

Predicting risk of Alzheimer's disease using polygenic risk scores developed for Parkinson's disease

Britney E. GRAHAM, PhD¹, Scott M. WILLIAMS, PhD² and Jason H. MOORE, PhD¹

¹ *Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA*

² *Departments of Population and Quantitative Health Sciences and Genetics and Genome Sciences, Cleveland Institute for Computational Biology, Case Western Reserve University, Cleveland, Ohio 44106, USA*

Abstract:

Background and Objectives:

The two most common neurodegenerative diseases are Alzheimer's disease (AD) and Parkinson's disease (PD), both related to age and affect millions of people across the world, especially as life expectancy increases in certain countries. Here, we explore the potential predictiveness of the genetic risk of AD and PD separately and then the extent of the underlying shared genetics of AD and PD.

Methods:

The population genetic risk estimates for AD and PD were derived using a previously developed population specific polygenic risk score (psPRS) and regression-based SNP filtering method. To test the overlap between AD and PD, we ran a regression of the AD psPRSs versus the population PD prevalences for both the filtered and unfiltered AD PRS and vice versa. We then assessed gene-gene interaction and pathway involvement using the Alzheimer's KnowledgeBase ([AlzKB](#)) and STRINGdb, respectively.

Results:

The unfiltered PD psPRS was moderately predictive, while the AD psPRS was not. After filtering, both the AD and PD, psPRSs improved to strongly predictive, explained most of the genetic variation. The ability of the unfiltered AD psPRS to predict PD, and vice versa, is poor. However, the filtered AD, and PD, psPRS were highly predictive.

Discussion:

Our results suggest that there is a correlation between AD, and PD, specific allele frequency and prevalence, as well as an overlap of AD and PD generally. However, the AD psPRS is a better predictor of PD, than the PD psPRS is of AD. Our results call for further research into the general overlap of Alzheimer's disease and Parkinson's disease, despite the previous lack of evidence.

Introduction

Neurodegenerative diseases affect millions of people across the world, and the two most common are Alzheimer's disease (AD) and Parkinson's disease (PD), both of which are age-related. Other neurodegenerative diseases include Huntington's disease, Amyotrophic lateral sclerosis (ALS) and Motor neuron disease. All of these diseases are hallmarked by loss of nerve cell function in the brain or the peripheral nervous system¹ and have some genetic overlap with each other². These disorders, among other neurodegenerative diseases, are genetically complex with multiple causes, including genes and environment. Prevalence of these diseases varies among countries, due, possibly, to differences in socioeconomic status and life expectancy across populations. These diseases are becoming more prevalent everywhere as populations age. However, even age adjusted prevalence has increased, in PD more than in AD^{3,4}.

Identifying relevant risk factors for neurodegenerative diseases is important for early detection and prevention⁵. These diseases are relatively complex and have multiple components, including environmental and genetic. As multiple loci have been found that associate with these neurodegenerative diseases, it might be of use to combine these loci into a single genetic predictor. This can be done quantitatively by building a polygenic risk score (PRS).

PRS

Polygenic risk scores (PRS) were developed with the intent of predicting diseases or phenotypes and usually incorporate, without interaction, multiple known risk associated loci. This follows the idea that effects are additive as opposed to multiplicative, i.e. epistatic. Often, the PRS is weighted by average estimated effects of risk alleles and, occasionally, the population allele frequency^{6,7}. However, as effect sizes can vary among populations, this method may limit generalizability⁸ and accurate prediction of understudied populations⁹. Modeling phenotypes

with a PRS is more difficult when there is complex genetic architecture and environmental factors¹⁰.

Disease risk prediction using genetic factors is usually done for individuals. Examples of PRS use include predicting obesity and type 2 diabetes mellitus, various cancers and heart conditions, and IQ and mental illness at the individual level using population-derived associations. However, although population-level prevalence prediction is relatively new, we have previously shown that population-level, risk allele frequency-based, polygenic risk scores (PRSs) can explain prevalence variation, to varying degrees, among geographic populations¹¹. PRSs have been used in both AD and PD in the past but have mostly been able to account for only a small amount of heritability, although, there is some evidence that AD PRSs are transferable across populations¹². Here, using our previous method, we explore the potential predictiveness of the genetic risk of AD and PD separately¹¹. We then attempt to explain the extent of the underlying shared genetics of AD and PD.

AD-PD Similarities

Although AD and PD are distinct diseases, they have many similarities. Neither AD, nor PD, are curable and only palliative treatments are available. Both are progressive, neurodegenerative, related to advanced age, characterized by protein buildup in the brain and affect movement and cognition. Olfactory impairment is an early symptom in both AD and PD¹³

Despite phenotypic similarities, AD and PD are rarely co-morbid¹⁴, but, interestingly, family history of AD increases the risk of developing PD and vice versa^{15, 16}. While there has been indication of only minor genetic overlap between AD and PD, both AD and PD share differentially expressed genes and pathways¹⁷. Indeed, the KEGG PD pathway has a 78.2%

overlap with the KEGG AD pathway¹⁸. They also share some disease mechanisms¹⁹, treatments²⁰ and environmental risk factors²¹

Alzheimer's disease

First described in 1906 by Alois Alzheimer²², Alzheimer's disease (AD), the most common neurodegenerative disease, is marked by increasing loss of memory, reasoning and bodily functions, leading to death. There are two main forms of AD. The most common is late onset (LOAD, 95% of cases), where patients are diagnosed in their mid-60s or older with an average survival time after diagnosis of four to eight years. LOAD is, notably, more common in women²³.

Treatments for AD are palliative, except for aducanumab (targets the amyloid beta protein), which has had mixed results²⁴. Aside from age, risk factors include vascular disease, family history of AD, mid-life hypertension, obesity and diabetes²³. While AD does have an environmental component, AD heritability, estimated by twin studies, is 58%-79%²⁵. However, even when the strongest genetic risk factor for LOAD, *APOE* $\epsilon 4$ allele²⁶, is combined with other known genetic factors, the variance explained is still only 24–33%²⁷.

There is an unequal distribution of AD worldwide, even taking the average age of a population into account³. Western Europe has the highest prevalence of AD and dementia and Sub-Saharan Africa the lowest. The fastest increase in population AD prevalence is in China, India, and other south Asian and western Pacific countries³. This is especially impactful as the proportion of people over 65, and cost of care, increases across the world and the size of the workforce decreases in some areas²⁸.

Parkinson's disease

Parkinson's disease (PD), first described in 1817 by James Parkinson²⁹, is the second most common neurodegenerative disorder after AD¹. PD has a range of symptoms including uncontrollable movement, shaking, difficulty speaking and walking. As with AD, PD is usually diagnosed after the age of 60³⁰, with only 3–5% of cases occurring earlier³¹. Contrasting with AD, PD occurs more frequently in men than in women³². Although PD tends to primarily affect the motor system, 20% to 40% of PD patients develop PD specific dementia, but only after the onset of the other symptoms³³.

There is no cure for PD, but there are a variety of palliative treatments for symptoms, including a range of drugs, primarily levodopa³⁴. Occupational, physical, and speech therapy have been used to help with balance, gait speed and voice and speech disorders. Surgeries to achieve deep brain stimulation have been somewhat successful in improving motor function and help to decrease motor fluctuations, dyskinesia, and antiparkinsonian medication use³⁵.

While the actual cause of PD is still uncertain, there seem to be both environmental and genetic risk factors³⁶. The heritable portion of PD is estimated to be 22%³⁶, but only about 16–36% of that heritability has been explained³⁷. Confirmed dominant genes include *SNCA*, *LRRK2*, *GBA*, and *VPS35*, and recessive genes include *Parkin*, *PINK1*, *DJI*. Prevalence of PD is higher in people of European descent and appears to be lower in Africa⁴. However, prevalence data by ethnicity is sometime contradictory³⁸.

The goal of this study was to assess the overlap of PRS components for AD and PD and the cross-predictive value. We did this by assessing to what extent PRSs could predict population prevalence of each disease, and then each disease PRS for the other.

Materials and Methods

Allele and Prevalence Data Collection

We assessed the role of PRSs in predicting population prevalence among the populations included in the International Genome Sample Resource (IGSR) from the 1000 Genomes Project³⁹. Prevalence data for each population in the 1000 Genomes (Table 1) was collected from the literature, principally from country level data³. Using a threshold of $p < 1 \times 10^{-5}$, we extracted AD associated alleles from the GWAS catalog using Alzheimer's disease as a search term (MONDO_0004975). We then mined the allele frequencies for each SNP from the 1000 Genomes project for each population (Table S1).

For the prediction of PD, we followed the same procedure as for AD. Using Parkinson's disease as a search term (MONDO_0005180), we extracted the associated alleles at a threshold of $p < 1 \times 10^{-5}$ from the GWAS catalog. The corresponding allele frequencies were taken from the 1000 Genomes project for each population. Additionally, we collected prevalence data for PD from literature⁴.

Table 1. Relative Alzheimer's and Parkinson's disease prevalence among populations.			
1000 Genomes Super population	1000 Genomes Population	Alzheimer's disease prevalence ³	Parkinson's disease prevalence ⁴
African (AFR)	African Caribbean in Barbados (ACB)	7.78×10^{-3}	
	Americans of African Ancestry in SW USA (ASW)	1.37×10^{-2}	1.04×10^{-2}
	Esan in Nigeria (ESN)	8.21×10^{-4}	
	Gambian in Western Divisions in the Gambia (GWD)	7.89×10^{-4}	8.65×10^{-5}
	Luhya in Webuye, Kenya (LWK)	1.25×10^{-3}	1.34×10^{-4}
	Mende in Sierra Leone (MSL)	7.66×10^{-4}	8.69×10^{-5}
	Yoruba in Ibadan, Nigeria (YRI)	8.21×10^{-4}	1.12×10^{-4}
Ad Mixed American (AMR)	Colombians from Medellin, Colombia (CLM)	5.16×10^{-3}	5.38×10^{-4}
	Mexican Ancestry from Los Angeles USA (MXL)	5.51×10^{-3}	5.57×10^{-4}
	Peruvians from Lima, Peru (PEL)	3.49×10^{-3}	5.47×10^{-4}
	Puerto Ricans from Puerto Rico (PUR)	9.50×10^{-3}	1.26×10^{-3}
East Asian (EAS)	Chinese Dai in Xishuangbanna, China (CDX)	7.37×10^{-3}	1.01×10^{-3}
	Han Chinese in Beijing, China (CHB)	7.37×10^{-3}	1.01×10^{-3}
	Southern Han Chinese (CHS)	7.37×10^{-3}	1.01×10^{-3}
	Japanese in Tokyo, Japan (JPT)	2.76×10^{-2}	2.02×10^{-3}
	Kinh in Ho Chi Minh City, Vietnam (KHV)	7.00×10^{-3}	6.88×10^{-4}
European (EUR)	Utah Residents (CEPH) with Northern and Western European Ancestry (CEU)	1.69×10^{-2}	1.67×10^{-2}
	Finnish in Finland (FIN)	1.53×10^{-2}	1.87×10^{-3}
	British in England and Scotland (GBR)	1.27×10^{-2}	1.77×10^{-3}
	Iberian Population in Spain (IBS)	1.78×10^{-2}	2.00×10^{-3}
	Toscans in Italia (TSI)	2.26×10^{-2}	2.39×10^{-3}
South Asian (SAS)	Bengali from Bangladesh (BEB)	2.54×10^{-3}	3.43×10^{-4}
	Gujarati Indian from Houston, Texas (GIH)	2.22×10^{-3}	4.35×10^{-4}
	Indian Telugu from the UK (ITU)	2.22×10^{-3}	4.35×10^{-4}
	Punjabi from Lahore, Pakistan (PJL)	1.70×10^{-3}	2.76×10^{-4}
	Sri Lankan Tamil from the UK (STU)	6.06×10^{-3}	8.40×10^{-4}

Polygenic risk scores

The population genetic risk estimate was derived using a previously developed population specific polygenic risk score (psPRS)¹¹. This method uses the frequency of all disease risk alleles for a population to predict that population's disease prevalence. As effect size is sometimes unavailable and usually only based on a limited number of populations, the psPRS uses only the population risk allele frequencies. Using the simplest assumption that the genetic architecture of the phenotype of interest is additive, the psPRS was a summation of the population risk allele frequencies in each population. For a population in the 1000 Genomes database, psPRS is the PRS for that population and p_i is the allele frequency of SNP_i :

$$psPRS = \sum_{k=1}^i p_k$$

We then performed regression analyses between the psPRSs and population prevalences, establishing the relationship between the psPRS and the population prevalence of each disease.

Maximization of the coefficient of determination sensitivity analysis

To obtain an optimal and likely universal PRS, we then filtered the risk associated SNPs based on a maximization of the coefficient of determination (r^2) of the regression of the psPRS with the population prevalence. We eliminated SNPs that reduced the predictiveness of each psPRS by assessing the effect of removing individual SNPs from the psPRS.

As an example, the equation for the sensitivity analysis for SNP_2 would be:

$$psPRS = SNP_1 + SNP_3 + SNP_4 + \dots + SNP_k$$

Each SNP was annotated with the r^2 value calculated for the linear regression between the psPRS, without that SNP, and the population disease prevalence. SNPs that negatively affected the predictability of the model were removed. We then repeated this process until only the most predictive SNPs remained. Assuming additivity, we expect the model with the largest r^2 to include all true disease associating SNPs with universal effects, i.e., those that align with disease prevalence across a broad array of populations. This method, as well as the psPRS algorithm, was described in our previous work¹¹.

Intersection

Given the known similarities between AD and PD, we sought to determine if AD and PD are similar enough genetically that the AD PRS would be able to predict PD population prevalence and vice versa. We ran a regression of the AD psPRSs versus the population PD prevalences for both the filtered and unfiltered AD PRS. We then ran the PD psPRS versus the population AD prevalences for both the filtered and unfiltered PD PRS.

Gene-Gene interaction

To identify possible genetic links between AD and PD, we explored the underlying interaction of the AD and PD specific genes using the Alzheimer's KnowledgeBase ([AlzKB](#)), a graph database that incorporates various AD resources into one graph. The AlzKB nodes include genes, pathways, drugs, and diseases, and the edges represent various relationships between nodes, including, but not limited to, co-expression, chemical effect and gene-gene interaction. We queried the AlzKB with the genes that have been associated with PD in several studies⁴⁰⁻⁴², and determined which of the AD genes interacted directly with the PD genes, based on the

relationship term in the graph (GENEINTERACTSWITHGENE), which is in turn based on protein-protein interaction (PPI)⁴³.

Pathway analysis

We then assessed the possible pathway involvement of the interacting AD and PD genes from AlzKB using the STRING database of Protein-Protein Interaction Networks Functional Enrichment Analysis. STRINGdb does not just contain PPI networks, but also incorporates enrichment analysis using other databases, including Gene Ontology and KEGG⁴⁴.

Results

Alzheimer's Disease

There are 1386 AD-associated SNPs in the GWAS catalog (Table S1). The AD psPRS using these SNPs had a positive but non-significant relationship, with AD prevalence (Figure 1A, slope: 0.0005, R^2 : 0.11, p-value: 0.091). Intracontinental populations similarly had weak, non-significant relationships, apart from the south Asian and Ad Mixed American population clusters (Figure S1A). Both superpopulations had a moderate r^2 value; 0.49 and 0.42, respectively, although still without a significant p-value. However, the south Asian population cluster showed a negative relationship between the psPRS and AD prevalence.

After maximization filtering, 387 SNPs remained in the PRS model (Table S2). The relationship between the filtered psPRSs and AD was strong, positive and highly significant (Figure 1B, Slope: 0.0057, r^2 : 0.99, p-value: $< 2.2 \times 10^{-16}$) (Figure 1B). Intracontinental populations also showed the same patterns of significance and strength, although the south Asian-AD relationship switched direction from negative to positive (Figure S1B).

Parkinson's disease

There were 450 PD associated SNPs in the GWAS catalog (Table S3). The PD psPRS using these SNPs had a significant association with PD prevalence (Figure 1C., slope: 9×10^{-5} , r^2 : 0.33, p-value: 0.0045). The correlation between the PD psPRSs and the PD prevalences was best in the European superpopulation ($r^2 = 0.64$) but was not statistically significant (Figure S2A; p-value: 0.20).

After maximization filtering, 242 SNPs remained (Table S4), yielding a positive, very significant and predictive, relationship (Figure 1D, slope = 0.0007, $r^2 = 0.99$, p-value: $< 2.2 \times 10^{-16}$). The superpopulation correlations improved in general, with the European populations, again, performing the best (Figure S2B, $r^2 = 0.99$, p-value: 0.0044).

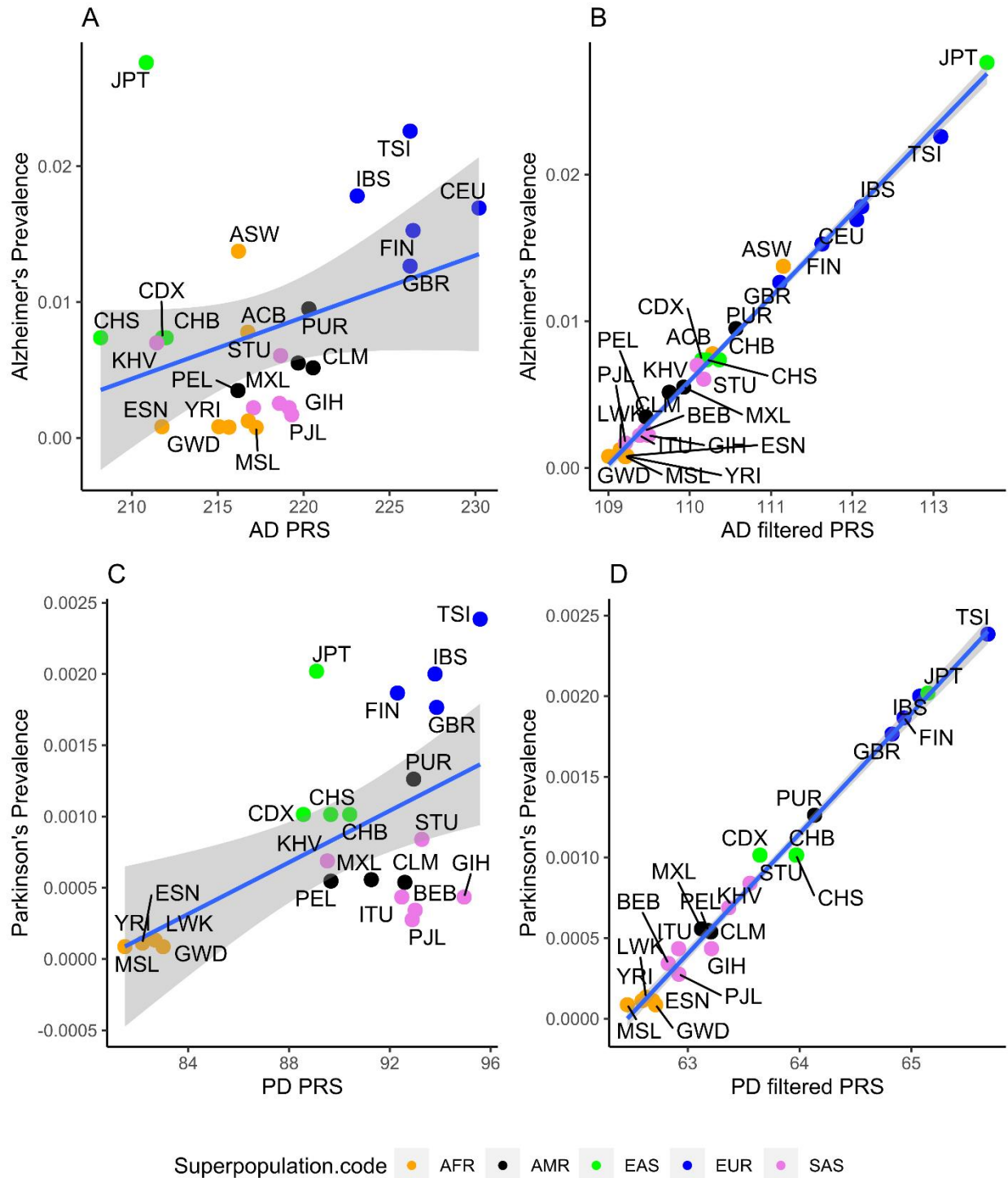


Figure 1. Correlation between diseases and the corresponding psPRSs. The data points are colored according to the super populations: AFR (orange), AMR (black), EAS (green), EUR (blue) and SAS (purple). The scale of the x-axis is not the same for both plots due to differing psPRS ranges. **A.** Full AD vs AD PRS model ($r^2=0.11$; p -value: 0.091). **B.** After maximization ($r^2=0.99$, p -value: $< 2.2 \times 10^{-16}$). **C.** Correlation between Parkinson's Disease and the PD full model psPRS ($r^2=0.33$; p -value: 0.0045). **D.** After maximization ($r^2=0.99$, p -value: $< 2.2 \times 10^{-16}$).

AD and PD Intersection

The regression analysis between the unfiltered AD psPRSs and PD prevalences showed a slightly positive, but only somewhat, predictive model (Figure 2A, slope: 6×10^{-5} , r^2 : 0.17, p-value: 0.052). Intracontinental populations mostly had weak, albeit non-significant relationships, apart from the European population cluster, which showed a strong, though non-significant r^2 value of 0.71. Additionally, the south Asian population cluster again showed a negative relationship between the psPRS and PD prevalence (Figure S3A).

When the same analysis was performed with the filtered AD psPRS values, the PD relationship was both highly significant and predictive (Figure 2B, slope: 0.0005, $r^2 = 0.90$, p-value: 3.56×10^{-12}). However, while most of the continental population clusters showed an improvement in strength and significance of the psPRS versus prevalence relationship, the European populations experienced a reduction of both significance and strength from the unfiltered comparison (Figure S3B).

The regression analysis between the unfiltered PD psPRSs and AD prevalences showed a slightly positive, significant, predictive model (Figure 2C and S4A, slope: 0.0006, r^2 : 0.16, p-value: 0.045). When the same analysis was performed with the filtered PD psPRS values, the AD relationship improved considerably in significance and predictiveness (Figure 2D and S4B, slope: 0.0068, $r^2 = 0.79$, p-value: 1.41×10^{-9}).

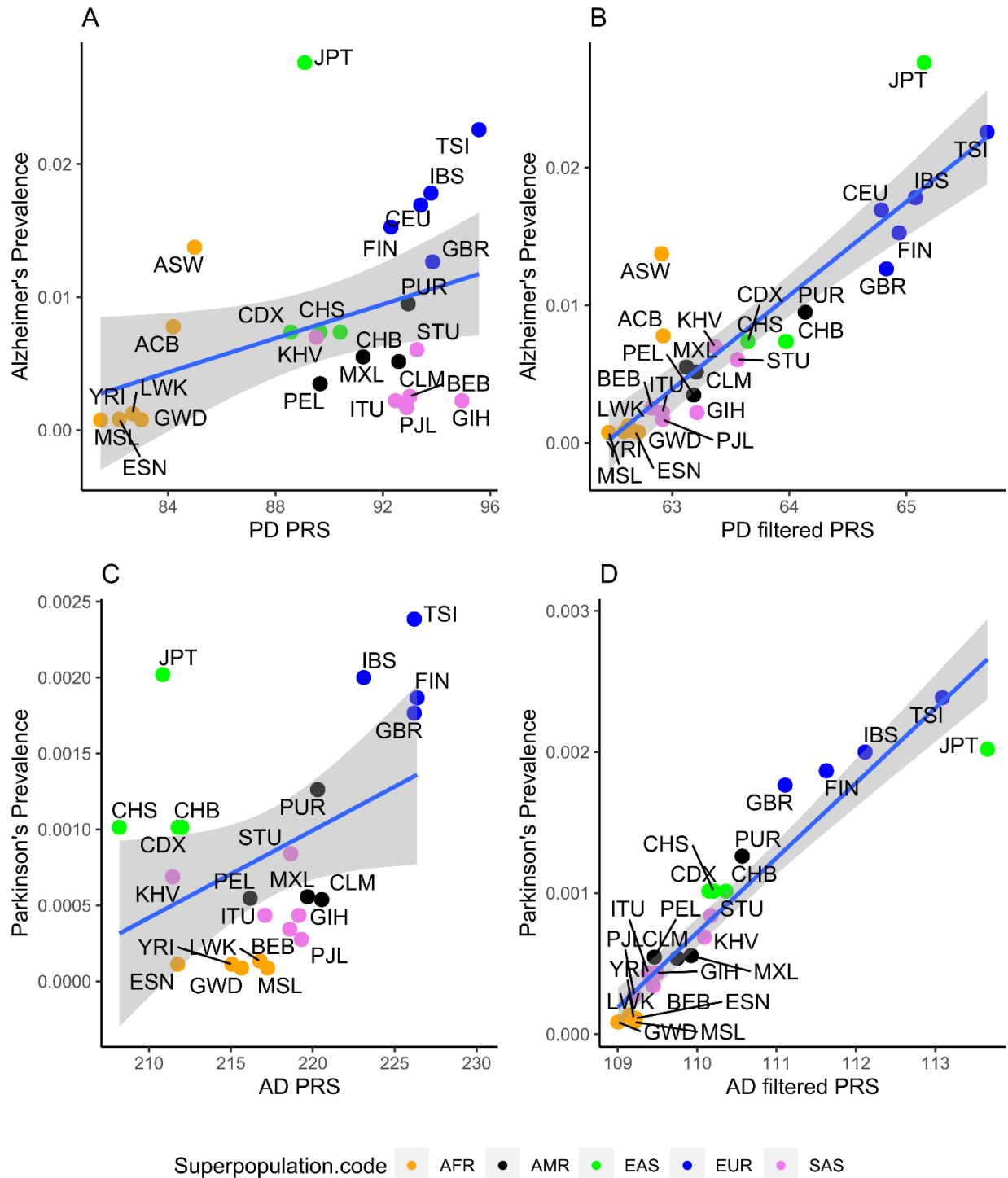


Figure 2. Correlation between diseases and the exchanged psPRS. The data points are colored according to the super populations: AFR (orange), AMR (black), EAS (green), EUR (blue) and SAS (purple). The scale of the x-axis is not the same for both plots due to differing psPRS ranges. **A.** Full PD prevalence and the AD psPRS model ($r^2 = 0.17$; p -value: 0.052). **B.** After maximization ($r^2 = 0.90$, p -value: 3.56×10^{-12}). **C.** AD prevalence and the PD psPRS full model ($r^2 = 0.16$; p -value: 0.045). **D.** After maximization ($r^2 = 0.79$, p -value: 1.41×10^{-9}).

Table 2. Interacting Alzheimer's and Parkinson's genes in AlzKB.	
Alzheimer's genes	Parkinson's genes
APOE	MAPT
APP	SNCA, GAK, RPS12, TRIM40, UCHL1, DDRGK1, MAPT, HTRA2, VPS35
ATP5F1A	RPS12
BAX	SNCA
BIN1	SYNJ1, ATXN2
CALM1	PAM, MAPT, DNAJC13, SNCA
CASP3	MAPT, FYN, DCTN1, KCNIP3
CD2AP	FYN
CHRNA7	FYN
ENO1	FYN, RPS12
ESR1	EIF4G1, UBTF, RPS12, GAK, TBP
GSK3B	DCTN1, MAPT, SNCA, DNAJC13
IGF1R	KCNIP3
IL1B	TBP
MAPT	DCTN1, FYN, PRKN, LRRK2, SNCA
PLCG2	CD19, FYN
PSEN1	MAPT, HTRA2, SCAF11, KCNIP3
PSEN2	KCNIP3
VEGFA	VPS35

AD and PD gene-gene interaction and pathway analysis

Interestingly, only two PD genes, *MAPT* and *HLA-DRB5*, were directly connected to AD in AlzKB. However, there were several interactions in AlzKB between the PD gene set and AD gene set (Figure 3, Table 2), but only *MAPT* is in the intersection of the two sets and interacts with 11 genes in total. In addition, we used STRINGdb to identify enrichment of *CASP3*, *ATP5F1A*, *SNCA* and *MAPT* (FDR: 6.14×10^{-7}) in the Kegg AD and PD pathways (Figure 4). *CASP3* and *ATP5F1A* are also in other neurodegenerative KEGG pathways, namely Huntington's disease and Amyotrophic Lateral Sclerosis. In fact, it has been observed that 91% of KEGG pathways are potentially connected to AD pathogenesis⁴⁵.

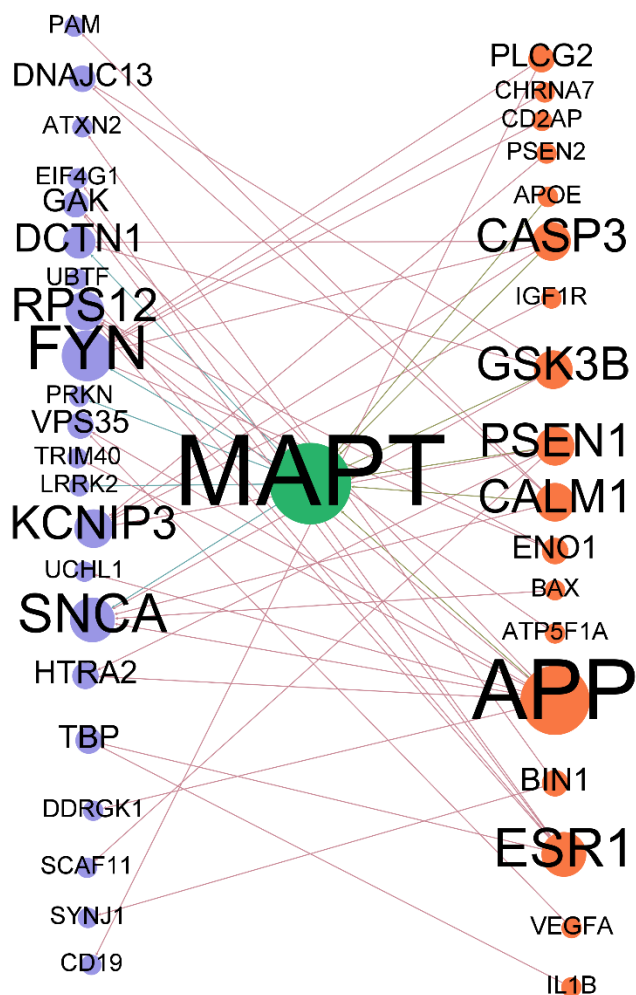


Figure 3. Bipartite network of AD gene and PD gene interactions from AlzKB. Interactions in AlzKB between the PD gene set (blue) and AD gene set (orange), but only *MAPT* (green) is in the intersection of the two sets and interacts with 11 genes in total.

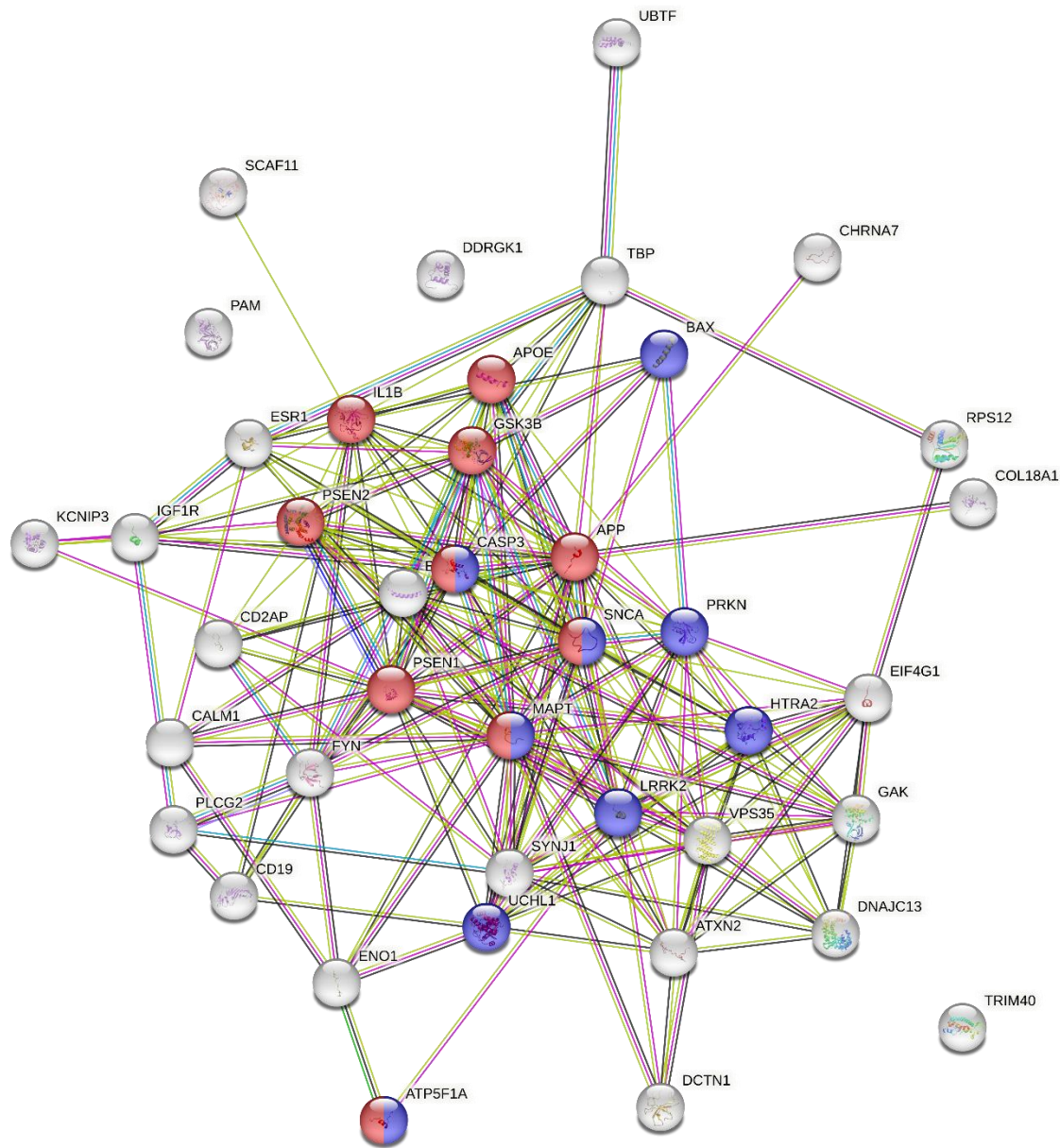


Figure 4. STRINGdb enrichment of *CASP3*, *ATP5F1A*, *SNCA* and *MAPT* in the Kegg AD (blue) and PD (red) pathways.

Discussion/Conclusion

AD

While the full unpruned relationship between the psPRS and AD prevalence was not significant, our pruned results were, suggesting that there is a correlation between AD specific allele frequency and prevalence. That the variance explained was low, supports an environmental and/or interactive piece of the overall etiology of AD. However, the significance of the filtered results suggests that there is a specific subset of AD loci that can predict the prevalence of AD with high accuracy in most populations.

AD-PD analyses

That the unfiltered AD PRS can, with statistical significance, predict the prevalence of PD in certain populations indicate that there may be a substantial shared underlying genetic risk for both diseases. As the AD specific filtering also improved the PD-AD-PRS comparison, it seems that same subset of AD loci can predict the prevalence of PD with high accuracy in many populations. Interestingly, the European population comparisons exchanged favorability between the filtered and unfiltered PRS. The strength and significance of AD unfiltered PRS-prevalence relationship was far less than the PD. The relationship for PD worsened with the filtering, possibly because African and European American PD prevalences are much higher than their ancestral groups. Perhaps this is due to different environmental conditions in the United States, than in ancestral lands.

This genetic link between AD and PD is further underlined by our analysis of the relationship between the PD psPRS and AD prevalence. Between the AD PRS and the PD PRSs

ability to predict AD prevalence, the best predictor over all is the filtered AD PRS. Yet, the PD PRS is at least significant and somewhat predictive before filtering and even more so after. The ability of the AD PRS to predict PD is still better than the ability of the PD PRS to predict AD. However, that the filtered PD PRS r^2 is so high, still indicates a, as yet unidentified, strong underlying interdependency of PD and AD.

Although AD and PD are distinct diseases, these results suggest that they are genetically similar enough that a PRS developed for one disease can be applied to the other to estimate population disease risk. This is supported by the gene-gene interaction observed in our AlzKB based analysis of the AD and PD specific genes. As only two genes, *MAPT* and *HLA-DRB5* are shared between AD and PD directly, this suggests that either there are shared genes yet to be characterized or the underlying common genetics are significantly more complex than previously thought.

The four AlzKB gene overlap of the KEGG AD and PD pathways shows a connection beyond AD and PD simply sharing two genes directly. The additional overlap with the Amyotrophic Lateral Sclerosis and Huntington's disease KEGG pathways, underlines the neurodegenerative role of these genes. Additionally, the 78.2% overlap of the full KEGG PD pathway with the full AD pathway supports a significant underlying genetic and functional connection between AD and PD in general.

Our analyses shows that *MAPT* is an important gene in both AD and PD and may be the key for understanding the overlap of AD and PD. Indeed, *MAPT* is involved in tau and is primarily expressed in the brain⁴⁶. *MAPT* is then a candidate for gene therapy in downregulating

tau protein⁴⁷. That *MAPT* is so central may not be surprising, as it has been suggested as a supergene^{48, 49}.

One thing to note, is that AD and PD prevalences are somewhat correlated in the 1000 Genomes populations. Indeed, the prevalences of Dementia, Parkinson's disease, multiple sclerosis, and motor neuron diseases are highly correlated worldwide⁵⁰. A regression analysis of the PD prevalences versus the AD prevalences shows a statistically significant moderate positive correlation (Slope: 0.23, p-value: 0.018, r^2 : 0.47). Therefore, we cannot rule out that this perceived link between AD and PD is based on shared environmental differences between populations. In conclusion, we have shown evidence supporting significant shared underlying biology for AD and PD, although it has continued to be elusive in other studies. Our results call for further research into the general overlap of Alzheimer's disease and Parkinson's disease, despite the previous lack of evidence.

References

1. Erkinen MG, Kim MO, Geschwind MD. Clinical Neurology and Epidemiology of the Major Neurodegenerative Diseases. *Cold Spring Harb Perspect Biol* 2018;10.
2. Wightman DP, Savage JE, Tissink E, Romero C, Jansen IE, Posthuma D. The genetic overlap between Alzheimer's disease, amyotrophic lateral sclerosis, Lewy body dementia, and Parkinson's disease. *Neurobiol Aging* 2023;127:99-112.
3. Collaborators GBDD. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019;18:88-106.
4. Collaborators GBDPSD. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2018;17:939-953.
5. Anstey KJ, Zheng L, Peters R, et al. Dementia Risk Scores and Their Role in the Implementation of Risk Reduction Guidelines. *Frontiers in Neurology* 2022;12.
6. Che R, Motsinger-Reif AA. A new explained-variance based genetic risk score for predictive modeling of disease risk. *Stat Appl Genet Mol Biol* 2012;11:Article 15.
7. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. *Hum Mol Genet* 2019;28:R133-R142.
8. Wojcik GL, Graff M, Nishimura KK, et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 2019;570:514-518.
9. Martin AR, Gignoux CR, Walters RK, et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am J Hum Genet* 2017;100:635-649.
10. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* 2018;19:581-590.
11. Graham BE, Plotkin B, Muglia L, Moore JH, Williams SM. Estimating prevalence of human traits among populations from polygenic risk scores. *Hum Genomics* 2021;15:70.
12. Jung SH, Kim HR, Chun MY, et al. Transferability of Alzheimer Disease Polygenic Risk Score Across Populations and Its Association With Alzheimer Disease-Related Phenotypes. *JAMA Netw Open* 2022;5:e2247162.
13. Walker IM, Fullard ME, Morley JF, Duda JE. Olfaction as an early marker of Parkinson's disease and Alzheimer's disease. *Handb Clin Neurol* 2021;182:317-329.
14. Rajput AH, Rozdilsky B, Rajput A. Alzheimer's disease and idiopathic Parkinson's disease coexistence. *J Geriatr Psychiatry Neurol* 1993;6:170-176.
15. Rosen AR, Steenland NK, Hanfelt J, Factor SA, Lah JJ, Levey AI. Evidence of shared risk for Alzheimer's disease and Parkinson's disease using family history. *Neurogenetics* 2007;8:263-270.
16. Stolp Andersen M, Tan M, Holtman IR, Hardy J, International Parkinson's Disease Genomics C, Pihlstrom L. Dissecting the limited genetic overlap of Parkinson's and Alzheimer's disease. *Ann Clin Transl Neurol* 2022;9:1289-1295.
17. Kelly J, Moyeed R, Carroll C, Albani D, Li X. Gene expression meta-analysis of Parkinson's disease and its relationship with Alzheimer's disease. *Mol Brain* 2019;12:16.
18. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* 2023;51:D587-D592.
19. Emon MA, Heinson A, Wu P, et al. Clustering of Alzheimer's and Parkinson's disease based on genetic burden of shared molecular mechanisms. *Sci Rep* 2020;10:19097.

20. Kelly J, Moyeed R, Carroll C, Luo S, Li X. Genetic networks in Parkinson's and Alzheimer's disease. *Aging (Albany NY)* 2020;12:5221-5243.
21. Campdelacreu J. Parkinson disease and Alzheimer disease: environmental risk factors. *Neurologia* 2014;29:541-549.
22. Vu M, Mangal R, Stead T, Lopez-Ortiz C, Ganti L. Impact of Alzheimer's Disease on Caregivers in the United States. *Health Psychol Res* 2022;10:37454.
23. 2022 Alzheimer's disease facts and figures. *Alzheimers Dement* 2022;18:700-789.
24. Haddad HW, Malone GW, Comardelle NJ, et al. Aduhelm, a novel anti-amyloid monoclonal antibody, for the treatment of Alzheimer's Disease: A comprehensive review. *Health Psychol Res* 2022;10:37023.
25. Gatz M, Reynolds CA, Fratiglioni L, et al. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* 2006;63:168-174.
26. Serrano-Pozo A, Das S, Hyman BT. APOE and Alzheimer's disease: advances in genetics, pathophysiology, and therapeutic approaches. *Lancet Neurol* 2021;20:68-80.
27. Wang H, Bennett DA, De Jager PL, Zhang QY, Zhang HY. Genome-wide epistasis analysis for Alzheimer's disease and implications for genetic risk prediction. *Alzheimers Res Ther* 2021;13:55.
28. Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M. World Alzheimer Report 2015-The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. 2015.
29. Reich SG, Savitt JM. Parkinson's Disease. *Med Clin North Am* 2019;103:337-350.
30. Aging NNI. Parkinson's Disease: Causes, Symptoms, and Treatments [online]. Available at: <https://www.nia.nih.gov/health/parkinsons-disease#:~:text=Parkinson%27s%20disease%20is%20a%20brain,have%20difficulty%20walking%20and%20talking>. Accessed 9/9/2022.
31. Post B, van den Heuvel L, van Prooije T, van Ruissen X, van de Warrenburg B, Nonnekes J. Young Onset Parkinson's Disease: A Modern and Tailored Approach. *J Parkinsons Dis* 2020;10:S29-S36.
32. Hayes MT. Parkinson's Disease and Parkinsonism. *Am J Med* 2019;132:802-807.
33. Poewe W, Gauthier S, Aarsland D, et al. Diagnosis and management of Parkinson's disease dementia. *Int J Clin Pract* 2008;62:1581-1587.
34. Armstrong MJ, Okun MS. Diagnosis and Treatment of Parkinson Disease: A Review. *JAMA* 2020;323:548-560.
35. Miocinovic S, Somayajula S, Chitnis S, Vitek JL. History, Applications, and Mechanisms of Deep Brain Stimulation. *JAMA Neurology* 2013;70:163-171.
36. Blauwendraat C, Nalls MA, Singleton AB. The genetic architecture of Parkinson's disease. *Lancet Neurol* 2020;19:170-178.
37. Bandres-Ciga S, Diez-Fairen M, Kim JJ, Singleton AB. Genetics of Parkinson's disease: An introspection of its journey towards precision medicine. *Neurobiol Dis* 2020;137:104782.
38. Ascherio A, Schwarzschild MA. The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol* 2016;15:1257-1272.
39. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526:68-74.
40. Nalls MA, Blauwendraat C, Vallerga CL, 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* 2019;18:1091-1102.

41. Day JO, Mullin S. The Genetics of Parkinson's Disease and Implications for Clinical Practice. *Genes (Basel)* 2021;12.
42. Nalls MA, Pankratz N, Lill CM, 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.
43. Himmelstein DS, Lizee A, Hessler C, et al. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *Elife* 2017;6.
44. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47:D607-D613.
45. Morgan SL, Naderi P, Koler K, et al. Most Pathways Can Be Related to the Pathogenesis of Alzheimer's Disease. *Frontiers in Aging Neuroscience* 2022;14.
46. Zhang W, Xu C, Sun J, Shen HM, Wang J, Yang C. Impairment of the autophagy-lysosomal pathway in Alzheimer's diseases: Pathogenic mechanisms and therapeutic potential. *Acta Pharm Sin B* 2022;12:1019-1040.
47. Chakravarthy M, Chen S, Wang T, Veedu RN. Development of Novel Chemically-Modified Nucleic Acid Molecules for Efficient Inhibition of Human MAPT Gene Expression. *Genes (Basel)* 2020;11.
48. Berdan EL, Flatt T, Kozak GM, Lotterhos KE, Wielstra B. Genomic architecture of supergenes: connecting form and function. *Philos Trans R Soc Lond B Biol Sci* 2022;377:20210192.
49. Campoy E, Puig M, Yakymenko I, Lerga-Jaso J, Caceres M. Genomic architecture and functional effects of potential human inversion supergenes. *Philos Trans R Soc Lond B Biol Sci* 2022;377:20210209.
50. James LM, Georgopoulos AP. High Correlations Among Worldwide Prevalences of Dementias, Parkinson's Disease, Multiple Sclerosis, and Motor Neuron Diseases Indicate Common Causative Factors. *Neurosci Insights* 2022;17:26331055221117598.