RESEARCH ARTICLE

Dysregulation of the Causative Genes for Hereditary Parkinsonism in the Midbrain in Parkinson's Disease

Yun Joong Kim, MD, PhD,[1,2,3]* Junbeom Jeon, MS,[4] Jaemoon Shin, MS,[4] Nan Young Kim, MS,[3] Jeong Hoon Hong, MS,[1] Jae-Min Oh, PhD,[1] SangKyoon Hong, PhD,[3] Yeo Jin Kim, MD, PhD,[5] Young-Eun Kim, MD,[2] Suk Yun Kang, MD, PhD,[6] Hyeo-Il Ma, MD, PhD,[2] Unjoo Lee, PhD,[7] and Jeehee Yoon, PhD[4]*

[1]*ILSONG Institute of Life Science, Hallym University, Anyang, Korea*
[2]*Department of Neurology, Hallym University Sacred Heart hospital, Hallym University College of Medicine, Hallym University, Anyang, Korea*
[3]*Hallym Institute of Translational Genomics & Bioinformatics, Hallym University Medical Center, Anyang, Korea*
[4]*Department of Computer Engineering, Hallym University, Chuncheon, Korea*
[5]*Department of Neurology, Chuncheon Sacred Heart Hospital, Hallym University College of Medicine, Hallym University, Chuncheon, Korea*
[6]*Department of Neurology, Hallym University Dongtan Sacred Heart Hospital, Hallym University College of Medicine, Hallym University, Dongtan, Korea*
[7]*Department of Electronic Engineering, Hallym University, Chuncheon, Korea*

ABSTRACT: Background and Objectives: Many hereditary movement disorders with complex phenotypes without a locus symbol prefix for familial PD present as parkinsonism; however, the dysregulation of genes associated with these phenotypes in the SNpc of PD patients has not been systematically studied.

Methods: Gene set enrichment analyses were performed using 10 previously published genome-wide expression datasets obtained by laser-captured microdissection of pigmented neurons in the SNpc. A custom-curated gene set for hereditary parkinsonism consisting of causative genes (n = 78) related to disorders with a parkinsonism phenotype, but not necessarily idiopathic or monogenic PD, was constructed from the Online Mendelian Inheritance in Man database.

Results: In 9 of the 10 gene expression data sets, gene set enrichment analysis showed that the disease-causing genes for hereditary parkinsonism were downregulated in the SNpc in PD patients compared to controls (nominal *P* values <0.05 in five studies). Among the 63 leading edge subset genes representing downregulated genes in PD, 79.4% were genes without a locus symbol prefix for familial PD. A meta–gene set enrichment analysis performed with a random-effect model showed an association between the gene set for hereditary parkinsonism and PD with a negative normalized enrichment score value (–1.40; 95% CI: –1.52~–1.28; *P* < 6.2E-05).

Conclusion: Disease-causing genes with a parkinsonism phenotype are downregulated in the SNpc in PD. Our study highlights the importance of genes associated with hereditary movement disorders with parkinsonism in understanding the pathogenesis of PD.

Key Words: Parkinson's disease; microarray; gene set enrichment analysis; hereditary movement disorders; *PPP2R2B*

Parkinson's disease (PD) is a clinically, pathologically, and genetically heterogeneous disorder.[1] Although most cases of PD are of sporadic onset, approximately 10% to 15% of patients with PD have a family history of the disease. Mutations in the causative genes have been found in cases of familial Parkinson's disease (FPD) and

*Correspondence to: Dr. Yun Joong Kim, ILSONG Institute of Life Science, Hallym University, 15 Gwanpyeong-ro 170 beon-gil, Anyang, 14066, Korea; E-mail: yunkim@hallym.ac.kr; or Dr. Jeehee Yoon, Department of Computer Engineering, Hallym University, Hallymdaehak-gil, Chuncheon, 24252 Gangwon-do, Korea; E-mail: jhyoon@hallym.ac.kr

Drs. Yun Joong Kim and Jeehee Yoon contributed equally to this work.

Funding agencies: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2014R1A2A1A11052141). This research was also supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1A02937313).

Relevant conflicts of interest/financial disclosures: Nothing to report. Full financial disclosures and author roles may be found in the online version of this article.

Received: 11 July 2016; **Revised:** 26 February 2017; **Accepted:** 17 March 2017

Published online 00 Month 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.27019

TABLE 1. Summary of the microarray data sets used in this study

No.	GEO_ID or Name of Data Set	Affymetrix Chip Model	No. of Samples			References
			PD	Control	Total	
1	GSE20333	HG-FOCUS	6	6	12	Grunblatt and colleagues[11]
2	GSE20292	HG_U133A	11	18	29	Zhang and colleagues[13]
3	GSE8397[a]	HG_U133A	24	15	39	Moran and colleagues[14]
4	GSE8397[b]	HG_U133A	9	7	16	Moran and colleagues[14]
5	GSE7621	HG-U133_Plus_2	16	9	25	Lesnick and colleagues[16]
6	GSE24378	U133_X3P	8	9	17	Cantuti-Castelvetri and colleagues,[15] Zheng and colleagues[18]
7	Simunovic and colleagues[c]	HG_U133A	10	9	19	Simunovic and colleagues[17]
8	GSE20141	HG-U133_Plus_2	10	8	18	Zheng eand colleagues[18]
9	GSE20164	HG_U133A	6	5	11	Hauser and colleagues[12]
10	GSE20163	HG_U133A	8	9	17	Durrenberger and colleagues[19]
		Total	108	95	203	

[a]Samples derived from the lateral and medial SN in GSE8397 (GSE8397_ Med_ Lat).
[b]Samples derived from the lateral SN in GSE8397 (GSE8397_Lat).
[c]Microarray data were not deposited in the GEO, but were generously provided by Dr. Kai-Christian Sonntag.

are referred to as PARKs. Although PARK gene mutations may not have been replicated in other populations, more than 20 PARKs have been identified.[2,3] Depending on the affected PARK genes, FPD patients may present with classical parkinsonism, early-onset parkinsonism, or atypical phenotypes.[4] On the other hand, parkinsonism is well recognized as a concomitant or isolated feature in many hereditary movement disorders caused by a single gene mutation (non-PARK genes).[5-7] Although the pathological substrates for parkinsonism caused by mutations in many non-PARK genes have not been documented, the degeneration of dopaminergic neurons regardless of Lewy body pathology is likely to be a common denominator. Functional studies on PARK genes have uncovered many molecular mechanisms and pathways for dopaminergic neuronal degeneration, and the results have provided important clues in the understanding of PD and FPD. Convergent molecular pathways include synaptic transmission, endosomal sorting and maturation, lysosomal degradation, and mitochondrial maintenance.[3,8,9] A number of the above mechanisms appear to be attributed to non-PARK parkinsonism genes, although functional studies on these genes have infrequently been performed in terms of dopaminergic neurodegeneration.[6,10]

To understand the molecular signatures related to dopaminergic neuronal death in the SNpc of patients with PD, several groups have performed genome-wide expression studies (GWESs) of microarray analyses using a laser-captured microdissection technique (Table 1).[11-19] A number of microarray studies adopting differential gene expression analysis revealed that certain PARK genes are downregulated in the brain of patients with PD.[12,14,20,21] Although the number of non-PARK parkinsonism genes has been increasing, a systematic study investigating the association of these genes with PD using gene expression data sets has never been performed. Gene set enrichment analyses (GSEAs) can statistically analyze the group-wise differential expression pattern of predefined gene sets to assess the degree of association with disease phenotypes. Only one GSEA and meta-GSEA study for PD[18] was performed for nine GWESs using 522 predefined gene sets of the Broad Institute's Molecular Signatures Database (MSigDB; www.broadinstitute.org/msigdb), which is not relevant to non-PARK parkinsonism genes.

We hypothesized that both non-PARK parkinsonism genes and PARK genes are dysregulated in the SNpc of PD patients. To evaluate our hypothesis, we performed a GSEA and a meta-GSEA on microarray data sets from PD patients and control subjects. We utilized a previously published manually curated parkinsonism gene set based on the Online Mendelian Inheritance in Man (OMIM) database and gene sets in the gene ontology (GO) database.[22] We also characterized our manually curated parkinsonism gene set by performing unbiased GO enrichment analyses.

Materials and Methods

Data Collection and Preprocessing

Ten data sets were selected based on the following inclusion criteria: (1) samples derived from the brain tissue of patients with PD and control subjects; (2) messenger (mRNA) collected using laser-captured microdissection; (3) lateral SNpc included (whole SNpc or lateral SNpc); and (4) Affymetrix chip used. Nine data sets were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) on 31 January 2015. An additional gene expression data set that satisfies the above criteria,[17] but was not deposited in GEO, was kindly provided by Dr. Kai-Christian Sonntag. A summary of

the data sets used in this study is shown in Table 1. Expression data for each data set were extracted from *.CEL files and normalized using the GeneChip Robust Multi-Array (GC-RMA) algorithm[23] implemented in the R/Bioconductor packages (http://www.bioconductor.org).

Constructing Gene Sets of Mendelian Genes Related to Disorders of Motor Systems

Custom gene sets consisting of causative genes related to disorders of motor systems were described in a previous report.[22] Briefly, the systematic retrieval of genes was performed using keywords referencing phenotypes of abnormal motor control in the OMIM database followed by the manual curation of all retrieved records. Keywords were searched in the text of the OMIM record with the prefix "#" (i.e., phenotype description with known molecular basis). For a gene set of parkinsonism (OMIM_Parkinsonism), a keyword, parkinson*, was used. OMIM_Parkinsonism included both PARK genes (n = 15) and non-PARK genes with parkinsonism phenotype (n = 63; Table S1). In mitochondrial disorders, the widespread involvement of the neuraxis with overlapping phenotypes of dysfunctional motor control is well established. Our custom-curated mitochondrial gene set (OMIM_Mitochondria) retrieved from OMIM consists of 82 nuclear-encoded genes that cause mitochondrial complex I, II, III, IV, and V deficiency syndrome, Leigh syndrome, combined oxidative phosphorylation deficiency, or mitochondrial DNA depletion syndrome. Because dysregulation of mitochondrial electron transport genes in PD was reported in a previous study,[18] we excluded these genes when constructing the OMIM_Mitochondria from the OMIM_Parkinsonism. However, among the genes for progressive external ophthalmoplegia with a parkinsonism phenotype, *SLC25A4*, *C10orf2*, and *POLG* were not excluded from the OMIM_Parkinsonism.

GSEA and Statistical Analysis in a Meta-GSEA

The association of our OMIM_Parkinsonism gene sets with gene expression changes in the SNpc in PD was tested by a GSEA (www.broadinstitute.org/gsea).[24,25] Briefly, a GSEA computationally determines whether predefined gene sets show significant differences between two distinct phenotypes (e.g., PD and control) using nonparametric Kolmogorov–Smirnov statistics. To this end, genes are ranked according to their differential expression between the phenotypes using an appropriate ranking metric. For each predefined gene set, the GSEA determines whether the members of the gene set are primarily distributed at the top or bottom of the ranked list of genes; thus, the gene set is correlated with the phenotypic distinction. GSEAs calculate an enrichment score (ES) that reflects the degree to which the gene set is over-represented at the top or bottom of the ranked list of genes. The core gene members that contribute most to the ES are called the leading edge subset (LES) genes. The ES is normalized to account for differences in the gene set size and correlations between the gene sets and the expression data set, thereby generating a normalized enrichment score (NES). The NES is used as an estimate of the effect size for comparing gene set enrichment results in which a positive (or negative) NES indicates gene set enrichment at the top (or bottom) of the ranked list of genes. The GSEA was performed using the Broad Institute JAVA command-line software Version 2.2.0 (http://www.broadinstitute.org/gsea/index.jsp) under the default parameters with the following exception: A *t* test was used as the ranking metric, which uses the difference of the means scaled by the standard deviation (SD) and the number of samples.

To integrate multiple microarray data sets from different research groups, we performed a meta-GSEA, which compares the effect size from an individual microarray data set and generates a summary effect size. For the meta-GSEA, the following gene sets from public databases were used: (1) the biological processes (GO_BP) gene sets in the gene ontology database (n = 825); (2) five disease gene sets from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.kegg.jp), which include KEGG_Parkinson's disease (KEGG_PD), KEGG_Alzheimer's disease (KEGG_AD), KEGG_Huntington's disease (KEGG_HD), KEGG_Amyotrophic lateral sclerosis (KEGG_ALS), and KEGG_Prion disease (KEGG_Prion); and (3) a gene set consisting of the 61 dysregulated genes that were reported to be shared among Alzheimer's disease, amyotrophic lateral sclerosis, Huntington's disease, multiple sclerosis, PD, and schizophrenia in the BrainNet Europe gene expression microarray study consisting of 113 subjects (BrainNet).[26]

For each gene set, the NES was combined across all data sets to obtain a summary effect size by adapting a random-effects model. We used the Comprehensive Meta-Analysis (CMA; www.meta-analysis.com, version 3.0) software package (Biostat, Englewood, NJ) to conduct the meta-analysis and interpret the associated statistics. To calculate the summary effect size under the random-effects model, we must consider two variance components: the within-study variance and the between-studies variance. Concretely, the weight is defined as the reciprocal of the sum of the within-study and between-studies variances. The summary effect size is then computed as the sum of the products (the effect size multiplied by the weight for each data set) divided by the sum of the weights of each data set[27]; see the Supplemental Method for more details.

We calculated the within-study variance for every gene set in every data set using a bootstrap method as suggested in the original GSEA papers.[18,25] To estimate the statistical distribution, a bootstrap method was used as follows: For each data set, we resampled the disease and control data to obtain the bootstrap resamples. We then recalculated the NES of each gene set for each iteration. We repeated this process 3,000 times. Given that the 3,000 NES values have a bimodal distribution, we randomly selected 1,000 positive (or negative) NESs from the 3,000 bootstrap resamples if the observed NES was positive (or negative), which is similar to performing a single-tail test. From the frequency distribution of the 1,000 NESs, we derived a within-study variance for each gene set. The between-studies variance was then computed using the within-study variances, the heterogeneity measure, and the total number of data sets.[27]

We assessed the significance of an observed summary effect size for a gene set by comparing it with the set of summary effect sizes of a null distribution calculated with randomly assigned weights. We used a permutation procedure as described previously[18]: We generated two 835 × 9 matrices: one contained the NES values and the other contained the weights of the 835 gene sets and 9 GWESs. We then randomly permuted the gene set labels of the weight matrix and recomputed the summary effect size of each gene set using the NES matrix and the permuted weight matrix, which generated a null distribution for the summary effect size. This procedure was repeated 1 million times, and the *P* value for an observed summary effect size was then calculated relative to this null distribution.

For the GSEA of the OMIM_Parkinsonism in the individual data sets, a *P* value <0.05 was considered statistically significant. For the meta-GSEA using the OMIM_Parkinsonism and OMIM_Mitochondria with other gene sets available from public databases, Bonferroni-corrected *P* values of less than 0.05 were used.

Gene Ontology Analysis of the PD Gene Set

A gene ontology analysis was performed with the functional annotation tools in the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource (https://david.ncifcrf.gov/).[28] Functional enrichment of our OMIM_Parkinsonism gene set was assessed using GO databases of biological process (BP), molecular function (MF), and cellular components (CC) by selecting GOTERM_BP_ALL, GOTERM_MF_ALL and GOTERM_CC_ALL.

Results

To investigate the association of our OMIM_Parkinsonism gene set in 10 GWESs, GSEAs were performed. In all of the GWESs except for one (GSE24378),[15] which was reported as a technical outlier in a previous GSEA study,[18] GSEAs revealed that many genes in the OMIM_Parkinsonism gene set are enriched at the right end of each rank-ordered list, indicating that these genes in the OMIM_Parkinsonism are underexpressed in PD (Fig. 1). In 5 of 10 GWESs, the GSEA of the OMIM_Parkinsonism was statistically significant (nominal *P* value: <0.05). As in a previous report,[18] we excluded GSE24378 in further analyses. The LES genes of the OMIM_Parkinsonism are represented in the ranked list subsequent to the peak negative enrichment score (Fig. 1). A list of the 63 genes in the LES identified from the nine GSEAs is shown in Figure 2. Fifty-two of 63 genes (82.5%) belonged to the leading edge subsets in five GWESs and showed statistical significance. To characterize 63 LES genes among the OMIM_Parkinsonism genes, we performed a GO analysis, which detects significantly over-represented gene ontologies. The enriched GO categories and their clusters identified with a Bonferroni-corrected *P* value of less than 0.05 are summarized in Table S2. Significantly enriched biological processes included synaptic transmission, locomotor behavior, nerve-nerve synaptic transmission, dopamine biosynthetic pathways, neurotransmitter transport, cellular ion homeostasis, negative regulation of neuronal apoptosis, vesicle-mediated transport, learning, and memory, and cell death. The enriched GO terms of cellular components with Bonferroni-corrected *P* values of less than 0.05 included neuron projection, cell projection, cell fraction, and mitochondrion.

A meta-GSEA of 827 gene sets (OMIM_Parkinsonism, OMIM_Mitochondria and the 825 gene sets of the biological processes in the GO database) with nine GWESs in a random-effects model showed that 41 of the 827 gene sets (4.96%) were significantly enriched in the PD brain (uncorrected $P < 0.05$; Table 2). Among these gene sets, the NES (summary effect size in Table 2) of the OMIM_Parkinsonism gene set, which passed the Bonferroni-corrected *P*-value threshold ($P <$ 6.2E-05), was the lowest followed by the OMIM_Mitochondria. A negative NES value (–1.40; 95% CI: –1.52∼–1.28) in the meta-GSEA of the OMIM_Parkinsonism suggests that many genes in the OMIM_Parkinsonism were downregulated in the SNpc of PD patients (Fig. 3). Expression levels of individual genes in the OMIM_Parkinsonism in each data set are shown in Table S3. We next expanded a meta-GSEA by dividing the OMIM_Parkinsonism gene set into two gene sets, namely PARK genes (OMIM_PARK) and non-PARK genes (OMIM_Non-PARK), and adding five gene sets for neurodegenerative disorders from the KEGG database and a BrainNet gene set consisting of 61 common dysregulated genes. The meta-GSEA showed that the

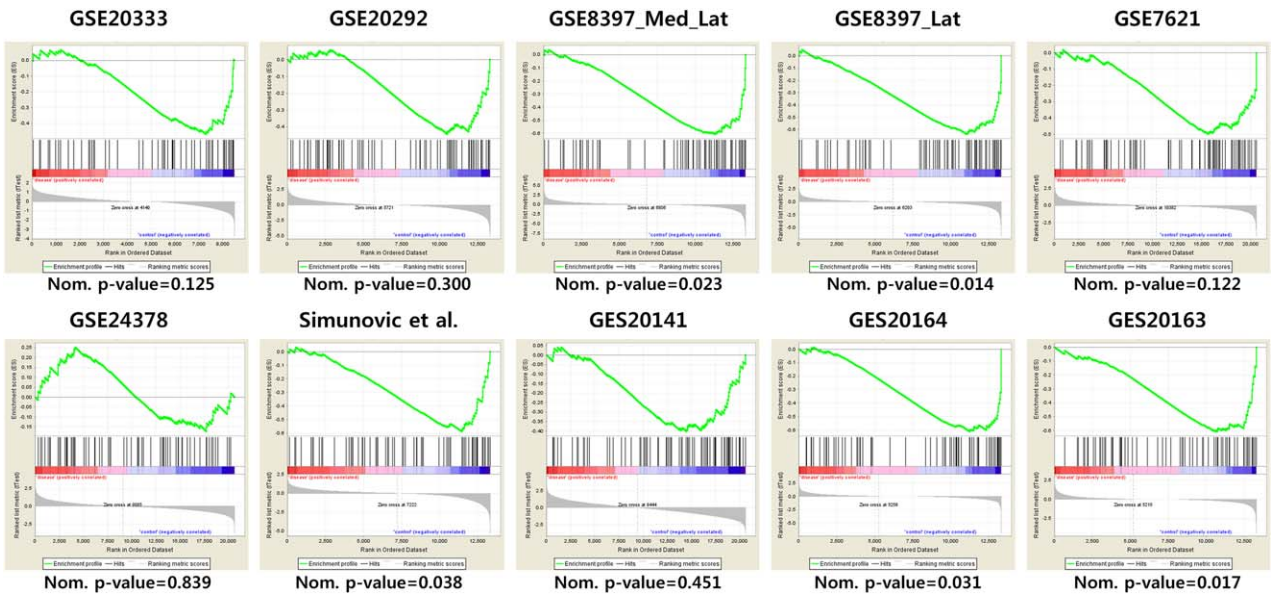

FIG. 1. Enrichment plots for the OMIM_Parkinsonism gene set in the 10 GWESs. Names of the datasets are shown in the upper center of each plot. The top portion of each plot shows the running enrichment score, in which the peak value represents the ES of the OMIM_Parkinsonism. The middle portion shows where each gene of the OMIM_Parkinsonism appears in the ranked list of genes. Each gene is marked as a vertical line. In the nine GWESs (excluding GSE_24378), the leading edge subset appears subsequent to the peak negative enrichment score. The bottom portion of each plot shows the ranking metric (*t* test) values of the ranked genes. Nom. p-value refers to the nominal *P* value. [Color figure can be viewed at wileyonlinelibrary.com]

genes in the OMIM_Non-PARK, OMIM_PARK, OMIM_Mitochondria, KEGG_PD, KEGG_HD, and KEGG_AD were downregulated in the PD brain (uncorrected $P < 0.05$), among which only the KEGG_PD passed the Bonferroni-corrected *P*-value threshold of less than 0.05 (Table S4). A comparison of the genes among these gene sets revealed that 94 genes were overlapped in the KEGG_PD, KEGG_HD, and KEGG_AD gene sets (Table S1). The majority of these overlapped genes were the 24 genes in the OMIM_Mitochondria that are related with mitochondrial ATP production. However, a total of 10 genes overlapped between the OMIM_Parkinsonism and KEGG_PD. On the other hand, there was no statistically significant enrichment of dysregulated genes in the KEGG_ALS, KEGG_Prion, and BrainNet gene sets.

Discussion

Our GSEA and meta-GSEA results showing an association of the OMIM_Parkinsonism gene set with PD with negative NESs strongly supports our hypothesis that both PARK genes and non-PARK parkinsonism genes are dysregulated in the SNpc of patients with PD. The trend of enrichment of the GO_BP gene sets related to mitochondria in our meta-GSEA is consistent with the results of a previous GSEA study supporting mitochondrial dysfunction in PD.[18] Given that genes related with mitochondria adenosine triphosphate (ATP) production are excluded in the OMIM_Parkinsonism gene set, we believe that our meta-GSEA with the OMIM_Parkinsonism genes is unique and possesses significant neurological implications. The dysregulation of non-PARK parkinsonism genes in PD has not been systematically studied. Our results showing that 79.4% of the LES genes in the OMIM_Parkinsonism gene set after the GSEA can be categorized as non-PARK parkinsonism genes provide the first evidence of a link between the causative genes for atypical parkinsonism and PD at the gene expression level. Although direct comparisons among the various disease gene sets were not the primary object of this study, the results of our meta-GSEA suggest that gene expression changes of OMIM_Parkinsonism genes may be specific to PD. Meta-GSEAs with PD-unrelated gene sets (KEGG_ALS or KEGG_Prion) and the BrainNET study gene set[26] showed negative results. We think that significant downregulation of genes in three KEGG gene sets (KEGG_PD, KEGG_HD, and KEGG_AD) in the meta-GSEA is attributed to the enrichment of genes related with mitochondrial ATP production in these genes sets. Unlike the disease gene sets in the KEGG database that may have incorporated knowledge on gene expression in the process of curating the gene sets, our custom-curated gene sets were constructed solely based on Mendelian inheritance of disease phenotypes, which are independent of knowledge of gene expression.

Mutations of Mendelian genes are generally understood to lead to derangement of cellular pathways or systems by protein dysfunction rather than transcriptional dysregulation of multiple genes. Mendelian genes

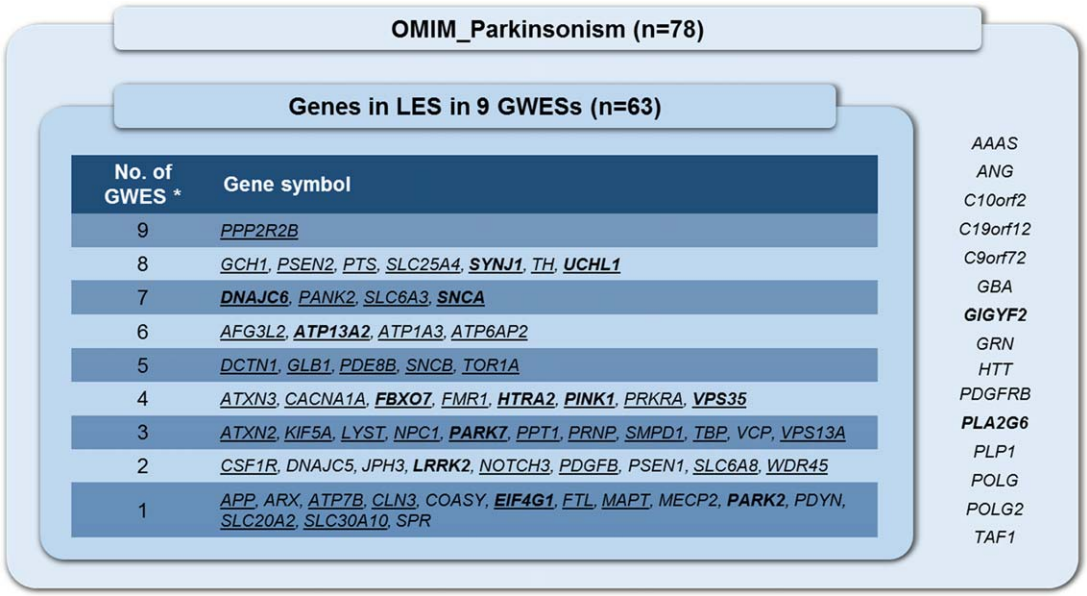

FIG. 2. Lists of genes within LES after a GSEA with the OMIM_Parkinsonism gene set in nine GWESs. The asterisk refers to the number of GWES in which a given gene was included in the LES after the GSEA. PARK genes are marked in bold. Genes belonging to the LES in the five data sets with $P < 0.05$ in GSEA (GSE8397_Med_Lat, GSE8397_Lat, Simunovic and colleagues, GSE20164, and GSE20163) are underlined (n = 52). The mRNA expression data for *C10orf2*, *HTT*, and *GIGYF2* are not available. [Color figure can be viewed at wileyonlinelibrary.com]

for inherited ataxia were used to elucidate a protein-protein interaction network (i.e., an interactome) and the pathogenesis of Purkinje cell degeneration.[29] Genome-wide gene expression studies usually identify disease networks consisting of dysregulated genes at the gene expression level. However, we observed convergence of genetic/heredity information of parkinsonism at the genomic level (i.e., OMIM_Parkinsonism genes) and the disease network consisting of dysregulated genes (i.e., downregulated LES genes) at the gene expression level in the SNpc in PD. These findings suggest that dysregulation of disease-causing genes presenting with parkinsonism, regardless of whether it is classic or atypical, contributes to pathogenesis in idiopathic PD. The inheritance of phenotypes in an autosomal-recessive manner or as a loss-of-function mutation was documented in many of our OMIM_Parkinsonism genes in the LES; thus, dopaminergic neuronal degeneration in PD may be related to the cumulative effect of a partial loss of function of variable degrees in multiple genes, which is caused by reduced mRNA expression. This assumption is consistent with recent genetic studies that demonstrate an increased risk for PD because of a single heterozygous mutation in autosomal-recessive disorders genes, which represents a partial loss of function. This group of genes includes PARK genes (*PARK2*, *PINK1*, *PLA2G6*, *ATP13A2*, and *SYNJ1*) and non-PARK genes (*GBA*, *SMPD1*, and *GCH1*).[30-41] Except for *GBA* and *PLA2G6*, all of these genes belonged to the LES (Fig. 2). A recent study showed that the enrichment of rare functional variants of PARK genes in PD cases compared to the controls.[42] Because the burden of a rare coding variant is likely to reflect a cumulative loss of function of a group of genes, the aforementioned study is consistent with our findings of the reduced expression of OMIM_Parkinsonism genes in the brains of patients with PD.

The results of the GO analyses of the OMIM_Parkinsonism genes in the LES are partially in agreement with the mechanisms suggested by experts and the results of previous GO analyses using differentially expressed genes from microarray data.[3,12,17,43,44] Remarkably, a substantial number of non-PARK parkinsonism genes are implicated in enriched ontologies. Downregulation of PARK and non-PARK parkinsonism genes is likely to affect multiple key biological processes and may lead to dopaminergic neuronal death in PD. Of note, we observed a downregulation of *PPP2R2B* in all nine GWESs, and this behavior has not been previously shown except for a recent study using transcriptome mapping software.[45] An expanded CAG repeat in the promoter region of *PPP2R2B* is causative for SCA12.[46] In unstable repeat disorders, a loss of function because of the inhibition of transcription and the gain-of-toxic function mutation because of expanded polyglutamine tracks or toxic RNA are well-documented mechanisms.[47] We did not exclude unstable repeat disorder genes from the OMIM_Parkinsonism as long as they showed parkinsonism because of the following reasons. Biological processes other than aggregation have been reported in unstable repeat disorder genes.[48] Considerable evidence also suggests a loss-of-function mutation in which unstable repeat disorder genes are knocked out, which enhances cell death.[49] Expanded CAG repeats in *PPP2R2B*

TABLE 2. List of the 41 biological processes gene sets that were significantly enriched in the PD brain after a meta-GSEA ($P < 0.05$)

Name of Gene Set	Summary Effect Size	Standard Error	Variance	Lower Limit	Upper Limit	*P* Value
OMIM_Parkinsonism	–1.3979	0.0619	0.0038	–1.5193	–1.2765	1.21E-05
OMIM_Mitochondria	–1.3693	0.0676	0.0046	–1.5017	–1.2369	1.90E-04
Energy derivation by oxidation of organic compounds	–1.3613	0.0701	0.0049	–1.4988	–1.2239	1.35E-03
Cellular respiration	–1.3612	0.0631	0.0040	–1.4849	–1.2375	1.35E-03
Aerobic respiration	–1.3475	0.0629	0.0040	–1.4707	–1.2242	4.88E-03
Purine nucleotide metabolic process	–1.3446	0.0596	0.0036	–1.4614	–1.2277	4.99E-03
Nucleobase nucleoside and nucleotide metabolic process	–1.3256	0.0574	0.0033	–1.4382	–1.2130	6.10E-03
Cofactor catabolic process	–1.3248	0.0726	0.0053	–1.4670	–1.1826	6.11E-03
Amine catabolic process	–1.3120	0.0632	0.0040	–1.4358	–1.1882	8.41E-03
Amino acid and derivative metabolic process	–1.3063	0.0556	0.0031	–1.4152	–1.1974	8.54E-03
Amino acid catabolic process	–1.3039	0.0647	0.0042	–1.4308	–1.1771	8.60E-03
Microtubule-based movement	–1.2998	0.0681	0.0046	–1.4332	–1.1664	8.70E-03
Amine metabolic process	–1.2949	0.0515	0.0027	–1.3959	–1.1940	8.92E-03
Tricarboxylic acid cycle intermediate metabolic process	–1.2906	0.0684	0.0047	–1.4246	–1.1566	1.15E-02
Nitrogen compound metabolic process	–1.2905	0.0540	0.0029	–1.3963	–1.1846	1.16E-02
Amine transport	–1.2898	0.0933	0.0087	–1.4727	–1.1069	1.20E-02
Purine ribonucleotide metabolic process	–1.2857	0.0670	0.0045	–1.4169	–1.1544	1.23E-02
Cellular carbohydrate metabolic process	–1.2792	0.0579	0.0034	–1.3927	–1.1656	1.44E-02
Amino acid metabolic process	–1.2766	0.0586	0.0034	–1.3914	–1.1618	1.62E-02
Amino acid transport	–1.2710	0.0688	0.0047	–1.4059	–1.1361	1.89E-02
Golgi vesicle transport	–1.2657	0.0659	0.0043	–1.3950	–1.1365	2.08E-02
Sphingolipid metabolic process	–1.2605	0.0626	0.0039	–1.3831	–1.1379	2.39E-02
Generation of precursor metabolites and energy	–1.2552	0.0571	0.0033	–1.3670	–1.1434	2.53E-02
Secretory pathway	–1.2542	0.0613	0.0038	–1.3743	–1.1341	2.54E-02
Heterocycle metabolic process	–1.2411	0.0643	0.0041	–1.3671	–1.1152	2.92E-02
Ribonucleotide metabolic process	–1.2389	0.0876	0.0077	–1.4105	–1.0673	3.11E-02
Nitrogen compound catabolic process	–1.2370	0.0688	0.0047	–1.3717	–1.1022	3.14E-02
Nucleotide metabolic process	–1.2329	0.0687	0.0047	–1.3675	–1.0982	3.16E-02
Cytoskeleton-dependent intracellular transport	–1.2273	0.0679	0.0046	–1.3604	–1.0943	3.19E-02
Secretion by cell	–1.2254	0.0561	0.0031	–1.3354	–1.1155	3.21E-02
Protein homo-oligomerization	–1.2163	0.0766	0.0059	–1.3665	–1.0662	3.56E-02
Amino acid derivative metabolic process	–1.2154	0.0622	0.0039	–1.3373	–1.0935	3.63E-02
Axon guidance	–1.2122	0.1007	0.0101	–1.4096	–1.0149	3.81E-02
Nucleosome assembly	–1.1993	0.0739	0.0055	–1.3440	–1.0545	4.37E-02
Peroxisome organization and biogenesis	–1.1968	0.0721	0.0052	–1.3382	–1.0555	4.60E-02
Vesicle-mediated transport	–1.1947	0.0725	0.0053	–1.3369	–1.0525	4.74E-02
Catabolic process	–1.1884	0.0636	0.0040	–1.3131	–1.0636	4.98E-02
G_1 phase	1.1938	0.0933	0.0087	1.0109	1.3768	4.79E-02
Inorganic anion transport	1.1973	0.0709	0.0050	1.0585	1.3362	4.55E-02
Meiotic recombination	1.1978	0.0617	0.0038	1.0768	1.3188	4.52E-02
Epithelial to mesenchymal transition	1.3023	0.0549	0.0030	1.1946	1.4100	8.65E-03

The summary effect size is the NES in the meta-GSEA.

are not translated, although they regulate the transcription of *PPP2R2B*.[50] We could not find studies that investigated the relationship between *PPP2R2B* and dopaminergic neuronal death. *PPP2R2B* encodes a protein phosphatase 2A (PP2A) regulatory subunit B that influences enzyme activity, substrate selectivity, and subcellular localization.[51] Interestingly, PP2A is a major phosphatase of the phosphorylated alpha-synuclein at S129 (alpha-syn-pS129), which is the major form of alpha-synuclein in Lewy bodies.[52] Taken together, the reduced expression of *PPP2R2B*, which leads to the reduced activity of PP2A, may be related to the accumulation of alpha-syn-pS129 in the brains of PD patients. Recent in vivo and in vitro studies demonstrated that the pharmacological activation of PP2A lowered the alpha-syn-pS129 burden.[53,54] We propose that our results showing *PPP2R2B* downregulation suggest exploring PP2A-enhancing drugs as potential therapeutic targets in PD, which is feasible considering that U.S. Food and Drug Administration–approved drugs activating PP2A are available.

Genetic and pathological evidence suggests that the overproduction and progressive accumulation of toxic forms of alpha-synuclein protein are key pathogenic mechanisms in PD.[55-57] Paradoxically, our results demonstrated a downregulation of *SNCA* expression in seven of nine GWESs. The direct measurement of *SNCA* mRNA expression in brain tissue from patients with PD revealed inconsistent results.[17,58-60] This

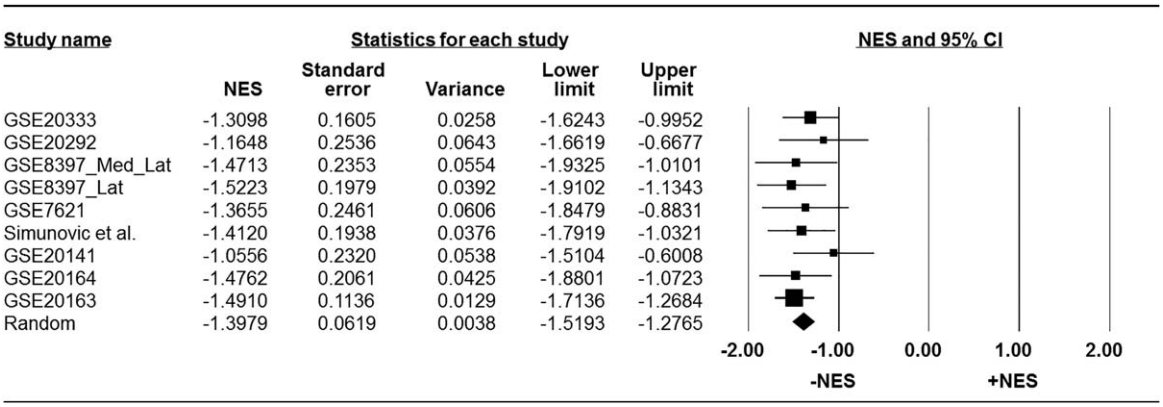

FIG. 3. Forest plot of the NES estimates for the OMIM_Parkinsonism gene set. The plot represents a meta-analysis result using random effects for the nine GWESs. The mid-point of each square represents the effect estimate (NES) for each data set, whereas the left and right extremes of the square represent the corresponding 95% confidence intervals (95% CI). The area of the square represents the weight given to the data set. The mid-point of the diamond represents the summary effect size, whereas the left and right extremes of the diamond represent the corresponding 95% confidence intervals.

inconsistency may have been caused by the different techniques used to measure mRNA (types of internal controls in real-time polymerase chain reaction [PCR], real-time PCR vs. microarray, etc.) or sampling bias (laser-captured microdissection vs. midbrain tissue block). In vivo silencing of *SNCA* was reported to induce nigral degeneration in rats.[61] A recent study showed that PD risk-related REP1 genotypes of *SNCA* were paradoxically associated with better motor and cognitive outcomes.[62] Taken together, the role of alpha-synuclein in the pathogenesis of PD may be more complex, which could include a dual and opposing role of *SNCA* depending on the clinical stages of PD. This issue should be explored further before developing a novel therapy aimed at reducing *SNCA* expression by RNA silencing.[62,63]

There are certain limitations to this study. Because microarray studies reveal a snapshot of gene expression changes, we cannot differentiate whether these changes are causes or results. Our custom-curated gene sets may be argued because we included genes that are not replicated by other groups (e.g., *UCHL1*, *HTRA1*). Genes without evidence of neurodegenerative phenotypes (e.g., *TOR1A*) were also included. We did not exclude these genes to avoid the selection bias caused by differences in personal knowledge. The presence of a parkinsonism phenotype was determined based on OMIM records, which may be biased. Nevertheless, we think that our approach in constructing a gene set is an ideal method for objective applications. GO analyses may be biased by well-researched protein or gene annotation deficits.[64] The results of GO analyses may not be specific to dopaminergic neurodegeneration because many annotations are from nondopaminergic cells. Our results are more suitable for a model of a loss-of-function mutation than a gain-of-toxic-function mutation, although both models may not be exclusive to each other. Whether a gain-of-toxic function mutation in PD might have led to the downregulation of a group of genes cannot be ruled out in this study.

In conclusion, we found that non-PARK parkinsonism genes and PARK genes were downregulated in the SNpc in PD. Enriched biological processes of these genes support previously suggested mechanisms of neurodegeneration. Our study highlights the importance of non-PARK parkinsonism genes in understanding the pathogenesis of PD. ■

Acknowledgments: The authors thank Dr. Song YH for her helpful discussions regarding this manuscript, and Drs. Choi WJ, Jang YJ, Kim KS, Kim Y, and Park WH for their participation in the manual curation of the OMIM-derived custom gene sets.

References

1. Kalia LV, Lang AE. Parkinson's disease. Lancet 2015;386:896-912.
2. Singleton AB, Farrer MJ, Bonifati V. The genetics of Parkinson's disease: progress and therapeutic implications. Mov Disord 2013; 28:14-23.
3. Volta M, Milnerwood AJ, Farrer MJ. Insights from late-onset familial parkinsonism on the pathogenesis of idiopathic Parkinson's disease. Lancet Neurol 2015;14:1054-1064.
4. Marras C, Lohmann K, Lang A, Klein C. Fixing the broken system of genetic locus symbols: Parkinson disease and dystonia as examples. Neurology 2012;78:1016-1024.
5. Klein C, Schneider SA, Lang AE. Hereditary parkinsonism: Parkinson disease look-alikes—an algorithm for clinicians to "PARK" genes and beyond. Mov Disord 2009;24:2042-2058.
6. Kara E, Hardy J, Houlden H. The pallidopyramidal syndromes: nosology, aetiology and pathogenesis. Curr Opin Neurol 2013;26: 381-394.
7. Stamelou M, Quinn NP, Bhatia KP. "Atypical" atypical parkinsonism: new genetic conditions presenting with features of progressive supranuclear palsy, corticobasal degeneration, or multiple system atrophy—a diagnostic guide. Mov Disord 2013;28:1184-1199.
8. Hardy J. Genetic analysis of pathways to Parkinson disease. Neuron 2010;68:201-206.
9. Pickrell AM, Youle RJ. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. Neuron 2015;85:257-273.
10. Small SA, Petsko GA. Retromer in Alzheimer disease, Parkinson disease and other neurological disorders. Nat Rev Neurosci 2015; 16:126-132.

11. Grunblatt E, Mandel S, Jacob-Hirsch J, et al. Gene expression profiling of parkinsonian substantia nigra pars compacta; alterations in ubiquitin-proteasome, heat shock protein, iron and oxidative stress regulated proteins, cell adhesion/cellular matrix and vesicle trafficking genes. J Neural Transm (Vienna) 2004;111:1543-1573.
12. Hauser MA, Li YJ, Xu H, et al. Expression profiling of substantia nigra in Parkinson disease, progressive supranuclear palsy, and frontotemporal dementia with parkinsonism. Arch Neurol 2005; 62:917-921.
13. Zhang Y, James M, Middleton FA, Davis RL. Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. Am J Med Genet B Neuropsychiatr Genet 2005;137B:5-16.
14. Moran LB, Duke DC, Deprez M, Dexter DT, Pearce RK, Graeber MB. Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. Neurogenetics 2006;7:1-11.
15. Cantuti-Castelvetri I, Keller-McGandy C, Bouzou B, et al. Effects of gender on nigral gene expression and parkinson disease. Neurobiol Dis 2007;26:606-614.
16. Lesnick TG, Papapetropoulos S, Mash DC, et al. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. PLoS Genet 2007;3:e98.
17. Simunovic F, Yi M, Wang Y, et al. Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. Brain 2009;132(Pt 7):1795-1809.
18. Zheng B, Liao Z, Locascio JJ, et al. PGC-1alpha, a potential therapeutic target for early intervention in Parkinson's disease. Sci Transl Med 2010;2:52ra73.
19. Durrenberger PF, Fernando FS, Magliozzi R, et al. Selection of novel reference genes for use in the human central nervous system: a BrainNet Europe Study. Acta Neuropathol 2012;124:893-903.
20. Moran LB, Croisier E, Duke DC, et al. Analysis of alpha-synuclein, dopamine and parkin pathways in neuropathologically confirmed parkinsonian nigra. Acta Neuropathol 2007;113:253-263.
21. Moran LB, Graeber MB. Towards a pathway definition of Parkinson's disease: a complex disorder with links to cancer, diabetes and inflammation. Neurogenetics 2008;9:1-13.
22. Kim YJ, Lyoo CH, Hong S, Kim NY, Lee MS. Neuroimaging studies and whole exome sequencing of PLA2G6-associated neurodegeneration in a family with intrafamilial phenotypic heterogeneity. Parkinsonism Related Disord 2015;21:402-406.
23. Wu Z, Irizarry RA. Preprocessing of oligonucleotide array data. Nat Biotechnol 2004;22:656-658; author reply, 658.
24. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267-273.
25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102:15545-15550.
26. Durrenberger PF, Fernando FS, Kashefi SN, et al. Common mechanisms in neurodegeneration and neuroinflammation: a BrainNet Europe gene expression microarray study. J Neural Transm (Vienna) 2015;122:1055-1068.
27. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Methods 2009;3:486-504.
28. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1-13.
29. Lim J, Hao T, Shaw C, et al. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. Cell 2006;125:801-814.
30. Hedrich K, Djarmati A, Schafer N, et al. DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. Neurology 2004;62:389-394.
31. Bonifati V, Rohe CF, Breedveld GJ, et al. Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes. Neurology 2005;65:87-95.
32. Lesage S, Lohmann E, Tison F, Durif F, Durr A, Brice A. Rare heterozygous parkin variants in French early-onset Parkinson disease patients and controls. J Med Genet 2008;45:43-46.
33. Lin CH, Tan EK, Chen ML, et al. Novel ATP13A2 variant associated with Parkinson disease in Taiwan and Singapore. Neurology 2008;71:1727-1732.
34. Di Fonzo A, Dekker MC, Montagna P, et al. FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. Neurology 2009;72:240-245.
35. Sidransky E, Nalls MA, Aasly JO, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med 2009;361:1651-1661.
36. Tan EK, Ho P, Tan L, Prakash KM, Zhao Y. PLA2G6 mutations and Parkinson's disease. Ann Neurol 2010;67:148.
37. Choi JM, Kim WC, Lyoo CH, et al. Association of mutations in the glucocerebrosidase gene with Parkinson disease in a Korean population. Neurosci Lett 2012;514:12-15.
38. Quadri M, Fang M, Picillo M, et al. Mutation in the SYNJ1 gene associated with autosomal recessive, early-onset Parkinsonism. Hum Mutat 2013;34:1208-1215.
39. Chu MK, Kim WC, Choi JM, et al. Analysis of dosage mutation in PARK2 among Korean patients with early-onset or familial Parkinson's disease. J Clin Neurol 2014;10:244-248.
40. Mencacci NE, Isaias IU, Reich MM, et al. Parkinson's disease in GTP cyclohydrolase 1 mutation carriers. Brain 2014;137(Pt 9): 2480-2492.
41. Gan-Or Z, Orr-Urtreger A, Alcalay RN, Bressman S, Giladi N, Rouleau GA. The emerging role of SMPD1 mutations in Parkinson's disease: implications for future studies. Parkinsonism Relat Disord 2015;21:1294-1295.
42. Spataro N, Calafell F, Cervera-Carles L, et al. Mendelian genes for Parkinson's disease contribute to the sporadic forms of the disease. Hum Mol Genet 2015;24:2023-2034.
43. Zhang B, Xia C, Lin Q, Huang J. Identification of key pathways and transcription factors related to Parkinson disease in genome wide. Mol Biol Rep 2012;39:10881-10887.
44. Perrett RM, Alexopoulou Z, Tofaris GK. The endosomal pathway in Parkinson's disease. Mol Cell Neurosci 2015;66(Pt A):21-28.
45. Mariani E, Frabetti F, Tarozzi A, Pelleri MC, Pizzetti F, Casadei R. Meta-analysis of Parkinson's disease transcriptome data using TRAM software: whole substantia nigra tissue and single dopamine neuron differential gene expression. PLoS One 2016;11:e0161567.
46. Holmes SE, O'Hearn EE, McInnis MG, et al. Expansion of a novel CAG trinucleotide repeat in the 5′ region of PPP2R2B is associated with SCA12. Nat Genet 1999;23:391-392.
47. Nelson DL, Orr HT, Warren ST. The unstable repeats—three evolving faces of neurological disease. Neuron 2013;77:825-843.
48. Cattaneo E, Zuccato C, Tartari M. Normal huntingtin function: an alternative approach to Huntington's disease. Nat Rev Neurosci 2005;6:919-930.
49. Dragatsis I, Levine MS, Zeitlin S. Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. Nat Genet 2000;26:300-306.
50. Lin CH, Chen CM, Hou YT, et al. The CAG repeat in SCA12 functions as a cis element to up-regulate PPP2R2B expression. Hum Genet 2010;128:205-212.
51. Chen KF, Liu CY, Lin YC, et al. CIP2A mediates effects of bortezomib on phospho-Akt and apoptosis in hepatocellular carcinoma cells. Oncogene 2010;29:6257-6266.
52. Anderson JP, Walker DE, Goldstein JM, et al. Phosphorylation of Ser-129 is the dominant pathological modification of alpha-synuclein in familial and sporadic Lewy body disease. J Biol Chem 2006;281:29739-29752.
53. Lee KW, Chen W, Junn E, et al. Enhanced phosphatase activity attenuates alpha-synucleinopathy in a mouse model. J Neurosci 2011;31:6963-6971.
54. Perez-Revuelta BI, Hettich MM, Ciociaro A, et al. Metformin lowers Ser-129 phosphorylated alpha-synuclein levels via mTOR-dependent protein phosphatase 2A activation. Cell Death Dis 2014;5:e1209.
55. Braak H, Del Tredici K, Rub U, de Vos RA, Jansen Steur EN, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 2003;24:197-211.
56. Dehay B, Bourdenx M, Gorry P, et al. Targeting alpha-synuclein for treatment of Parkinson's disease: mechanistic and therapeutic considerations. Lancet Neurol 2015;14:855-866.

57. Maraganore DM, de Andrade M, Elbaz A, et al. Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. JAMA 2006;296:661-670.

58. Chiba-Falek O, Lopez GJ, Nussbaum RL. Levels of alpha-synuclein mRNA in sporadic Parkinson disease patients. Mov Disord 2006;21:1703-1708.

59. Papapetropoulos S, Adi N, Mash DC, Shehadeh L, Bishopric N. Expression of alpha-synuclein mRNA in Parkinson's disease. Mov Disord 2007;22:1057-1059; author reply, 1057.

60. Grundemann J, Schlaudraff F, Haeckel O, Liss B. Elevated alpha-synuclein mRNA levels in individual UV-laser-microdissected dopaminergic substantia nigra neurons in idiopathic Parkinson's disease. Nucleic Acids Res 2008;36:e38.

61. Gorbatyuk OS, Li S, Nash K, et al. In vivo RNAi-mediated alpha-synuclein silencing induces nigrostriatal degeneration. Mol Ther 2010;18:1450-1457.

62. Markopoulou K, Biernacka JM, Armasu SM, et al. Does alpha-synuclein have a dual and opposing effect in preclinical vs. clinical Parkinson's disease? Parkinsonism Relat Disord 2014;20:584-589; discussion, 584.

63. Maraganore DM. Rationale for therapeutic silencing of alpha-synuclein in Parkinson's disease. J Mov Disord 2011;4:1-7.

64. Foulger RE, Denny P, Hardy J, Martin MJ, Sawford T, Lovering RC. Using the gene ontology to annotate key players in Parkinson's disease. Neuroinformatics 2016;14:297-304.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.