<div style="text-align:center">

RESEARCH ARTICLE

The Endocytic Membrane Trafficking Pathway Plays a Major Role in the Risk of Parkinson's Disease

</div>

Sara Bandres-Ciga, PhD,[1,2] Sara Saez-Atienzar, PhD,[3] Luis Bonet-Ponce, PhD,[4] Kimberley Billingsley, MSc,[1,5,6] Dan Vitale, MSc,[7] Cornelis Blauwendraat, PhD,[1] Jesse Raphael Gibbs, PhD,[7] Lasse Pihlstrøm, MD, PhD,[8] Ziv Gan-Or, MD, PhD,[9,10] [iD] The International Parkinson's Disease Genomics Consortium (IPDGC), Mark R. Cookson, PhD,[4] Mike A. Nalls, PhD,[1,11] and Andrew B. Singleton, PhD[1*]

[1]Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA
[2]Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain
[3]Transgenics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA
[4]Cell Biology and Gene Expression Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA
[5]Department of Molecular and Clinical Pharmacology, Institute of Translational Medicine, University of Liverpool, Liverpool, United Kingdom
[6]Department of Pathophysiology, University of Tartu, Tartu, Estonia
[7]Computational Biology Group, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA
[8]Department of Neurology, Oslo University Hospital, Oslo, Norway
[9]Department of Neurology and Neurosurgery, Department of Human Genetics, McGill University, Montréal, Quebec, Canada
[10]Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montréal, Quebec, Canada
[11]Data Tecnica International, Glen Echo, Maryland, USA

ABSTRACT

Background: PD is a complex polygenic disorder. In recent years, several genes from the endocytic membrane-trafficking pathway have been suggested to contribute to disease etiology. However, a systematic analysis of pathway-specific genetic risk factors is yet to be performed.

Objectives: To comprehensively study the role of the endocytic membrane-trafficking pathway in the risk of PD.

Methods: Linkage disequilibrium score regression was used to estimate PD heritability explained by 252 genes involved in the endocytic membrane-trafficking pathway including genome-wide association studies data from 18,869 cases and 22,452 controls. We used pathway-specific single-nucleotide polymorphisms to construct a polygenic risk score reflecting the cumulative risk of common variants. To prioritize genes for follow-up functional

studies, summary-data based Mendelian randomization analyses were applied to explore possible functional genomic associations with expression or methylation quantitative trait loci.

Results: The heritability estimate attributed to endocytic membrane-trafficking pathway was 3.58% (standard error = 1.17). Excluding previously nominated PD endocytic membrane-trafficking pathway genes, the missing heritability was 2.21% (standard error = 0.42). Random heritability simulations were estimated to be 1.44% (standard deviation = 0.54), indicating that the unbiased total heritability explained by the endocytic membrane-trafficking pathway was 2.14%. Polygenic risk score based on endocytic membrane-trafficking pathway showed a 1.25 times increase of PD risk per standard deviation of genetic risk. Finally, Mendelian randomization identified 11 endocytic membrane-trafficking

---

*__Correspondence to:__ Dr. Andrew Singleton, Porter Neuroscience Center, 35 Convent Drive, Bethesda, MD 20892, USA; E-mail: singleta@mail.nih.gov

Dr. Bandres-Ciga, Dr. Saez-Atienzar, and Dr. Bonet-Ponce contributed equally to the work.

pathway genes showing functional consequence associated to PD risk.

Conclusions: We provide compelling genetic evidence that the endocytic membrane-trafficking pathway plays a relevant role in disease etiology. Further research on this pathway is warranted given that critical effort should be made to identify potential avenues within this biological process suitable for therapeutic interventions. © 2019 International Parkinson and Movement Disorder Society

Key Words: endocytosis; genetic risk; heritability; Parkinson's disease; polygenic risk score

Parkinson's disease (PD) is a progressive neurodegenerative disorder clinically characterized by the manifestation of motor disturbances such as bradykinesia, rigidity, tremor, and a wide range of nonmotor symptoms. Disease is caused by the death of the dopaminergic neurons in the SN with the consequent deficit of striatal dopamine.[1] The etiology of PD is complex, including genetic, epigenetic, and environmental factors. Familial PD accounts for less than 10% of the total cases, mostly attributed to monogenic forms with autosomal-dominant (including *SNCA*, *LRRK2*, and *VPS35*) or autosomal-recessive pattern of inheritance, such as *PARKIN*, *PINK1*, and *DJ-1*, among others.[2,3] However, the vast majority of the disease is considered to be sporadic, likely caused by the cumulative effect of multiple common or rare variants of small or moderate effect sizes in addition to other unknown environmental and stochastic factors.

Over the last two decades, extensive work in genetics has uncovered up to 92 common risk loci associated with idiopathic PD.[4] As expected from complex diseases, the PD-related signals identified to date only account for a small proportion of the total phenotypic variation, defined as heritability.[5] It is estimated that only around 10% to 15% of the overall heritable risk of PD is explained by current genome-wide association studies (GWAS).[4] The most plausible explanation for the "missing heritability" might be small effect sizes of the yet to be discovered risk loci which do not reach GWAS significance, as well as structural genomic and rare variation poorly covered by current technology.[6] Undoubtedly, the estimated heritability in PD suggests that a significant amount of risk has not been uncovered yet.

Aligned with the genetic discoveries, a big effort has been made to identify how those genes interact within biological pathways to elucidate the cellular and molecular processes that could explain the neurodegenerative mechanism.[7] Recent findings point out that dysfunction in the endosomal membrane-trafficking pathway (EMTP) could be a driving force in the pathogenesis of PD.[8] The EMTP is a complex multistep pathway that involves different process and organelles. First, cells uptake macromolecules and other cargo from the extracellular space to a group of organelles called early endosomes. From this point, internalized molecules can recycle back to the plasma membrane directly or through recycling endosomes. Alternatively, endocytosed cargo can be retained in early endosomes that will eventually mature into late endosomes and fuse with lysosomes to ensure cargo degradation. In parallel, proteins can also traffic between endosomes and the Golgi apparatus (for a review, see a previous work[9]). A substantial amount of genetic variation in PD and parkinsonism previously nominated genes has been associated with vesicle trafficking by endosomal gene alterations,[10] including monogenic genes (*VPS35*,[11] *DNAJC6*[13,18]) and GWAS-nominated genes (*RAB7L1*,[12,13] *SH3GL2*,[14] *GAK*,[15,16] and *CHMP2B*[15,17]). In addition, other well-known PD genes, such as *LRRK2*[12,13] and *PLA2G6*,[18] seem to be linked to the aforementioned pathway through their biological interaction with EMPT genes. LRRK2 binds to RAB7L1[12,13] and PLA2G6 interacts with VPS35 and VPS26, although a clear role of these genes in the EMTP is yet to be elucidated. The genes involved in the EMTP play different roles within the pathway. For instance, *DNAJC6*, *SYNJ1*, *GAK*, and *SH3GL2* are linked to clathrin-coated vesicles whereas *VPS35*, *DNAJC13*, or *RAB7L1* participate in recycling components from the endosomes to the Golgi. Hence, it seems plausible that other genes in that pathway could be involved in PD risk and therefore account for the nonreported heritability.

Based on these premises, we aimed to benefit from a cellular perspective to systematically explore the genetic contribution of 252 endosomal membrane-trafficking–related genes on the susceptibility for PD. We set out to study to what extent common variation in these genes might play a role on the genetic architecture of PD by performing genome-wide heritability estimations and polygenic risk versus disease status and age at onset in the International Parkinson's Disease Genomics Consortium (IPDGC) GWAS data set. We also applied Mendelian randomization to attempt to investigate possible functional genomic associations between nominated variants of interest within this pathway versus expression and methylation quantitative trait loci. Finally, we used biological pathway analysis and expression data from murine SN to point out the biological contribution of the genetic findings.

Patients and Methods

Subjects

The IPDGC GWAS data set consisted of 41,321 individuals (18,869 cases and 22,452 controls) of European

ancestry. Clinical and demographic characteristics of the cohorts under study are shown in Supporting Information Table S1. Additional details of these cohorts can be found in Nalls and colleagues.[4]

No individuals with known pathogenic mutations were excluded from following analyses. When the case age at onset or control age at last exam was not provided, individuals were likewise included.

Genes Selection and Pathway Analysis

The EMTP is a general and extense pathway that includes several subpathways. EMTP-related genes under study were selected by using The Molecular Signatures Database (MSigDB) of annotated gene set for "endocytosis"[19,20] based on the Kyoto Encyclopedia of Genes and Genomes information (KEGG-endocytosis). However, the gene list provided by KEGG was not completely accurate and failed to represent some relevant genes within the EMTP. In order to validate whether each gene flagged as endocytosis was involved in the EMTP, a detailed literature search was performed. In addition, because the current research in the EMTP is very dynamic and there are constantly new discovered proteins involved in some aspects of the pathway, we included relevant genes that have been proved to participate in the EMPT, but were not nominated by KEGG. In addition, a thorough review of the literature was performed using the terms "endocytosis," "endosome to Golgi retrograde transport," "endosomal maturation," and "endosomal recycling." We have provided a table (Supporting Information Table S2) that includes each gene involved in our analysis with the literature source that refers to. A total of 262 genes were flagged as "endocytic membrane-trafficking" (Supporting Information Table S2), of which 252 were autosomal and 10 spanned in chromosome X. The latter were excluded for the following analyses because of lack of data. The PANTHER web-based Classification System was used to obtain the Top 10 significant enriched Reactome Pathways with an adjusted false discovery rate (FDR) P value <0.05. Additionally, we obtained expression data from single-cell RNA sequencing performed in 243 putative dopaminergic neurons at postnatal days 28 to 56 from Slc6a3-Cre/tdTomato mice (GSE76381).[21] We applied an arbitrary selected threshold of expression. Only genes with total expression values across the 243 cells higher than 100 were nominated as expressed. Human genes were converted to their ortholog genes using Mouse Genome Informatics (MGI; http://www.informatics.jax.org; Supporting Information Table S3). For Venn diagram visualization, the online-based website http://bioinformatics.psb.ugent.be/webtools/Venn/ and Inkscape software (version 0.92.2; https://inkscape.org) were used.

Quality Control
Genome-Wide Association Data

Quality-control (QC) procedures in the IPDGC GWAS data set have been previously described elsewhere.[4] In brief, samples with call rates <95% and whose genetically determined sex from X chromosome heterogeneity did not match that from clinical data were excluded. Individuals exhibiting heterozygosity greater than 6 standard deviations (SDs) from the population mean and samples of non-European ancestry were not included. Individuals related at more than the level of first cousin estimated from linkage pruned genetic data on common variants were excluded (1 proband from any family cluster was extracted randomly). Regarding variant QC, variants exhibiting Hardy–Weinberg equilibrium P value <1E-5 and missingness rates >5% as well as palindromic single-nucleotide polymorphisms (SNPs) were excluded from further analyses. Data were imputed using the Haplotype Reference Consortium (http://www.haplotype-reference-consortium.org), under default settings with phasing using the EAGLE option (https://www.nature.com/articles/ng.3643 https://www.ncbi.nlm.nih.gov/pubmed/27571263). VCFs were filtered for imputation quality >0.30 for inclusion in analyses.

Heritability Analysis

In order to estimate the phenotypic variance of PD explained by genes in the extended EMTP pathway, we estimated heritability using linkage disequilibrium (LD) score regression (LDSC).[22,23] We used LDSC to estimate heritability in this study because it requires only summary-level data and is more computationally efficient for larger data sets. We extracted EMTP regions of interest from a meta-analysis of GWAS data and then ran LDSC under default settings. Reference LD scores were computed with the European ancestry subset of the 1000 Genomes data for SNPs within 500 kilobase pairs (kb) of the SNP to be scored.

First, we estimated heritability explained by variants in the whole set of EMTP genes and missing heritability by excluding previously PD-associated, EMTP-related genes. Subsequently, to evaluate significant enrichment of heritability within our pathway of interest compared to any similar set of genes, we resampled 50 random sets of 252 genes in the same sample series (after excluding EMTP genes and regions ±250 kb from PD GWAS loci). During resampling, we implemented the exact sample LDSC analyses and extracted the heritability per sampling iteration. Next, we calculated the mean and SDs of the heritability across sampling iterations. Finally, the mean heritability of our resampled series was compared to the test set to determine whether our previous estimates were significantly

divergent from random sets of genes of similar size using a t test.

In an attempt to explore the genes that contribute the most to the estimated heritability, we performed the Sequence Kernel Association Test (SKAT) in imputed data adjusting by age, sex, and data subset membership without applying any frequency filtering. Significance was evaluated using 10,000 permutations at alpha = 0.001.

Polygenic Risk Profiles Versus Disease Status and Age at Onset

For the risk-profiling analysis, we mirrored the workflow used in the most recent PD meta-analysis to date,[4] with the exception that before running this analysis pipeline, we extracted only genotypes from the pathway of interest. In brief, PRS profiling is calculated traditionally based on weighted allele dose manner as per previous work.[25-32] We used the R package PRSice2.[33] Permutation testing and P-value aware LD pruning was used to identify best P thresholds in external GWAS data to construct the PRS, allowing us to utilize variants below the common GWAS significance threshold of 5E-08. External summary statistics utilized in this phase of analysis included data from leave-one-out meta-analyses that exclude the study in which the PRS was being tested, avoiding overfitting/circularity to some degree. LD clumping was implemented under default settings (window size = 250 kb; $r^2 > 0.1$), and for each data set, 10,000 permutations were used to generate empirical P estimates for each GWAS-derived P threshold ranging from 5E-08 to 0.5, at a minimum increment of 5E-08. Each permutation test in each data set provided a Nagelkerke's pseudo r^2 after adjustment for an estimated prevalence of 0.005 (aged population estimate as per Gasser and colleagues)[33] and study-specific eigenvector 1 to 5, age, and sex as covariates. GWAS-derived P threshold with the highest pseudo r^2 was selected for further analysis. Summary statistics were meta-analyzed using random effects (restricted/residual maximum likelihood) per study-specific data set by using the R package PRSice-2.[34]

For the age at onset risk profiling, we performed similar analyses with the exception that we used age at onset as a continuous variable.

Mendelian Randomization Analysis to Explore Pathway-Specific Quantitative Trait Loci

We utilized two-sample Mendelian randomization to investigate possible functional genomic associations between nominated variants of interest in the EMTP pathway versus expression and methylation patterns. Quantitative trait loci (QTL) association summary statistics across well-curated methylation and expression data sets from the *SMR* (http://cnsgenomics.com/software/smr)[35] were compared to PD outcome summary statistics[4] after extracting the pathway-specific variants considered as the instrumental variables. These include estimates for methylation and cis-expression from the Genotype-Tissue Expression (GTEx) Consortium (v6; whole blood and 10 brain regions),[36] the CommonMind Consortium (CMC; dorsolateral prefrontal cortex),[37] the Religious Orders Study and Memory and Aging Project (ROSMAP),[38] and the Brain eQTL Almanac project (Braineac; 10 brain regions),[39] as reported elsewhere. Furthermore, we studied expression patterns in blood from the largest eQTL meta-analysis so far.[40]

Wald ratios were generated for each of the instrumental variable SNPs tagging a cis-QTL (probes within each gene and meeting a QTL P value of at least 5E-8 in the original QTL study) and for a methylation or expression probe with a nearby gene. Linkage pruning and clumping were carried out using default *SMR* protocols. With each data type, P values per instrument substrate (gene-level expression summaries for eQTLs and cis-gene probe level for mQTLs) were adjusted by FDR.

Results

Pathway Analysis and Expression of EMTP-Related Genes

EMTP variants from imputed data were extracted from 252 genes selected as described in the Patients and Methods section. Distribution of variants within the EMTP in the IPDGC cohort are highlighted in Supporting Information Table S4. Gene Ontology analysis was performed on these genes showing an enrichment for trafficking-related biological processes such as lysosomal biogenesis, clathrin-mediated endocytosis, retrograde transport at the *trans*-Golgi network or vesicle-mediated transport among others (Fig. 1A). Remarkably, 151 genes among the analyzed genes were found expressed in dopaminergic neurons in adult mouse brains (Fig. 1B).

Heritability Analysis

We measured heritability explained by the 252 genes under study. EMTP genetic contribution to PD heritability was estimated to be 3.58% (standard error [SE] = 1.17). After excluding previously nominated PD and parkinsonism genes related to the EMTP, including *DNAJC13*, *VPS35*, *DNAJC6*, *SYNJ1*, *SYT4*, *SYT11*, *GAK*, *RAB7L1*, *CHMP2B*, and *SH3GL2*, missing heritability explained by EMTP-related genes was estimated to be 2.21% (SE = 0.42). Reported heritability estimates might be underestimated given that there are possibly rare variants within these genes poorly covered by current genotyping technologies or lost after imputation. These estimates were significantly different from 50 randomly generated sets of 252 genes (t test, $P < 2.2E-16$) giving a heritability estimate of 1.44%
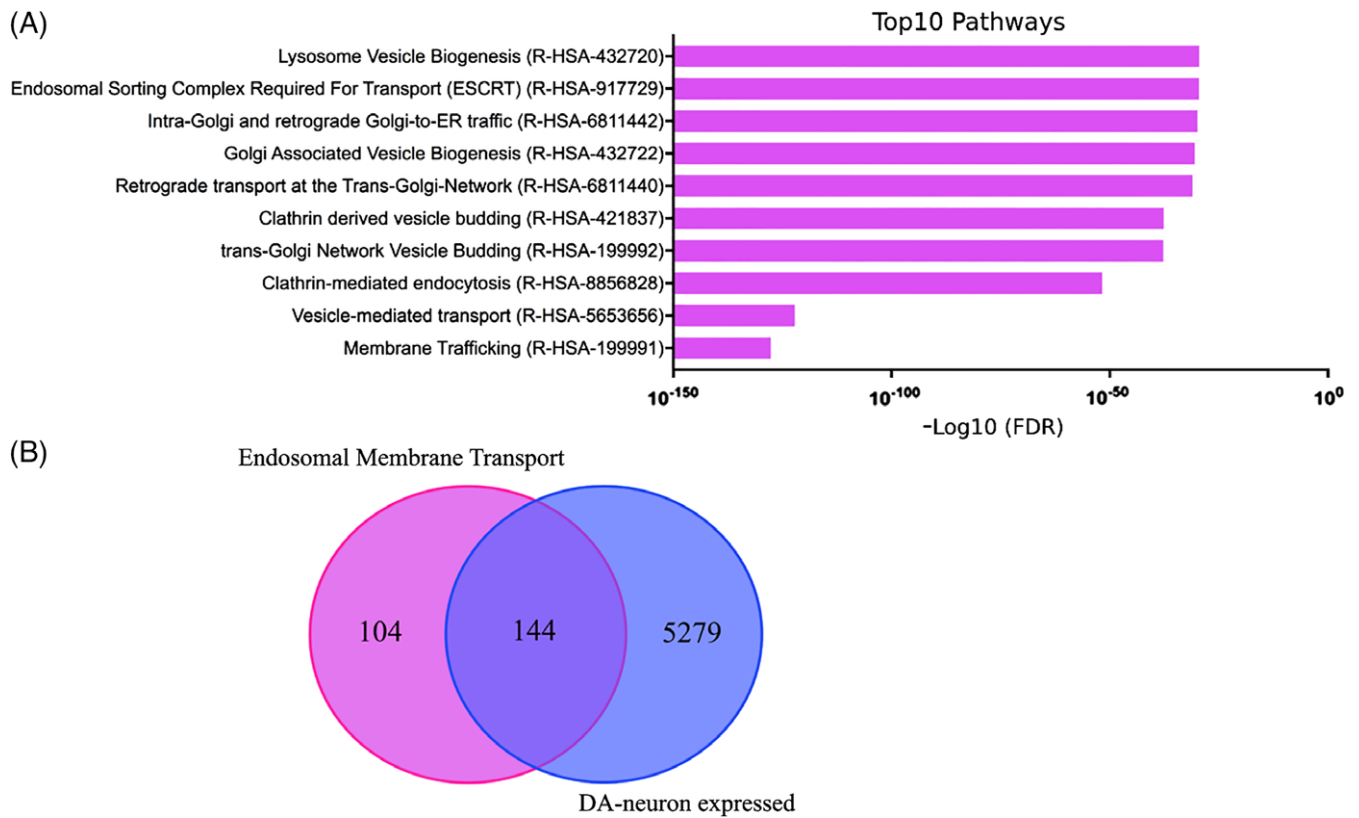
(A)

(B)

FIG. 1. Biological implication of EMPT genes in dopaminergic (DA) neurons. (A) Top 10 significant Enriched Reactome Pathways with an adjusted FDR *P* value <0.05. Pathway enrichment analysis was performed using Panther. Only the top 10 enriched pathways were selected for visualization. (B) Overlap between genes expressed in DA neurons and EMTP gene set. The EMTP human genes were transformed in their mouse homolog genes. From the 252 total human genes, a total of 244 were found in mouse (Supporting Information Table S3). DA neuron gene expression was obtained from GSE76381)[21] after an arbitrary selected threshold of expression. [Color figure can be viewed at wileyonlinelibrary.com]

(SD = 0.54) and showing that the heritability enrichment in the EMTP pathway is not likely by chance (Supporting Information Fig. S1). After subtracting the randomly generated heritability estimates (1.44%) that might have biased our initial EMTP estimates, the EMTP genetic contribution to PD heritability was estimated to be 2.14%. When considering 22.9% as the latest estimate of total heritability for PD using the same methodological approach,[4] the EMTP explains roughly 9.34% of the overall heritability of PD.

Additionally, we explored what genes contributed to a greater extent to the PD heritability estimates within the EMTP. Nineteen genes surpassed multiple test correction after the SKAT analysis with FDR-adjusted *P* values <0.05 (Supporting Information Table S5). Only *GAK* has been previously associated with PD.

Polygenic Risk Profiles Versus Disease Status and Age at Onset

EMTP risk profiling versus disease status was found to be significantly associated with PD (random-effects *P* value = 2.55E-12; beta = 0.227; SE = 0.032) and an odds ratio of 1.20 per SD increase in the PRS from the population mean (see Fig. 2 for forest plot showing

effect estimates within gene sets). Heterogeneity was estimated at an I^2 of 59.32% across all arrays. In all data sets, the directionality of effect was similar. The data obtained from the different data sets strongly enhance the general conclusion for our study. EMTP risk profiling versus age at onset was not found to be significantly associated (random-effects *P* value = 0.182).

Mendelian Randomization Analysis to Explore Pathway-Specific QTL

We also tested a number of QTL resources across varied tissues of interest to make functional inferences relating to biological features that might be involved in the genetic etiology of the disease. After adjustment for FDR correction, 11 EMTP-related genes showed functional consequence by two-sample Mendelian randomization in expression and methylation datasets (see Table 1). Increased blood expression of *VAMP4*, *ARL8B*, and *GAK* was found to be inversely associated to PD risk, whereas a positive risk association was found for *RABGEF1*, *VAMP8*, *CLTCL1*, and *ITSN1*. Increased brain expression of *SH3GL2* and *GAK* was causally linked to PD risk whereas *ARL8B* expression
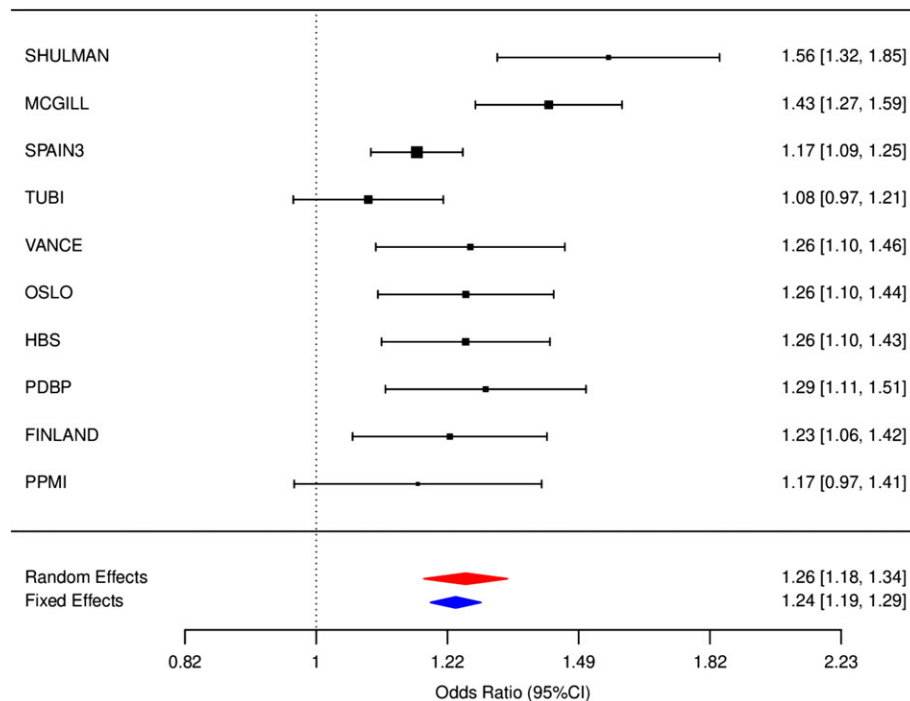
FIG. 2. Polygenic Risk Score across EMTP. Forest plots of PRS estimates across cohorts. For each effect estimate, the size of the square is proportional to the size of the study, with 95% confidence intervals as horizontal error bars. The random-effects estimates are represented by the red diamond, with fixed effects in blue, and the centerline of each diamond representing the summary PRS estimate. The number of samples per cohort are reported as follows: SHULMAN: 789 cases and 195 controls; MCGILL: 583 cases and 906 controls; SPAIN3: 2,120 cases and 1,333 controls; TUBI: 741 cases and 944 controls; VANCE: 621 cases and 303 controls; OSLO: 476 cases and 462 controls; HBS: 541 cases and 743 controls; PDBP: 543 cases and 284 controls; FINLAND: 386 cases and 493 controls; and PPMI: 363 cases and 165 controls. [Color figure can be viewed at wileyonlinelibrary.com]

exhibited a negative association. mQTL Mendelian randomization analyses revealed eight CpG sites linked to PD risk after multiple test correction. Increased methylation of *GAK* and *HSPA1B* was found to be positively associated with disease whereas *PLEKHM1* and *VPS39* showed a negative directionality versus PD risk.

Discussion

The present study provides compelling evidence for EMTP variation in PD risk. Using a set of pathway-specific genes/variants, we demonstrate that the phenotypic variance attributed to this complex process contributes to the heritability of PD outside of what is explained by GWAS.

Genetic variance in PD has been estimated to be ~22.9%,[4] and to date only a small proportion of the total heritability has been explained.[4] Our data show that approximately 2.14% of the common variant heritability in PD can be explained by genetic variants in the EMTP, explaining roughly 9.34% of the overall estimates in PD. Here, we have only included the transcribed genes assuming the limitation that the heritability residing in regulatory regions adjacent to the up- or downstream of these genes is not included in this estimate. The reported estimates suggest that there are additional PD genes within the EMTP yet to be discovered. Our results are in concordance with the largest and most recent meta-analysis in PD, which tested up to 10,651 biological pathways. Among these, only 27 were enriched for PD associations after multiple testing correction,[4] and, notably, endosomal trafficking was included with significant *P* value.

PRS has previously demonstrated how cumulative small effect variants can contribute to PD risk and age at onset.[30] PRS has emerged as a probabilistic tool to infer how independent loci with small effects can be collapsed into a single numerical value useful in predicting the risk of a disease. It has been successfully used to dissect the genetic complexity of brain disorder and quantify risk in many common complex conditions[42,43] for early detection, prevention and therapeutics.[44,45] Here, we assessed the overall cumulative contribution of EMTP-specific loci/genes in PD risk through PRS calculations. Our data show that the PRS based on 252 EMTP gene regions is associated with a 1.2 times increased of PD risk per SD, confirming a cumulative pathogenic effect of EMPT variants in PD.

When assessing whether the EMTP had any effect on the age at onset (AAO), we did not find any significant association. PD genetic studies using AAO as a predictor report greater genetic burden with earlier AAO. [29,46] Linking specific phenotypic aspects of PD to pathways constitutes a critical effort, but remains challenging given that larger longitudinal cohorts of well clinically

TABLE 1. Significant functional associations for EMTP-related genes by two-sample Mendelian randomization

Gene	Probe	Beta	SE	P Adjusted	Data Source	Analyte	Top QTL SNP	Chr	Bp
VAMP4	ILMN_1761363	−0.121	0.022	5.21E-06	Expression[40]	Blood	rs10913529	1	171679465
VAMP4	ILMN_1804676	−0.156	0.029	5.21E-06	Expression[40]	Blood	rs11580522	1	171720978
ARL8B	ILMN_1752837	−0.174	0.045	4.80E-03	Expression[40]	Blood	rs4499578	3	5167618
GAK	ILMN_1813775	−0.090	0.025	1.12E-02	Expression[40]	Blood	rs6964	4	843184
RABGEF1	ILMN_2230577	0.266	0.080	1.94E-02	Expression[40]	Blood	rs4718424	7	66376095
VAMP8	ILMN_2190084	0.041	0.012	2.20E-02	Expression[40]	Blood	rs1058588	2	85808871
ITSN1	ILMN_1718769	0.077	0.025	3.66E-02	Expression[40]	Blood	rs2251854	21	35056247
CLTCL1	ILMN_1694584	0.066	0.022	3.88E-02	Expression[40]	Blood	rs2854643	22	19189706
SH3GL2	ENSG00000107295	0.252	0.045	2.45E-06	Expression[41]	Brain	rs10756899	9	17684784
GAK	ENSG00000178950	0.508	0.103	3.47E-05	Expression[41]	Brain	rs11248057	4	906131
ARL8B	ENSG00000134108	−0.033	0.010	1.90E-02	Expression[41]	Brain	rs6787725	3	5201054
GAK	cg14517359	0.158	0.030	1.51E-05	Methylation[41]	Brain	rs3775127	4	886061
HSPA1B	cg00970279	0.048	0.010	4.99E-05	Methylation[41]	Brain	rs506770	6	31785228
HSPA1B	cg16237409	0.051	0.011	4.99E-05	Methylation[41]	Brain	rs506770	6	31785228
PLEKHM1	cg06925179	−0.094	0.026	4.92E-03	Methylation[41]	Brain	rs4523963	17	43570226
GAK	cg23992470	0.093	0.027	1.20E-02	Methylation[41]	Brain	rs114066161	4	825635
PLEKHM1	cg04703951	−0.144	0.043	1.20E-02	Methylation[41]	Brain	rs4523963	17	43570226
VPS39	cg04726019	−0.092	0.028	1.28E-02	Methylation[41]	Brain	rs57393446	15	42494536
GAK	cg26978381	0.119	0.037	1.38E-02	Methylation[41]	Brain	rs2306253	4	884475

P adjusted = adjusted false discovery rate *P* value after performing Mendelian randomization between QTL resources (exposures) versus Nalls and colleagues[4] (outcome).
Abbreviations: Chr, chromosome; Bp, base pair.

characterized patients are necessary. We believe that the PRS approach we report on here should be further explored in other clinical outcomes related to PD.

Growing evidence has suggested that abnormalities in endosomes or dysregulation in their trafficking plays an important role in several neurological diseases such as Alzheimer's disease or Lewy body dementia.[47,48] This pathway is emerging as key to understanding the mechanisms underlying both protein degradation and neurodegeneration. We performed SKAT analysis in order to nominate the genes that contribute the most to the risk of PD within the EMTP. Nineteen genes were significant after FDR correction, including *PLEKHM1*, *IQSEC3*, *GAK*, *VPS39*, *AMPH*, *SMURF1*, *CHMP1A*, *AP2B1*, *VPS28*, *VAMP8*, *AP3B2*, *RABGEF1*, *PLEKHG7*, *TBC1D15*, *AP4E1*, *AP2M1*, *STX8*, *ARF-GAP2*, and *AP3M1*. Interestingly, nine of them are involved in cargo degradation and eight of them play a role in cargo uptake, which strongly suggests that these two subpathways are specially relevant for PD risk. Only *GAK* has already been associated with PD risk through GWAS,[4,30,31] and for the remaining 18 genes not identified in previous studies, only five, including *IQSEC3*, *VPS39*, *CHMP1A*, *AP3B2*, and *AP4E1*, have been associated with neurological disorders. For instance, IQSEC3 protein expression is downregulated in the brain of patients that suffer from autism and schizophrenia,[49] and loss-of-function mutations in *CHMP1A* cause pontocerebellar hypoplasia and microcephaly.[50] Autosomal-recessive mutations in *AP3B2* cause early-onset epileptic encephalopathy,[51] and a mutation in *AP4E1* has been found in a familiar case of spastic paraplegia.[52] Furthermore, QTL analysis using

Mendelian randomization indicates that five genes among the ones nominated by SKAT showed functional relevance associated with PD (*PLEKHM1*, *VPS39*, *GAK*, *VAMP8*, and *RABGEF1*).

The present study shows that EMTP genes contribute to PD risk. However, we are aware of some limitations. First, our work is restricted by the sensitivity of array-based genotyping methods used to calculate heritability. Likewise, heritability estimates might be biased, appearing lower than what should be expected. The remaining variance in PD etiology not ascribed to the EMTP-identified heritable factors suggests a contribution of joint effects of rare and structural variants poorly covered by current genotyping technology. Therefore, further comprehensive studies are needed to identify novel variants within the EMTP that are relevant for disease. Additionally, there may be some selection bias in the PRS, although the meta-analysis suggests consistent effect estimates across studies.

The endosomal trafficking membrane system is a complex and dynamic process where internalized proteins, receptors, and ligands interact and move between organelles in order to fulfill the cell's needs. Here, we point to PD, at least to some extent, as an endosomal membrane-trafficking disease. This study demonstrates the contribution of this unknown pathway to the oligogenic nature of PD and highlights the importance of pathway-based comprehensive approaches as a way to prioritize functional studies. Following the most recent meta-analysis,[4] pathway analysis established that the nominated candidate genes were linked to autophagy, mitochondrial biology, immune response, and lysosomal function. We envisage that our methodology

could be applied to remaining relevant pathways in PD as a way to better understand genetic risk factors. Unraveling the genetics underlying PD will provide with therapeutic options for drug discovery and ultimately could lead to the development of effective interventions. ∎

References

1. Surmeier DJ, Obeso JA, Halliday GM. Selective neuronal vulnerability in Parkinson disease. Nat Rev Neurosci 2017;18:101–113.

2. Hernandez DG, Reed X, Singleton AB. Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. J Neurochem 2016;139(Suppl 1):59–74.

3. Billingsley KJ, Bandres-Ciga S, Saez-Atienzar S, Singleton AB. Genetic risk factors in Parkinson's disease. Cell Tissue Res 2018; 373:9–20.

4. Nalls MA, et al. Parkinson's disease genetics: identifying novel risk loci, providing causal insights and improving estimates of heritable risk. bioRxiv 2018 Aug 9. https://doi.org/10.1101/388165.

5. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88: 76–82.

6. Maher B. Personal genomes: the case of the missing heritability. Nature 2008;456:18–21.

7. Fujita KA, et al. Integrating pathways of Parkinson's disease in a molecular interaction map. Mol Neurobiol 2014;49:88–102.

8. Abeliovich A, Gitler AD. Defects in trafficking bridge Parkinson's disease pathology and genetics. Nature 2016;539:207–216.

9. Schreij AM, Fon EA, McPherson PS. Endocytic membrane trafficking and neurodegenerative disease. Cell Mol Life Sci 2016;73: 1529–1545.

10. Perrett RM, Alexopoulou Z, Tofaris GK. The endosomal pathway in Parkinson's disease. Mol Cell Neurosci 2015;66:21–28.

11. McGough IJ, et al. Retromer binding to FAM21 and the WASH complex is perturbed by the Parkinson disease-linked VPS35 (D620N) mutation. Curr Biol 2014;24:1670–1676.

12. Beilina A, et al. Unbiased screen for interactors of leucine-rich repeat kinase 2 supports a common pathway for sporadic and familial Parkinson disease. Proc Natl Acad Sci U S A 2014;111:2626–2631.

13. MacLeod DA, et al. RAB7L1 interacts with LRRK2 to modify intraneuronal protein sorting and Parkinson's disease risk. Neuron 2013; 77:425–439.

14. Ambroso MR, Hegde BG, Langen R. Endophilin A1 induces different membrane shapes using a conformational switch that is regulated by phosphorylation. Proc Natl Acad Sci U S A 2014;111: 6982–6987.

15. Song L, et al. Auxilin underlies progressive locomotor deficits and dopaminergic neuron loss in a Drosophila model of Parkinson's disease. Cell Rep 2017;18:1132–1143.

16. Yim YI, et al. Endocytosis and clathrin-uncoating defects at synapses of auxilin knockout mice. Proc Natl Acad Sci U S A 2010;107: 4412–4417.

17. Stuchell-Brereton MD, et al. ESCRT-III recognition by VPS4 ATPases. Nature 2007;449:740–744.

18. Lin G, et al. Phospholipase PLA2G6, a parkinsonism-associated gene, affects Vps26 and Vps35, retromer function, and ceramide levels, similar to α-synuclein gain. Cell Metab 2018;28:605–618.e6.

19. Subramanian A, 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.

20. Liberzon A, et al. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems 2015;1:417–425.

21. La Manno G, et al. Molecular diversity of midbrain development in mouse, human, and stem cells. Cell 2016;167:566–580.e19.

22. Zheng J, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 2017;33:272–279.

23. Bulik-Sullivan BK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–295.

24. Finucane HK, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 2015;47:1228–1235.

25. International Parkinson Disease Genomics Consortium, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet 2011;377:641–649.

26. International Parkinson's Disease Genomics Consortium (IPDGC) & Wellcome Trust Case Control Consortium 2 (WTCCC2). A two-stage meta-analysis identifies several new loci for Parkinson's disease. PLoS Genet 2011;7:e1002142.

27. Hernandez DG, et al. Genome wide assessment of young onset Parkinson's disease from Finland. PLoS One 2012;7:e41859.

28. Ripatti S, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. Lancet 2010; 376:1393–1400.

29. Nalls MA, et al. Genetic risk and age in Parkinson's disease: continuum not stratum. Mov Disord 2015;30:850–854.

30. Nalls MA, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet 2014;46:989–993.

31. Chang D, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat Genet 2017;49: 1511–1516.

32. Bandrés-Ciga S, et al. Genome-wide assessment of Parkinson's disease in a Southern Spanish population. Neurobiol Aging 2016;45: 213.e3–213.e9.

33. Gasser T. Genetics of parkinson's disease. Ann. Neurol. 1998;44: S53–S57.

34. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2014;31:1466–1468.

35. Zhu Z, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016;48: 481–487.

36. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648–660.

37. Fromer M, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci 2016;19:1442–1453.

38. Ng B, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat Neurosci 2017; 20:1418–1426.

39. Trabzuni, D. et al. Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. J Neurochem 2011;119:275–282.

40. Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013;45: 1238–1243.

41. Qi, T, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nature Communications 2018;9:2282.

42. Wang T, et al. Polygenic risk for five psychiatric disorders and cross-disorder and disorder-specific neural connectivity in two independent populations. Neuroimage Clin 2017;14:441–449.

43. Reus LM, et al. Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in UK Biobank. Sci Rep 2017;7:42140.

44. Natarajan P, et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. Circulation 2017; 135:2091–2101.

45. Maas P, et al. Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States. JAMA Oncol 2016;2:1295–1302.

46. Lill CM, et al. Impact of Parkinson's disease risk loci on age at onset. Mov Disord 2015;30:847–850.

47. Gao S, Casey AE, Sargeant TJ, Mäkinen VP. Genetic variation within endolysosomal system is associated with late-onset Alzheimer's disease. Brain 2018;141:2711–2720.

48. Neefjes J, van der Kant R. Stuck in traffic: an emerging theme in diseases of the nervous system. Trends Neurosci 2014;37:66–76.

49. Ellis SE, Panitch R, West AB, Arking DE. Transcriptome analysis of cortical tissue reveals shared sets of downregulated genes in autism and schizophrenia. Transl Psychiatry 2016;6:e817.

50. Mochida GH, et al. CHMP1A encodes an essential regulator of BMI1-INK4A in cerebellar development. Nat Genet 2012;44:1260–1264.

51. Assoum M, et al. Autosomal-recessive mutations in AP3B2, adaptor-related protein complex 3 beta 2 subunit, cause an early-onset epileptic encephalopathy with optic atrophy. Am J Hum Genet 2016;99:1368–1376.

52. Kong XF, et al. A novel homozygous p.R1105X mutation of the AP4E1 gene in twins with hereditary spastic paraplegia and mycobacterial disease. PLoS One 2013;8:e58286.

53. Robak LA, et al. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain 2017;140: 3191–3203.

Supporting Data

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