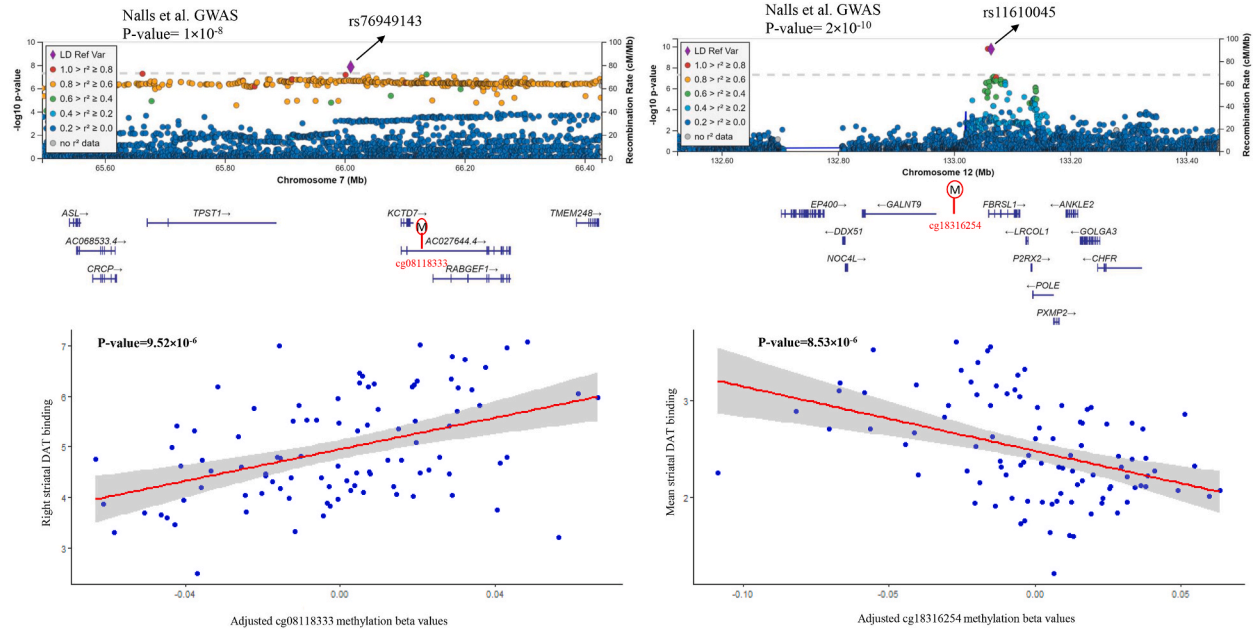


Fig. 2. The figure shows two top significant methylation probes of our EWAS on striatal DAT binding and their positional position and vicinity to two previously reported GWAS hits of the largest PD GWAS to date:
a) This panel zooms into one of the top GWAS hits of the Nalls et al. study in the 7q11.21 locus and its linkage disequilibrium (LD) structure. The top SNP rs76949143 in this locus lies between the *TPST1* and *KCTD7* genes. It also shows the location of our top EWAS result (cg08118333 methylation probe) between the *KCTD7* and *RABGEF1* genes. The lower part shows the direct linear correlation of cg08118333 methylation with right striatal DAT binding among 96 healthy individuals obtained from our EWAS analysis.
b) This panel zooms into one of the other top GWAS hits of the Nalls et al. study in the 12q24.33 locus and its LD structure. The top SNP rs11610045 in this locus is located between the *GALNT9* and *FBRSL1* genes. It also shows the location of our top EWAS result (cg18316254 methylation probe) between the *GALNT9* and *FBRSL1* genes. The lower part shows the inverse linear correlation of cg18316254 methylation with mean striatal DAT binding among 96 healthy individuals obtained from our EWAS analysis.

individuals.

We found several DMPs lying at loci potentially related to PD pathology. For instance, we found that hypermethylation at the promoter of the Potassium Voltage-Gated Channel Subfamily Q member 2 (*KCNQ2*) gene correlated with reduced striatal DAT binding among healthy individuals. Growing evidence has shown that potassium channels regulate the firing of dopaminergic neurons in the striatum. The expression of two members of the *KCNQ* family (*KCNQ2* and *KCNQ4*) are mostly confined to the substantia nigra and ventral tegmental area of the midbrain based on immunohistochemistry studies [29]. Interestingly, modulation of the *KCNQ2* has been proposed as a potential therapeutic target for PD [30]. Regarding the cg20494062 located between the *OTOS* and *GPC1* genes, a previous study has shown that the rs4417745 between *OTOS* and *GPC1* genes was associated with frontotemporal dementia disease [31]. However, due to the intergenic nature of most of these DMPs, future functional studies are needed to unravel the exact target genes in these loci.

Interestingly, we found that among our top seven suggestive DMPs, three were located in close vicinity of three GWAS hits previously reported in the largest PD GWAS conducted by Nalls et al. [25]. For instance, regarding the 08118333 probe, both *KCTD7* and *RABGEF1* genes can be involved in PD. By regulating K^+ conductance through K^+ channels, the *KCTD7* gene is involved in neuronal stimulation and activation [32]. Using mammalian cultured cells, Yamano and colleagues have shown that the RABGEF1 protein directly interacts with Parkin, one of the most famous mutated genes implicated in familial PD, in mitophagy-related pathways [33]. Our study could replicate these two known PD GWAS loci at the methylome level, this time in association with striatal DAT binding among healthy individuals. These findings can shed light on the biological pathways involved in early PD.

We found two suggestive DMRs associated with all striatal DAT binding measures. The first one was located on 19q13.31 in the *CEACAM20* gene in the *APOE* region, which was reported previously in Alzheimer's disease (AD) GWAS [34]. The other region was located on 5q35.2 in the *UNC5A* gene. The UNC5A gene encodes one of the netrin receptors and is involved in axonal guidance and nervous system development. Apoptotic netrin UNC5 signaling has been implicated in AD and PD pathology [35].

Another DMP identified in our EWAS, cg14458991, is located next to the previously reported *CTSB* gene in the largest PD GWAS. Several interesting genes exist in this region, such as the *FDFT1*, *CTSB*, *NEIL2,* and *GATA4*. A recent study has comprehensively examined this region and shown experimentally that the *CTSB* gene is the main driver of PD in this locus by inhibiting the formation of alpha-synuclein fibrils in the dopaminergic neurons [36]. Expression quantitative trait loci (eQTL) studies have shown that the GWAS hit (rs1293298) affects the expression of the *FDFT1*, *CTSB*, and *NEIL2* genes in blood tissue (https://fivex.sph.umich.edu/). Using the ABC model, we found that across B-cells, the *FDFT1*, *CTSB*, and *NEIL2* were the potential target genes of this potential enhancer region.

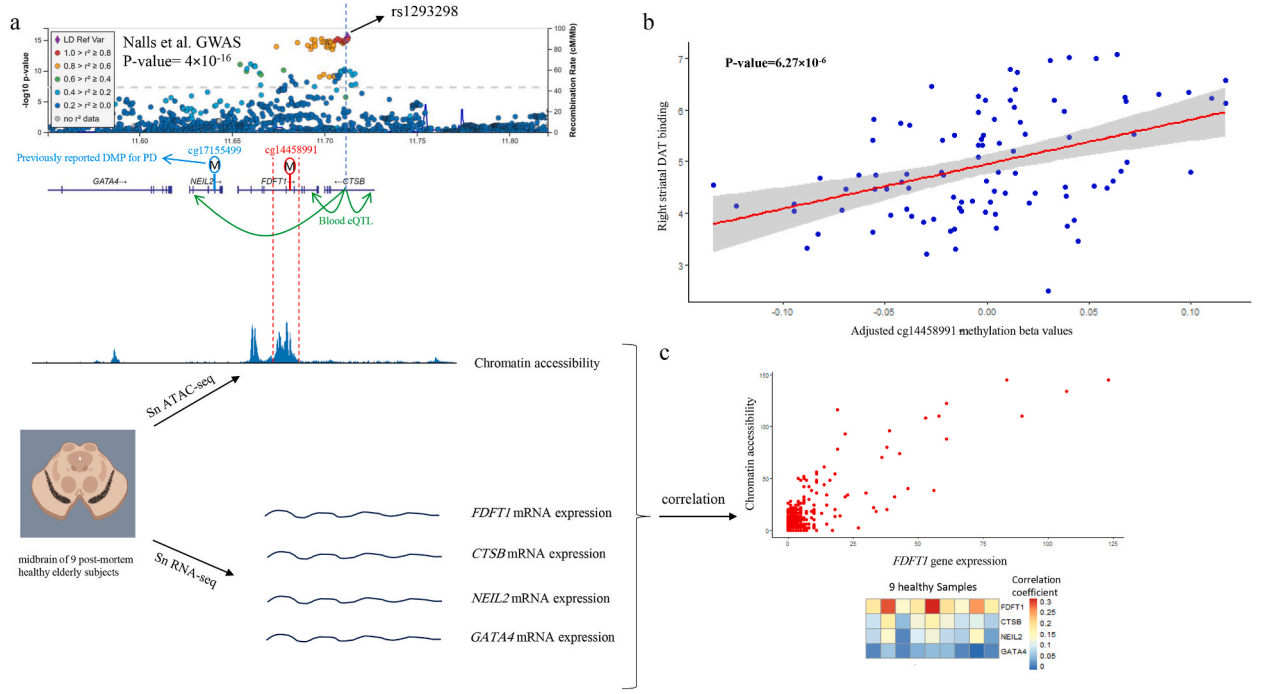


Fig. 3. The upper part of the figure shows the genomic position of one of the suggestive DMPs associated with right striatal DAT binding and its vicinity to the *CTSB* locus previously reported in the largest PD GWAS. The lower parts illustrate the overlap of this DMP with the chromatin-accessible region of the midbrain of nine healthy elderly individuals. It also shows our enhancer-gene mapping strategy using a midbrain single-nucleus multimodal dataset, which prioritized the *FDFT1* gene as the potential candidate gene in this PD GWAS locus.

a) This panel zooms into the *CTSB* GWAS locus, previously reported in the largest PD GWAS conducted by Nalls et al., and its LD structure and P-value. The top SNP rs1293298 in this locus is located in the *CTSB* gene. It also shows the location of our top EWAS DMP (cg14458991 methylation probe) in the *FDFT1* gene next to the *CTSB* gene and a previously reported DMP in PD EWAS (cg17155499) in the *NEIL2* gene. The lower part shows the overlap of the cg14458991 with the ATAC peaks obtained from the single-nucleus multimodal dataset of midbrain of 9 post-mortem healthy elderly individuals.

b) This panel shows the direct linear correlation of the cg14458991 methylation probe with right striatal DAT binding among 96 healthy individuals obtained from our EWAS analysis.

c) The lower part of the panel shows the heatmap result of our enhancer-gene mapping analysis. The rows are four potential genes in this locus. The columns show nine multimodal samples obtained from the midbrain of post-mortem healthy elderly individuals. Each cell's colors depict the degree to which ATAC peak regions correlate with gene expression. The upper part shows the scatter plot of the correlation between the expression of the top gene, *FDFT1*, and the chromatin accessibility of that region.

Our enhancer-gene mapping strategy prioritized the *FDFT1* gene as the potential target gene of this region in the midbrain. The *FDFT1* gene product is the first specific enzyme in cholesterol synthesis, one of the pathways linked to PD and alpha-synuclein aggregation [37–39]. Interestingly, a blood-based methylome-wide association study with 1132 patients with PD and 999 control subjects indicated that differential methylation of the cg17155499 locus in the *NEIL2* gene, a neighbor of the *FDFT1* gene, is associated with PD [40]. Future experimental studies are needed to unravel the exact driver of PD in this region. However, the association of altered methylation of this previously reported locus in PD with striatal DAT binding, even among healthy elderly individuals, can be insightful for understanding early pathological alterations in PD.

Assessing clinical cognitive functions, we found that striatal DAT binding is markedly lower in healthy individuals with subtle impairments in RAVLT delayed recall memory, irrespective of the effect of age, gender, and education. It was reported that impairments of memory and executive function domains are the prominent cognitive deficits in patients with newly diagnosed PD [41]. A few studies have shown a connection between striatum presynaptic dopamine function and memory performance [42]. In line with our findings, Chen et al. showed that delayed verbal memory and working memory explain most of the variation of the D2/D3 receptor densities in the striatum of 62 healthy individuals [43]. However, due to the scarcity of clinical imaging studies on non-PD individuals (healthy and prodromal elderly individuals), more studies are needed to investigate impairments in which the cognitive domain manifests earlier during the PD course.

Our study has some limitations: Our sample size was modest. However, since our primary goal was to generate novel hypotheses on the pathophysiological mechanisms involved in early PD, our population can be unique, considering the high costs of obtaining DaTscan and blood methylome profiles from healthy individuals. However, more sample sizes are needed in future studies to confirm or rule out our findings. Although we did not validate the result of blood EWAS in other healthy cohorts, replicating three PD GWAS loci in our EWAS may indicate the robustness of our findings. We hope that the suggestive results of our study, at molecular and clinical

Table 2
Association of striatal DAT binding with a battery of neuropsychological tests across 5 cognitive domains.

Cognitive domains	Number of subjects	Mean score (±SD)	Age		Education		Gender		Striatal DAT binding	
Executive-working Memory			P-value	Effect size (SE)	P-value	Effect size (SE)	P-value	Effect size (SE)	P-value	Effect size (SE)
Letter number sequencing	251	10.86 (±2.71)	**3.00×10^{-6}**	−0.30 (0.07)	**0.004**	0.18 (0.07)	0.68	0.02 (0.06)	0.72	−0.02 (0.05)
Semantic fluency animal sub score	251	22.16 (±5.32)	0.37	−0.05 (0.06)	0.09	0.10 (0.06)	0.25	−0.07 (0.06)	0.41	0.05 (0.06)
Trail making test B	61	74.42 (±42.26)	0.09	0.22 (0.22)	0.96	−0.005 (0.08)	0.46	−0.10 (0.14)	0.93	−0.01 (0.14)
Language-executive										
Lexical fluency	61	44.73 (±12.71)	0.71	−0.04 (0.10)	**0.02**	0.30 (0.10)	0.56	0.08 (0.14)	0.72	0.05 (0.14)
Modified Boston naming test	61	55.75 (±7.85)	0.64	0.05 (0.10)	0.05	0.25 (0.10)	0.19	0.18 (0.14)	0.11	0.22 (0.14)
Attention-processing Speed										
Symbol digit modalities test	251	47.31 (±10.37)	**5.68×10^{-19}**	−0.53 (0.06)	**0.006**	0.15 (0.03)	0.14	−0.08 (0.05)	0.86	0.01 (0.10)
Trail making test A	61	32.98 (±11.99)	**0.02**	0.30 (0.15)	0.65	0.05 (0.16)	0.46	−0.10 (0.14)	0.97	0.005 (0.14)
Visuospatial processing										
Clock drawing test	62	6.38 (±1.23)	0.31	−0.13 (0.13)	0.33	−0.13 (0.13)	0.51	0.09 (0.14)	0.44	0.11 (0.14)
Benton judgment of line orientation test	250	13.14 (±2.04)	**0.04**	−0.12 (0.06)	**0.001**	0.20 (0.06)	**5.90×10^{-5}**	0.25 (0.05)	0.44	−0.04 (0.05)
Memory										
HVLT immediate recall	251	26.01 (±4.58)	**2.97×10^{-4}**	−0.23 (0.05)	0.12	0.09 (0.06)	0.05	−0.12 (0.06)	0.91	0.007 (0.07)
HVLT delayed recall	251	9.24 (±2.34)	**0.001**	−0.20 (0.10)	**0.04**	0.12 (0.06)	**0.04**	−0.12 (0.06)	**0.04**	0.13 (0.06)
HVLT discrimination recognition	251	10.14 (±2.64)	0.61	−0.03 (0.06)	0.83	−0.01 (0.05)	0.27	−0.07 (0.06)	0.05	0.12 (0.06)

levels, bring novel insights for uncovering the pathological mechanisms underlying the earliest stages of PD in future studies.

The loci identified in healthy individuals may serve as biomarkers for early detection of PD. Moreover, by understanding the epigenetic changes that occur before the onset of symptoms, we can develop screening tools to identify at-risk populations. Notably, the loci highlighted in our study can provide insights into the biological pathways that contribute to PD. We can better understand the mechanisms underlying PD by examining how these epigenetic changes affect gene expression and neuroimaging endophenotypes. Finally, findings from EWAS can guide translational research efforts to develop therapeutic strategies.

Future research should focus on longitudinal studies that track healthy individuals over time to see if those with altered epigenetic profiles in our identified loci will develop PD. This strategy could validate the predictive power of our identified markers. Additionally, combining EWAS findings with genomic, transcriptomic, and proteomic data can provide a holistic view of how genetic and epigenetic factors interact in the context of PD and prioritize the most likely causal genes. Identifying specific epigenetic changes associated with neuroimaging endophenotypes (DaT scan) and their effect on the expression of nearby genes could lead to future experimental studies investigating these genes' effects in PD pathology in vitro.

In conclusion, our blood EWAS revealed that altered methylations of seven distinct CpG sites were associated with lower striatal DAT binding among healthy individuals, highlighting the replication of three previously reported PD GWAS loci at the methylome level, even among healthy individuals. Using a midbrain single-nucleus multimodal dataset, we also prioritized the *FDFT1* gene as the potential target gene in the *CTSB* GWAS locus. We also showed that lower striatal DAT binding can manifest as subtle impairments in delayed memory function among healthy individuals.

CRediT authorship contribution statement

Arash Yaghoobi: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Homa Seyedmirzaei:** Writing – review & editing, Writing – original draft. **Marzie Jamaat:** Writing – review & editing, Writing – original draft. **Moein Ala:** Writing – review & editing, Writing – original draft.

Ethics approval and consent to participate

This study was exempted from IRB approval as data from a pre-existing database was used.

Consent for publication

Not applicable.

Availability of data and materials

We used data from the PPMI cohort for this study. The details of PPMI cohort cohort can be found on its website (https://www.ppmi-info.org/). By sending our research proposal form, we had permission to access the database of the PPMI cohort via the LONI database (https://ida.loni.usc.edu/).

Funding

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

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e40618.

References

[1] D. Bega, et al., Clinical utility of DaTscan in patients with suspected Parkinsonian syndrome: a systematic review and meta-analysis, NPJ Parkinsons Dis 7 (1) (2021) 43.
[2] L.M. de Lau, M.M. Breteler, Epidemiology of Parkinson's disease, Lancet Neurol. 5 (6) (2006) 525–535.
[3] O.B. Tysnes, A. Storstein, Epidemiology of Parkinson's disease, J. Neural. Transm. 124 (8) (2017) 901–905.
[4] N. Heng, et al., Striatal dopamine loss in early Parkinson's disease: systematic review and novel analysis of dopamine transporter imaging, Mov Disord Clin Pract 10 (4) (2023) 539–546.
[5] L.M. Chahine, et al., Proposal for a biologic staging system of Parkinson's disease, J. Parkinsons Dis. 13 (3) (2023) 297–309.
[6] T. Ogawa, et al., Role of neuroimaging on differentiation of Parkinson's disease and its related diseases, Yonago Acta Med. 61 (3) (2018) 145–155.
[7] S. Mayo, et al., Recent evidence in epigenomics and proteomics biomarkers for early and minimally invasive diagnosis of alzheimer's and Parkinson's diseases, Curr. Neuropharmacol. 19 (8) (2021) 1273–1303.
[8] S.K.L. Darweesh, et al., Association between poor cognitive functioning and risk of incident parkinsonism: the Rotterdam study, JAMA Neurol. 74 (12) (2017) 1431–1438.
[9] C. Fang, et al., Cognition Deficits in Parkinson's Disease: Mechanisms and Treatment, vol. 2020, Parkinsons Dis, 2020 2076942.
[10] F.J. Siepel, et al., Cognitive executive impairment and dopaminergic deficits in de novo Parkinson's disease, Mov. Disord. 29 (14) (2014) 1802–1808.
[11] X.X. Gao, et al., Prevalence and risk factors of intrahepatic cholestasis of pregnancy in a Chinese population, Sci. Rep. 10 (1) (2020) 16307.
[12] L.M. Chahine, et al., Cognition in individuals at risk for Parkinson's: Parkinson associated risk syndrome (PARS) study findings, Mov. Disord. 31 (1) (2016) 86–94.
[13] J. Seibyl, et al., 123-I Ioflupane SPECT measures of Parkinson disease progression in the Parkinson Progression Marker Initiative (PPMI) trial, J. Nucl. Med. 54 (supplement 2) (2013), 190-190.
[14] L. Pihlstrøm, et al., Epigenome-wide association study of human frontal cortex identifies differential methylation in Lewy body pathology, Nat. Commun. 13 (1) (2022) 4932.
[15] M. Meijer, et al., Epigenome-wide DNA methylation in externalizing behaviours: a review and combined analysis, Neurosci. Biobehav. Rev. 145 (2023) 104997.
[16] Y. Sommerer, et al., Epigenome-wide association study in peripheral tissues highlights DNA methylation profiles associated with episodic memory performance in humans, Biomedicines 10 (11) (2022).
[17] A. Fernández-Sanlés, et al., Physical Activity and Genome-wide DNA Methylation: The REgistre GIroní del COR Study, Med. Sci. Sports Exerc. 52 (3) (2020) 589–597.
[18] S. Sayols-Baixeras, et al., Identification and validation of seven new loci showing differential DNA methylation related to serum lipid profile: an epigenome-wide approach. The REGICOR study, Hum. Mol. Genet. 25 (20) (2016) 4556–4565.
[19] O. Story Jovanova, et al., DNA methylation signatures of depressive symptoms in middle-aged and elderly persons: meta-analysis of multiethnic epigenome-wide studies, JAMA Psychiatr. 75 (9) (2018) 949–959.
[20] W. Zhang, et al., Distinct CSF biomarker-associated DNA methylation in Alzheimer's disease and cognitively normal subjects, Alzheimer's Res. Ther. 15 (1) (2023) 78.
[21] I.K. Karlsson, et al., Epigenome-wide association study of level and change in cognitive abilities from midlife through late life, Clin. Epigenet. 13 (1) (2021) 85.
[22] A. Starnawska, et al., Epigenome-wide association study of cognitive functioning in middle-aged monozygotic twins, Front. Aging Neurosci. 9 (2017) 413.
[23] M. van Iterson, E.W. van Zwet, B.T. Heijmans, Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution, Genome Biol. 18 (1) (2017) 19.
[24] M. Suderman, et al., dmrff: identifying differentially methylated regions efficiently with power and control, bioRxiv (2018) 508556.
[25] M.A. Nalls, et al., 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 (12) (2019) 1091–1102.

[26] J. Nasser, et al., Genome-wide enhancer maps link risk variants to disease genes, Nature 593 (7858) (2021) 238–243.
[27] L. Adams, et al., Single-nuclei paired multiomic analysis of young, aged, and Parkinson's disease human midbrain reveals age- and disease-associated glial changes and their contribution to Parkinson's disease, medRxiv (2022) 2022, 01.18.22269350.
[28] H.R. Gosselt, et al., Epigenome wide association study of response to methotrexate in early rheumatoid arthritis patients, PLoS One 16 (3) (2021).
[29] X. Chen, et al., Potassium channels: a potential therapeutic target for Parkinson's disease, Neurosci. Bull. 34 (2) (2018) 341–348.
[30] D. Ma, et al., Ligand activation mechanisms of human KCNQ2 channel, Nat. Commun. 14 (1) (2023) 6632.
[31] R. Ferrari, et al., Genetic architecture of sporadic frontotemporal dementia and overlap with Alzheimer's and Parkinson's diseases, J. Neurol. Neurosurg. Psychiatry 88 (2) (2017) 152–164.
[32] J. Sharma, et al., Calpain activity is negatively regulated by a KCTD7-Cullin-3 complex via non-degradative ubiquitination, Cell Discov 9 (1) (2023) 32.
[33] K. Yamano, et al., Endosomal Rab cycles regulate Parkin-mediated mitophagy, Elife 7 (2018).
[34] I.E. Jansen, et al., Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk, Nat. Genet. 51 (3) (2019) 404–413.
[35] J. Yuasa-Kawada, et al., Neuronal guidance genes in health and diseases, Protein Cell 14 (4) (2023) 238–261.
[36] J. Jones-Tabah, et al., The Parkinson's disease risk gene cathepsin B promotes fibrillar alpha-synuclein clearance, lysosomal function and glucocerebrosidase activity in dopaminergic neurons, bioRxiv (2023), 2023.11.11.566693.
[37] P. García-Sanz, M.F.G. A. J, R. Moratalla, The role of cholesterol in α-synuclein and lewy body pathology in GBA1 Parkinson's disease, Mov. Disord. 36 (5) (2021) 1070–1085.
[38] H. Xicoy, B. Wieringa, G.J.M. Martens, The role of lipids in Parkinson's disease, Cells 8 (1) (2019).
[39] C. Birkenbihl, et al., Artificial intelligence-based clustering and characterization of Parkinson's disease trajectories, Sci. Rep. 13 (1) (2023) 2897.
[40] C.L. Vallerga, et al., Analysis of DNA methylation associates the cystine-glutamate antiporter SLC7A11 with risk of Parkinson's disease, Nat. Commun. 11 (1) (2020) 1238.
[41] D. Muslimovic, et al., Cognitive profile of patients with newly diagnosed Parkinson disease, Neurology 65 (8) (2005) 1239–1245.
[42] L.H.T. Fornari, et al., Striatal dopamine correlates to memory and attention in Parkinson's disease, Am J Nucl Med Mol Imaging 11 (1) (2021) 10–19.
[43] P.S. Chen, et al., Correlation between different memory systems and striatal dopamine D2/D3 receptor density: a single photon emission computed tomography study, Psychol. Med. 35 (2) (2005) 197–204.