

## **A genetic and transcriptomic assessment of the *KTN1* gene in Parkinson's disease risk.**

Anni Moore<sup>1</sup>, BSc, Sara Bandres-Ciga<sup>2</sup>, PhD, Cornelis Blauwendraat<sup>2</sup>, PhD, Monica Diez-Fairen<sup>3</sup>, PhD, on behalf of the International Parkinson's Disease Genomics Consortium (IPDGC)

<sup>1</sup> Computational Biology Group, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA

<sup>2</sup> Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA

<sup>3</sup> Fundació Docència i Recerca MútuaTerrassa and Movement Disorders Unit, Department of Neurology, University Hospital MútuaTerrassa, Terrassa 08221, Barcelona, Spain

**Corresponding author:** Dr. Monica Diez-Fairen, email: [mdiez@mutuaterrassa.es](mailto:mdiez@mutuaterrassa.es). University Hospital MútuaTerrassa, Sant Antoni 19, Terrassa 08221, Barcelona, Spain.

**Keywords:** Parkinson's disease; risk factor; KTN1; putamen; substantia nigra; GWAS; WGS

**Word count:** 2898

**Abstract word count:** 165

### **Abstract**

Parkinson's disease (PD) is a progressive neurological disorder caused by both genetic and environmental factors. A recent finding has suggested an association between *KTN1* genetic variants and changes in its expression in the putamen and substantia nigra brain regions and an increased risk for PD. Here, we examine the link between PD susceptibility and *KTN1* using individual-level genotyping data and summary statistics from the most recent genome-wide association studies (GWAS) for PD risk and age at onset from the International Parkinson's Disease Genomics Consortium (IPDGC), as well as whole-genome sequencing data from the Accelerating Medicines Partnership Parkinson's disease (AMP-PD) initiative. To investigate the potential effect of changes in *KTN1* expression on PD compared to healthy individuals, we further assess publicly available expression quantitative trait loci (eQTL) results from GTEx v8 and BRAINEAC and transcriptomics data from AMP-PD. Overall, we found no genetic associations between *KTN1* and PD in our cohorts but found potential evidence of differences in mRNA expression, which needs to be further explored.

## Introduction

Parkinson's disease (PD) is a neurodegenerative disorder caused by both genetic and environmental risk factors. Its neuropathological hallmark is the loss of midbrain dopaminergic neurons in the substantia nigra pars compacta (SNpc) that innervate the basal ganglia, a set of subcortical nuclei that play a central role in motor control, motor learning, executive functions, and emotion regulation <sup>1</sup>. In PD, the progressive degeneration of these neurons leads to severe dopamine (DA) deficiency in the striatum and the subsequent appearance of the cardinal motor features of the disease, including tremors, rigidity, hypokinesia, and postural imbalance <sup>2,3</sup>. Several treatments for PD are known to target the SNpc as well as the putamen such as the pharmacological substitution of striatal dopamine using L-DOPA, the immediate biosynthetic precursor of DA <sup>4</sup>, and surgical interventions like deep brain stimulation for those patients developing intractable L-DOPA-related motor complications <sup>5</sup>.

Recently, Mao *and colleagues* reported nominal associations between common intergenic variants downstream and upstream of the *KTNI* gene and PD risk, as well as an increased *KTNI* mRNA expression in the putamen and SNpc, which resulted in a compensatory increase in the gray matter volumes (GMV) of these two brain regions in PD patients, suggesting that *KTNI* may play a functional role in the development of PD <sup>6</sup>. *KTNI* encodes the integral membrane protein kinectin 1, a receptor that is involved in organelle transport within the cell by allowing vesicle binding to kinesin <sup>7</sup>. Impairment of these molecular motors has been increasingly linked to neurological diseases with a subcortical component, such as PD <sup>8</sup>. In addition, rs945270, an intergenic locus downstream of *KTNI*, has been associated with larger putamen GMV and an increased *KTNI* expression in the frontal cortex and putamen brain regions and blood tissue <sup>9</sup>. More recently, variants in close proximity to *KTNI* were also associated with the volume of the nucleus accumbens, caudate nucleus, and globus pallidus, suggesting that this genomic region may have an important role in determining multiple subcortical brain volumes during development <sup>10</sup>.

In the study by Mao *et al*, a total of eight intergenic variants near *KTNI* were nominally associated with PD. This included six variants associated with PD across at least two independent cohorts and two other GMV-related variants that were not found in the discovery dataset but nominally associated with PD in one of the two replication cohorts. Of note, none of the reported *KTNI* associated variants were consistently replicated across all cohorts. In addition, six of these risk variants were found to regulate either the mRNA expression in putamen, SNpc or putamen GMV, and *KTNI* was differentially expressed in the SNpc between PD and controls in one cohort. As part of the International Parkinson's Disease Genomics Consortium (IPDGC) efforts to investigate and replicate reported risk factors for PD, we

sought to assess the role of *KTNI* as a genetic risk factor for PD within large European cohorts. Here, we considered all potential influencers of *KTNI* by including variants in the entire cis-*KTNI* region to account for their potential role in regulating *KTNI* mRNA expression in the putamen and SNpc.

## Methods

### *Genotype data*

Imputed individual-level genotyping data from genome-wide association studies (GWAS) were obtained from the IPDGC and contained 14,671 PD samples and 17,667 healthy controls, all of European ancestry. All datasets underwent quality control procedures as previously reported<sup>11,12</sup>. Of note, one of Mao *et al* replication datasets, “*lng\_coriell\_pd*” (phs001172.v1.p2), is partially included in IPDGC genotype data we used. Data were based on the hg19 genome assembly, expanding the *KTNI* region (chr14: 56,025,790-56,168,244) by 1MB on each end to include the entire cis-*KTNI* region (chr14: 55,025,790-57,168,244), and were extracted with PLINK1.9<sup>13</sup>. Then, PLINK1.9 was used to perform a logistic regression adjusted by sex, age, 10 principal components (PCs) to account for population stratification, and cohort to control for chip bias. Variants were annotated using ANNOVAR<sup>14</sup>.

We further assessed data from the most recent GWAS meta-analyses to date (excluding 23andMe data) for PD risk consisting of 15,056 PD cases, 18,618 UK Biobank proxy-cases (i.e., subjects with a first degree relative with PD) and 449,056 controls<sup>11</sup>, as well as for age at onset (AAO) comprising 17,996 PD cases<sup>12</sup>. All participants were of European ancestry and details on genotyping quality and imputation are described in the original sources. Summary statistics used in this study are public and available via <https://pdgenetics.org>.

### *Whole-genome sequencing data*

Whole-genome sequencing (WGS) data were retrieved from the Accelerating Medicines Partnership - Parkinson’s Disease Initiative (AMP-PD; [www.amp-pd.org](http://www.amp-pd.org)) which contains three cohorts including Parkinson’s Progression Markers Initiative (PPMI; <https://www.ppmi-info.org/>), Parkinson’s Disease Biomarkers Program (PDBP; <https://pdbp.ninds.nih.gov/>), and BioFIND (<https://biofind.loni.usc.edu/>). In total, the set contained 1,647 PD patients and 1,050 healthy controls of European ancestry, with an average AAO of 59.7 years in cases and an average age of 60.1 years in controls. More detailed cohort characteristics, as well as quality control procedures, are further described in <https://amp-pd.org/whole-genome-data>. Again we used the 1MB expanded cis-*KTNI* region, (chr14: 54,559,072-56,701,526) based on the hg38 genome assembly, and extracted it with PLINK1.9. Results were annotated with ANNOVAR. Association analyses of *KTNI* variants and risk for PD were done using

single-variant score tests in RVTESTS package v2.1.0<sup>15</sup>, with a minimum allele count (MAC) threshold of 3 applied, and adjusted by sex, age, 10 PCs, and origin cohort to account for underlying variables.

To assess the cumulative effect of multiple rare variants on the risk for PD, we performed gene-based burden analyses on WGS data. Including all variants within the region, a minimum allele count (MAC) threshold of 1 was applied, and variants with a minor allele frequency (MAF) < 1% were filtered out before running burden tests (SKAT, sequence Kernel association test, and SKAT-O, optimized SKAT) with RVTESTS package v2.1.0 following default parameters<sup>15</sup>. Age, gender, 10 PCs, and the dataset were again accounted for covariates. All code used in this study can be found at <https://github.com/ipdgc/IPDGC-Trainees/blob/master/KTN1.md>.

### *Expression Quantitative Trait Loci (eQTL)*

We chose to investigate eQTL values of all genes found within the cis-*KTN1* region to account for potential effects on expression by any variants within the region. Variant eQTL results of all genes were obtained from the GTEx Analysis Release V8 ([dbGaP Accession phs000424.v8.p2](https://dbgap.ncbi.nlm.nih.gov/oa/GET.cgi?acc=phs000424.v8.p2)) based on analyses from WGS genotypes and RNA-seq expression data from putamen and SNpc tissue samples. GTEx subjects are predominantly of European ancestry (84.6%) but do contain other ancestries as well. Putamen eQTL results were calculated from 170 total samples with genotype and expression data. SNpc eQTL values were calculated from 114 patient samples<sup>16</sup>. We also obtained values from BRAINEAC (UKBEC)<sup>17</sup>, which used genotype data and reads per kilobase per million (RPKM) normalized and covariate-adjusted expression data from 117 post-mortem putamen samples of European ancestry to calculate eQTL values<sup>18</sup>.

### *Transcriptomics data*

Whole blood time progression gene expression data (gencode v29) from 0-24 months was accessed from AMP-PD. We used the PPMI, PDBP, and BioFIND cohorts at the baseline, 0 month time point (note BioFIND 'Baseline' is at '0.5' month), including a total of 1,886 cases and 1,285 control samples. Samples with sex mismatch were excluded. Raw mRNA expression of *KTN1* in all datasets can be seen in **Figure 1**. Expression data were quantified as transcripts per kilobase million (TPM) and quantile normalized before *KTN1* values were extracted using Ensembl gene id ENSG00000126777.17 (**Figure 2**). Using the scikit-learn Python packages<sup>19</sup>, residuals were calculated using linear regression, and the data were adjusted using age, sex, race, 20 genotype PCs, cohort, and missingness rate as covariates (**Figure 2**). To test for significant differences in expression between the cases and controls, we performed a t-test

with the residuals. We also used an additional dataset taken from the 24-month time point in the PDBP and PPMI cohorts of 841 cases and 386 controls to confirm our findings.

### *Ethics Statement*

We confirm that our study does not need ethical approval from the institutions involved since it involves information freely available in the public domain, such as the analysis of open-source datasets, where the data was properly anonymized and informed consent was obtained at the time of original data collection.

## **Results**

### *Association between *KTNI* variants and PD risk*

Using imputed genotype data from GWAS, we identified a total of 6,188 variants present in the cis-*KTNI* region ([Supplementary Table 1](#)). We performed a case-control association analysis using logistic regression which showed no variants significantly associated with PD ( $p < 5 \cdot 10^{-8}$ ). Among the 6,188 variants found, 35 were exonic, with 19 of them being synonymous and 16 nonsynonymous variants. The remaining 6,153 non-coding variants included 2,331 intronic, 54 in the 3'-untranslated region (3'-UTR) and 8 in the 5'-untranslated region (5'-UTR) ([Supplementary Table 1](#)). All variants found by Mao *and colleagues* were noncoding, including five variants (rs12880292, rs8017172, rs17253792, rs945270, rs1188184) that were intergenic, two in non-coding exonic regions (rs7142488, rs4144657), and one that was downstream (rs7157819). We further analyzed summary statistics from the above-mentioned risk and AAO PD GWAS meta-analyses. No evidence for an association between *KTNI* common genetic variation and PD was detected ( $p < 5 \cdot 10^{-8}$ ) for either PD risk or AAO (**Figure 3**).

Similarly, using WGS data, we identified 17,382 variants in the cis-*KTNI* region ([Supplementary Table 2](#)), 337 of which were coding including 137 synonymous and 193 non-synonymous variants. Upon performing a case-control analysis using a single score variant test, no significant difference was determined ( $p < 5 \cdot 10^{-8}$ ) in the frequency of *KTNI* variants between PD and healthy individuals. Gene-based burden analyses showed no enrichment of rare variants ( $p < 5 \cdot 10^{-8}$ ) in cases compared to controls ([Supplementary Table 3](#)).

Of note, the original eight variants reported by Mao *and colleagues* were not significantly associated with PD ( $p < 5 \cdot 10^{-8}$ ) in our genotyping nor sequencing datasets (**Table 1**).

### *KTNI and mRNA expression*

To test the potential regulatory effect of cis 1MB region for *KTNI* risk variants affecting the level of expression of *KTNI* mRNA in the putamen and substantia nigra brain regions, we first considered cis-eQTL results from the GTEyv8 cohort. 18 genes were found within the cis-*KTNI* region including *DLGAP5*, *KTNI-AS1*, *LGALS3*, *SOCS4*, *SAMD4A*, *MIR4308*, *TBPL2*, *KTNI*, *PRL13AP3*, *LINC00520*, *PELI2*, *LOC101927690*, *FBXO34*, *MAPK1IP1L*, *TMEM260*, *GCHI*, *WDHD1*, and *ATG14*. After filtering results from all genes by p-value for both brain tissue types, we only found two significant values for the gene *ATG14* in the putamen. Both variants, rs112424815 ( $p=6.10e^{-8}$ ) and rs79509284 ( $p=6.90e^{-8}$ ) fall within our cis-*KTNI* region. No significant variants were found in the substantia nigra. We also consulted eQTL results from BRAINEAC derived from UKBEC, which showed no significance from any gene variants for putamen or substantia nigra tissue.

Using mRNA gene expression data from whole blood, we assessed the difference in *KTNI* expression between 1,886 cases and 1,285 control samples from AMP-PD. We performed a t-test between normalized, covariate-adjusted *KTNI* expression in cases vs controls, and found a significant difference in *KTNI* mRNA expression in whole blood ( $p=0.0012$ ) at the 0 month time point (**Table 2**). However, when sub-cohorts PPMI and PDBP were separated within the dataset and run individually to test if any single sub-cohort was driving overall significance, neither cohort indicated a significant difference between *KTNI* expression in cases vs controls ([Supplementary Table 4](#)). The initial test was also not consistent in additional whole blood expression data from a follow-up visit at 24 months from AMP-PD for 841 cases and 386 controls. We calculated a p-value of 0.8112 for the t-test, indicating a non-significant difference (**Table 2**). Therefore we cannot conclude that a significant difference exists in *KTNI* mRNA expression between PD patients and controls in whole blood, and further analyses should be done.

## Discussion

As a substantial proportion of PD risk is driven by genetic factors, efforts are continually being made to determine new genes that may be linked to the disease. While genes like *LRRK2*, *SNCA*, and *GBA* are well-known to have a substantial impact on one's likelihood of developing PD, they only explain a small fraction of PD heritability, suggesting that additional genes may harbor rare and common variation associated with the disease<sup>11,20</sup>. In addition, in most cases, it is still unknown how the identified genetic risk variants disrupt molecular and cellular processes underlying the pathobiology of the disease. With many of the known PD-related mutations affecting genes involved in intracellular transport systems<sup>21</sup>, *KTNI* may be a relevant gene to investigate given its role in organelle transport and vesicle binding<sup>7</sup>.

In our study, we were not able to replicate the findings by Mao *and colleagues* and cannot conclude that the *KTNI* gene is significantly implicated in PD etiology. Using large European cohorts, we were not able

to find any variants in the cis-*KTNI* region significantly associated with PD when compared to healthy individuals. Analysis of summary statistics from the most recent and largest GWAS meta-analyses to date for PD risk and AAO also showed no significant variants ( $p < 5 \cdot 10^{-8}$ ). We were not able to find a higher frequency of *KTNI* rare variants among PD patients. All eight *KTNI* variants found by Mao *et al.* were found outside the *KTNI* gene, one upstream and seven downstream, which we covered by considering the entire gene cis-region. None of the variants were significantly associated with PD in our analyses. Therefore we could not find genetic evidence of any association between *KTNI*-related variants and PD. We recognize that these findings only represent patients of European ancestry and we cannot exclude association between *KTNI* and PD in other populations.

Despite not finding genetic evidence, we continued to consider any potential effects of *KTNI* expression on PD risk. We consulted public databases, GTEx and BRAINEAC, consisting of hundreds of genotype and expression samples from the SNpc and putamen brain regions. Of note, one of the 18 genes within the region, *GCHI*, has been previously found to be associated with PD<sup>11</sup>. *GCHI* encodes the enzyme GTP cyclohydrolase 1, which is necessary for dopamine synthesis in nigrostriatal cells<sup>22,23</sup> and dopamine deficiency in the nigrostriatal region is a core biological feature of PD<sup>24</sup>. However, we found no significant eQTL values for *GCHI*. We were only able to locate two eQTL values of significance within the cis-*KTNI* gene region for the putamen - both variants affecting the expression of the gene *ATG14*. No eQTL values were found to be significant in SNpc tissue for any gene within the region. *ATG14* function is required for both basal and inducible autophagy but also determines the localization of the autophagy-specific PI3-kinase complex and contributes to autophagosome formation<sup>25</sup>. Interestingly, there is genetic, pathological, and experimental evidence suggesting an important role for autophagy impairment in PD<sup>26,27</sup>.

To test for differences in *KTNI* expression we used transcriptomic data from large European cohorts to perform a test on overall expression in whole blood. We did initially note a significant difference ( $p=0.0012$ ) in *KTNI* mRNA expression in PD patients versus control samples, however upon replication in an additional dataset, we could not reproduce significant results, warranting further investigation in this area. Unfortunately, we only had access to expression data from whole blood and we acknowledge that expression may be different in brain tissues. However, it has been shown that blood samples act as reliable alternatives for expression when brain tissues are not available<sup>28</sup>. Using blood also allowed us to increase the overall number of samples considerably, contrary to brain samples which are often harder to acquire.

*KTNI* has been previously associated with other neurological disorders, especially impacting putamen volume and structure, providing further evidence that it potentially plays a significant role in the brain. A previously mentioned study by Hilbar and colleagues found that a specific variant of *KTNI*, rs945270, was significantly linked to *KTNI* expression in blood, frontal cortex, and putamen tissues, ultimately resulting in a larger putamen volume<sup>9</sup>, though we were not able to conclude a similar association in our blood datasets. *KTNI* has also been implicated in alcohol and drug dependency, where the putamen GMV has been shown to be decreased in alcohol-dependent patients<sup>29</sup> but increased in stimulant drug-dependent patients<sup>30,31</sup>. Luo *et al* found significant *KTNI* variants that increased risk for substance abuse disorders in Europeans, as well as increased *KTNI* mRNA expression levels in the putamen and putamen GMVs that led them to conclude a positive relationship existed among variants of *KTNI* and substance use disorders risk<sup>32</sup>. Thus, despite our inability to conclude a casual increased risk for PD, *KTNI* remains of interest in brain disorders, especially those affecting the putamen area.

In conclusion, based on the current data presented, our analyses do not support a role for *KTNI* variants in PD genetic susceptibility. However, differences in *KTNI* mRNA expression between PD patients and healthy individuals should be further studied. The vast majority of our datasets are of European ancestry, therefore we cannot rule out a potential role of *KTNI* rare variants in PD among other populations. Additional large-scale familial and case-control studies in non-European ancestry populations are necessary to further evaluate the role of *KTNI* in PD etiology.

## **Acknowledgments**

We would like to thank all of the subjects who donated their time and biological samples to be a part of this study. We also would like to thank all members of the International Parkinson's Disease Genomics Consortium (IPDGC). For a complete overview of members, acknowledgments and funding, please see <http://pdgenetics.org/partners>. This work was supported in part by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute on Aging (NIA), and the National Institute of Environmental Health Sciences both part of the National Institutes of Health, Department of Health and Human Services; project numbers 1ZIA-NS003154, Z01-AG000949-02 and Z01-ES101986. In addition, this work was supported by the Department of Defense (award W81XWH-09-2-0128), and The Michael J Fox Foundation for Parkinson's Research. Data used in the preparation of this article were obtained from the AMP PD Knowledge Platform. For up-to-date information on the study, visit <https://www.amp-pd.org>. AMP PD – a public-private partnership – is managed by the FNIH and funded by Celgene, GSK, the Michael J. Fox Foundation for Parkinson's Research, the National Institute of Neurological Disorders and Stroke, Pfizer, and Verily. We would like

to thank AMP-PD for the publicly available whole-genome sequencing data, including cohorts from the Fox Investigation for New Discovery of Biomarkers (BioFIND), the Parkinson’s Progression Markers Initiative (PPMI), and the Parkinson’s Disease Biomarkers Program (PDBP). This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

### Competing interests

The authors declare that they have no conflict of interest.

### Author contributions

A.M. contributed to conceptualization and design of the work, formal analysis, curation and interpretation of data, and drafting the manuscript. M.D.F. contributed to conceptualization and design of the work, interpretation of the data, supervision and reviewing the manuscript. S.B.C. and C.B. contributed in reviewing and editing the final manuscript.

### Data availability

The summary statistics from PD meta-analyses that support the findings of this study are available from IPDGC consortium website, <https://pdgenetics.org/resources>. The whole genome sequencing and transcriptomics data that support the findings of this study are available from AMP-PD website, <https://amp-pd.org/>. Individual-level genotyping data that support the findings of this study are available from the corresponding author upon reasonable request. All code used in this study is publicly available at <https://github.com/ipdgc/IPDGC-Trainees/blob/master/KTN1.md>.

### Figures and Tables

**Table 1. Variants from Mao *and colleagues* in all datasets.**

Position (hg19)	Position (hg38)	rsID	Ref	Alt	Frequency					P-values		
					IPDGC Case	IPDGC Control	GWAS Summary Statistics*	AMP-PD	gnomAD NFE	IPDGC	GWAS Summary Statistics*	AMP-PD
14:55952864	14:55486146	rs12880292	G	A	0.3059	0.3250	0.321	0.310	0.338	$1.90 \cdot 10^{-7}$	0.001	0.649
14:56199048	14:55732330	rs8017172	G	A	0.4033	0.4182	0.414	0.422	0.418	$1.39 \cdot 10^{-4}$	0.011	0.045
14:56205030	14:55738312	rs17253792	T	C	0.0873	0.0892	0.905	0.085	0.079	0.384	0.343	0.108
14:56200473	14:55733755	rs945270	C	G	0.4041	0.4185	0.585	0.423	0.418	$2.06 \cdot 10^{-4}$	0.015	0.050

14:56247593	14:55780875	rs7157819	C	T	0.2989	0.2974	0.303	0.295	0.290	0.691	0.524	0.859
14:56248010	14:55781292	rs7142488	T	C	0.2997	0.2987	0.697	0.296	0.291	0.784	0.452	0.852
14:56248455	14:55781737	rs4144657	G	A	0.0657	0.0688	0.073	0.069	0.075	0.115	0.055	0.780
14:56442412	14:55975694	rs1188184	G	A	0.3649	0.3655	0.371	0.380	0.608	0.872	0.837	0.469

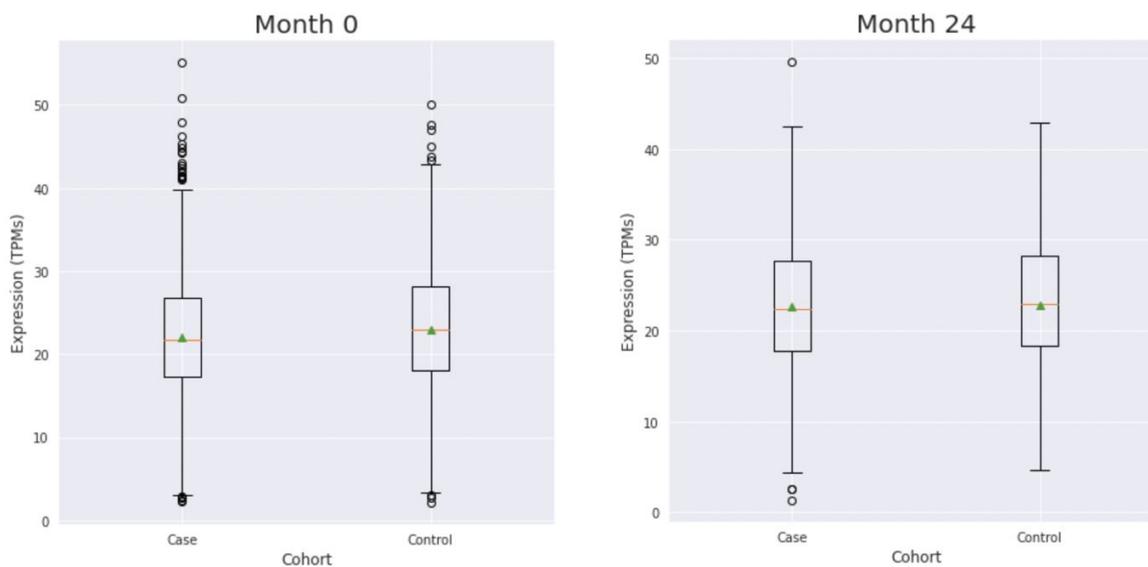
Key: AMP-PD, Accelerating Medicines Partnership Parkinson's disease; IPDGC, International Parkinson's Disease Genomics Consortium; GWAS, Genome Wide Association Study; NFE, Non-Finnish Europeans.

\*Summary statistics frequency from Nalls et al. 2019<sup>11</sup>

**Table 2. T-test in AMP-PD whole blood mRNA expression.**

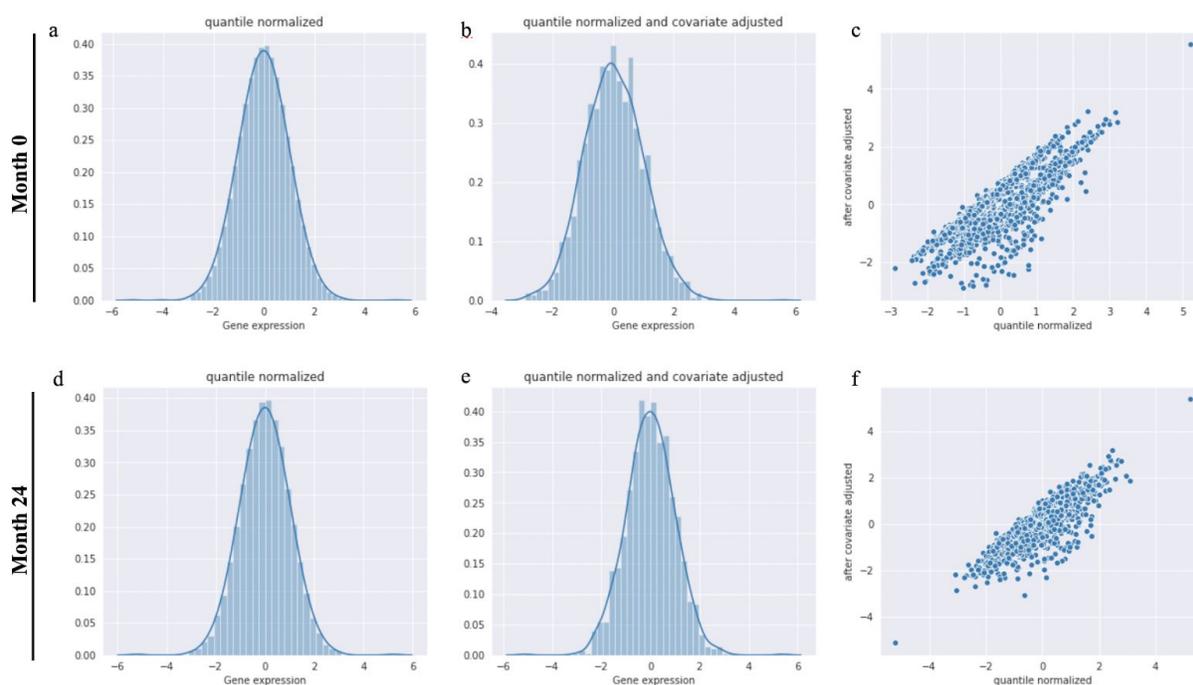
Cohorts	Month 0		Month 24	
	BioFind, PDBP, PPMI		PDBP, PPMI	
Case status	Case	Control	Case	Control
Samples	1568	1042	785	328
Mean age (yrs)	63.9	60.1	65	63.3
<i>KTNI</i> expression (TPMs)	21.98	23.01	22.61	22.78
p-value	0.0012		0.8112	

**Figure 1**



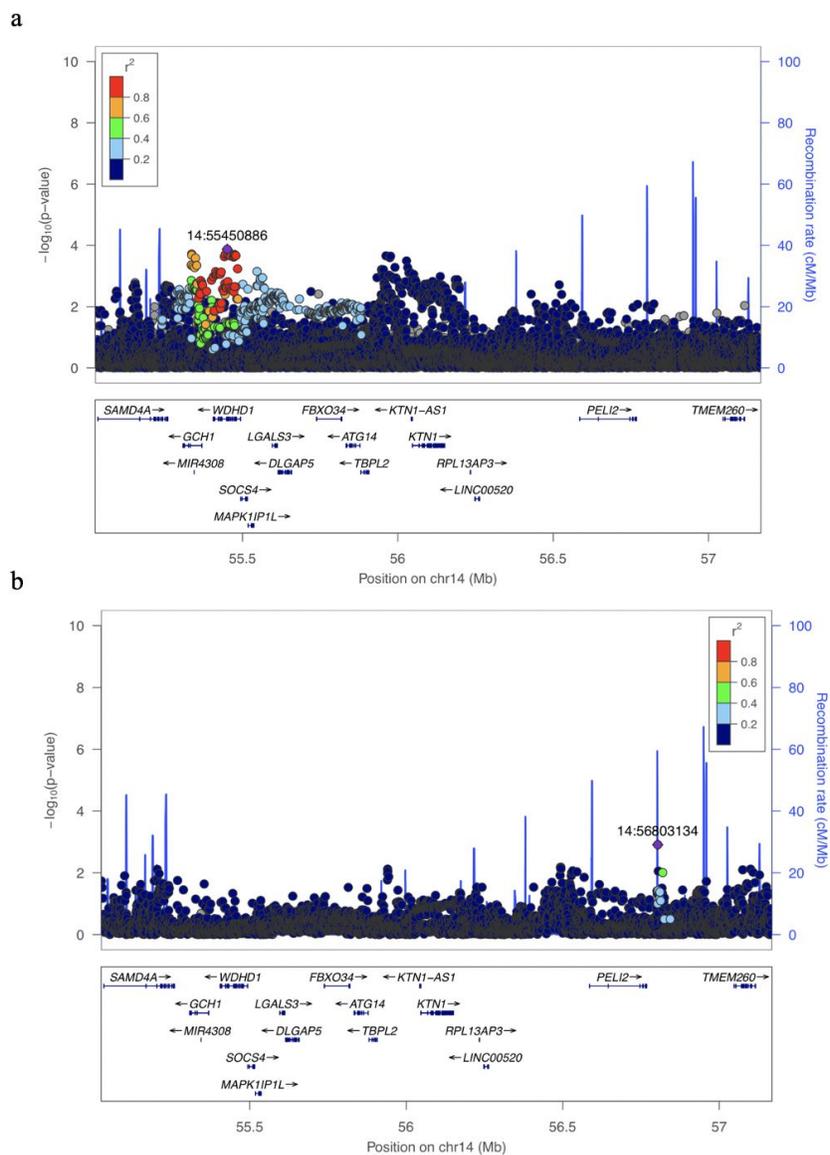
**Figure 1.** Raw *KTN1* mRNA expression - expressed in transcripts per kilobase million (TPMs) - by cohort in both tested AMP-PD expression datasets.

**Figure 2**



**Figure 2.** KTN1 mRNA expression processing. **a.** Distribution of all gene expression after quantile normalization at Month 0. **b.** Distribution of all gene expression after quantile normalization and covariate adjustment at Month 0. **c.** Linear regression of quantile normalized values versus quantile normalized and covariate adjusted at Month 0. **d.** Distribution of all gene expression after quantile normalization at Month 24. **e.** Distribution of all gene expression after quantile normalization and covariate adjustment at Month 24. **f.** Linear regression of quantile normalized values versus quantile normalized and covariate adjusted at Month 24.

**Figure 3**



**Figure 3.** *KTN1* locus zoom plots for PD risk and age at onset (AAO) plotted as the base-pair position of each variant (x-axis) versus the  $-\log_{10}(\text{p-value})$  of variants on or near *KTN1* (y-axis). Variants are colored by their  $R^2$  linkage disequilibrium value, which is relative to the variant with the lowest p-value, shown in purple. **a.** Plotting of p-values from PD risk GWAS meta-analysis of cis-*KTN1* region, excluding

23andMe data <sup>11</sup>. **b.** Plotting of p-values from PD AAO GWAS meta-analysis of cis-*KTN1* region, excluding 23andMe data <sup>12</sup>.

## References

1. Lanciego, J. L., Luquin, N. & Obeso, J. A. Functional neuroanatomy of the basal ganglia. *Cold Spring Harb. Perspect. Med.* **2**, a009621 (2012).
2. Barter, J. W., Castro, S., Sukharnikova, T., Rossi, M. A. & Yin, H. H. The role of the substantia nigra in posture control. *Eur. J. Neurosci.* **39**, 1465–1473 (2014).
3. Poewe, W. *et al.* Parkinson disease. *Nat Rev Dis Primers* **3**, 17013 (2017).
4. LeWitt, P. A. & Fahn, S. Levodopa therapy for Parkinson disease: A look backward and forward. *Neurology* **86**, S3–12 (2016).
5. Limousin, P. & Foltynie, T. Long-term outcomes of deep brain stimulation in Parkinson disease. *Nature Reviews Neurology* vol. 15 234–242 (2019).
6. Mao, Q. *et al.* KTN1 Variants Underlying Putamen Gray Matter Volumes and Parkinson’s Disease. *Frontiers in Neuroscience* vol. 14 (2020).
7. Kumar, J., Yu, H. & Sheetz, M. P. Kinectin, an essential anchor for kinesin-driven vesicle motility. *Science* **267**, 1834–1837 (1995).
8. Liu, X.-A., Rizzo, V. & Puthanveetil, S. V. PATHOLOGIES OF AXONAL TRANSPORT IN NEURODEGENERATIVE DISEASES. *Transl. Neurosci.* **3**, 355–372 (2012).
9. Hibar, D. P. *et al.* Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224–229 (2015).
10. Satizabal, C. L. *et al.* Genetic architecture of subcortical brain structures in 38,851 individuals. *Nat. Genet.* **51**, 1624–1636 (2019).
11. Nalls, M. A. *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**, 1091–1102 (2019).

12. Blauwendraat, C. *et al.* Parkinson's disease age at onset genome-wide association study: Defining heritability, genetic loci, and  $\alpha$ -synuclein mechanisms. *Mov. Disord.* **34**, 866–875 (2019).
13. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
14. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* vol. 38 e164–e164 (2010).
15. Zhan, X., Hu, Y., Li, B., Abecasis, G. R. & Liu, D. J. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* **32**, 1423–1426 (2016).
16. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
17. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosci.* **17**, 1418–1428 (2014).
18. Guelfi, S. *et al.* Regulatory sites for splicing in human basal ganglia are enriched for disease-relevant information. *Nat. Commun.* **11**, 1041 (2020).
19. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. and Grisel, O. and Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
20. Bandres-Ciga, S., Diez-Fairen, M., Kim, J. J. & Singleton, A. B. Genetics of Parkinson's disease: An introspection of its journey towards precision medicine. *Neurobiol. Dis.* **137**, 104782 (2020).
21. Bandres-Ciga, S. *et al.* The endocytic membrane trafficking pathway plays a major role in the risk of Parkinson's disease. *Mov. Disord.* **34**, 460–468 (2019).
22. Rudakou, U. *et al.* Common and rare GCH1 variants are associated with Parkinson's disease. *Neurobiol. Aging* **73**, 231.e1–231.e6 (2019).
23. Yoshino, H. *et al.* GCH1 mutations in dopa-responsive dystonia and Parkinson's disease. *J. Neurol.* **265**, 1860–1870 (2018).

24. Pan, H.-X. *et al.* GCH1 variants contribute to the risk and earlier age-at-onset of Parkinson's disease: a two-cohort case-control study. *Transl. Neurodegener.* **9**, 31 (2020).
25. Itakura, E., Kishi, C., Inoue, K. & Mizushima, N. Beclin 1 forms two distinct phosphatidylinositol 3-kinase complexes with mammalian Atg14 and UVRAG. *Mol. Biol. Cell* **19**, 5360–5372 (2008).
26. Moors, T. E. *et al.* Therapeutic potential of autophagy-enhancing agents in Parkinson's disease. *Mol. Neurodegener.* **12**, 11 (2017).
27. Bandres-Ciga, S. *et al.* Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. *Acta Neuropathol.* **140**, 341–358 (2020).
28. Rollins, B., Martin, M. V., Morgan, L. & Vawter, M. P. Analysis of whole genome biomarker expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153B**, 919–936 (2010).
29. Mackey, S. *et al.* Mega-Analysis of Gray Matter Volume in Substance Dependence: General and Substance-Specific Regional Effects. *Am. J. Psychiatry* **176**, 119–128 (2019).
30. Jacobsen, L. K., Giedd, J. N., Gottschalk, C., Kosten, T. R. & Krystal, J. H. Quantitative morphology of the caudate and putamen in patients with cocaine dependence. *Am. J. Psychiatry* **158**, 486–489 (2001).
31. Ersche, K. D. *et al.* Abnormal brain structure implicated in stimulant drug addiction. *Science* **335**, 601–604 (2012).
32. Luo, X. *et al.* KTN1 variants and risk for attention deficit hyperactivity disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **183**, 234–244 (2020).