Ageing Research Reviews 79 (2022) 101648

Contents lists available at ScienceDirect

Ageing Research Reviews

journal homepage: www.elsevier.com/locate/arr

Review

What have we learned from genome-wide association studies (GWAS) in Parkinson's disease?

Rubén Fernández-Santiago [a,b,c,d,*,1], Manu Sharma [e,**,1]

[a] Lab of Parkinson's Disease & Other Neurodegenerative Movement Disorders, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Institut de Neurociències, Universitat de Barcelona, ES-08036 Barcelona, Spain
[b] Parkinson's Disease and Movement Disorders Unit, Neurology Service, Institut Clínic de Neurociències, Hospital Clínic de Barcelona, ES-08036 Barcelona, Spain
[c] Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED: CB06/05/0018-ISCIII), ES-08036 Barcelona, Spain
[d] Histology Unit, Department of Biomedicine, Faculty of Medicine, Universitat de Barcelona, ES-08036 Barcelona, Spain
[e] Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, University of Tübingen, DE-72076 Tübingen, Germany

ARTICLE INFO

Keywords:
Parkinson's disease (PD)
Advanced genome-wide association studies (GWAS)
Prediction of PD risk and progression
Translational PD genetics

ABSTRACT

After fifteen years of genome-wide association studies (GWAS) in Parkinson's disease (PD), what have we learned? Addressing this question will help catalogue the progress made towards elucidating disease mechanisms, improving the clinical utility of the identified loci, and envisioning how we can harness the strides to develop translational GWAS strategies. Here we review the advances of PD GWAS made to date while critically addressing the challenges and opportunities for next-generation GWAS. Thus, deciphering the missing heritability in underrepresented populations is currently at the reach of hand for a truly comprehensive understanding of the genetics of PD across the different ethnicities. Moreover, state-of-the-art GWAS designs hold a true potential for enhancing the clinical applicability of genetic findings, for instance, by improving disease prediction (PD risk and progression). Lastly, advanced PD GWAS findings, alone or in combination with clinical and environmental parameters, are expected to have the capacity for defining patient enriched cohorts stratified by genetic risk profiles and readily available for neuroprotective clinical trials. Overall, envisioning future strategies for advanced GWAS is currently timely and can be instrumental in providing novel genetic readouts essential for a true clinical translatability of PD genetic findings.

1. Introduction

An estimated 6 million people worldwide affected by PD are likely to double by 2040, (Dorsey et al., 2018) along with derived medical expenses due to increases in life expectancy. The annual economic burden of PD was estimated at $14.4 billion in the U.S.A (Kowal et al., 2013). and €14 billion in Europe (Gustavsson et al., 2011). Such a public health liabilities on the societies urgently requires strategies that can provide a roadmap to develop therapeutic interventions at an early stage to stem the disease's progression. Overall, in the last two decades, classic genetic

Abbreviations: AAO, age-at-onset; ALS, amyotrophic lateral sclerosis; AMD, age-related macular degeneration; AUC, area under the ROC curve; *APOE*, apolipoprotein E; CD, Crohn's disease; CDCV, common-disease common-variant hypothesis; CDCV, common-disease rare-variant hypothesis; CSF, cerebrospinal fluid; *GBA*, glucocerebrosidase; GRS, genetic risk score; GWAS, genome-wide association studies; G x E, the gene-by-environment interactions; H3K27ac, acetylated lysine 27 at histone 3; H3K4me1, monomethylated lysine 4 at histone 3; H3K4me3, trimethylated lysine 4 at histone 3; LD, linkage disequilibrium; *LRRK2*, leucine-rich repeat kinase 2; LSD, lysosomal storage disorders; *MAPT*, microtubule-associated protein Tau; MDR, multi-dimension reduction analysis; MDS, Movement Disorders Society; MR, Mendelian randomization; PCR, polymerase chain reaction; PD, Parkinson's disease; *PRKN*, parkin; PRS, polygenic risk score; PRP, personalized risk prediction; QTL, quantitative trait loci; RCT, randomized controlled trial; RF, random forest; ROC, receiver operating characteristic; RBD, REM sleep behavior disorder; SN, substantia nigra; *SNCA*, α-synuclein;; SNP, single nucleotide polymorphism; SVM, support vector machine; TF, transcription factor.

* Correspondence to: Laboratory of Parkinson Disease & Other Neurodegenerative Movement Disorders, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hospital Clínic de Barcelona, University of Barcelona, Casanova 143, Floor 3B, 08036 Barcelona, Spain.
** Correspondence to: Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry, Eberhard-Karls University, Silcherstrasse 5, Tübingen, Baden, Württemberg 72076, Germany.
E-mail addresses: ruben.fernandez.santiago@clinic.cat (R. Fernández-Santiago), manu.sharma@uni-tuebingen.de (M. Sharma).
[1] Both authors contributed equally

https://doi.org/10.1016/j.arr.2022.101648
Received 3 January 2022; Received in revised form 11 April 2022; Accepted 11 May 2022
Available online 18 May 2022

discoveries have greatly helped to broaden the arc of our understanding of PD etiopathogenesis. With the advent of high-throughput technologies, after the first genome-wide association study (GWAS) in age-related macular degeneration (AMD) in 2005, (Klein et al., 2005) GWAS's initial wave in different complex diseases, including PD, took off and started revealing novel loci associated with these maladies. Like other multifactorial diseases, (Tam et al., 2019) the genomic research of PD by array-based approaches has offered great opportunities and challenges, e.g., given that most of the studies have been performed in the European population (Xue et al., 2018). The inclusion of ethnic diversity in PD genomic research is essential to enhance our understanding of biological PD mechanisms. The various initiatives such as the "All of US" program launched by NIH, the ethnic diversity program by the Michael J. Fox Foundation for Parkinson's Research (MJFF), and the recently launched Global Parkinson's Genetics Program "GP2" have provided the impetus to include ethnic diversity not only for PD but also to map various multifactorial disorders. These initiatives will help promote awareness in the underrepresented populations and play an essential role in democratising genomic research (Tam et al., 2019). Rather than reviewing all genetic work done in PD, we specifically focus on novel perspectives for future PD GWAS. Briefly, we envision that following PD genetic efforts pursuing GWAS approaches should be aimed at (i) further enhancing our understanding in deciphering the missing heritability of disease in underrepresented populations as to gain a truly comprehensive understanding of the disease across ethnicities; (Keller et al., 2012) (ii) using state-of-the-art novel approaches helpful in clinical practice for predicting disease risk and progression, and for anticipating specific clinical outcomes until the limit of the possible; and (iii) defining patient enriched cohorts stratified by genetic risk profiles, alone or in combination with other parameters, that should be available for future clinical trials. Following the PRISMA criteria as a reference, we used the Pubmed searching engine and the terms "Parkinson's", "GWAS", and "meta-analysis" to select the literature reviewed in this study. From the retrieved total of 168 results from 2006 until 2022, the strict inclusion criteria included, but were not restricted to, GWAS meta-analysis and related genome-wide studies performed in PD until date.

2. Discussion

2.1. Advances

2.1.1. Genetics beyond monogenic PD

Identifying disease-causing mutations in the α-synuclein (*SNCA*) gene altered the PD landscape and firmly established the genetic component in PD pathogenesis (Polymeropoulos et al., 1997). The list of PD causative genes has grown steadily since, with more than 20 causative genes of PD or atypical parkinsonism identified to date (Del Rey et al., 2018). Of these, the autosomal-dominant genes *SNCA*, (Polymeropoulos et al., 1997) leucine-rich repeat kinase 2 (*LRRK2*),

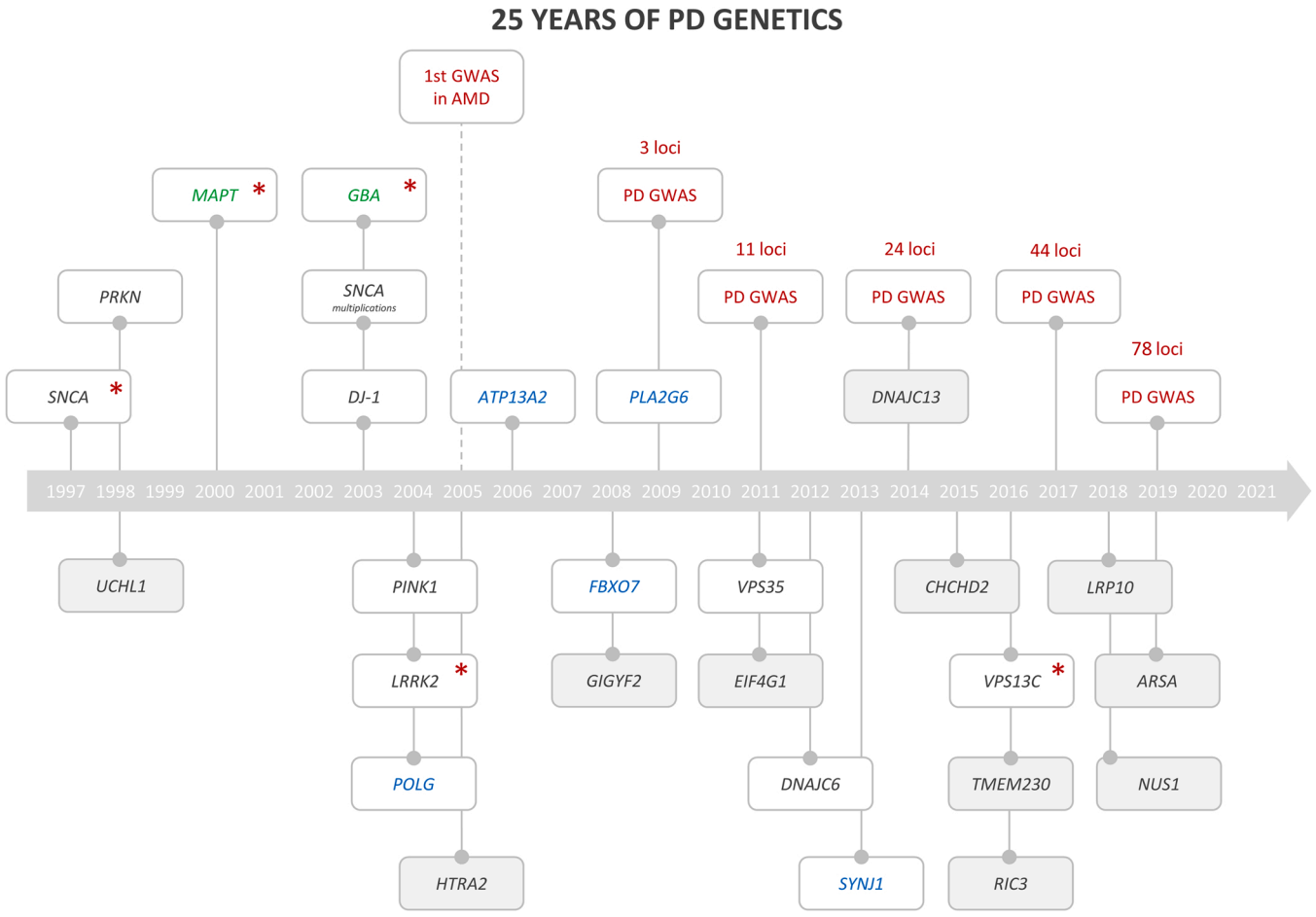

Fig. 1. : Timeline of the 25 years of genetic research in PD. The first genome-wide association study (GWAS) in age-related macular degeneration (AMD) in 2005 was a stepping stone for the genomic research of complex multifactorial diseases such as PD. Since then, PD GWAS took off in the so-called GWAS era, revealing novel loci associated with the disease (red). This timeline also shows genes with reported mutations causative of Mendelian forms of PD (black) or atypical parkinsonism (blue), either validated (white rectangles) or pending validation (grey-shadowed rectangles); genes nominated by GWAS (red asterisks); and well-established PD risk genes without Mendelian segregation of disease such as *MAPT*, and *GBA* (green).

(Zimprich et al., 2004; Paisán-Ruíz et al., 2004) and *VPS35*, (Vilariño-Güell et al., 2011; Zimprich et al., 2011) and the autosomal-recessive parkin (*PRKN*), (Kitada et al., 1998) *PINK1*, (Valente et al., 2004), *DJ-1* (Bonifati et al., 2003), *DNAJC6*, (Edvardson et al., 2012) and *VPS13C* (Lesage et al., 2016) have been convincingly associated with PD (Fig. 1). However, *PRKN* mutations cause early-onset PD - age-at-onset (AAO) below 50 years – without Lewy body (LB) pathology, (Doherty et al., 2013) whereas other recessive genes such as *PINK1* or *DJ-1* are often linked, but not always, to juvenile PD. Moreover, genetic variants in other genes, including the glucocerebrosidase (*GBA*) (Sidransky et al., 2009) and the microtubule-associated protein Tau (*MAPT*) (Pastor et al., 2000) genes, are well-established risk factors for PD, the former only in Europeans but not in Asians. Other genes have been associated with atypical parkinsonism-related syndromes, including *POLG*, (Luoma et al., 2004) *ATP13A2*, (Ramirez et al., 2006) *FBXO7*, (Fonzo et al., 2009) *PLA2G6,* (Paisan-Ruiz et al., 2008) and *SYNJ1* (Krebs et al., 2013). Lastly, other genes have been linked to typical PD (*UCHL1*, *HTRA2*, *GIGYF2*, *EIF4G1*, *SMPD1*, *DNAJC13*, *CHCHD2*, (Funayama et al., 2015) *TMEM230*, *RIC3*, *LRP10*, *NUS1*, and *ARSA*) (Bandres-Ciga et al., 2020a) yet some require further validation given that they lack reports in multiple independent families or functional validation, or that negative reports have been published (Blauwendraat et al., 2020). Overall, the genetic discoveries in familial forms have proven instrumental in uncovering most of our current knowledge on pathogenic molecular processes in PD (Przedborski, 2017). Though there has been a relative success, the overall genetic burden explained by monogenic forms of the disease remains marginal, with unequivocal Mendelian causation in only 5–10% of the patients (Singleton and Hardy, 2019).

Variant gene-mapping through association studies was proposed to further explain the disease variance in complex diseases. The first-generation approaches were hampered due to the lack of technologies to process a large number of samples. The development of polymerase chain reaction (PCR)-based approaches made it feasible to perform genetic association studies cost-effectively. However, as these studies grew exponentially, so did the non-replicability of the findings. Many of the early association studies were eventually flawed by the design issues, including a lack of statistical power, incorrect handling of multiple testing, insufficient appreciation of population structuring, inadequate genetic coverage, or non-robust genotyping methods (Ioannidis, 2018). It became clear later that the effect size associated with risk variants is small and that large samples are needed to observe a given effect. Moreover, beyond technical and statistical issues, the lack of consistency in the clinical-based and primarily symptomatic diagnosis of PD and the potential 'contamination' with other misdiagnosed movement disorders, especially at the initial stages, contributed to the lack of reproducibility (Ioannidis, 2018). Therefore, it was not surprising to render a lack of replication for most early genetic association studies (Ioannidis, 2018).

In this context, the human genome mapping project resulted in the extensive cataloguing of millions of single nucleotide polymorphisms (SNPs) to overcome such limitations (Collins and McKusick, 2001). This deluge of data enabled researchers to define the haplotypic structure of the human genome (Gabriel et al., 2002). The haplotypic structure revealed that SNPs are not randomly distributed, a process defined as linkage disequilibrium (LD) and that SNPs allocated nearby on a chromosome often segregate together. This progress led to defining the haplotypic structure in the genome, and because of the extent of linkage disequilibrium, few SNPs could be used to capture the genetic information (Slatkin, 2008). Such SNPs were later on defined as tag SNPs (Johnson et al., 2001). This seminal observation that the genome is organised in haplotypic blocks and that few SNPs within each block can capture the neighbouring genetic information made a considerable impact in designing and developing arrays, which eventually led to starting the first wave of GWAS.

2.1.2. GWAS and heritability

One of the earlier GWAS in PD provided unequivocal evidence regarding the role of *SNCA* and *MAPT* in PD (Table 1) (Simon-Sanchez et al., 2009). Subsequent GWAS with larger sample sizes confirmed these loci and suggested population-specific effects (IPDGC and WTCCC2, 2011; Nalls et al., 2014, 2019; Chang et al., 2017). These studies highlighted three main aspects: (i) associated common variants in at-risk loci have only a modest effect, often with an odds ratio below 1.4; (ii) large samples are required to discover common variants with smaller effects; and (iii) the identified loci also include genes found by linkage studies (*SNCA*, *LRRK2*) but also reveal new loci suggesting that pathways involved in monogenic and idiopathic forms of PD are not mutually exclusive; (IPDGC and WTCCC2, 2011) GWAS-meta analyses with ever-increasing sample size have increased the number of genetic loci implicated in sporadic PD to 78, explaining an overall heritability estimated at 16–36% (Nalls et al., 2019). Interestingly, this estimate is in line with the first genetic studies reporting that 10–30% of patients with PD report a first-degree relative with parkinsonism (Rocca et al., 2004; Sveinbjörnsdóttir et al., 2000). Although cross-sectional twin studies have mainly argued against heritability in PD, (Przedborski, 2017) concordance rates in longitudinal twin studies between monozygotic and dizygotic disease have also reported modest inheritability rates for PD similar to PD GWAS e.g., 34% in Sweden (Wirdefeldt et al., 2011), or 27% in the U.S.A (Goldman et al., 2019). Another derivative from PD GWAS meta-analyses, which are based on the common-disease common-variant (CDCV) hypothesis, is the alternative that additional multiple rare DNA variants, each with relative high penetrance, i.e., common-disease rare-variant (CDRV), (Schork et al., 2009) might also contribute to genetic susceptibility and missing heritability of the disease.

Despite the success, little progress has been made to infer the role of genetic variability in influencing disease progression, for instance: AAO (early vs late-onset), predominant phenotype (tremor vs non-tremor), adverse response to L-DOPA (dyskinesias), cognitive decline, or non-motor symptoms. One study using the first wave of PD GWAS risk loci generated the cumulative genetic risk score (GRS) to predict a disease progression (Pihlstrøm et al., 2016). Subsequently, it has been shown that the cumulative burden of common risk loci, albeit small, is strongly associated with the AAO in PD, and when compared with patients with a late-onset, PD patients with an early AAO have a significantly higher polygenic risk score (PRS) (Nalls et al., 2015a). Recently published GWASs have shown a heritability attributed to a disease progression of 11%, revealing that not all PD GWAS SNPs at risk loci differentially influence the AAO, thus indicating that there can be underlying mechanistic divergences between risk and disease progression in PD (Blauwendraat et al., 2019). Indeed, another recent study used principal component analysis (PCA) and PD GWAS data to identify genetic variants associated with PD progression, i.e., motor and cognitive decline (Tan et al., 2021). Interestingly, they reported no significant overlap between variants associated with PD risk and PD progression and identified the apolipoprotein E (*APOE*) ε4 allele as the main driver for progressive cognitive impairment. Nevertheless, additional studies are needed to validate and expand these findings. Taken together, the evidence generated so far established the polygenic architecture of PD.

In addition, attempts to use cerebrospinal fluid (CSF) proteins as biomarkers combined with GWAS data to establish disease onset have

Table 1
Summary list of top-10 hits from all PD GWAS meta-analyses, all performed in cohorts of European ancestry, ranked by most significant P-values, with highlighted gene functions. Only shown the top-10 hits per study.

Nalls et al. Lancet Neurol. 2019 // European ancestry, 37.688 PD, 18.618 relatives, and 1,4 million controls // Reference: GRCh38 / hg38 * Novel hits

Chr.	SNP	Proxy	Gene symbol	Gene name	Gene function	Function expanded
3	rs6808178	*LINC00693*	*LINC00693*	Long Intergenic Non-Protein Coding RNA 693	lncRNA	long non-coding RNA
3	rs55961674	*KPNA1*	*KPNA1*	Karyopherin Subunit Alpha 1	nuclear transport	belongs to importin alpha family, and is involved in nuclear protein import
6	rs75859381	*RPS12*	*RPS12*	Ribosomal Protein S12	protein homeostasis	ribosomal protein that is a component of the 40S subunit
6	rs12528068	*RIMS1*	*RIMS1*	Regulating Synaptic Membrane Exocytosis 1, Rab-3-Interacting Prot. 2	synapsis	regulates synaptic vesicle exocytosis
4	rs34025766	*LCORL*	*LCORL*	Ligand Dependent Nuclear Receptor Corepressor Like	transcription	transcription factor that appears to function in spermatogenesis
17	rs2269906	*UBTF*	*UBTF*	Upstream Binding Transcription Factor	transcription	HMG-box DNA-binding protein family. Chromatin remodeling and pre-rRNA processing
16	rs2904880	*CD19*	*CD19*	CD19 B-Lymphocyte Surface Antigen B4	immunity	activation of downstream signaling pathways and for triggering B-cell responses to antigens
10	rs10748818	*GBF1*	*GBF1*	Golgi Brefeldin A Resistant Guanine Nucleotide Exchange Factor 1	vesicle-mediated transport	Golgi guanine nucleotide exchange factor, GTP-dependent recruitment of proteins to membranes
13	rs4771268	*MBNL2*	*MBNL2*	Muscleblind Like Splicing Regulator 2	transcription	zinc finger protein that modulates alternative splicing of pre-mRNAs
16	rs6500328	*NOD2*	*NOD2*	Nucleotide Binding Oligomerization Domain Containing 2	immunity	leukocyte protein involve in immune response to intracellular bacterial lipopolysaccharides

Chang et al. Nat Genet. 2017 // European ancestry, 26.035 PD and 403.190 controls // Reference: GRCh37 / hg19 Feb 2009

Chr.	SNP	Proxy	Gene symbol	Gene name	Gene function	Function expanded
4	rs356182	*SNCA*	*SNCA*	a-Synuclein	synapsis	neurotransmitter release
17	rs17649553	*MAPT*	*MAPT*	Microtubule Associated Protein Tau	cytoskeleton	microtubule assembly / manteinance of neural polarity
4	rs34311866	*TMEM175 / DGKQ*	*TMEM175*	Transmembrane Protein 175	endosome / lysosome	endosomal / lysosomal potassium channel
4	rs34311866	*TMEM175 / DGKQ*	*DGKQ*	Diacylglycerol Kinase Theta	metabolism	mediates the regeneration of phosphatidylinositol (PI) from diacylglycerol during cell signal transduction
1	rs35749011	*GBA*	*GBA*	Glucosylceramidase Beta	endosome / lysosome	degradation of complex lipids and the turnover of cellular membranes
3	rs12637471	*MCCC1*	*MCCC1*	Methylcrotonoyl-CoA Carboxylase 1	metabolism	valine, leucine and isoleucine degradation and metabolism of water-soluble vitamins and cofactors
2	rs1474055	*STK39*	*STK39*	Serine/Threonine Kinase 39	stress-activated cell response	mediator of stress-activated signals
2	rs6430538	*TMEM163 / CCNT2*	*TMEM163*	Transmembrane Protein 163	metabolism	may bind zinc and other divalent cations and recruit them to vesicular organelles
2	s6430538	*TMEM163 / CCNT2*	*CCNT2*	Cyclin T2	cell cycle	mitosis, regulates CDK kinases
12	rs11060180	*OGFOD2*	*OGFOD2*	2-Oxoglutarate & Iron Dependent Oxygenase Domain Containing 2	metabolism	iron ion binding and oxidoreductase activity
12	rs76904798	*LRRK2*	*LRRK2*	Leucine-rich repeat kinase 2	autophagy control?	PD gene
4	rs11724635	*FAM200B / CD38*	*FAM200B*	Family With Sequence Similarity 200 Member B	unknown	nucleic acid binding?
4	rs11724635	*FAM200B / CD38*	*CD38*	ADP-Ribosyl Cyclase 1	metabolism	transmembrane glycoprotein that synthesizes and hydrolyzes cADP-Ribose

Nalls et al. Nat Genet. 2014 // European ancestry, 13.708 PD and 95.282 controls // Reference: GRCh37 / hg19 Feb 2009

Chr.	SNP	Proxy	Gene symbol	Gene name	Gene function	Function expanded
4	rs356182	*SNCA*	*SNCA*	a-Synuclein	synapsis	neurotransmitter release
17	rs17649553	*MAPT*	*MAPT*	Microtubule Associated Protein Tau	cytoskeleton	microtubule assembly / manteinance of neural polarity
4	rs34311866	*TMEM175 / GAK / DGKQ*	*TMEM175*	Transmembrane Protein 175	endosome / lysosome	endosomal / lysosomal potassium channel
4	rs34311866	*TMEM175 / GAK / DGKQ*	*GAK*	Cyclin G Associated Kinase	vesicle-mediated transport	regulates cell cycle, Clathrin derived vesicle budding and vesicle-mediated transport
4	rs34311866	*TMEM175 / GAK / DGKQ*	*DGKQ*	Diacylglycerol Kinase Theta	metabolism	mediates the regeneration of phosphatidylinositol (PI) from diacylglycerol during cell signal transduction
1	rs35749011 / rs71628662	*GBA-SYT11*	*GBA*	Glucosylceramidase Beta	endosome / lysosome	degradation of complex lipids and the turnover of cellular membranes
1	rs35749011 / rs71628662	*GBA-SYT11*	*SYT11*	Synaptotagmin 11	vesicle-mediated transport	calcium-dependent regulation of membrane trafficking in synaptic transmission
3	rs12637471	*MCCC1*	*MCCC1*	Methylcrotonoyl-CoA Carboxylase 1	metabolism	valine, leucine and isoleucine degradation and Metabolism of water-soluble vitamins and cofactors
2	rs1474055 / rs1955337	*STK39*	*STK39*	Serine/Threonine Kinase 39	stress-activated cell response	mediator of stress-activated signals
2	rs6430538	*ACMSD / TMEM163*	*ACMSD*	Aminocarboxymuconate Semialdehyde Decarboxylase	metabolism	de novo synthesis pathway of NAD from tryptophan
2	rs6430538	*ACMSD / MEM163*	*TMEM163*	Transmembrane Protein 163	metabolism	may bind zinc and other divalent cations and recruit them to vesicular organelles
4	rs11724635	*BST1*	*BST1*	Bone Marrow Stromal Cell Antigen 1	immunity	innate immnune system
1	rs823118	*RAB7L1 / NUCKS1*	*RAB7L1*	RAB29, Member RAS Oncogene Family	vesicle-mediated transport	ligand of RAB7A, key regulator in endo-lysosomal trafficking. Fusion of phagosomes with lysosomes
1	rs823118	*RAB7L1 / NUCKS1*	*NUCKS1*	Nuclear Casein Kinase & Cyclin Dependent Kinase Substrate 1	transcription	two putative nuclear localization signals, and a basic DNA-binding domain
12	rs76904798	*LRRK2*	*LRRK2*	Leucine-rich repeat kinase 2	autophagy control?	PD gene

IPDGC et al. PLoS Genet. 2011 // European ancestry, 13.708 PD and 95.282 controls // Reference: NCBI36 / hg18

Chr.	SNP	Proxy	Gene symbol	Gene name	Gene function	Function expanded
4	rs356219	*SNCA*	*SNCA*	a-Synuclein	synapsis	neurotransmitter release
17	rs2942168	*MAPT*	*MAPT*	Microtubule Associated Protein Tau	cytoskeleton	microtubule assembly / manteinance of neural polarity
4	chr4:911311	*GAK*	*GAK*	Cyclin G Associated Kinase	vesicle-mediated transport	cell cycle, cyclin-dependent protein kinases (CDKs)
2	rs2102808	*STK39*	*STK39*	Serine/Threonine Kinase 39	stress-activated cell response	mediator of stress-activated signals
2	rs6710823	*ACMSD*	*ACMSD*	Aminocarboxymuconate Semialdehyde Decarboxylase	metabolism	de novo synthesis pathway of NAD from tryptophan
12	rs12817488	*CCDC62/HIP1R*	*CCDC62*	Coiled-Coil Domain Containing 62	transcription	enhances estrogen receptors ESR1 and ESR2 transactivation, and also glucocorticoid receptor NR3C1
12	rs12817488	*CCDC62/HIP1R*	*HIP1R*	Huntingtin Interacting Protein 1 Related	vesicle-mediated transport	Clathrin derived vesicle budding
1	chr1:154105678	*SYT11*	*SYT11*	Synaptotagmin 11	vesicle-mediated transport	calcium-dependent regulation of membrane trafficking in synaptic transmission
4	rs11724635	*BST1*	*BST1*	Bone Marrow Stromal Cell Antigen 1	immunity	innate immnune system
3	rs11711441	*MCCC1/LAMP3*	*MCCC1*	Methylcrotonoyl-CoA Carboxylase 1	metabolism	critical step for leucine and isovaleric acid catabolism
3	rs11711441	*MCCC1/LAMP3*	*LAMP3*	Lysosomal Associated Membrane Protein 3	immunity	antigen-presenting in dendritic cell function for adaptive immunity.
6	chr6:32588205	*HLA-DRB5*	*HLA-DRB5*	Major Histocompatibility Complex, Class II, DR Beta 5	immunity	presents antigens in antigen presenting cells (APC) (B lymphocytes, dendritic cells, macrophages)
12	rs1491942	*LRRK2*	*LRRK2*	Leucine-rich repeat kinase 2	autophagy control?	PD gene

been unsuccessful to date (Ibanez et al., 2020). Interestingly, a recent study performed a GWAS in REM sleep behaviour disorder (RBD), and found RBD-specific polygenic risk score (PRS) has different effects in individuals with idiopathic RBD (iRBD) or PD plus probable RBD (pRBD) compared to PD without pRBD. This study suggested RBD as a unique subpopulation that will allow future early intervention for synucleinopathies.(Krohn L et al. medRxiv 2021.09.08.21254232).

2.1.3. Curated GWAS unravelling novel genetic pathways

GWAS meta-analyses with a larger sample size were well-powered to identify variants that could not have been identified in individual GWAS alone (Table 1) (IPDGC and WTCCC2, 2011; Nalls et al., 2014, 2019; Chang et al., 2017). Though this approach proved successful, a majority of loci, which may be genuinely associated with PD, could not pass the statistical threshold as required in GWAS. Pathway-based strategies were considered to address this issue. Such approaches examine whether test statistics for a group of related genes have a consistent yet moderate deviation from chance (Wang et al., 2010). The underlying rationale is that genetic variants are unlikely to act randomly, but multiple genes work together in functional gene sets based on prior biological knowledge. The significance of each pathway can be summarised based on markers in or near genes from that pathway (Wang et al., 2010). Illustratively, functional studies in monogenic PD genes showed that the autosomal-recessive genes *PRKN*, *PINK-1*, and *DJ-1* regulate the mitophagy processes, recycling damaged mitochondria and oxidative stress responses (Cookson, 2012). Moreover, *FBXO7* acts together with *PRKN* and *PINK1*, regulating mitochondrial maintenance, (Burchell et al., 2013) and *CHCHD2* mutations cause mitochondrial dysfunction (Zhou et al., 2019). Autosomal-dominant genes such as *LRRK2* participate in the endolysosomal pathway, involving endocytosis, vesicle trafficking, and lysosomal degradation, and *LRRK2* is regulated by *VPS35 (*Erb and Moore, 2020*)*. Moreover, it is well-defined that SNCA accumulation impacts the autophagy lysosomal pathway (Xilouri et al., 2016) and that reduced activity of the risk factor GBA impairs autophagy increasing SNCA levels in PD (Murphy et al., 2014).

One early success of pathway-based GWAS was Crohn's disease (CD), which highlighted the role of the immunity component in CD (Wang et al., 2009). This intuitive approach provided novel insight for other complex diseases including PD (Holmans et al., 2013). Thus, a PD GWAS leveraged pathways implicated in PD and provided evidence of additional PD susceptibility loci that failed to meet the stringent thresholds of previous GWAS (Holmans et al., 2013). identified pathways involved leukocyte/lymphocyte regulation and cytokine-mediated signalling (Holmans et al., 2013). Interestingly, these immune pathways showed significant enrichment even after removing the top loci, including a human leukocyte antigen (*HLA*) locus nominated in GWAS, thus providing unequivocal evidence of an immune component in PD related to innate or adaptative immunity and inflammation (Holmans et al., 2013; Hamza et al., 2010). Other studies highlighted the role of the endolysosomal pathway (vesicle trafficking, lysosomes, and autophagy) (Robak et al., 2017; Bandres-Ciga et al., 2019a, 2020b). One study analysed if a genetic burden of variants from lysosomal storage disorders (LSD) genes was associated with PD risk (Robak et al., 2017). Beyond *GBA*, a genetic burden of rare lysosomal storage disorder genes was associated with PD susceptibility (*SMPD1*, *CTSD*, *SLC17A5*, and *ASAH1*). By Mendelian randomisation (MR), another study performed a pathway-based analysis of 252 genes involved in endocytic trafficking using GWAS data (Bandres-Ciga et al., 2019a). They found 19 significant genes associated with PD (9 genes involved in cargo degradation and 8 in cargo uptake), among which only the cyclin G associated kinase (*GAK*) gene was previously linked to PD.

A systematic study applied a high-throughput hypothesis-free approach to detect PD genetic risk linked to any particular biological pathway using PRS and MR (Bandres-Ciga et al., 2020b). They identified multiple biological pathways associated with PD through common genetic variation, including protein aggregation and post-translational modifications, immune response, lipid metabolism, synaptic transmission, endosomal-lysosomal dysfunction, and apoptosis. Interestingly, many of these reflect the nominal biological functions of individual GWAS hits alone (Table 1 shows gene functions of top-10 hits from PD GWAS). Overall, the main lesson learned is that the convergence of monogenic and sporadic PD pathways indicates that both disease forms are not mutually exclusive. An illustrative example is the physical interaction at the protein level between LRRK2 and the GWAS hits RAB7L1 and GAK, showing coordinated disease networks in which rare mutations and common risk alleles can act in the same pathway (Singleton, 2015). In summary, further pathway-based research holds the potential for a comprehensive understanding of disease mechanisms and advanced target discovery.

2.1.4. Oligogenic inheritance and missing heritability

Assessing the contribution of additional rare alleles in Mendelian PD genes can provide evidence of the overall "pathogenic burden" in PD pathogenesis and contribute to explaining the missing heritability in PD (Lubbe et al., 2016). Conceptually, in-depth resequencing of 391 amyotrophic lateral sclerosis (ALS) cases revealed that 64% familial and 27% sporadic cases carry additional pathogenic variants (Cady et al., 2015). Likewise, another study in PD performed a comprehensive screening in 7900 patients comprising cases with and without a known primary pathogenic genetic cause and/ or *GBA* and 6166 controls and estimated the additional burden of rare loci in known Mendelian genes responsible for PD familial forms (Lubbe et al., 2016). They found that at least 30% of cases with known pathogenic PD mutation had at least one additional rare variant in Mendelian PD genes, compared with 17% in not known mutation PD cases or 16% controls. They also observed that known mutation PD samples with additional variants had 6-year younger AAO using the NeuroX array or 4-years earlier AAO compared to cases with not known PD mutations using exome data (Lubbe et al., 2016). These findings indicate that oligogenic inheritance of rare Mendelian variants can be relevant in cases with a known primary pathogenic cause of disease.

Alternative strategies were applied to determine the burden of rare variants in top PD loci nominated in GWAS. Using exome sequencing data and restricting rare variants analysis to 56 PD GWAS nominated loci, (Chang et al., 2017; Nalls et al., 2019) a study identified additional rare coding variants within these genes in *STAB1, NOD2*, and *SH3GL2*. This study showed that the effect of additional rare variants could further modulate some associations detected in PD risk loci (Germer et al., 2019). Yet, other studies reported no rare variant enrichment in PD risk loci (Gaare et al., 2020). Another study restricting rare variants analysis only to newly defined PD GWAS loci provided further evidence of other associations of rare variants signals in putative causal genes underneath previously identified PD GWAS peaks, including *LRRK2*, *STBD1*, and *SPATA19 (*Jansen et al., 2017*)*. In the most recent PD GWAS meta-analysis, seven genes under GWAS peaks contained two or more rare coding variants after Bonferroni's correction (*LRRK2*, *GBA*, *CATSPER3*, *LAMB2*, *LOC442028*, *NFKB2*, and *SCARB2*) (Nalls et al., 2019). These results align with other reports showing that several rare and common genetic variants in *LRRK2* can affect disease risk independently (Ross et al., 2011). Altogether, these studies also support that some of the risk associated with these loci may be due to additional rare coding variants. Moreover, it has been shown that some typically monogenic PD loci such as *SNCA*, *LRRK2*, *GBA* and *VPS13C* can harbour both rare large-effect mutations and, simultaneously, common smaller effect

variants such as those typically identified by GWAS. These loci are considered as pleomorphic risk loci (Singleton and Hardy, 2011) and can also contribute to the oligogenic inheritance of PD.

In multi-ethnic cohorts, in-depth sequencing analysis of coding variants in *LRRK2* identified novel putative variants for PD (Ross et al., 2011; Kishore et al., 2019) showing that using these approaches in ethnically diverse cohorts can reveal hitherto novel variants that help understand the pathogenetic mechanisms underlying the PD pathogenesis (Ross et al., 2011; Kishore et al., 2019). In addition, another study determined the burden of additional unknown variants in genes from the lysosomal pathway such as *GBA* and the newly identified candidate genes such as *CTSD*, *SLC17A5*, and *ASAH1 (*Robak et al., 2017*)*.

Interestingly, they observed that most PD cases in their cohort have at least one putative damaging variant in a lysosomal storage disorder gene, and 21% carry multiple alleles, thus providing compelling evidence of the oligogenic inheritance affecting specific pathways involved in PD pathogenesis such as the lysosomal pathway (Robak et al., 2017). Lastly, additional loss-of-function (Bobbili et al., 2020) or extreme rare variants (Bustos et al., 2020 BioRxiv doi. org/10.1101/2020.06.06.137299) identified by exome sequencing have also been recently suggested to contribute to the genetic burden of PD. In summary, data from available studies support a potential oligogenic inheritance in some PD cases encompassing additional rare variants related to PD pathogenic causes and/ or to PD GWAS peaks, but further large-scale unbiased studies are warranted to address this question fully.

2.1.5. *From risk association to causality*

As the number of risk loci implicated in PD increases, the critical challenge will be to understand how these variants mechanistically contribute to the disease risk, given that the majority of variants are located in non-coding regions and have small effect estimates. Thus, GWAS have proven to have a limited scope to identify the "causal" gene in multi-gene loci because neither the variant with the strongest association (lead SNP) nor the nearest gene are necessarily causal. For that reason, it is becoming common practice to avoid naming a genomic locus for the “candidate” gene assigned in the original research article. Here, the GWAS-associated candidate loci need to be mapped to variants and genes through functional genomics studies that combine annotation of variants, gene expression, and gene-based or pathway-based analyses. For instance, the loci containing the GWAS signals around *MAPT* of *GAK* are examples of variants requiring functional validation. In such cases, the functional modelling of candidate loci is instrumental in assessing true causality.

Several approaches are being used to prioritise putative causal variants and genes pinpointed by GWAS, and PD OMICs open resources are already available (Schilder et al., 2022). One strategy to prioritise functional variants is to filter GWAS hits by overlapping with additional quantitative trait loci (QTL) thus providing further functional evidence. QTLs can typically include gene expression (transcriptomics and RNA-seq) or alternative splicing data, epigenomic data (DNA methylation or histone marks), protein levels, and virtually post-translational modifications (e.g., protein phosphorylation). Using QTLs at proxy genes (DNA methylation or RNA expression), the 305 genes under GWAS peaks from the latest PD GWAS could be narrowed down to 78 loci (Nalls et al., 2019). This inference of functional variants was made by MR comparing the local polygenic risk of the exposure (methylation or expression) to the polygenic risk in an outcome (PD), assuming no additional confounders and that the association is not due to LD. Yet, alternative methods such as Bayesian colocalisation allow the overlap of GWAS hits and QTL while mitigating the confounding effects of LD (Schilder and Raj, 2022).

Moreover, along with the simultaneous spur in GWAS, genome-wide functional studies have shown that the risk loci identified from GWAS are primarily enriched in disease-relevant cell types. Thus, identifying disease-specific cell types will expectably provide a better understanding of the disease’s pathophysiology (Hannon et al., 2019; Eichler et al., 2010). For instance, a study integrating gene expression data and publicly available GWAS summary statistics from several traits found significant tissue-specific enrichment for 34 diseases (Finucane et al., 2015). Such approaches are expected to enhance the biological interpretability of GWAS signals (Finucane et al., 2015; Fernando et al., 2020). More specifically, one single-nuclei RNA-seq study of post-mortem substantia nigra (SN) reported distinct sets of neuron-specific genes for different neuropsychiatric disorders yet converged onto shared loci within oligodendrocytes (Agarwal et al., 2020). Similarly, a mouse single-cell transcriptomic study overlapping with PD GWAS data revealed enrichment of not only cholinergic and monoaminergic neurons but also enteric neurons and oligodendrocytes (Bryois et al., 2020).

A single-cell multi-omic study in post-mortem midbrain also identified biological pathways relevant to PD SN (neuroinflammation, immune response activation, mitochondrial and synaptic dysfunction) (Caldi Gomes et al., 2022). Such differential patterns are also likely to reveal cell-type-specific deregulation at different levels. Here, 3D models of the chromatin, including DNA methylation and key functional histone marks, e.g. H3K27ac (active enhancer), H3K4me1 (poised enhancer), and H3K4me3 (promoter), can also be instrumental in identifying functional GWAS variants (Schilder et al., 2022). Thus, a study showed that one non-coding distal enhancer element regulates the expression levels of *SNCA* and is related to the activity of brain-specific transcription factors (TF) such as EMX2 and NKX6–1 (Soldner et al., 2016). Lastly, in-vitro and in-vivo functional modelling of the identified putative causal variants, also by CRISPR-Cas9, shall ultimately reveal true causality of variants identified by GWAS. In addition, various ongoing efforts such as AMP-PD, FOUNDIN PD, and GP2 have been ongoing to develop a comprehensive multi-omics profile that aims to address the role of common/rare and the impact of structural variants on PD pathogenesis.

While the studies integrating GWAS findings and functional data offer mechanistic clues, other approaches such as MR are being applied in GWAS settings to understand how one phenotype (exposure) is causally related to the other (outcome) (Pingault et al., 2018). In this context, MR strengthens causal inference by using genetic variants to mimic a randomised controlled trial (RCT) (Pingault et al., 2018). The availability of GWAS datasets across a broad spectrum of phenotypes "unlocks" the potential to understand the relationship between the exposure and the outcome (Bandres-Ciga et al., 2019b). For example, one of the earlier studies using PD GWAS showed a causal relationship between increased iron levels in serum and reduced PD risk (Pichler et al., 2013). Since then, there has been an increase in MR studies in PD, thus enhancing our understanding of various exposures that may or may not be causally associated with PD (Nalls et al., 2019; Bandres-Ciga et al., 2019a; Bandres-Ciga et al., 2019; Noyce et al., 2017a). It is anticipated that the availability of deeply phenotyped clinical cohorts will help determine novel relationships between exposures and outcomes and help to prioritise targets for clinical trials (Noyce et al., 2019).

In summary, integration of multi-omic data, including epigenetics, i.e., methylation and chromatin immunoprecipitation ChIP-seq data, bulk/single-cell gene expression data, and QTLs, among other approaches aimed at uncovering and elucidating the mechanisms underlying any associated variant(s) are expected to expand our understanding of PD GWAS hits.

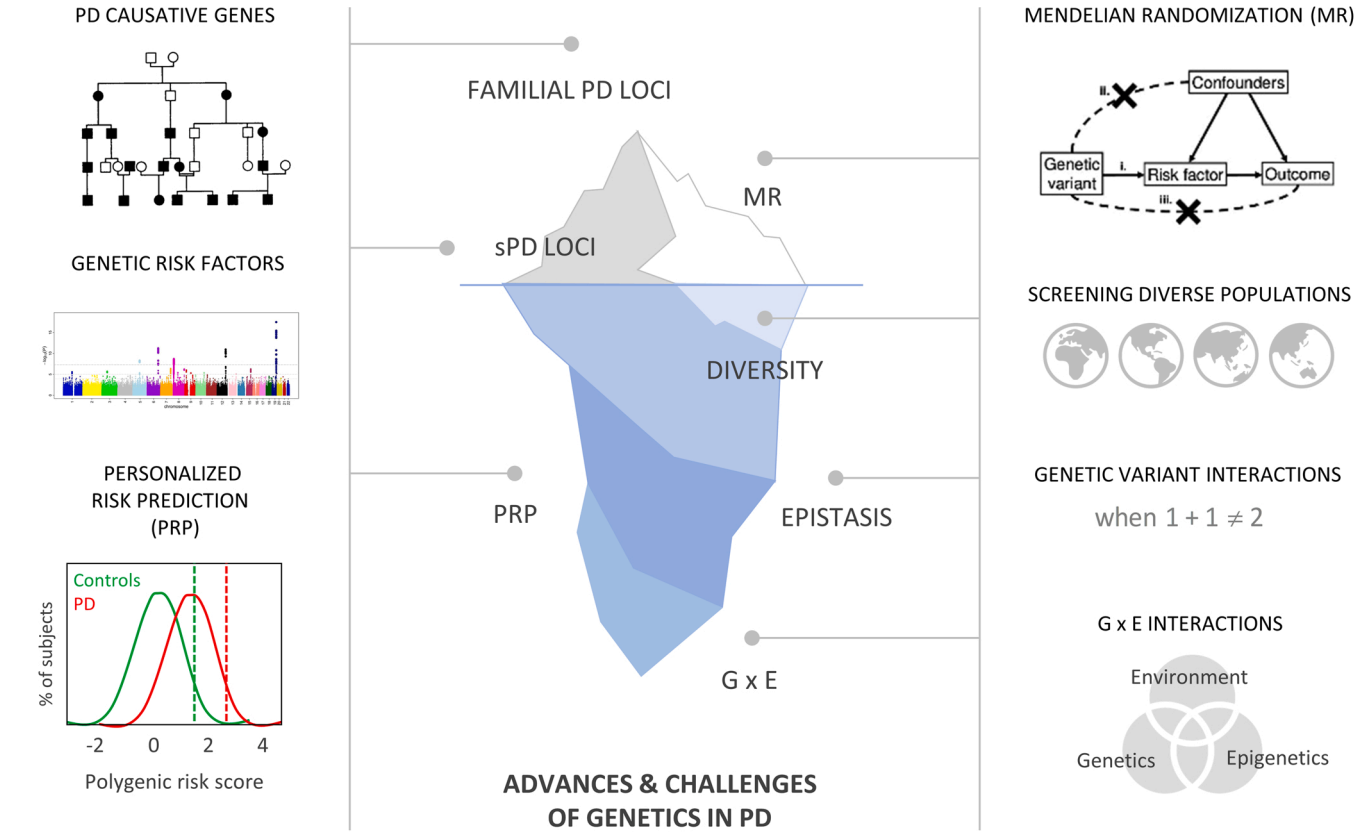

Fig. 2. : The iceberg analogy of PD genetic research. The visible part of the iceberg represents the current PD genetic discoveries status, with familial PD genetics, sporadic PD loci discovered by GWAS, or Mendelian Randomization (MR) as a promising approach to decipher risk. However, the central part of the missing heritability in PD is still submerged. For instance, diversifying PD genomic research by GWAS and uncovering genetic risk in the different ancestries is within reach. Regarding personalised risk prediction (PRP) models, it is aimed that the combination of comprehensive genomic characterisation with detailed clinical data, especially longitudinal progression data, shall eventually improve the predictive capacity. Lastly, gene-gene (epistasis) and gene-environmental (G x E) interactions are still challenging and underexplored. The advent of new methods and approaches shall help broaden the overall genetic understanding of PD.

2.2. Challenges & opportunities

2.2.1. From risk association to risk prediction

While the number of loci found by GWAS grows exponentially, there remains a need for inferring clinically relevant genetic readouts (Fig. 2). Two approaches are being applied for risk prediction models, i.e., PRS and machine learning (van Rheenen et al., 2019; Guyon and Elisseeff, 2003). Based on a fixed model approach, PRS are modelled by considering the contribution of risk alleles, either weighted or unweighted. In weighted PRS, the contribution of each risk allele is computed by their effect estimates, whereas unweighted PRS assumes that each risk allele is contributing to the same effect size (van Rheenen et al., 2019). The latter limits the efficacy for modelling risk prediction based on GWAS data since most of the genetic loci identified in complex diseases such as PD often show the different magnitude of effect estimates (Jakobsdottir et al., 2009). In these diseases, PRS have shown limited portability across groups of different genetic ancestries, but also of socio-economic status, age, or gender, (Mostafavi et al., 2020) yet it is unclear whether this applies explicitly to PD. First PRS-based studies in Asia (Foo et al., 2020) and Latin America (Loesch et al., 2021) have successfully started to elucidate the genetic heterogeneity of PD genetic risk factors across different populations. Further studies aimed at detecting population-specific variation and minimising potential predictive disparities by PRS are promising.

Alternatively, machine learning approaches such as support vector machine (SVM) or random forest (RF) can be applied to develop more robust predictive models, either using supervised or unsupervised methods for risk prediction modelling (Guyon and Elisseeff, 2003). Unlike number-of-risk-allele-based approaches, SVMs discriminate between "case" and "control" by finding a separating hyperplane for the data points by transforming input data, i.e., SNP genotypes, into a higher dimensional feature space. However, it should be noted that the data feature selection is the main factor influencing the predictive behaviour of the models. For instance, machine learning models using GWAS top hits are considered the most informative approach for developing risk models. By taking into account the weight of top loci, the initial disease prediction models were predicated on the assumption that top GWAS loci can also be effective classifiers and valuable for clinical decision making (Mihaescu et al., 2010). Yet, these studies have shown only a marginal increase in area under the area under the receiver operating characteristic (ROC) curve (AUC) of performance, thus hampering the translation of findings to the clinical practice to date.

To develop accurate disease prediction models, either by PRS or machine learning, the common denominator determining the predictive power of a model in an independent cohort is the sample size of the training data set, based on which different predictive models can be built. As the number of PD risk loci will scale up with the ever-increasing sample size in future GWAS, it would be anticipated that prediction models based on training data sets will progressively yield more accurate predictive models, as shown in the most recently published meta-GWAS (Nalls et al., 2019). Nevertheless, it has to be acknowledged that the application of such approaches in developing prediction models

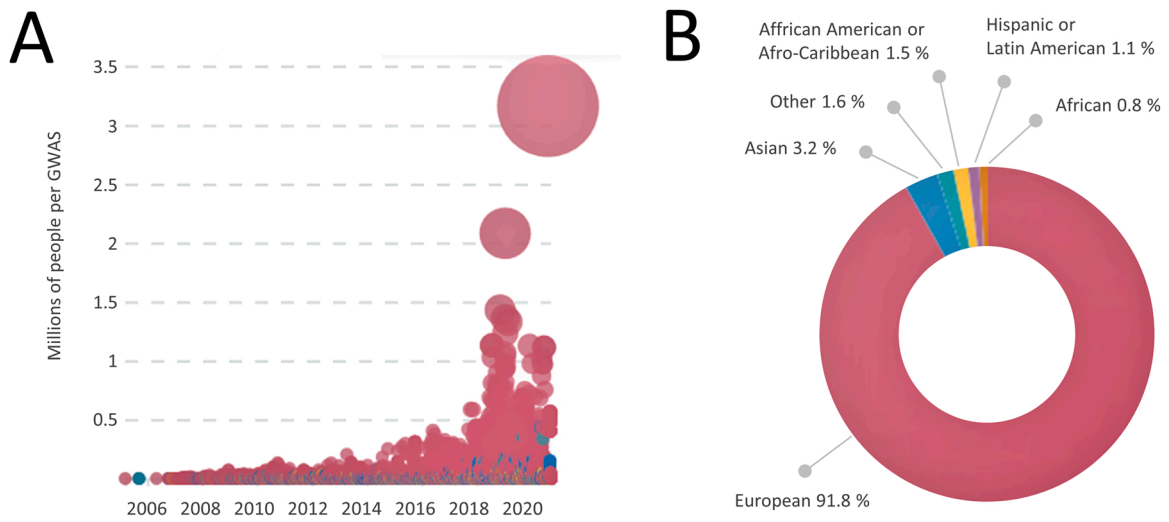

Fig. 3. : The worldwide diversity of GWAS. The genetic discoveries in PD are driven by European ancestry. (**A**) Graphical representation of GWAS by time and ancestry where the X-axis represents the years and the Y-axis represents the number of samples for a GWAS study. (B) doughnut plot that describes the ancestry of GWAS participants as adapted from http://gwasdiversitymonitor.com.

based on genetic risk variants alone may not be able to estimate the actual risk entirely in complex diseases, as most the complex diseases, including PD, are an interplay between genetic and environmental risk factors. However, the clinical utility of disease prediction models in PD might progressively be improved as long as future approaches use patient enriched cohorts with detailed clinical and environmental exposure data. Lastly, including age as one of the most consistent predicting factors in PD prediction models seems to be highly advisable (Dorsey et al., 2018).

2.2.2. *PD GWAS in underrepresented populations*

The decentralisation and democratisation of genomic research are expected to provide further impetus for identifying novel genetic discoveries hitherto not captured by current approaches (International Parkinson Disease Genomics Consortium (IPDGC), 2020). Akin to other diseases, PD GWAS has been centred mainly on European ancestry (Fig. 3) (Boeve, 2010). However, cancer GWAS in non-European ethnicities has identified ancestry-specific variation relevant to disease, describing genetic risk factors relevant to specific populations (Park et al., 2018). As an example, compared to Europeans, PD GWAS from the East Asian population have shown *SNCA* and *LRRK2, PARK16*, and *BST1* as commonly shared risk loci but identified population-specific hits such as *MAPT* or *GBA* in Europeans but not in Asians (Nalls et al., 2014; Foo et al., 2017; Satake et al., 2009). In addition, for *GBA*, population-specific variants have been identified in non-Europeans (Velez-Pardo et al., 2019; Zhang et al., 2018). Thus, uncovering and characterising PD risk loci in diverse populations is essential for understanding the underlying biological mechanisms of disease and developing effective genetic risk prediction models in non-Europeans (Singleton and Hardy, 2019). Recognising this unmet need to diversify genomic research in PD, large multi-national consortia have recently initiated an effort to catalogue genomic diversity (International Parkinson Disease Genomics Consortium (IPDGC), 2020). These include the LARGE-PD consortium (Latin American Research Consortium on the Genetics of Parkinson's Disease), (Zabetian and Mata, 2017) the Luxembourg-German-Indian Alliance on Neurodegenerative diseases (Lux-GIANT)[106], or recent efforts in Southeast Asia, (Foo et al., 2017) China, and Africa spearheaded by the International Parkinson Disease Genomics Consortium[98] (IPDGC) and the Genetic Epidemiology of Parkinson disease (GEoPD) consortium (Wang et al., 2017). Such initiatives are making significant progress in integrating ethnic diversity in PD GWAS. Applying the transversal strategies proposed here to these populations, along with the novel opportunities they offer for advanced GWAS, is expected to comprehensively characterise genetic susceptibility for PD in underrepresented populations and maximise the potential clinical correlation (Zabetian and Mata, 2017; Rajan et al., 2020). Moreover, deciphering the genetic architecture of PD across diverse populations hold direct implications for the reproducibility and validation of the European-descendent loci in admixture populations. Lastly, recent studies have also suggested an essential role of genetic ancestry in population pharmacogenetics (Yang et al., 2021).

2.2.3. *Epistasis*

Epistasis is non-linear complex gene-gene interactions among two or more genetic variants in determining specific traits of the phenotypic variation, (Gros et al., 2009), e.g., disease susceptibility (Moore, 2003). In complex diseases, epistasis can likely explain part of the missing heritability (Fig. 2). PD GWAS have addressed the missing heritability[7] but not yet by unbiased, comprehensive multi-locus assessment of epistasis. Challenges lay in the vast number of statistical tests, the limiting computational capacity, and the biological interpretation. Methods reducing such complexity are being developed, e.g., by screening two-locus interactions filtered by genotype independence test and calculating pairs with non-equilibrium frequencies by logistic regression (Pecanka et al., 2017; Cordell, 2009). Such approaches have successfully identified the interaction of two calcium channel subunits in bipolar disorder (Prabhu and Pe'er, 2012). Alternatively, hypothesis-based approaches based on functional interaction have been proposed (Bochdanovits et al., 2008). At the multi-locus level, the multi-dimension reduction analysis (MDR) method has been developed to uncover higher-order gene-gene interactions underlying complex disorders (Ritchie et al., 2001). This method assesses the relation of multi-locus "haplotypic" combinations of SNPs with disease risk or specific clinical outcomes. Using MDR and hypothesis-based centred on the Akt/mTOR pathway, two epistasis studies in PD identified up to 4-loci interactions modulating the PD risk and AAO determined by *SNCA*

variants (Fernández-Santiago et al., 2019). Also, they increased susceptibility to dyskinesia's in L-DOPA treated PD patients (Martín-Flores et al., 2019). Advanced PD GWAS should also explore such strategies reducing epistasis complexity. They could maximise significant previous economic investments by re-analysing the available PD GWAS large datasets to provide knowledge of the missing heritability in PD and identify multi-locus epistatic effects modulating, e.g., penetrance or AAO of *LRRK2* mutations, or explaining differential susceptibility in *GBA*.

2.2.4. Environmental cues

Beyond genetic susceptibility factors, environmental cues influencing the risk of PD are to be considered in advanced GWAS for a comprehensive overview of disease risks (Fig. 2). However, the unavailability of exposome data has been a limiting issue for most cohorts. Thus, pesticide exposure, rural living, and agriculture occupation have been linked to increased PD risk in populational studies, (Przedborski, 2017) and smoking and coffee drinking correlated inversely (Hernán et al., 2002). Regarding industry chemicals related to urbanisation, a recent study in South Korea assessing the effect of particulate matters (ozone, nitrogen dioxide, sulfur dioxide, and carbon monoxide) reported increased PD risk for nitrogen dioxide (Jo et al., 2021). If validated, these findings highlight the role of air pollution impacting the quality of life and its adverse impact specifically on PD. Another aspect that has been often overlooked in complex diseases is the gene-by-environment (G x E) interactions (Hunter, 2005) reflecting the combined synergistic effects of loci and environment in modulating disease outcomes. However, G x E interactions need to be considered to minimise over-or under-estimates of the overall heritability (Gage et al., 2016). As a successful example, a study showed that PD risk variants identified by GWAS such as rs660895 (A/G) at the *HLA* locus are further modulated by smoking, possibly through common inflammatory pathways (Chuang et al., 2017). This study showed that anti-inflammatory protective effects of the G allele were specifically restricted to never smokers. Nevertheless, conceptually, what will be the possible way to decipher the G x E in PD? GWAS will continue to reveal new loci in the foreseeable future, and it might well be that some environmental cues will exert an indirect effect on the disease risk rather than direct, mediated by exposure. However, G x E approaches are a nascent approach to complex diseases, and their effects, although biologically relevant, may remain hidden (Dahl et al., 2020). Thus, despite the collection of environmental data being sparse so far, a recent study provided a framework that can be applied in future GWAS to discern the role of G x E in prediction models (Li et al., 2019). The availability of large-scale cohorts with systematic environmental data in PD, such as the UK biobank, are expected to provide novel opportunities to directly assess the role of G x E, as shown in recent studies using PRS (Jacobs et al., 2020). In summary, combined analyses of GWAS and exposome data, when available, including G x E interactions, may result in a holistic biological understanding of the overall PD risk, but for the moment, we are somehow far from this scenario.

2.2.5. Clinically enriched, harmonised, multi-centre cohorts for advanced PD GWAS

Genetic discoveries are expected to provide individual risk profiles to the clinic (Fig. 4) (Kim and Ober, 2019). In turn, as shown in cardiovascular disease and rheumatoid arthritis (RA), better availability of clinical data improves disease prediction (Weng et al., 2017). Yet, the limited accessibility of detailed clinic-demographic data in PD GWAS (often disease status, age, gender, ethnicity, AAO, and disease duration) has partially limited potential clinical correlation of risk profiles (Tam et al., 2019). Earlier prediction studies integrating in-depth clinical information have increased disease prediction (Liu et al., 2017). Accordingly, it is conceivable that integration of, for instance, non-motor symptoms and genetic risk profiles may help increase the predictive utility of GWAS and provide stratified cohorts based on trials. Thus, the Movement Disorders Society (MDS) has established research criteria for prodromal stages of PD (Berg et al., 2015; Heinzel et al., 2019). Although RBD patients show up to 75% pheno-conversion rates 10 years after the RBD onset, (Iranzo et al., 2013) pheno-conversion can vary up to 20 years (Ponsen et al., 2004). An strategy to minimise the uncertainty of motor pheno-conversion would be to combine clinical markers indicative of prodromal disease (RBD, hyposmia, altered DaT-SPECT imaging) with genetic risk profiles from advanced GWAS to provide informative multimodal individual risk predictors (Singleton and Hardy, 2019).

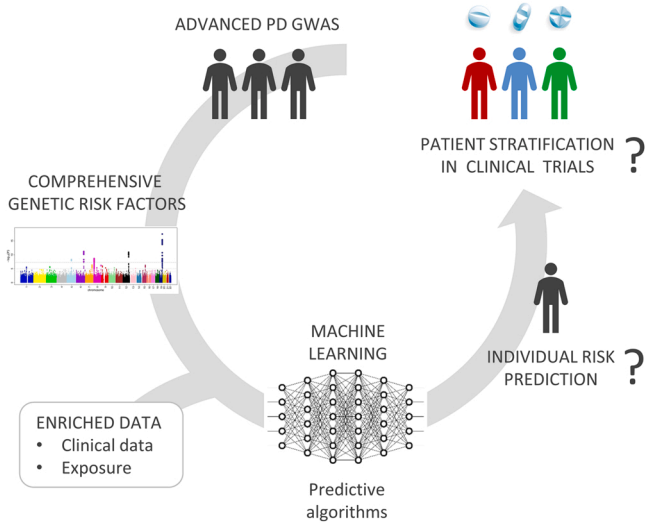

Fig. 4. : Advanced PD GWAS and clinical translatability: challenges and opportunities. Moving from genetic risk association to individual risk prediction is a natural derivate of current PD genetic research aiming at translational clinical applicability. Expectably, meta-analysing GWAS data will eventually provide novel hits for PD. The integration of genomic findings with enriched clinical exposure can be used to construct high-dimensional genomic matrixes for machine learning. If successful, predictive algorithms shall be able to profile individual genetic risk burdens. Beyond the excitement of advancing our knowledge of PD at the genome level, it is how we will design advanced GWAS in the coming years that will maximise its genuine potential for a higher clinical translatability.

Despite constraints of patient recruitment, a recent study performed a longitudinal genome-wide survival study, identified *RIMS2* as a progression locus, and provided suggestive evidence for *TMEM108* and *WWOX* as progression loci (Liu et al., 2021). Although further validation studies are warranted, this study underscores a comprehensive spectrum of PD genetic architecture that entails the underlying divergent genetic makeup for progression and susceptibility, respectively.

Another emerging theme is whether integrating in-depth clinical information will help predict AAO. For example, can the simultaneous use of enriched clinical cohort and the GWAS data provide a novel clinical-genetic readout to predict AAO in non-affected carriers of pathogenic mutations in the *LRRK2* gene? *LRRK2* mutation carriers are at-risk of PD, but given that penetrance is reduced, it is difficult to predict if and when they will develop the disease. After initial reports, (Healy et al., 2008) subsequent studies established penetrance estimates

at 26–42.5% at 80 years in Ashkenazi (Marder et al., 2015) and non-Ashkenazi cohorts (Lee et al., 2017), up to 47% in Cantabria (Sierra et al., 2011) or 83.4% in the Basque Country, (Ruiz-Martínez et al., 2010) in Northern Spain. The *LRRK2* risk factors are different in the Chinese population, thus adding more complexity (Wang et al., 2012).

Illustratively, the penetrance of *LRRK2* mutations can vary as much, for example, 60% at 60 years in Tunisia vs 20% in Norway (Hentati et al., 2014) and their expressivity can be modulated by additional factors. Indeed, risk variants in dynamin 3 (*DNM3*) have been shown to modulate the AAO in Arab Berbers (12·5 years younger AAO for rs2421947 G/G), (Trinh et al., 2016) or in *SNCA* in Europeans (11 years earlier AAO for rs356219 G/G) (Fernández-Santiago et al., 2018), yet with apparent ethnic or population-specific effects (Brown et al., 2021). Accordingly, a comprehensive PD GWAS screening in diverse ethnicities, (International Parkinson Disease Genomics Consortium (IPDGC) (IPDGC), 2020) using detailed clinical records from longitudinal cohorts described earlier, (Nalls et al., 2015b) can hopefully contribute to identifying genetic modifiers at the genome-wide level and explain reduced penetrance of *LRRK2* mutations missing heritability in sporadic PD[7]. A recent study addressed the impact of reduced penetrance in *LRRK2* carriers (Lai et al., 2021). They performed GWAS of penetrance and AAO of PD in *LRRK2* mutation carriers and assessed whether a PRS derived from the previously published PD GWAS helped to explain the variance in *LRRK2* carriers. At a genome-wide level, they found a genetic locus, *CORO1C*, that modifies the penetrance of *LRRK2* mutations, and further, it was reported that the common variants identified so far in GWAS increase the penetrance of *LRRK2* carriers. Although this is the first study that aims to understand the penetrance of *LRRK2*, findings from this kind of study will expectably provide genetic readouts to understand mechanistically why some carriers of *LRRK2* pathogenic mutation develop PD and others do not.

Similar to *LRRK2*, heterozygous mutations at *GBA* predispose to PD with highly variable penetrance. Thus, at an advanced age of around 80 years, GBA mutations' penetrance was 11% in the U.S.A., (Rosenbloom et al., 2011) 15% in the U.K., (McNeill et al., 2012) 30% in French, (Anheim et al., 2012) 7.7% in Ashkenazi Jews (Alcalay et al., 2014) or 19% in Italians (Balestrino et al., 2020). In summary, gathering and harmonising clinical-genetic information across sites from different countries will require close multidisciplinary collaboration between clinicians and researchers. In this regard, various efforts to collect non-motor symptoms data such as hyposmia and RBD, among others, are ongoing (Nalls et al., 2015b; Noyce et al., 2017b). Hopefully, in the not-so-distant future, the derived potential outcomes based on genomics data should benefit all stakeholders fostering the advance of their research approaches, therapeutic strategies, and drug screenings (Weng et al., 2017).

3. Conclusions

It is remarkable that a disease such as PD, once considered a non-genetic entity continues providing mechanistic insight primarily driven by genetic discoveries. Translating these approaches into improved clinical care in the post-GWAS era is essential to developing novel strategies for longitudinal cohort studies. Such strategies should integrate electronic medical records, wearable sensors recording real-time information, technological advances in the collection, and analyses of biological specimens. Big data management platforms and analytical methods such as HAIL could also be implemented (https://hail.is/). Measures described above can be taken transversally, combined in such a way that PRS or machine learning strategies, including SVM or RF, could be newly undertaken in enriched clinical populations with detailed clinical data accompanying DNA samples from genetically diverse populations, not only from European ancestry, or with clinical markers (RBD, hyposmia, or *LRRK2* mutations). Advanced PD GWAS will provide better clinical-genomics readouts for developing disease prediction models with these integrative approaches. Indeed, the ongoing push to streamline longitudinal cohorts for PD will help develop personalised medicine therapeutics in the future (Lerche et al., 2015). Lastly, clinical trials currently undertaken in *LRRK2* or *GBA* patients (https://clinicaltrials.gov/) will hopefully provide the first targeted therapeutic interventions for PD. Such integrative approaches are also expected to lead to cost-effective therapeutic strategies for PD patients.

Detailed funding sources

R.F.-S. received funding from the Michael J. Fox Foundation for Parkinson's Research (MJFF) (grant #MJFF-000858 and grant #11159.01). R.F.-S. was also supported by a Jóvenes Investigadores (JIN) grant of the Spanish Ministry of Economy and Competitiveness (MINECO) / Agencia Estatal de Investigación (AEI) (AEI/FEDER/EU) (grant #SAF2015-73508-JIN), a Miguel Servet grant from the Instituto de Salud Carlos III (ISCIII) co-funded by the European Union (grant #CP19/00048), and the FIS programme from Instituto de Salud Carlos III (ISCIII) co-funded by the European Union (grant #PI20/00659). M.S. received funding from the Michael J Fox Foundation for Parkinson's Research (MJFF) Genetic Diversity in PD Program: GAP-India (grant #17473). M.S. is further supported by the German Research Council (grant #DFG/SH 599/6-1) and the MSA Coalition (grant #2017-10-006).

CRediT authorship contribution statement

R.F.-S. and M.S. contributed equally to all CRediT roles in this manuscript: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing – original draft, Writing – review & editing.

Submission declaration & verification

We declare that the work described has not been published previously (nor in the form of an abstract, a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright holder.

Acknowledgments

R.F.-S. received funding from the Michael J. Fox Foundation for Parkinson's Research (MJFF) (USA) (grant #MJFF-000858 and grant #11159.01). R.F.-S. was also supported by a Jóvenes Investigadores (JIN) grant of the Spanish Ministry of Economy and Competitiveness (MINECO) / Agencia Estatal de Investigación (AEI) (Spain) (AEI/FEDER/EU) (grant #SAF2015-73508-JIN), a Miguel Servet grant from the Instituto de Salud Carlos III (ISCIII) (Spain) co-funded by the European Union (grant #CP19/00048), and the FIS programme from the Instituto de Salud Carlos III (ISCIII) (Spain) co-funded by the European Union (grant #PI20/00659). M.S. received funding from the Michael J Fox Foundation for Parkinson's Research (MJFF) (USA) Genetic Diversity in PD Program: GAP-India (grant #17473). M.S. is further supported by the German Research Council (Germany) (grant #DFG/SH 599/6-1) and the MSA Coalition (USA) (grant #2017-10-006).

References

Agarwal, D., Sandor, C., Volpato, V., et al., 2020. A single-cell atlas of the human substantia nigra reveals cell-specific pathways associated with neurological disorders. Nat. Commun. 11 (1), 4183. https://doi.org/10.1038/s41467-020-17876-0.

Alcalay, R.N., Dinur, T., Quinn, T., et al., 2014. Comparison of parkinson risk in ashkenazi jewish patients with gaucher disease and gba heterozygotes. JAMA Neurol. 71 (6), 752–757. https://doi.org/10.1001/jamaneurol.2014.313.

Anheim, M., Elbaz, A., Lesage, S., et al., 2012. Penetrance of Parkinson disease in glucocerebrosidase gene mutation carriers. Neurology 78 (6), 417–420. https://doi.org/10.1212/WNL.0b013e318245f476.

Balestrino, R., Tunesi, S., Tesei, S., Lopiano, L., Zecchinelli, A.L., Goldwurm, S., 2020. Penetrance of glucocerebrosidase (GBA) mutations in Parkinson's disease: a kin cohort study. Mov. Disord. 35 (11), 2111–2114. https://doi.org/10.1002/mds.28200.

Bandres-Ciga, S., Saez-Atienzar, S., Bonet-Ponce, L., et al., 2019a. The endocytic membrane trafficking pathway plays a major role in the risk of Parkinson's disease. Mov. Disord. 34 (4), 460–468. https://doi.org/10.1002/mds.27614.

Bandres-Ciga, S., Noyce, A.J., Traynor, B.J., 2019b. Mendelian randomization-a journey from obscurity to center stage with a few potholes along the way. JAMA Neurol. 77 (1), 7–8. https://doi.org/10.1001/jamaneurol.2019.3419.

Bandres-Ciga, S., Diez-Fairen, M., Kim, J.J., Singleton, A.B., 2020a. Genetics of Parkinson's disease: an introspection of its journey towards precision medicine. Neurobiol. Dis. 137, 104782 https://doi.org/10.1016/j.nbd.2020.104782.

Bandres-Ciga, S., Saez-Atienzar, S., Kim, J.J., et al., 2020b. Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. Acta Neuropathol. 140 (3), 341–358. https://doi.org/10.1007/s00401-020-02181-3.

Bandres-Ciga, S., Ahmed, S., Sabir, M.S., et al., 2019. The genetic architecture of Parkinson disease in Spain: characterizing population-specific risk, differential haplotype structures, and providing etiologic insight. Mov. Disord. 34 (12), 1851–1863. https://doi.org/10.1002/mds.27864.

Berg, D., Postuma, R.B., Adler, C.H., et al., 2015. MDS research criteria for prodromal Parkinson's disease. Mov. Disord. 30 (12), 1600–1611. https://doi.org/10.1002/mds.26431.

Blauwendraat, C., Heilbron, K., Vallerga, C.L., et al., 2019. Parkinson's disease age at onset genome-wide association study: defining heritability, genetic loci, and α-synuclein mechanisms. Mov. Disord. 34 (6), 866–875. https://doi.org/10.1002/mds.27659.

Blauwendraat, C., Nalls, M.A., Singleton, A.B., 2020. The genetic architecture of Parkinson's disease. Lancet Neurol. 19 (2), 170–178. https://doi.org/10.1016/S1474-4422(19)30287-X.

Bobbili, D.R., Banda, P., Krüger, R., May, P., 2020. Excess of singleton loss-of-function variants in Parkinson's disease contributes to genetic risk. J. Med. Genet. 57 (9), 617–623. https://doi.org/10.1136/jmedgenet-2019-106316.

Bochdanovits, Z., Sondervan, D., Perillous, S., van Beijsterveldt, T., Boomsma, D., Heutink, P., 2008. Genome-wide prediction of functional gene-gene interactions inferred from patterns of genetic differentiation in mice and men. PLoS One 3 (2), e1593. https://doi.org/10.1371/journal.pone.0001593.

Boeve, B.F., 2010. REM sleep behavior disorder: updated review of the core features, the REM sleep behavior disorder-neurodegenerative disease association, evolving concepts, controversies, and future directions. NYAS5115[pii] Ann. N. Y. Acad. Sci. 1184, 15–54. https://doi.org/10.1111/j.1749-6632.2009.05115.x.

Bonifati, V., Rizzu, P., Van Baren, M.J., et al., 2003. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science 80–299 (5604), 256–259. https://doi.org/10.1126/science.1077209.

Brown, E.E., Blauwendraat, C., Trinh, J., et al., 2021. Analysis of DNM3 and VAMP4 as genetic modifiers of LRRK2 Parkinson's disease. Neurobiol. Aging 97, 148.e17–148.e24. https://doi.org/10.1016/j.neurobiolaging.2020.07.002.

Bryois, J., Skene, N.G., Hansen, T.F., et al., 2020. Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. Nat. Genet 52 (5), 482–493. https://doi.org/10.1038/s41588-020-0610-9.

Burchell, V.S., Nelson, D.E., Sanchez-Martinez, A., et al., 2013. The Parkinson's disease-linked proteins Fbxo7 and Parkin interact to mediate mitophagy. Nat. Neurosci. 16 (9), 1257–1265. https://doi.org/10.1038/nn.3489.

Cady, J., Allred, P., Bali, T., et al., 2015. Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. Ann. Neurol. 77 (1), 100–113. https://doi.org/10.1002/ana.24306.

Caldi Gomes, L., Galhoz, A., Jain, G., et al., 2022. Multi-omic landscaping of human midbrains identifies disease-relevant molecular targets and pathways in advanced-stage Parkinson's disease. Clin. Transl. Med. 12 (1), e692 https://doi.org/10.1002/ctm2.692.

Chang, D., Nalls, M.A., Hallgrímsdóttir, I.B., et al., 2017. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nat. Genet. 49 (10), 1511–1516. https://doi.org/10.1038/ng.3955.

Chuang, Y.H., Lee, P.C., Vlaar, T., et al., 2017. Pooled analysis of the HLA-DRB1 by smoking interaction in Parkinson disease. Ann. Neurol. 82 (5), 655–664. https://doi.org/10.1002/ana.25065.

Collins, F.S., McKusick, V.A., 2001. Implications of the human genome project for medical science. J. Am. Med. Assoc. 285 (5), 540–544. https://doi.org/10.1001/jama.285.5.540.

Cookson, M.R., 2012. Parkinsonism due to mutations in PINK1, Parkin, and DJ-1 and oxidative stress and mitochondrial pathways. Cold Spring Harb. Perspect. Med. 2 (9), a009415 https://doi.org/10.1101/cshperspect.a009415.

Cordell, H.J., 2009. Detecting gene-gene interactions that underlie human diseases. Nat. Rev. Genet. 10 (6), 392–404. https://doi.org/10.1038/nrg2579.

Dahl, A., Nguyen, K., Cai, N., Gandal, M.J., Flint, J., Zaitlen, N., 2020. A robust method uncovers significant context-specific heritability in diverse complex traits. Am. J. Hum. Genet. 106 (1), 71–91. https://doi.org/10.1016/j.ajhg.2019.11.015.

Del Rey, N.L.-G., Quiroga-Varela, A., Garbayo, E., 2018. Advances in Parkinson's disease: 200 years later. Front. Neuroanat. 12, 113. https://doi.org/10.3389/fnana.2018.00113.

Doherty, K.M., Silveira-Moriyama, L., Parkkinen, L., et al., 2013. Parkin disease: a clinicopathologic entity? JAMA Neurol. 70 (5), 571–579. https://doi.org/10.1001/jamaneurol.2013.172.

Dorsey, E.R., Sherer, T., Okun, M.S., Bloemd, B.R., 2018. The emerging evidence of the Parkinson pandemic. J. Parkinson's Dis. 8 (s1), S3–S8. https://doi.org/10.3233/JPD-181474.

Edvardson, S., Cinnamon, Y., Ta-Shma, A., et al., 2012. A deleterious mutation in DNAJC6 encoding the neuronal-specific clathrin-uncoating Co-chaperone auxilin, is associated with juvenile parkinsonism. PLoS One 7 (5), e36458. https://doi.org/10.1371/journal.pone.0036458.

Eichler, E.E., Flint, J., Gibson, G., et al., 2010. Missing heritability and strategies for finding the underlying causes of complex disease. Nat. Rev. Genet. 11 (6), 446–450. https://doi.org/10.1038/nrg2809.

Erb, M.L., Moore, D.J., 2020. LRRK2 and the endolysosomal system in Parkinson's disease. J. Parkinson's Dis. 10 (4), 1271–1291. https://doi.org/10.3233/JPD-202138.

Fernández-Santiago, R., Garrido, A., Infante, J., et al., 2018. synuclein (SNCA) but not dynamin 3 (DNM3) influences age at onset of leucine-rich repeat kinase 2 (LRRK2) Parkinson's disease in Spain. Mov. Disord. 33 (4), 637–641. https://doi.org/10.1002/mds.27295.

Fernández-Santiago, R., Martín-Flores, N., Antonelli, F., et al., 2019. SNCA and mTOR pathway single nucleotide polymorphisms interact to modulate the age at onset of Parkinson's disease. Mov. Disord. 34 (9), 1333–1344. https://doi.org/10.1002/mds.27770.

Fernando, M.B., Ahfeldt, T., Brennand, K.J., 2020. Modeling the complex genetic architectures of brain disease. Nat. Genet. 52 (4), 363–369. https://doi.org/10.1038/s41588-020-0596-3.

Finucane, H.K., Bulik-Sullivan, B., Gusev, A., et al., 2015. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. 47 (11), 1228–1235. https://doi.org/10.1038/ng.3404.

Fonzo, A.D., Dekker, M.C.J., Montagna, P., et al., 2009. FBXO7 mutations cause autosomal recessive, early-onset parkinsonian- pyramidal syndrome. Neurology 72 (3), 240–245. https://doi.org/10.1212/01.wnl.0000338144.10967.2b.

Foo, J.N., Tan, L.C., Irwan, I.D., 2017. Genome-wide association study of Parkinson's disease in East Asians. Hum. Mol. Genet. 26 (1), 226–232. https://doi.org/10.1093/hmg/ddw379.

Foo, J.N., Chew, E.G.Y., Chung, S.J., et al., 2020. Identification of risk loci for Parkinson disease in Asians and comparison of risk between Asians and Europeans: a genome-wide association study. JAMA Neurol. 77 (6), 746–754. https://doi.org/10.1001/jamaneurol.2020.0428.

Funayama, M., Ohe, K., Amo, T., et al., 2015. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. Lancet Neurol. 14 (3), 274–282. https://doi.org/10.1016/S1474-4422(14)70266-2.

Gaare, J.J., Nido, G., Dölle, C., et al., 2020. Meta-analysis of whole-exome sequencing data from two independent cohorts finds no evidence for rare variant enrichment in Parkinson disease associated loci. PLoS One 15 (10), e0239824. https://doi.org/10.1371/journal.pone.0239824.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., et al., 2002. The structure of haplotype blocks in the human genome. Science 80–296 (5576), 2225–2229. https://doi.org/10.1126/science.1069424.

Gage, S.H., Davey Smith, G., Ware, J.J., Flint, J., Munafò, M.R., 2016. G = E: What GWAS Can Tell Us about the environment. PLoS Genet. 12 (2) https://doi.org/10.1371/journal.pgen.1005765.

Germer, E.L., Imhoff, S., Vilariño-Güell, C., et al., 2019. The role of rare coding variants in Parkinson's disease GWAS loci. Front. Neurol. 10, 1284. https://doi.org/10.3389/fneur.2019.01284.

Goldman, S.M., Marek, K., Ottman, R., et al., 2019. Concordance for Parkinson's disease in twins: a 20-year update. Ann. Neurol. 85 (4), 600–605. https://doi.org/10.1002/ana.25441.

Gros, P.-A., Le Nagard, H., Tenaillon, O., 2009. The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. Genetics 182 (1), 277–293. https://doi.org/10.1534/genetics.108.099127.

Gustavsson, A., Svensson, M., Jacobi, F., et al., 2011. Cost of disorders of the brain in Europe 2010. Eur. Neuropsychopharmacol. 21 (10), 718–779. https://doi.org/10.1016/j.euroneuro.2011.08.008.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3 (Mar), 1157–1182. ⟨http://www.jmlr.org/papers/v3/guyon03a.html⟩.

Hamza, T.H., Zabetian, C.P., Tenesa, A., et al., 2010. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. Nat. Genet. 42 (9), 781. https://doi.org/10.1038/ng.642.
Hannon, E., Marzi, S.J., Schalkwyk, L.S., Mill, J., 2019. Genetic risk variants for brain disorders are enriched in cortical H3K27ac domains. Mol. Brain 12 (1). https://doi.org/10.1186/s13041-019-0429-4.
Healy, D.G., Falchi, M., O'Sullivan, S.S., et al., 2008. Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. Lancet Neurol. 7 (7), 583–590. https://doi.org/10.1016/S1474-4422(08)70117-0.
Heinzel, S., Berg, D., Gasser, T., Chen, H., Yao, C., Postuma, R.B., 2019. Update of the MDS research criteria for prodromal Parkinson's disease. Mov. Disord. 34 (10), 1464–1470. https://doi.org/10.1002/mds.27802.
Hentati, F., Trinh, J., Thompson, C., Nosova, E., Farrer, M.J., Aasly, J.O., 2014. LRRK2 parkinsonism in Tunisia and Norway: a comparative analysis of disease penetrance. Neurology 83 (6), 568–569. https://doi.org/10.1212/WNL.0000000000000675.
Hernán, M.A., Takkouche, B., Caamaño-Isorna, F., Gestal-Otero, J.J., 2002. A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease. Ann. Neurol. 52 (3), 276–284. https://doi.org/10.1002/ana.10277.
Holmans, P., Moskvina, V., Jones, L., et al., 2013. A pathway-based analysis provides additional support for an immune-related genetic susceptibility to Parkinson's disease. Hum. Mol. Genet. 22 (5), 1039–1049. https://doi.org/10.1093/hmg/dds492.
Hunter, D.J., 2005. Gene-environment interactions in human diseases. Nat. Rev. Genet. 6 (4), 287–298. https://doi.org/10.1038/nrg1578.
Ibanez, L., Bahena, J.A., Yang, C., et al., 2020. Functional genomic analyses uncover APOE-mediated regulation of brain and cerebrospinal fluid beta-amyloid levels in Parkinson disease. Acta Neuropathol. Commun. 8 (1), 196. https://doi.org/10.1186/s40478-020-01072-8.
International Parkinson Disease Genomics Consortium (IPDGC) Singleton A. Ten Years of the International Parkinson Disease Genomics Consortium: Progress and Next Steps, J. Park. Dis., 10, 1, 2020, pp. 19–30 doi: 10.3233/JPD-191854.
Ioannidis JPA. Why most published research findings are false. In: Getting to Good: Research Integrity in the Biomedical Sciences. Vol 2. PLoS Med; 2018:2–8. doi:⟨10.1371/journal.pmed.0020124⟩.
IPDGC WTCCC2 IPDGC, International Parkinson's Disease Genomics C, Wellcome Trust Case Control C. A two-stage meta-analysis identifies several new loci for Parkinson's disease, PLoS Genet, 7, 6, 2011, e1002142 doi: 10.1371/journal.pgen.1002142PGENETICS-D-11-00446.([pii]).
Iranzo, A., Tolosa, E., Gelpi, E., et al., 2013. Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. S1474-4422(13)70056-5 [pii] Lancet Neurol. 12 (5), 443–453. https://doi.org/10.1016/S1474-4422(13)70056-5.
Jacobs, B.M., Belete, D., Bestwick, J., et al., 2020. Parkinson's disease determinants, prediction and gene-environment interactions in the UK Biobank. J. Neurol. Neurosurg. Psychiatry 91 (10), 1046–1054. https://doi.org/10.1136/jnnp-2020-323646.
Jakobsdottir, J., Gorin, M.B., Conley, Y.P., Ferrell, R.E., Weeks, D.E., 2009. Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers. PLoS Genet. 5 https://doi.org/10.1371/journal.pgen.1000337.
Jansen, I.E., Gibbs, J.R., Nalls, M.A., et al., 2017. Establishing the role of rare coding variants in known Parkinson's disease risk loci. Neurobiol. Aging 59, 220.e11–220.e18. https://doi.org/10.1016/j.neurobiolaging.2017.07.009.
Jo, S., Kim, Y.J., Park, K.W., et al., 2021. Association of NO_2 and other air pollution exposures with the risk of Parkinson disease. JAMA Neurol. 78 (7), 800–808. https://doi.org/10.1001/jamaneurol.2021.1335.
Johnson, G.C.L., Esposito, L., Barratt, B.J., 2001. Haplotype tagging for the identification of common disease genes. Nat. Genet. 29 (2), 233–237. https://doi.org/10.1038/ng1001-233.
Keller, M.F., Saad, M., Bras, J., et al., 2012. Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. Hum. Mol. Genet. 21 (22), 4996–5009. https://doi.org/10.1093/hmg/dds335.
Kim, K.W., Ober, C., 2019. Lessons learned from GWAS of asthma. Allergy, Asthma Immunol. Res. 11 (2), 170–187. https://doi.org/10.4168/aair.2019.11.2.170.
Kishore, A., Ashok Kumar Sreelatha, A., Sturm, M., et al., 2019. Understanding the role of genetic variability in LRRK2 in Indian population. Mov. Disord. 34 (4), 496–505. https://doi.org/10.1002/mds.27558.
Kitada, T., Asakawa, S., Hattori, N., et al., 1998. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 392 (6676), 605–608. https://doi.org/10.1038/33416.
Klein, R.J., Zeiss, C., Chew, E.Y., et al., 2005. Complement factor H polymorphism in age-related macular degeneration. Science 80–308 (5720), 385–389. https://doi.org/10.1126/science.1109557.
Kowal, S.L., Dall, T.M., Chakrabarti, R., Storm, M.V., Jain, A., 2013. The current and projected economic burden of Parkinson's disease in the United States. Mov. Disord. 28 (3), 311–318. https://doi.org/10.1002/mds.25292.
Krebs, C.E., Karkheiran, S., Powell, J.C., et al., 2013. The sac1 domain of SYNJ1 identified mutated in a family with early-onset progressive parkinsonism with generalised seizures. Hum. Mutat. 34 (9), 1200–1207. https://doi.org/10.1002/humu.22372.
Lai, D., Alipanahi, B., Fontanillas, P., et al., 2021. Genomewide association studies of LRRK2 modifiers of Parkinson's disease. Ann. Neurol. 90 (1), 76–88. https://doi.org/10.1002/ana.26094.
Lee, A.J., Wang, Y., Alcalay, R.N., et al., 2017. Penetrance estimate of LRRK2 p.G2019S mutation in individuals of non-Ashkenazi Jewish ancestry. Mov. Disord. 32 (10), 1432–1438. https://doi.org/10.1002/mds.27059.
Lerche, S., Liepelt-Scarfone, I., Alves, G., et al., 2015. Methods in neuroepidemiology characterization of European longitudinal cohort studies in Parkinson's disease - report of the JPND Working Group BioLoC-PD. Neuroepidemiology 45 (4), 282–297. https://doi.org/10.1159/000439221.
Lesage, S., Drouet, V., Majounie, E., et al., 2016. Loss of VPS13C function in autosomal-recessive Parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent mitophagy. Am. J. Hum. Genet. 98 (3), 500–513. https://doi.org/10.1016/j.ajhg.2016.01.014.
Li, J., Li, X., Zhang, S., Snyder, M., 2019. Gene-environment interaction in the era of precision medicine. Cell 177 (1), 38–44. https://doi.org/10.1016/j.cell.2019.03.004.
Liu, G., Locascio, J.J., Corvol, J.C., et al., 2017. Prediction of cognition in Parkinson's disease with a clinical–genetic score: a longitudinal analysis of nine cohorts. Lancet Neurol. 16 (8), 620–629. https://doi.org/10.1016/S1474-4422(17)30122-9.
Liu, G., Peng, J., Liao, Z., et al., 2021. Genome-wide survival study identifies a novel synaptic locus and polygenic score for cognitive progression in Parkinson's disease. Nat. Genet. 53 (6), 787–793. https://doi.org/10.1038/s41588-021-00847-6.
Loesch, D.P., Horimoto, A.R.V.R., Heilbron, K., et al., 2021. Characterising the genetic architecture of Parkinson's disease in Latinos. Ann. Neurol. 90 (3), 353–365. https://doi.org/10.1002/ana.26153.
Lubbe, S.J., Escott-Price, V., Gibbs, J.R., et al., 2016. Additional rare variant analysis in Parkinson's disease cases with and without known pathogenic mutations: evidence for oligogenic inheritance. Hum. Mol. Genet. 25 (24), 5483–5489. https://doi.org/10.1093/hmg/ddw348.
Luoma, P., Melberg, A., Rinne, J.O., et al., 2004. Parkinsonism, premature menopause, and mitochondrial DNA polymerase γ mutations: clinical and molecular genetic study. Lancet 364 (9437), 875–882. https://doi.org/10.1016/S0140-6736(04)16983-3.
Marder, K., Wang, Y., Alcalay, R.N., et al., 2015. Age-specific penetrance of LRRK2 G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 consortium. Neurology 85 (1), 89–95. https://doi.org/10.1212/WNL.0000000000001708.
Martín-Flores, N., Fernández-Santiago, R., Antonelli, F., et al., 2019. MTOR pathway-based discovery of genetic susceptibility to L-DOPA-induced dyskinesia in Parkinson's disease patients. Mol. Neurobiol. 56 (3), 2092–2100. https://doi.org/10.1007/s12035-018-1219-1.
McNeill, A., Duran, R., Hughes, D.A., Mehta, A., Schapira, A.H.V., 2012. A clinical and family history study of Parkinson's disease in heterozygous glucocerebrosidase mutation carriers. J. Neurol. Neurosurg. Psychiatry 83 (8), 853–854. https://doi.org/10.1136/jnnp-2012-302402.
Mihaescu, R., Van Zitteren, M., Van Hoek, M., et al., 2010. Improvement of risk prediction by genomic profiling: reclassification measures versus the area under the receiver operating characteristic curve. Am. J. Epidemiol. 172 (3), 353–361. https://doi.org/10.1093/aje/kwq122.
Moore JH. The ubiquitous nature of epistasis in determining susceptibility to common human diseases. In: Human Heredity. Vol 56. Hum Hered; 2003:73–82. doi:⟨10.1159/000073735⟩.
Mostafavi, H., Harpak, A., Agarwal, I., Conley, D., Pritchard, J.K., Przeworski, M., 2020. Variable prediction accuracy of polygenic scores within an ancestry group. Elife 9. https://doi.org/10.7554/eLife.48376.
Murphy, K.E., Gysbers, A.M., Abbott, S.K., et al., 2014. Reduced glucocerebrosidase is associated with increased α-synuclein in sporadic Parkinson's disease. Brain 137 (3), 834–848. https://doi.org/10.1093/brain/awt367.
Nalls, M.A., Pankratz, N., Lill, C.M., et al., 2014. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. doi:ng.3043 [pii] Nat. Genet. 46 (9), 989–993. https://doi.org/10.1038/ng.3043.
Nalls, M.A., Escott-Price, V., Williams, N.M., et al., 2015a. Genetic risk and age in Parkinson's disease: continuum not stratum. Mov. Disord. 30 (6), 850–854. https://doi.org/10.1002/mds.26192.
Nalls, M.A., McLean, C.Y., Rick, J., et al., 2015b. Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: a population-based modelling study. Lancet Neurol. 14 (10), 1002–1009. https://doi.org/10.1016/S1474-4422(15)00178-7.
Nalls, M.A., Blauwendraat, C., Vallerga, C.L., et al., 2019. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 18 (12), 1091–1102. https://doi.org/10.1016/S1474-4422(19)30320-5.
Noyce, A.J., Kia, D.A., Hemani, G., et al., 2017a. Estimating the causal influence of body mass index on risk of Parkinson disease: a Mendelian randomisation study. PLoS Med. 14 (6) https://doi.org/10.1371/journal.pmed.1002314.
Noyce, A.J., R'Bibo, L., Peress, L., et al., 2017b. PREDICT-PD: an online approach to prospectively identify risk indicators of Parkinson's disease. Mov. Disord. 32 (2), 219–226. https://doi.org/10.1002/mds.26898.

Noyce, A.J., Bandres-Ciga, S., Kim, J., 2019. The Parkinson's disease Mendelian randomization research portal. Mov. Disord. 34 (12), 1864–1872. https://doi.org/10.1002/mds.27873.
Paisan-Ruiz, C., Bhatia, K.P., Li, A., et al., 2008. Characterisation of PLA2G6 as a locus for dystonia-parkinsonism. Ann. Neurol. 65 (1), 19–23. https://doi.org/10.1002/ana.21415.
Paisán-Ruíz, C., Jain, S., Evans, E.W., et al., 2004. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. doi:S0896627304006890 [pii] Neuron 44 (4), 595–600. https://doi.org/10.1016/j.neuron.2004.10.023.
Park, S.L., Cheng, I., Haiman, C.A., 2018. Genome-wide association studies of cancer in diverse populations. Cancer Epidemiol. Biomark. Prev. 27 (4), 405–417. https://doi.org/10.1158/1055-9965.EPI-17-0169.
Pastor, P., Ezquerra, M., Munoz, E., et al., 2000. Significant association between the tau gene A0/A0 genotype and Parkinson's disease. Ann. Neurol. 47 (2), 242–245. https://doi.org/10.1002/1531-8249(200002)47:2<242::AID-ANA16>3.0.CO;2-L.
Pecanka, J., Jonker, M.A., Bochdanovits, Z., Van Der Vaart, A.W., 2017. A powerful and efficient two-stage method for detecting gene-to-gene interactions in GWAS. Biostatistics 18 (3), 477–494. https://doi.org/10.1093/biostatistics/kxw060.
Pichler, I., Del Greco, M.F., Gögele, M., et al., 2013. Serum iron levels and the risk of Parkinson disease: a mendelian randomization study. PLoS Med. 10 (6) https://doi.org/10.1371/journal.pmed.1001462.
Pihlstrøm, L., Morset, K.R., Grimstad, E., Vitelli, V., Toft, M., 2016. A cumulative genetic risk score predicts progression in Parkinson's disease. Mov. Disord. 31 (4), 487–490. https://doi.org/10.1002/mds.26505.
Pingault, J.B., O'Reilly, P.F., Schoeler, T., Ploubidis, G.B., Rijsdijk, F., Dudbridge, F., 2018. Using genetic data to strengthen causal inference in observational research. Nat. Rev. Genet. 19 (9), 566–580. https://doi.org/10.1038/s41576-018-0020-3.
Polymeropoulos, M.H., Lavedan, C., Leroy, E., et al., 1997. Mutation in the α-synuclein gene identified in families with Parkinson's disease. Science 80–276 (5321), 2045–2047. https://doi.org/10.1126/science.276.5321.2045.
Ponsen, M.M., Stoffers, D., Booij, J., Van Eck-Smit, B.L.F., Wolters, Ec.E.C., Berendse, H. W., 2004. Idiopathic hyposmia as a preclinical sign of Parkinson's disease. Ann. Neurol. 56 (2), 173–181. https://doi.org/10.1002/ana.20160.
Prabhu, S., Pe'er, I., 2012. Ultrafast genome-wide scan for SNP-SNP interactions in common complex disease. Genome Res. 22 (11), 2230–2240. https://doi.org/10.1101/gr.137885.112.
Przedborski, S., 2017. The two-century journey of Parkinson disease research. Nat. Rev. Neurosci. 18 (4), 251–259. https://doi.org/10.1038/nrn.2017.25.
Rajan, R., Divya, K.P., Kandadai, R.M., et al., 2020. Genetic architecture of parkinson's disease in the indian population: harnessing genetic diversity to address critical gaps in parkinson's disease research. Front. Neurol. 11, 1–11. https://doi.org/10.3389/fneur.2020.00524.
Ramirez, A., Heimbach, A., Gründemann, J., et al., 2006. Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. Nat. Genet. 38 (10), 1184–1191. https://doi.org/10.1038/ng1884.
Ritchie, M.D., Hahn, L.W., Roodi, N., et al., 2001. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am. J. Hum. Genet. 69 (1), 138–147. https://doi.org/10.1086/321276.
Robak, L.A., Jansen, I.E., Rooij, J. van, et al., 2017. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. Brain 140 (12), 3191–3203. https://doi.org/10.1093/brain/awx285.
Rocca, W.A., McDonnell, S.K., Strain, K.J., et al., 2004. Familial aggregation of Parkinson's disease: the mayo clinic family study. Ann. Neurol. 56 (4), 495–502. https://doi.org/10.1002/ana.20228.
Rosenbloom, B., Balwani, M., Bronstein, J.M., et al., 2011. The incidence of parkinsonism in patients with type 1 Gaucher disease: data from the ICGG Gaucher registry. Blood Cells Mol. Dis. 46 (1), 95–102. https://doi.org/10.1016/j.bcmd.2010.10.006.
Ross, O.A., Soto-Ortolaza, A.I., Heckman, M.G., et al., 2011. Association of LRRK2 exonic variants with susceptibility to Parkinson's disease: a case-control study. Lancet Neurol. 10 (10), 898–908. https://doi.org/10.1016/S1474-4422(11)70175-2.
Ruiz-Martínez, J., Gorostidi, A., Ibañez, B., et al., 2010. Penetrance in Parkinson's disease related to the LRRK2 R1441G mutation in the Basque country (Spain). Mov. Disord. 25 (14), 2340–2345. https://doi.org/10.1002/mds.23278.
Satake, W., Nakabayashi, Y., Mizuta, I., et al., 2009. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. Nat. Genet. 41 (12), 1303–1307. https://doi.org/10.1038/ng.485.
Schilder, B.M., Raj, T., 2022. Fine-mapping of Parkinson's disease susceptibility loci identifies putative causal variants. Hum. Mol. Genet. 31 (6), 888–900. https://doi.org/10.1093/hmg/ddab294.
Schilder, B.M., Navarro, E., Raj, T., 2022. Multi-omic insights into Parkinson's disease: from genetic associations to functional mechanisms. Neurobiol. Dis. 163, 105580 https://doi.org/10.1016/j.nbd.2021.105580.
Schork, N.J., Murray, S.S., Frazer, K.A., Topol, E.J., 2009. Common vs. rare allele hypotheses for complex diseases. Curr. Opin. Genet. Dev. 19 (3), 212–219. https://doi.org/10.1016/j.gde.2009.04.010.
Sidransky, E., Nalls, M.A., Aasly, J.O., et al., 2009. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. New Engl. J. Med. 361 (17), 1651–1661. https://doi.org/10.1056/nejmoa0901281.
Sierra, M., González-Aramburu, I., Sánchez-Juan, P., et al., 2011. High frequency and reduced penetrance of LRRK2 G2019S mutation among Parkinson's disease patients in Cantabria (Spain). Mov. Disord. 26 (13), 2343–2346. https://doi.org/10.1002/mds.23965.
Simon-Sanchez, J., Schulte, C., Bras, J.M., et al., 2009. Genome-wide association study reveals genetic risk underlying Parkinson's disease. doi:ng.487[pii] Nat. Genet. 41 (12), 1308–1312. https://doi.org/10.1038/ng.487.
Singleton, A., 2015. A new gene for Parkinson's disease: should we care? Lancet Neurol. 14 (3), 238–239. https://doi.org/10.1016/S1474-4422(14)70270-4.
Singleton, A., Hardy, J., 2011. A generalisable hypothesis for the genetic architecture of disease: Pleomorphic risk loci. Hum. Mol. Genet. 20 (R2), 158–162. https://doi.org/10.1093/hmg/ddr358.
Singleton, A., Hardy, J., 2019. Progress in the genetic analysis of Parkinson's disease. Hum. Mol. Genet. 28 (R2), R215–R218. https://doi.org/10.1093/hmg/ddz183.
Slatkin, M., 2008. Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nat. Rev. Genet. 9 (6), 477–485. https://doi.org/10.1038/nrg2361.
Soldner, F., Stelzer, Y., Shivalila, C.S., et al., 2016. Parkinson-associated risk variant in distal enhancer of α-synuclein modulates target gene expression. Nature 533 (7601), 95–99. https://doi.org/10.1038/nature17939.
Sveinbjörnsdóttir, S., Hicks, A.A., Jónsson, T., et al., 2000. Familial aggregation of Parkinson's disease in Iceland. New Engl. J. Med. 343 (24), 1765–1770. https://doi.org/10.1056/nejm200012143432404.
Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., Meyre, D., 2019. Benefits and limitations of genome-wide association studies. Nat. Rev. Genet. 20 (8), 467–484. https://doi.org/10.1038/s41576-019-0127-1.
Tan, M.M.X., Lawton, M.A., Jabbari, E., et al., 2021. Genome-wide association studies of cognitive and motor progression in Parkinson's disease. Mov. Disord. 36 (2), 424–433. https://doi.org/10.1002/mds.28342.
Trinh, J., Gustavsson, E.K., Vilariño-Güell, C., et al., 2016. DNM3 and genetic modifiers of age of onset in LRRK2 Gly2019Ser parkinsonism: a genome-wide linkage and association study. Lancet Neurol. 15 (12), 1248–1256. https://doi.org/10.1016/S1474-4422(16)30203-4.
Valente, E.M., Abou-Sleiman, P.M., Caputo, V., et al., 2004. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304 (5674), 1158–1160. https://doi.org/10.1126/science.1096284.
van Rheenen, W., Peyrot, W.J., Schork, A.J., Lee, S.H., Wray, N.R., 2019. Genetic correlations of polygenic disease traits: from theory to practice. Nat. Rev. Genet. 20 (10), 567–581. https://doi.org/10.1038/s41576-019-0137-z.
Velez-Pardo, C., Lorenzo-Betancor, O., Jimenez-Del-Rio, M., et al., 2019. The distribution and risk effect of GBA variants in a large cohort of PD patients from Colombia and Peru. Parkinsonism Relat. Disord. 63, 204–208. https://doi.org/10.1016/j.parkreldis.2019.01.030.
Vilariño-Güell, C., Wider, C., Ross, O.A., et al., 2011. VPS35 mutations in parkinson disease. Am. J. Hum. Genet 89 (1), 162–167. https://doi.org/10.1016/j.ajhg.2011.06.001.
Wang, C., Cai, Y., Zheng, Z., et al., 2012. Penetrance of LRRK2 G2385R and R1628P is modified by common PD-associated genetic variants. Parkinsonism Relat. Disord. 18 (8), 958–963. https://doi.org/10.1016/j.parkreldis.2012.05.003.
Wang, K., Zhang, H., Kugathasan, S., 2009. Diverse genome-wide association studies associate the IL12/IL23 pathway with Crohn disease. Am. J. Hum. Genet. 84 (3), 399–405. https://doi.org/10.1016/j.ajhg.2009.01.026.
Wang, K., Li, M., Hakonarson, H., 2010. Analysing biological pathways in genome-wide association studies. Nat. Rev. Genet. 11 (12), 843–854. https://doi.org/10.1038/nrg2884.
Wang, L., Heckman, M.G., Aasly, J.O., et al., 2017. Evaluation of the interaction between LRRK2 and PARK16 loci in determining risk of Parkinson's disease: analysis of a large multicenter study. Neurobiol. Aging 49, 217.e1–217.e4. https://doi.org/10.1016/j.neurobiolaging.2016.09.022.
Weng SF, Reps J., Kai J., Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? Published online 2017. doi:⟨10.1371/journal.pone.0174944⟩.
Wirdefeldt, K., Gatz, M., Reynolds, C.A., Prescott, C.A., Pedersen, N.L., 2011. Heritability of Parkinson disease in Swedish twins: a longitudinal study. Neurobiol. Aging 32 (10), 1923.e1–1923.e8. https://doi.org/10.1016/j.neurobiolaging.2011.02.017.
Xilouri, M., Brekk, O.R., Stefanis, L., 2016. Autophagy and alpha-synuclein: relevance to Parkinson's disease and related synucleopathies. Mov. Disord. 31 (2), 178–192. https://doi.org/10.1002/mds.26477.
Xue, A., Wu, Y., Zhu, Z., et al., 2018. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat. Commun. 9 (1), 1–14. https://doi.org/10.1038/s41467-018-04951-w.
Yang, H.C., Chen, C.W., Lin, Y.T., Chu, S.K., 2021. Genetic ancestry plays a central role in population pharmacogenomics. Commun. Biol. 4 (1), 171. https://doi.org/10.1038/s42003-021-01681-6.
Zabetian, C.P., Mata, I.F., 2017. Latin American Research Consortium on the genetics of PD (LARGE-PD) LARGE-PD: examining the genetics of Parkinson's disease in Latin America. Mov. Disord. 32 (9), 1330–1331. https://doi.org/10.1002/mds.27081.
Zhang, Y., Shu, L., Sun, Q., et al., 2018. Integrated genetic analysis of racial differences of common GBA variants in Parkinson's disease: a meta-analysis. Front. Mol. Neurosci. 11, 43. https://doi.org/10.3389/fnmol.2018.00043.

Zhou, W., Ma, D., Sun, A.X., et al., 2019. PD-linked CHCHD2 mutations impair CHCHD10 and MICOS complex leading to mitochondria dysfunction. Hum. Mol. Genet. 28 (7), 1100–1116. https://doi.org/10.1093/hmg/ddy413.

Zimprich, A., Biskup, S., Leitner, P., 2004. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. doi:S0896627304007202[pii] Neuron 44 (4), 601–607. https://doi.org/10.1016/j.neuron.2004.11.005.

Zimprich, A., Benet-Pagès, A., Struhal, W., et al., 2011. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset parkinson disease. Am. J. Hum. Genet. 89 (1), 168–175. https://doi.org/10.1016/j.ajhg.2011.06.008.