



Review

Genetic Risk Profiling in Parkinson's Disease and Utilizing Genetics to Gain Insight into Disease-Related Biological Pathways

Ashley Hall ¹, Sara Bandres-Ciga ², Monica Diez-Fairen ³, John P. Quinn ¹ and Kimberley J. Billingsley ^{2,*}

¹ Department of Pharmacology and Therapeutics, Institute of Systems, Molecular & Integrative Biology, University of Liverpool, L69 7BE, UK; ashley.hall@liverpool.ac.uk (A.H.); jquinn@liverpool.ac.uk (J.P.Q.)

² Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA; sara.bandresciga@nih.gov

³ Neurogenetics Group, University Hospital MutuaTerrassa, Sant Antoni 19, 08221 Terrassa, Barcelona, Spain; monicadifa@gmail.com

* Correspondence: kimberley.billingsley@nih.gov

Received: 1 September 2020; Accepted: 1 October 2020; Published: 4 October 2020



Abstract: Parkinson's disease (PD) is a complex disorder underpinned by both environmental and genetic factors. The latter only began to be understood around two decades ago, but since then great inroads have rapidly been made into deconvoluting the genetic component of PD. In particular, recent large-scale projects such as genome-wide association (GWA) studies have provided insight into the genetic risk factors associated with genetically "complex" PD (PD that cannot readily be attributed to single deleterious mutations). Here, we discuss the plethora of genetic information provided by PD GWA studies and how this may be utilized to generate polygenic risk scores (PRS), which may be used in the prediction of risk and trajectory of PD. We also comment on how pathway-specific genetic profiling can be used to gain insight into PD-related biological pathways, and how this may be further utilized to nominate causal PD genes and potentially druggable therapeutic targets. Finally, we outline the current limits of our understanding of PD genetics and the potential contribution of variation currently uncaptured in genetic studies, focusing here on uncatalogued structural variants.

Keywords: Parkinson's disease; genetic risk; biological pathways

1. Parkinson's Disease Genetics

Parkinson's disease (PD) is characterized by a loss of dopaminergic neurons, particularly in the substantia nigra, and widespread accumulation of intracellular α -synuclein protein aggregates such as Lewy bodies. Symptoms include motor features such as bradykinesia (slowness of movement), rigidity, resting tremor and loss of balance, amongst others [1]. A majority of PD patients also have non-motor symptoms, including REM (rapid eye movement) sleep perturbation, cognitive impairment, mood disorders, constipation, sensory symptoms such as hyposmia (reduced sense of smell), and chronic pain [2]. PD is a heterogeneous disorder in which these symptoms may manifest in a spectrum of varying combinations, highlighting that clinical presentation and progression is likely influenced by a mixture of genetic and environmental factors.

The global prevalence of PD is expected to double from 6.2 million cases in 2015 to 12.9 million by 2040 [3]. PD therefore represents a growing challenge for society, since current treatments only manage symptoms. It is imperative that a better understanding of the etiology of PD is achieved to lay the groundwork for treatments that may halt or reverse its progression.

Until only two decades ago, PD was considered to be wholly caused by environmental factors. Early epidemiology studies pointed to exposure to viruses and neurotoxins such as MPTP [4]; perhaps most famous was the strong association between the 1918 influenza pandemic and the increased rates of post-encephalitic parkinsonism that followed. This non-genetic basis of PD was supported by the first cross-sectional twin studies of the disease [5]. It is now known, however, that PD is a complex disorder influenced by both genetic and environmental factors. Indeed, 5–10% of PD follows a classical Mendelian inheritance pattern, and around 15% of PD patients have family history of PD [6]. The first direct evidence of a heritable component of PD came in 1997 with the identification of rare mutations in SNCA (encoding α -synuclein) that were responsible for a form of monogenic PD [7]. This was shortly followed by the discovery of additional rare recessive forms of PD caused by deleterious mutations in PRKN [8], PINK1 (also known as PARK6) [9] and DJ-1 [10], and the identification of autosomal dominant PD resulting from mutations in LRRK2 [11]. To date, mutations within several additional genes have been associated with monogenic PD and have been discussed previously [12].

The majority of PD cases cannot be attributed to a single penetrant, deleterious mutation. Considering this, many studies have pursued the “common disease common variant” (CDCV) hypothesis, in which the genetic component of PD is considered to be the cumulative result of many common, low-risk alleles [13]. Genome-wide association (GWA) studies have been instrumental in addressing this hypothesis. Briefly, the premise of a GWA study is to compare the genotypes of many individuals, often in a case vs. control setup, and determine which common genetic variants, typically single nucleotide variants (SNVs), are consistently associated with a trait. The power of this analysis increases with the number of participants, as this allows the contributions of relatively low-risk disease-associated variants to be detected. Accordingly, PD GWA studies of increasing size have been performed across several populations [14–19]. The most recent and largest PD GWA study meta-analysis identified 90 independent genetic signals associated with genetically complex PD and involved approximately 37,700 cases, 18,600 ‘proxy’ cases and 1.4 million controls [20]. However, it has proven difficult to directly ascribe function to these disease-associated SNVs because most occur in non-coding DNA. Any regulatory effects upon genes may, therefore, be subtle and difficult to study, and SNVs may not necessarily exert their effects on the nearest gene. Furthermore, a genetic signal may simply be in linkage disequilibrium (LD) with the true disease-causing variant. To help researchers prioritize candidate genes associated with GWA study signals, a Parkinson’s disease GWA study locus browser application has been recently developed that compiles functional data, genetic data and literature reviews for PD-associated loci [21] (<https://pdgenetics.shinyapps.io/GWASBrowser/>).

Here, we comment on how this wealth of genomic information can be utilized to build genetic risk scores, which can be useful in the prediction of disease risk and progression. Further, we note how large-scale pathway-specific genetic risk profiling can give insight into disease-related biological pathways. Utilizing this data and leveraging new analytical approaches can further identify specific genes potentially linked to PD. Finally, despite the huge success over the last decade in identifying PD genetic risk factors, we outline the current gaps in our understanding of PD genetics, with a focus on the contribution of the uncharacterized parts of the human genome, namely, uncatalogued structural variants.

2. Genetic risk Profiling in Parkinson’s Disease

Each risk allele identified by GWA studies confers relatively little risk when considered individually. Polygenic risk scores (PRS) are simple models that sum the contribution of multiple risk variants of variable effect sizes, as determined by GWA studies summary statistics (Figure 1). In this way, the information representing an individual’s genetic architecture is aggregated to form a picture of their total known genetic risk for disease. PRS can therefore be potentially used to identify at-risk individuals and allow implementation of early lifestyle changes or prophylaxis, although most commonly PRS is used in concert with other disease predictors. As the sample sizes involved in GWA studies have

grown, PRS have been used to analyze the genetic risk of a number of complex disorders such as schizophrenia [22], Alzheimer's disease [23] and PD [20].

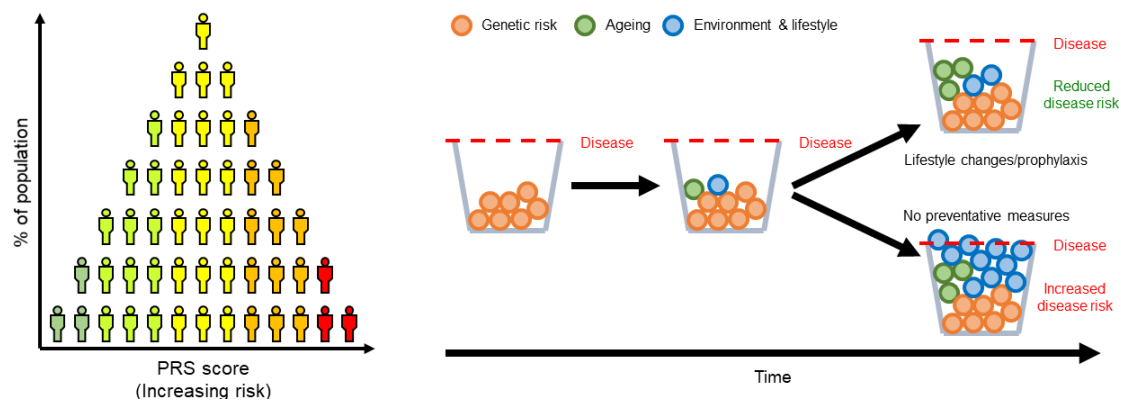


Figure 1. Polygenic risk scores based on the weighted sum of genetic variants can be used to estimate an individual's predisposition to complex polygenic disorders such as Parkinson's disease (PD), and allow those at greatest risk in the population to be identified (left). Those at risk can be recommended lifestyle changes or preventative treatments that may combat their increased likelihood of developing a disease (right).

PRS have shown potential as a tool for prediction, with scores having been associated with PD risk [24], age at onset (AAO) [24,25], and increased cognitive and motor decline [26]. The 90 risk loci identified in the largest, most recent PD meta-analysis are associated with higher relative risk of developing PD, with those in the top 10% of PRS being nearly six-fold more likely to develop PD than those in the bottom 10% [20]. It has also been demonstrated that by factoring in other variables such as age, sex, family history and hyposmia, a combined risk score can be produced with considerable sensitivity and specificity [27]. However, at present PRS are not without limitations. PRS may incorporate genetic variants that are not perfectly correlated with the causal factor, leading to uncertainty of variant effect sizes. This reduces the applicability of PRS across populations besides the training population, even if they actually share large effect-size causal variants [28,29]. It is also imperative that a PRS and any associated uncertainty is effectively interpreted, as it may unduly confer stress, a false sense of security, or the wrong decision regarding preventative measures.

It is important to note that several loci associated with PD risk variants are pleomorphic, capable of harboring both rare, deleterious mutations and common variants that confer risk. For example, point mutations and multiplications of SNCA are causative of early-onset autosomal-dominant PD [30,31], while variation in non-coding regions of the locus have been repeatedly associated with increased risk of genetically complex PD [14,18,19]. Additionally, recent GWA studies for PD AAO suggest that not all PD risk alleles influence AAO [32], indicating that overall PD risk and PD AAO may be driven by processes that only partially overlap. Similarly, GWA studies for PD progression (i.e., identifying variants that are associated with how PD changes over time) have identified variants at SLC44A1 and ADRA2A, two loci not currently associated with PD risk, with increased disease progression and development of insomnia, respectively [33]. Taken together, these observations highlight the heterogeneous and multifaceted nature of PD and reinforce the importance of refining PRS; more accurate disease modelling will allow the various molecular and pathological subtypes of PD to be delineated, and disease trajectories to be predicted. In this way, certain PRS may potentially serve as inclusion criteria for clinical trials, thereby enabling more subtype-targeted therapies to be developed. For example, a recent study observed that those with lower PD PRS were at greater risk of developing PD if they were diabetic, but conversely diabetes may have had a protective effect among those with higher PRS for PD [34]. If this stems from the effects of anti-diabetic drugs being modified by genetic risk of PD, then individuals who may benefit from their use could be identified on the basis of their PRS, potentially resulting in more targeted clinical trials.

3. Utilizing Genetic Risk Profiling to Identify Biological Pathways Involved in Disease

As noted, there has been great success at identifying genetic risk factors for PD and using this information to build genetic risk scores as a tool for predicting disease risk, AAO and increased cognitive and motor decline [26]. However, there are still major gaps in our understanding of the underlying molecular and cellular processes that are involved in PD. In light of this, recent studies have started to address the latter by constructing pathway-specific genetic risk scores to gain insight into the potential disease-related biological pathways.

First, building on previous work that suggested that the endocytic membrane trafficking pathway (EMTP) could be a driving force in PD pathogenesis [35], Bandres-Ciga et al. ran a systematic analysis to identify the contribution of genetic variants within genes associated with the EMTP to risk of PD. An extensive list of genes that are involved in the EMTP was constructed using the Molecular Signatures Database (MSigDB) of an annotated gene set for “endocytosis” based on the Kyoto Encyclopedia of Genes and Genomes information (KEGG-endocytosis) and a detailed literature search. SNVs within the 262 EMTP-associated genes were extracted from 41,321 individuals (18,869 cases and 22,452 controls) of European ancestry, and an EMTP-specific PRS was built reflecting the cumulative risk of common variants in this pathway. The EMTP risk profiling versus disease status was found to be significantly associated with PD (random-effects p -value = 2.55×10^{-12} , beta = 0.227, SE (standard error) = 0.032) with an odds ratio of 1.20 per standard deviation increase in the PRS from the population mean, confirming a cumulative pathogenic effect of EMTP variants in PD [36].

Another pathway-specific risk profiling study focused on the role of the mitochondrial function pathway in PD. Although mitochondrial dysfunction has been strongly implicated in the etiology of monogenic forms of PD, the role that mitochondrial processes played in the risk of genetically complex disease was less clear. Two gene lists were curated to encompass different levels of evidence for the involvement of the respective protein products in disease phenotypes that relate to mitochondrial function. The “primary” gene list consisted of genes mutated in mitochondrial disorders ($n = 196$). The “secondary” gene list was a broader list ($n = 1487$), curated using the OMIM API to identify all genes for which the word “mitochondria” (or derivatives) appeared in the free-text description, and this was combined with MitoCarta v2.0 genes with no OMIM phenotype. A mitochondrial function pathway-specific PRS was constructed for both lists, again from 41,321 individuals (18,869 cases and 22,452 controls) of European ancestry. Even when the known PD risk loci [20] were removed from the primary gene list, the pathway was still significantly associated with PD with an odds ratio of 1.12 per standard deviation increase in the PRS from the population mean (random-effects p -value = 6.00×10^{-4} , beta = 0.11, SE = 0.03). The secondary gene list was also associated with PD with an odds ratio of 1.28 per standard deviation increase in the PRS from the population mean (random-effects p -value = 1.9×10^{-22} , beta = 0.25, SE = 0.03). Further, the mitochondrial function pathway-specific PRS was identified to be significantly associated with later AAO of PD [37]. Overall, this study provided robust evidence that mitochondrial processes are drivers of genetically complex PD.

Most significantly, a recent large-scale analysis implemented a high-throughput and hypothesis-free approach to assess the role of PD risk in over two thousand curated and well-defined gene sets. The analysis was split into a discovery and replication phase. The discovery dataset consisted of 7218 PD cases and 9424 controls, and the replication dataset, which was used to validate the results, consisted of 5429 PD cases and 5814 controls. Summary statistics from Chang et al. 2017 [19] were used as the reference dataset to perform the analysis. The MSigDB was used to extract gene sets representative of biological pathways, and a pathway-specific polygenic effect score was constructed for each dataset. In support of the previous pathway-specific analyses [36,37], the study found a significant association between PD risk and the endocytic membrane trafficking and mitochondrial function pathways [38]. Most notably, this approach identified significant associations between risk of PD and novel biological pathways. Overall, the study concluded that the main contributors to PD etiology are molecular processes underlying protein misfolding and aggregation, post-translational protein modification, immune response, membrane and intracellular trafficking,

lipid metabolism, synaptic transmission, endosomal–lysosomal dysfunction and apoptosis mediated by initiator and executioner caspases (Figure 2). This was the first study to test the contribution of thousands of molecular processes to PD risk in an unbiased and data-driven manner and provide a foundational resource for the PD community through a publicly available pathways browser [38] (<https://pdgenetics.shinyapps.io/pathwaysbrowser/>).

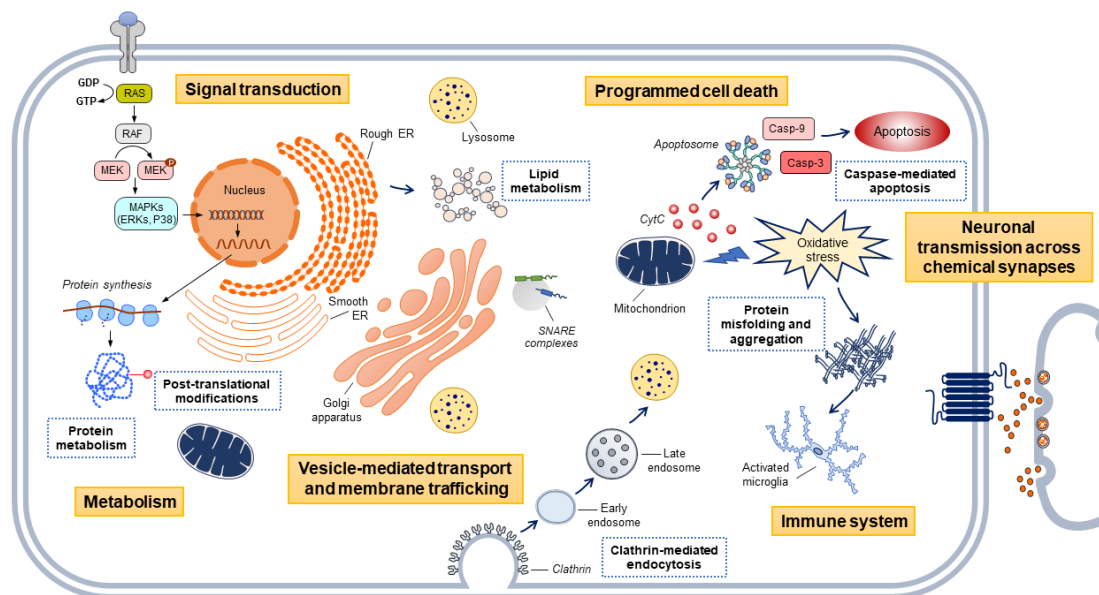


Figure 2. Main biological pathways associated with PD risk through genetic risk profiling. Yellow squares indicate the hierarchies in which the pathways are clustered according to published literature and public curated databases. Blue squares indicate the pathways that contribute the most to PD etiology. Key: CytC, Cytochrome C; ER, endoplasmic reticulum.

In addition to relating PD-related SNVs to functional pathways, a key step in unravelling PD etiology is to determine in which cell types these variants are biologically relevant. The genes highlighted in the most recent PD GWAS meta-analysis were found to be enriched for expression exclusively in brain-derived tissues, including the substantia nigra, frontal cortex and cerebellum [20]. Recent single-nuclei transcriptomics in the substantia nigra support association of PD with dopaminergic neuron-specific gene expression and, interestingly, implicate oligodendrocyte-specific expression [39]. Similar single-cell transcriptomics in the mouse nervous system also confirm the association between dopaminergic neurons and PD, and reiterate the novel association with oligodendrocytes [40]. Additionally, the pathway analysis study mentioned previously integrated single-cell expression data to link PD risk to expression patterns in dopaminergic neurons, serotonergic neurons, hypothalamic GABAergic neurons and neural progenitors [38].

A caveat to these pathway approaches, and indeed much research utilizing GWAS signals, is that it is assumed that SNVs in noncoding regions affect the nearest gene, although it is increasingly acknowledged that they can regulate genes over 10 kb away [41,42]. Recently, H-MAGMA was developed from MAGMA, a tool for linking GWAS risk signals to cognate genes, by incorporating HiC chromatin structure information in order to map SNVs to distal genes they interact with via chromatin looping [43]. This approach also revealed that these genes underwent developmental stage-specific and cell type-specific interactions in several neurobiological disorders, demonstrating how greater functional understanding of GWAS signals can be achieved when they are considered in their cellular context. Another recent study identified PD-relevant pathways via analysis of co-expressed modules of genes, the expression of which was perturbed in brain regions susceptible to PD [44].

4. From Pathway to Possible Therapeutic Target: Analytical Methods for Gene Prioritization

Although GWA studies have been a key approach to identify loci linked to PD etiology, it is widely assumed that the top genetic variants underlying such associations are not usually the disease-causing variants. GWA study signals are often non-coding and lie on intergenic or regulatory regions of the genome, making interpretation complex. The mapping resolution of a GWA study is limited by the LD structure of the genome [45], and therefore understanding what the true causal variants are and how they disrupt molecular mechanisms contributing to disease remains challenging.

As previously mentioned, pathway-specific genetic risk profiling is one approach to identify general biological processes involved in disease. Expanding on this, it is possible to further leverage this data and newly developed statistical tooling to then nominate specific genes in the associated pathways that are related to PD risk. Several integrative analytic approaches have been developed in an attempt to prioritize functional genes or regulatory elements that may reveal druggable targets amenable to therapeutic intervention. Summary-data-based Mendelian randomization (SMR) arose as a promising methodology that allows the use of GWA studies and expression or methylation quantitative trait loci (eQTL and mQTL) data from independent studies to explore the possibility of GWA study variants affecting phenotype through genetic regulation of a transcriptional output (changes either in expression or methylation at CpG sites) [46]. The premises are that if the expression or methylation of a gene is influenced by a genetic variant (QTL), then there will be differences in gene expression or methylation levels among individuals carrying different genotypes of the genetic variant. Perhaps one of the strengths of this methodology is that it has the ability to distinguish whether changes in gene expression or methylation and the phenotype of interest are associated owing to a single shared genetic variant (pleiotropic model), or in contrast, whether there are two or more distant genetic variants in LD affecting gene expression and the phenotype of interest independently (linkage model).

In the PD genetics arena, SMR has been widely applied to prioritize genes from GWA study signals. A curious example is illustrated in the last PD GWA study meta-analysis [20] in which a novel multi-signal locus (UBTF-GRN-FAM171A2) comprising independent genome-wide associated SNVs was identified. SMR analyses showed that GRN, encoding progranulin, is the most putative causal gene within this locus, and that disease risk might be explained by modulation of progranulin expression in this instance. GRN has previously been associated with frontotemporal dementia [47], and mutations within this gene have been shown to be connected with lysosomal storage disorders and neuronal ceroid lipofuscinosis [48], making this target an interesting PD candidate to further study.

Furthermore, SMR has also been successfully used to explore enrichment of eQTLs or mQTLs linked to certain biological processes involved in PD, such as the mitochondrial pathway [37] and the EMTP [36]. Among the most recent state-of-the-art functional genomics methods implemented in the PD genetics field, the use of colocalization and fine-mapping should be highlighted. Unlike SMR, this type of analysis consists of a Bayesian framework that uses summary data (usually from eQTL and GWA studies) to estimate signals that colocalize with expression changes by applying posterior probabilities [49]. In PD, a recent colocalization study of publicly available brain eQTL and PD GWA study summary statistics reinforced the RAB7L1 gene as the priority candidate for the chromosome 1q32 locus PD risk association [50]. In fact, additional studies have provided further functional evidence linking RAB7L1 to LRRK2 [51] by implicating RAB7L1 as a substrate for LRRK2 kinase activity [52–54].

Lastly, transcriptome-wide association studies (TWAS) have been powerful strategies to integrate large-scale functional genomics data with GWA studies to characterize the functional effects of associated variants [55]. TWAS integrates SNV-expression correlation (cis-SNV effect sizes), GWA study summary statistics and LD reference panels to impute RNA expression levels onto large cohorts of individuals to identify putative genes involved in disease. In the PD genetics field, TWAS has been applied to dissect the effect of genetic variation on RNA expression and splicing in order to prioritize disease-relevant genes. As an example, a recent PD TWAS has prioritized potential targets whose predicted expression or splicing levels in peripheral monocytes cells and in dorsolateral

prefrontal cortex are significantly associated with PD risk, advancing our understanding of PD in a genomic context [56]. Additionally, dysregulation of small noncoding transcripts, particularly microRNAs, is increasingly recognized as a potential mediator of reactive oxygen species production and mitochondrial dysfunction in PD and other neurodegenerative diseases [57]. Recent RNA-seq experiments in longitudinal PD cohorts associated PD with dysregulation of miRNAs implicated in function of mitochondria and inflammatory immune cells such as leukocytes [58].

Despite these advanced methodological approaches, the number of PD loci that are functionally validated remains very low and mostly includes genes that are known to cause monogenic forms of PD. A limitation to SMR, TWAS and fine-mapping/colocalization methods is that they operate under different assumptions. Additionally, it is likely that the genes prioritized with one methodology and using data from a particular tissue do not replicate when using a different approach in another tissue.

5. The Limitations in Our Current Understanding of Parkinson's Disease Genetics, Focusing on the Contribution of Structural Variants

By providing insight into disease-related biological processes, the ultimate aim of genetic research is to support the development of therapeutic treatments that stop or slow disease. Gaining a full understanding of the genetic factors that contribute to PD is essential for achieving the latter. As previously noted, the field has seen great success in identifying genetic variants that influence PD risk, onset, and progression. Despite these efforts, our understanding of the genetic factors that contribute to PD is incomplete. The largest PD GWA study to date identified that when leveraging current GWA study datasets and genetic risk profiling methods, only around 16–36% of the heritable component of PD can be explained. This large gap in our understanding of PD genetics is often referred to as the “missing heritability”. Many factors likely constitute the “missing heritability”, including uncaptured genetic variation represented by (1) rare variants, (2) epigenetics and (3) structural variants (SV). As the contributions of both (1) and (2) have been explored in great detail in recent reviews [59,60], here we focus instead on the contribution of (3) SVs.

Unlike SNVs (a change in only one nucleotide), SV are somewhat arbitrarily defined as DNA rearrangements that involve at least fifty nucleotides. They can be divided into subclasses that consist of unbalanced copy number variants (CNVs), which include deletions, duplications and insertions, of novel sequence, as well as balanced rearrangements, such as interchromosomal and intrachromosomal translocations and inversions. SVs also include transposable element (TE) insertions, segmental duplications, multi-allelic CNVs of highly variable copy number and complex rearrangements that consist of multiple combinations of these described events. SV can have a huge phenotypic impact by either disrupting gene function and regulation or modifying gene dosage. Further, numerous studies have highlighted the role of SV in functional changes across human populations and cell and tissue types [61–63]. Detection of SV from short-read sequencing data is problematic, as the evidence for SV resembles common sequencing and alignment artefacts. However, following the recent rapid development of improved sequencing and bioinformatic tooling that can more accurately call SV, it is now apparent that human genomes differ more as a consequence of SV than as a result of a point mutation [64–68]. For comparison, based on SNVs alone, on average the genomic variation between two individuals is 0.1%; however, when SV are incorporated, this increases over ten-fold, to 1.5% [69]. SV therefore represent the majority of the genetic variation in the human genome, yet this form of genetic variation remains mainly uncatalogued.

In regards to the role of SV in PD, our current understanding of the genetic basis of genetically complex PD has been predominantly gained through GWA studies that utilize genome-wide SNV-chip based datasets. Hence SV represent a substantial and important mutational force that has not yet been systematically addressed. On the other hand, our understanding of the genetic basis of monogenic forms of PD has been primarily obtained using more targeted approaches, such as linkage analysis and sequencing of genes in candidate intervals. As a result of these analyses, it has been repeatedly

identified that SV are causal variants of many forms of monogenic PD and Parkinsonism, with examples including causative CNVs, repeat expansions and TE insertions.

CNV SV causative of monogenic forms of PD have been discussed in great detail in recent reviews [70,71]. In brief, the first major discovery to highlight the role of SV in PD was in 2003, whereby Singleton et al. identified a causative triplication (CNV) of the entire genomic region that encompassed the gene *SNCA* in a large family with autosomal dominant PD [31]. Multiplications of *SNCA* were later reported in several other families from different ancestral backgrounds, including other *SNCA* triplications [72–80] and, more commonly, duplications of the gene [81–87]. Overall, the severity of the clinical phenotype of *SNCA* multiplications is associated with gene dosage and mRNA/protein expression levels in brain [73]. Hence, patients with *SNCA* triplication have rapidly progressive symptoms (generally in the fourth decade) whereas phenotypically *SNCA* duplication patients usually resemble late-onset genetically complex PD cases [81,82]. Further, causative CNVs have been reported in other familiar PD genes, such as the genes *PARK2* [8], *PINK1* [88] and *DJ-1* [10].

Other classes of SV have been associated with forms of PD and Parkinsonism. One prominent example is the TE SINE-VNTR-Alu (SVA) insertion in intron 32 of the Transcription initiation factor TFIID subunit 1 (*TAF1*) gene, which is causative of X-Linked Dystonia Parkinsonism (XDP). The XDP TE SVA insertion is not only variable in its presence/absence in the genome but between individuals with the XDP TE SVA, it is also variable in the size of one of the SVAs repeat domain. The variation in the XDP TE SVAs CT elements is disease-modifying, as the size of the SVAs hexanucleotide CT repeat domain inversely correlates with XDP age at onset [89]. The XDP TE SVA insertion was previously found to alter sequence within *TAF1* introns causing abnormal mRNA expression and significant dysregulation of a neural-specific *TAF1* isoform—N-*TAF1*—in XDP causative relative to control brain tissue [90]. In addition, generated XDP and matched control induced pluripotent stem cell (iPSC) lines confirmed *TAF1* transcript dysregulation and also revealed a significant decrease in expression of *TAF1* transcript fragments that span the region of the SVA (intron 32–36). Remarkably CRISPR/Cas9 excision of the SVA rescued the aberrant transcriptional signature and normalized expression of *TAF1* in patient-derived iPSCs [91].

Despite the evidence that SV are causative of many forms of monogenic PD and Parkinsonism, SV are yet to be directly assessed in a large-scale and whole-genome manner in genetically complex PD cases. However, one recent example that highlights the importance of characterizing a structural variant with such an approach identified that deletions at 22q11.2 are associated with genetically complex PD. Deletions at 22q11.2 were previously associated with a wide range of clinical syndromes, including 22q11.2 deletion syndrome, Schizophrenia [92] and rare reported cases of Parkinsonism and PD [93–95]. Consequently, Mok et al. screened data from independent GWA studies to systematically establish the frequency of 22q11.2 deletions in ~9k PD cases and ~14k controls. This large-scale analysis identified that deletions at 22q11.2 increase the risk of PD, particularly early-onset PD (onset age <45 years) [96]. Overall, few other complex trait GWA studies have directly assessed SV, although rare and de novo SV have been implicated in the genetics of other complex genetic diseases such as autism [97–99], and schizophrenia [100–103]. The current lack of large-scale SV studies for complex genetic diseases is mainly due to two reasons. The first is that previous tooling has not been able to accurately call SV at scale. The second is that there has been a lack of a comprehensive SV reference set. Fortunately, several groups are pioneering large-scale efforts to further develop structural variant detection tooling and to establish population-scale structural variant reference sets for the broader research community. One of the largest SV studies to date detected and genotyped SV in 14,891 genomes from diverse populations. Following this, Collins et al. discovered over 400,000 novel SV and identified that SV are responsible for 25–29% of all rare protein-truncating events per genome. Further, to provide an SV reference map, these data were made publicly available through the publicly available gnomAD browser (<https://gnomad.broadinstitute.org/downloads/>) [104].

In sum, as we enter an era of whole-genome and long-read sequencing, it is evident that to gain a complete understanding of the genetics of PD, SV will need to be routinely incorporated into GWA

studies. The current state of SV calling has been compared to SNV calling a decade ago [105]. Therefore, in the context of PD it is important to reflect on how the great success of the last decade in PD genetics was made with the SNV based studies, i.e., through the use of well-powered unbiased GWA studies. Hence, moving forward, to systematically address SV in PD it will be crucial to leverage the newly developed accurate SV callers and perform large-scale SV oriented genetic studies.

6. Future Directions

Over the last decade, tremendous advances have been made in PD genetics, improving our ability to understand and define the cumulative risk of disease. The new era holds promise, and success will only be possible through working collaboratively and openly sharing data, processes and results. In this context, the Global Parkinson's Genetics Program (GP2) (<https://parkinsonsroadmap.org/roadmap/>) will further our understanding of the genetic architecture of PD by first expanding the sample size of GWA studies to include over 150,000 individuals from all over the world. This will have a special focus on under-represented populations for fine-mapping and trans-ethnic admixture analysis. Secondly, GP2 will facilitate the application of bioinformatic methodologies by promoting open science and connecting researchers worldwide. Finally, GP2 will help unmask the contribution of currently uncaptured genetic variation, such as SV and rare variation, to the risk of PD.

Author Contributions: Initial manuscript preparation: A.H, S.B.-C, M.D.-F, and K.J.B; manuscript editing and commentary: A.H, S.B.-C, M.D.-F, K.J.B., and J.P.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Intramural Research Program of the National Institutes of Health (National Institute on Aging) 1ZIA-NS003154 and in part by the Wellcome Trust PhD Programme, grant number 215192/Z/19/Z.

Acknowledgments: We acknowledge funding from the Wellcome Trust PhD Programme, grant number 215192/Z/19/Z.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

PD	Parkinson's Disease
REM	Rapid eye movement
GWA	Genome-wide association
PRS	Polygenic risk score
CDCV	Common disease common variant
SNV	Single nucleotide variants
LD	Linkage disequilibrium
AAO	Age at onset
EMTP	Endocytic membrane trafficking pathway
MSigDB	Molecular Signatures Database
KEGG	Kyoto Encyclopedia of Genes and Genomes
SE	Standard error
SMR	Summary-data-based Mendelian randomization
QTL	Quantitative trait loci
TWAS	Transcriptome-wide association studies
SV	Structural variants
CNV	Copy number variant
GP2	Global Parkinson's Genetics Program

References

1. Poewe, W.; Seppi, K.; Tanner, C.M.; Halliday, G.M.; Brundin, P.; Volkman, J.; Schrag, A.E.; Lang, A.E. Parkinson disease. *Nat. Rev. Dis Primers* **2017**, *3*, 17013. [[CrossRef](#)] [[PubMed](#)]
2. Chaudhuri, K.R.; Schapira, A.H. Non-motor symptoms of Parkinson's disease: Dopaminergic pathophysiology and treatment. *Lancet Neurol.* **2009**, *8*, 464–474. [[CrossRef](#)]
3. Dorsey, E.R.; Bloem, B.R. The Parkinson Pandemic-A Call to Action. *JAMA Neurol.* **2018**, *75*, 9–10. [[CrossRef](#)] [[PubMed](#)]
4. Langston, J.W.; Ballard, P.; Tetrud, J.W.; Irwin, I. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* **1983**, *219*, 979–980. [[CrossRef](#)]
5. Eldridge, R.; Ince, S.E. The low concordance rate for Parkinson's disease in twins: A possible explanation. *Neurology* **1984**, *34*, 1354–1356. [[CrossRef](#)]
6. Lesage, S.; Brice, A. Parkinson's disease: From monogenic forms to genetic susceptibility factors. *Hum. Mol. Genet.* **2009**, *18*, R48–R59. [[CrossRef](#)]
7. Polymeropoulos, M.H.; Lavedan, C.; Leroy, E.; Ide, S.E.; Dehejia, A.; Dutra, A.; Pike, B.; Root, H.; Rubenstein, J.; Boyer, R.; et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **1997**, *276*, 2045–2047. [[CrossRef](#)]
8. Kitada, T.; Asakawa, S.; Hattori, N.; Matsumine, H.; Yamamura, Y.; Minoshima, S.; Yokochi, M.; Mizuno, Y.; Shimizu, N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **1998**, *392*, 605–608. [[CrossRef](#)]
9. Valente, E.M.; Bentivoglio, A.R.; Dixon, P.H.; Ferraris, A.; Ialongo, T.; Frontali, M.; Albanese, A.; Wood, N.W. Localization of a novel locus for autosomal recessive early-onset parkinsonism, PARK6, on human chromosome 1p35-p36. *Am. J. Hum. Genet.* **2001**, *68*, 895–900. [[CrossRef](#)]
10. Bonifati, V.; Rizzu, P.; Squitieri, F.; Krieger, E.; Vanacore, N.; van Swieten, J.C.; Brice, A.; van Duijn, C.M.; Oostra, B.; Meco, G.; et al. DJ-1 (PARK7), a novel gene for autosomal recessive, early onset parkinsonism. *Neurol. Sci.* **2003**, *24*, 159–160. [[CrossRef](#)]
11. Zimprich, A.; Biskup, S.; Leitner, P.; Lichtner, P.; Farrer, M.; Lincoln, S.; Kachergus, J.; Hulihan, M.; Uitti, R.J.; Calne, D.B.; et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **2004**, *44*, 601–607. [[CrossRef](#)]
12. Hernandez, D.G.; Reed, X.; Singleton, A.B. Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. *J. Neurochem.* **2016**, *139*, 59–74. [[CrossRef](#)] [[PubMed](#)]
13. Lohmueller, K.E.; Pearce, C.L.; Pike, M.; Lander, E.S.; Hirschhorn, J.N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* **2003**, *33*, 177–182. [[CrossRef](#)]
14. Simón-Sánchez, J.; Schulte, C.; Bras, J.M.; Sharma, M.; Gibbs, J.R.; Berg, D.; Paisan-Ruiz, C.; Lichtner, P.; Scholz, S.W.; Hernandez, D.G.; et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat. Genet.* **2009**, *41*, 1308–1312. [[CrossRef](#)]
15. Kara, E.; Xiromerisiou, G.; Spanaki, C.; Bozi, M.; Koutsis, G.; Panas, M.; Dardiotis, E.; Ralli, S.; Bras, J.; Letson, C.; et al. Assessment of Parkinson's disease risk loci in Greece. *Neurobiol. Aging* **2014**, *35*, 442.e9–442.e16. [[CrossRef](#)] [[PubMed](#)]
16. Satake, W.; Nakabayashi, Y.; Mizuta, I.; Hirota, Y.; Ito, C.; Kubo, M.; Kawaguchi, T.; Tsunoda, T.; Watanabe, M.; Takeda, A.; et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* **2009**, *41*, 1303–1307. [[CrossRef](#)]
17. Nalls, M.A.; Plagnol, V.; Hernandez, D.G.; Sharma, M.; Sheerin, U.M.; Saad, M.; Simón-Sánchez, J.; Schulte, C.; Lesage, S.; Sveinbjörnsdóttir, S.; et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: A meta-analysis of genome-wide association studies. *Lancet* **2011**, *377*, 641–649. [[CrossRef](#)] [[PubMed](#)]
18. Nalls, M.A.; Pankratz, N.; Lill, C.M.; Do, C.B.; Hernandez, D.G.; Saad, M.; DeStefano, A.L.; Kara, E.; Bras, J.; Sharma, M.; et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat. Genet.* **2014**, *46*, 989–993. [[CrossRef](#)]
19. Chang, D.; Nalls, M.A.; Hallgrímsdóttir, I.B.; Hunkapiller, J.; van der Brug, M.; Cai, F.; Kerchner, G.A.; Ayalon, G.; Bingol, B.; Sheng, M.; et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* **2017**, *49*, 1511–1516. [[CrossRef](#)]

20. Nalls, M.A.; Blauwendraat, C.; Vallerga, C.L.; Heilbron, K.; Bandres-Ciga, S.; Chang, D.; Tan, M.; Kia, D.A.; Noyce, A.J.; Xue, A.; et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: A meta-analysis of genome-wide association studies. *Lancet Neurol.* **2019**, *18*, 1091–1102. [[CrossRef](#)]
21. Grenn, F.P.; Kim, J.J.; Makarios, M.B.; Iwaki, H.; Illarionova, A.; Brodin, K.; Kluss, J.H.; Schumacher-Schuh, A.F.; Leonard, H.; Faghri, F.; et al. The Parkinson's Disease Genome-Wide Association Study Locus Browser. *Mov. Disord.* **2020**. [[CrossRef](#)] [[PubMed](#)]
22. Purcell, S.M.; Wray, N.R.; Stone, J.L.; Visscher, P.M.; O'Donovan, M.C.; Sullivan, P.F.; Sklar, P.; Consortium, I.S. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **2009**, *460*, 748–752. [[CrossRef](#)] [[PubMed](#)]
23. Chaudhury, S.; Patel, T.; Barber, I.S.; Guetta-Baranes, T.; Brookes, K.J.; Chappell, S.; Turton, J.; Guerreiro, R.; Bras, J.; Hernandez, D.; et al. Polygenic risk score in postmortem diagnosed sporadic early-onset Alzheimer's disease. *Neurobiol. Aging* **2018**, *62*, 244.e241–244.e248. [[CrossRef](#)] [[PubMed](#)]
24. Ibanez, L.; Dube, U.; Saef, B.; Budde, J.; Black, K.; Medvedeva, A.; del-Aguila, J.L.; Davis, A.A.; Perlmutter, J.S.; Harari, O.; et al. Parkinson disease polygenic risk score is associated with Parkinson disease status and age at onset but not with alpha-synuclein cerebrospinal fluid levels. *BMC Neurol.* **2017**, *17*, 198. [[CrossRef](#)] [[PubMed](#)]
25. Escott-Price, V.; Nalls, M.A.; Morris, H.R.; Lubbe, S.; Brice, A.; Gasser, T.; Heutink, P.; Wood, N.W.; Hardy, J.; Singleton, A.B.; et al. Polygenic risk of Parkinson disease is correlated with disease age at onset. *Ann. Neurol.* **2015**, *77*, 582–591. [[CrossRef](#)]
26. Paul, K.C.; Schulz, J.; Bronstein, J.M.; Lill, C.M.; Ritz, B.R. Association of Polygenic Risk Score With Cognitive Decline and Motor Progression in Parkinson Disease. *JAMA Neurol.* **2018**, *75*, 360–366. [[CrossRef](#)]
27. Nalls, M.A.; McLean, C.Y.; Rick, J.; Eberly, S.; Hutten, S.J.; Gwinn, K.; Sutherland, M.; Martinez, M.; Heutink, P.; Williams, N.M.; et al. Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: A population-based modelling study. *Lancet Neurol.* **2015**, *14*, 1002–1009. [[CrossRef](#)]
28. Crouch, D.J.M.; Bodmer, W.F. Polygenic inheritance, GWAS, polygenic risk scores, and the search for functional variants. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 18924–18933. [[CrossRef](#)]
29. Torkamani, A.; Wineinger, N.E.; Topol, E.J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **2018**, *19*, 581–590. [[CrossRef](#)]
30. Chartier-Harlin, M.C.; Kachergus, J.; Roumier, C.; Mouroux, V.; Douay, X.; Lincoln, S.; Levecque, C.; Larvor, L.; Andrieux, J.; Hulihan, M.; et al. Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **2004**, *364*, 1167–1169. [[CrossRef](#)]
31. Singleton, A.B.; Farrer, M.; Johnson, J.; Singleton, A.; Hague, S.; Kachergus, J.; Hulihan, M.; Peuralinna, T.; Dutra, A.; Nussbaum, R.; et al. alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **2003**, *302*, 841. [[CrossRef](#)] [[PubMed](#)]
32. Blauwendraat, C.; Heilbron, K.; Vallerga, C.L.; Bandres-Ciga, S.; von Coelln, R.; Pihlstrøm, L.; Simón-Sánchez, J.; Schulte, C.; Sharma, M.; Krohn, L.; et al. Parkinson's disease age at onset genome-wide association study: Defining heritability, genetic loci, and α -synuclein mechanisms. *Mov. Disord.* **2019**, *34*, 866–875. [[CrossRef](#)] [[PubMed](#)]
33. Iwaki, H.; Blauwendraat, C.; Leonard, H.L.; Kim, J.J.; Liu, G.; Maple-Grødem, J.; Corvol, J.C.; Pihlstrøm, L.; van Nimwegen, M.; Hutten, S.J.; et al. Genomewide association study of Parkinson's disease clinical biomarkers in 12 longitudinal patients' cohorts. *Mov. Disord.* **2019**, *34*, 1839–1850. [[CrossRef](#)]
34. Jacobs, B.M.; Belete, D.; Bestwick, J.; Blauwendraat, C.; Bandres-Ciga, S.; Heilbron, K.; Dobson, R.; Nalls, M.A.; Singleton, A.; Hardy, J.; et al. Parkinson's disease determinants, prediction and gene-environment interactions in the UK Biobank. *J. Neurol. Neurosurg. Psychiatry* **2020**, *91*, 1046–1054. [[CrossRef](#)] [[PubMed](#)]
35. Abeliovich, A.; Gitler, A.D. Defects in trafficking bridge Parkinson's disease pathology and genetics. *Nature* **2016**, *539*, 207–216. [[CrossRef](#)]
36. Bandres-Ciga, S.; Saez-Atienzar, S.; Bonet-Ponce, L.; Billingsley, K.; Vitale, D.; Blauwendraat, C.; Gibbs, J.R.; Pihlstrøm, L.; Gan-Or, Z.; Cookson, M.R.; et al. The endocytic membrane trafficking pathway plays a major role in the risk of Parkinson's disease. *Mov. Disord.* **2019**, *34*, 460–468. [[CrossRef](#)]
37. Billingsley, K.J.; Barbosa, I.A.; Bandrés-Ciga, S.; Quinn, J.P.; Bubb, V.J.; Deshpande, C.; Botia, J.A.; Reynolds, R.H.; Zhang, D.; Simpson, M.A.; et al. Mitochondria function associated genes contribute to Parkinson's Disease risk and later age at onset. *NPJ Parkinsons Dis.* **2019**, *5*, 8. [[CrossRef](#)]

38. Bandres-Ciga, S.; Saez-Atienzar, S.; Kim, J.J.; Makarious, M.B.; Faghri, F.; Diez-Fairen, M.; Iwaki, H.; Leonard, H.; Botia, J.; Ryten, M.; et al. Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. *Acta Neuropathol.* **2020**, *140*, 341–358. [[CrossRef](#)]
39. Agarwal, D.; Sandor, C.; Volpato, V.; Caffrey, T.; Monzon-Sandoval, J.; Bowden, R.; Alegre-Abarrategui, J.; Wade-Martins, R.; Webber, C. A human single-cell atlas of the Substantia nigra reveals novel cell-specific pathways associated with the genetic risk of Parkinson's disease and neuropsychiatric disorders. *bioRxiv* **2020**. [[CrossRef](#)]
40. Bryois, J.; Skene, N.G.; Hansen, T.F.; Kogelman, L.J.A.; Watson, H.J.; Liu, Z.; Brueggeman, L.; Breen, G.; Bulik, C.M.; Arenas, E.; et al. Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **2020**, *52*, 482–493. [[CrossRef](#)]
41. Wang, D.; Liu, S.; Warrell, J.; Won, H.; Shi, X.; Navarro, F.C.P.; Clarke, D.; Gu, M.; Emani, P.; Yang, Y.T.; et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* **2018**, *362*. [[CrossRef](#)] [[PubMed](#)]
42. Won, H.; de la Torre-Ubieta, L.; Stein, J.L.; Parikshak, N.N.; Huang, J.; Opland, C.K.; Gandal, M.J.; Sutton, G.J.; Hormozdiari, F.; Lu, D.; et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **2016**, *538*, 523–527. [[CrossRef](#)] [[PubMed](#)]
43. Sey, N.Y.A.; Hu, B.; Mah, W.; Fauni, H.; McAfee, J.C.; Rajarajan, P.; Brennand, K.J.; Akbarian, S.; Won, H. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat. Neurosci.* **2020**, *23*, 583–593. [[CrossRef](#)] [[PubMed](#)]
44. Keo, A.; Mahfouz, A.; Ingrassia, A.M.T.; Menebo, J.P.; Villenet, C.; Mutez, E.; Comptdaer, T.; Lelieveldt, B.P.F.; Figeac, M.; Chartier-Harlin, M.C.; et al. Transcriptomic signatures of brain regional vulnerability to Parkinson's disease. *Commun. Biol.* **2020**, *3*, 101. [[CrossRef](#)]
45. Wu, Y.; Zheng, Z.; Visscher, P.M.; Yang, J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biol.* **2017**, *18*, 86. [[CrossRef](#)] [[PubMed](#)]
46. Zhu, Z.; Zhang, F.; Hu, H.; Bakshi, A.; Robinson, M.R.; Powell, J.E.; Montgomery, G.W.; Goddard, M.E.; Wray, N.R.; Visscher, P.M.; et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **2016**, *48*, 481–487. [[CrossRef](#)] [[PubMed](#)]
47. Cruts, M.; Gijssels, I.; van der Zee, J.; Engelborghs, S.; Wils, H.; Pirici, D.; Rademakers, R.; Vandenberghe, R.; Dermaut, B.; Martin, J.J.; et al. Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature* **2006**, *442*, 920–924. [[CrossRef](#)]
48. Smith, K.R.; Damiano, J.; Franceschetti, S.; Carpenter, S.; Canafoglia, L.; Morbin, M.; Rossi, G.; Pareyson, D.; Mole, S.E.; Staropoli, J.F.; et al. Strikingly different clinicopathological phenotypes determined by progranulin-mutation dosage. *Am. J. Hum. Genet.* **2012**, *90*, 1102–1107. [[CrossRef](#)]
49. Giambartolomei, C.; Zhenli Liu, J.; Zhang, W.; Hauberg, M.; Shi, H.; Boocock, J.; Pickrell, J.; Jaffe, A.E.; Pasaniuc, B.; Roussos, P.; et al. A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics* **2018**, *34*, 2538–2545. [[CrossRef](#)]
50. Kia, D.A.; Zhang, D.; Guelfi, S.; Manzoni, C.; Hubbard, L.; United; (UKBEC), K.B.E.C.; Disease, I.P.s.; (IPDGC), G.C.; Reynolds, R.H.; et al. Integration of eQTL and Parkinson's disease GWAS data implicates 11 disease genes. *bioRxiv* **2019**, 627216. [[CrossRef](#)]
51. Beilina, A.; Rudenko, I.N.; Kaganovich, A.; Civiero, L.; Chau, H.; Kalia, S.K.; Kalia, L.V.; Lobbstaal, E.; Chia, R.; Ndukwe, K.; et al. Unbiased screen for interactors of leucine-rich repeat kinase 2 supports a common pathway for sporadic and familial Parkinson disease. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2626–2631. [[CrossRef](#)] [[PubMed](#)]
52. Fujimoto, T.; Kuwahara, T.; Eguchi, T.; Sakurai, M.; Komori, T.; Iwatsubo, T. Parkinson's disease-associated mutant LRRK2 phosphorylates Rab7L1 and modifies trans-Golgi morphology. *Biochem. Biophys. Res. Commun.* **2018**, *495*, 1708–1715. [[CrossRef](#)] [[PubMed](#)]
53. Liu, Z.; Bryant, N.; Kumaran, R.; Beilina, A.; Abeliovich, A.; Cookson, M.R.; West, A.B. LRRK2 phosphorylates membrane-bound Rabs and is activated by GTP-bound Rab7L1 to promote recruitment to the trans-Golgi network. *Hum. Mol. Genet.* **2018**, *27*, 385–395. [[CrossRef](#)] [[PubMed](#)]
54. Purlyte, E.; Dhekne, H.S.; Sarhan, A.R.; Gomez, R.; Lis, P.; Wightman, M.; Martinez, T.N.; Tonelli, F.; Pfeffer, S.R.; Alessi, D.R. Rab29 activation of the Parkinson's disease-associated LRRK2 kinase. *EMBO J.* **2018**, *37*, 1–18. [[CrossRef](#)] [[PubMed](#)]

55. Gusev, A.; Ko, A.; Shi, H.; Bhatia, G.; Chung, W.; Penninx, B.W.; Jansen, R.; de Geus, E.J.; Boomsma, D.I.; Wright, F.A.; et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **2016**, *48*, 245–252. [[CrossRef](#)] [[PubMed](#)]
56. Li, Y.I.; Wong, G.; Humphrey, J.; Raj, T. Prioritizing Parkinson's disease genes using population-scale transcriptomic data. *Nat. Commun.* **2019**, *10*, 994. [[CrossRef](#)]
57. Catanesi, M.; d'Angelo, M.; Tupone, M.G.; Benedetti, E.; Giordano, A.; Castelli, V.; Cimini, A. MicroRNAs Dysregulation and Mitochondrial Dysfunction in Neurodegenerative Diseases. *Int. J. Mol. Sci.* **2020**, *21*. [[CrossRef](#)]
58. Kern, F.; Fehlmann, T.; Violich, I.; Alsop, E.; Hutchins, E.; Kahraman, M.; Grammes, N.L.; Guimarães, P.; Backes, C.; Poston, K.; et al. Deep sncRNA-seq of the PPMI cohort to study Parkinson's disease progression. *bioRxiv* **2020**. [[CrossRef](#)]
59. Ohnmacht, J.; May, P.; Sinkkonen, L.; Krüger, R. Missing heritability in Parkinson's disease: The emerging role of non-coding genetic variation. *J. Neural Transm.* **2020**, *127*, 729–748. [[CrossRef](#)]
60. Bandres-Ciga, S.; Diez-Fairen, M.; Kim, J.J.; Singleton, A.B. Genetics of Parkinson's disease: An introspection of its journey towards precision medicine. *Neurobiol. Dis.* **2020**, *137*, 104782. [[CrossRef](#)]
61. Sudmant, P.H.; Rausch, T.; Gardner, E.J.; Handsaker, R.E.; Abyzov, A.; Huddleston, J.; Zhang, Y.; Ye, K.; Jun, G.; Fritz, M.H.; et al. An integrated map of structural variation in 2,504 human genomes. *Nature* **2015**, *526*, 75–81. [[CrossRef](#)] [[PubMed](#)]
62. Jeffares, D.C.; Jolly, C.; Hoti, M.; Speed, D.; Shaw, L.; Rallis, C.; Balloux, F.; Dessimoz, C.; Bähler, J.; Sedlazeck, F.J. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat. Commun.* **2017**, *8*, 14061. [[CrossRef](#)] [[PubMed](#)]
63. Han, L.; Zhao, X.; Benton, M.L.; Perumal, T.; Collins, R.L.; Hoffman, G.E.; Johnson, J.S.; Sloofman, L.; Wang, H.Z.; Stone, M.R.; et al. Functional annotation of rare structural variation in the human brain. *Nat. Commun.* **2020**, *11*, 2990. [[CrossRef](#)] [[PubMed](#)]
64. Iafrate, A.J.; Feuk, L.; Rivera, M.N.; Listewnik, M.L.; Donahoe, P.K.; Qi, Y.; Scherer, S.W.; Lee, C. Detection of large-scale variation in the human genome. *Nat. Genet.* **2004**, *36*, 949–951. [[CrossRef](#)]
65. Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shapero, M.H.; Carson, A.R.; Chen, W.; et al. Global variation in copy number in the human genome. *Nature* **2006**, *444*, 444–454. [[CrossRef](#)]
66. Tuzun, E.; Sharp, A.J.; Bailey, J.A.; Kaul, R.; Morrison, V.A.; Pertz, L.M.; Haugen, E.; Hayden, H.; Albertson, D.; Pinkel, D.; et al. Fine-scale structural variation of the human genome. *Nat. Genet.* **2005**, *37*, 727–732. [[CrossRef](#)]
67. Kidd, J.M.; Cooper, G.M.; Donahue, W.F.; Hayden, H.S.; Sampas, N.; Graves, T.; Hansen, N.; Teague, B.; Alkan, C.; Antonacci, F.; et al. Mapping and sequencing of structural variation from eight human genomes. *Nature* **2008**, *453*, 56–64. [[CrossRef](#)]
68. Conrad, D.F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T.D.; Barnes, C.; Campbell, P.; et al. Origins and functional impact of copy number variation in the human genome. *Nature* **2010**, *464*, 704–712. [[CrossRef](#)]
69. Pang, A.W.; MacDonald, J.R.; Pinto, D.; Wei, J.; Rafiq, M.A.; Conrad, D.F.; Park, H.; Hurles, M.E.; Lee, C.; Venter, J.C.; et al. Towards a comprehensive structural variation map of an individual human genome. *Genome Biol.* **2010**, *11*, R52. [[CrossRef](#)]
70. Bandrés-Ciga, S.; Ruz, C.; Barrero, F.J.; Escamilla-Sevilla, F.; Pelegrina, J.; Vives, F.; Duran, R. Structural genomic variations and Parkinson's disease. *Minerva Med.* **2017**, *108*, 438–447. [[CrossRef](#)]
71. La Cognata, V.; Morello, G.; D'Agata, V.; Cavallaro, S. Copy number variability in Parkinson's disease: Assembling the puzzle through a systems biology approach. *Hum. Genet.* **2017**, *136*, 13–37. [[CrossRef](#)] [[PubMed](#)]
72. Ibáñez, P.; Lesage, S.; Janin, S.; Lohmann, E.; Durif, F.; Destée, A.; Bonnet, A.M.; Brefel-Courbon, C.; Heath, S.; Zelenika, D.; et al. Alpha-synuclein gene rearrangements in dominantly inherited parkinsonism: Frequency, phenotype, and mechanisms. *Arch. Neurol.* **2009**, *66*, 102–108. [[CrossRef](#)] [[PubMed](#)]
73. Farrer, M.; Kachergus, J.; Forno, L.; Lincoln, S.; Wang, D.S.; Hulihan, M.; Maraganore, D.; Gwinn-Hardy, K.; Wszolek, Z.; Dickson, D.; et al. Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. *Ann. Neurol.* **2004**, *55*, 174–179. [[CrossRef](#)]
74. Sekine, T.; Kagaya, H.; Funayama, M.; Li, Y.; Yoshino, H.; Tomiyama, H.; Hattori, N. Clinical course of the first Asian family with Parkinsonism related to SNCA triplication. *Mov. Disord.* **2010**, *25*, 2871–2875. [[CrossRef](#)]

75. Keyser, R.J.; Lombard, D.; Veikondis, R.; Carr, J.; Bardien, S. Analysis of exon dosage using MLPA in South African Parkinson's disease patients. *Neurogenetics* **2010**, *11*, 305–312. [[CrossRef](#)]
76. Kojovic, M.; Sheerin, U.M.; Rubio-Agusti, I.; Saha, A.; Bras, J.; Gibbons, V.; Palmer, R.; Houlden, H.; Hardy, J.; Wood, N.W.; et al. Young-onset parkinsonism due to homozygous duplication of α -synuclein in a consanguineous family. *Mov. Disord.* **2012**, *27*, 1827–1829. [[CrossRef](#)]
77. Olgiati, S.; Thomas, A.; Quadri, M.; Breedveld, G.J.; Graafland, J.; Eussen, H.; Douben, H.; de Klein, A.; Onofri, M.; Bonifati, V. Early-onset parkinsonism caused by alpha-synuclein gene triplication: Clinical and genetic findings in a novel family. *Parkinsonism Relat. Disord.* **2015**, *21*, 981–986. [[CrossRef](#)] [[PubMed](#)]
78. Ferese, R.; Modugno, N.; Campopiano, R.; Santilli, M.; Zampatti, S.; Giardina, E.; Nardone, A.; Postorivo, D.; Fornai, F.; Novelli, G.; et al. Four Copies of SNCA Responsible for Autosomal Dominant Parkinson's Disease in Two Italian Siblings. *Parkinsons Dis.* **2015**, *2015*, 546462. [[CrossRef](#)]
79. Byers, B.; Cord, B.; Nguyen, H.N.; Schüle, B.; Fenno, L.; Lee, P.C.; Deisseroth, K.; Langston, J.W.; Pera, R.R.; Palmer, T.D. SNCA triplication Parkinson's patient's iPSC-derived DA neurons accumulate α -synuclein and are susceptible to oxidative stress. *PLoS ONE* **2011**, *6*, e26159. [[CrossRef](#)]
80. Fuchs, J.; Nilsson, C.; Kachergus, J.; Munz, M.; Larsson, E.M.; Schüle, B.; Langston, J.W.; Middleton, F.A.; Ross, O.A.; Hulihan, M.; et al. Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication. *Neurology* **2007**, *68*, 916–922. [[CrossRef](#)]
81. Konno, T.; Ross, O.A.; Puschmann, A.; Dickson, D.W.; Wszolek, Z.K. Autosomal dominant Parkinson's disease caused by SNCA duplications. *Parkinsonism Relat. Disord.* **2016**, *22*, S1–S6. [[CrossRef](#)] [[PubMed](#)]
82. Ibáñez, P.; Bonnet, A.M.; Débarges, B.; Lohmann, E.; Tison, F.; Pollak, P.; Agid, Y.; Dürr, A.; Brice, A. Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet* **2004**, *364*, 1169–1171. [[CrossRef](#)]
83. Nishioka, K.; Hayashi, S.; Farrer, M.J.; Singleton, A.B.; Yoshino, H.; Imai, H.; Kitami, T.; Sato, K.; Kuroda, R.; Tomiyama, H.; et al. Clinical heterogeneity of alpha-synuclein gene duplication in Parkinson's disease. *Ann. Neurol.* **2006**, *59*, 298–309. [[CrossRef](#)] [[PubMed](#)]
84. Nishioka, K.; Ross, O.A.; Ishii, K.; Kachergus, J.M.; Ishiwata, K.; Kitagawa, M.; Kono, S.; Obi, T.; Mizoguchi, K.; Inoue, Y.; et al. Expanding the clinical phenotype of SNCA duplication carriers. *Mov. Disord.* **2009**, *24*, 1811–1819. [[CrossRef](#)] [[PubMed](#)]
85. Ikeuchi, T.; Kakita, A.; Shiga, A.; Kasuga, K.; Kaneko, H.; Tan, C.-F.; Idezuka, J.; Wakabayashi, K.; Onodera, O.; Iwatsubo, T.; et al. Patients Homozygous and Heterozygous for SNCA Duplication in a Family With Parkinsonism and Dementia. *Arch. Neurol.* **2008**, *65*, 514–519. [[CrossRef](#)] [[PubMed](#)]
86. Nuytemans, K.; Meeus, B.; Crosiers, D.; Brouwers, N.; Goossens, D.; Engelborghs, S.; Pals, P.; Pickut, B.; Van den Broeck, M.; Corsmit, E.; et al. Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population. *Hum. Mutat* **2009**, *30*, 1054–1061. [[CrossRef](#)]
87. Sironi, F.; Trotta, L.; Antonini, A.; Zini, M.; Ciccone, R.; della Mina, E.; Meucci, N.; Sacilotto, G.; Primignani, P.; Brambilla, T.; et al. alpha-Synuclein multiplication analysis in Italian familial Parkinson disease. *Parkinsonism Relat. Disord.* **2010**, *16*, 228–231. [[CrossRef](#)] [[PubMed](#)]
88. Bonifati, V.; Rohé, C.F.; Breedveld, G.J.; Fabrizio, E.; De Mari, M.; Tassorelli, C.; Tavella, A.; Marconi, R.; Nicholl, D.J.; Chien, H.F.; et al. Early-onset parkinsonism associated with PINK1 mutations: Frequency, genotypes, and phenotypes. *Neurology* **2005**, *65*, 87–95. [[CrossRef](#)]
89. Bragg, D.C.; Mangkalaphiban, K.; Vaine, C.A.; Kulkarni, N.J.; Shin, D.; Yadav, R.; Dhakal, J.; Ton, M.-L.; Cheng, A.; Russo, C.T.; et al. Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in TAF1. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E11020–E11028. [[CrossRef](#)]
90. Makino, S.; Kaji, R.; Ando, S.; Tomizawa, M.; Yasuno, K.; Goto, S.; Matsumoto, S.; Tabuena, M.D.; Maranon, E.; Dantes, M.; et al. Reduced neuron-specific expression of the TAF1 gene is associated with X-linked dystonia-parkinsonism. *Am. J. Hum. Genet.* **2007**, *80*, 393–406. [[CrossRef](#)]
91. Aneichyk, T.; Hendriks, W.T.; Yadav, R.; Shin, D.; Gao, D.; Vaine, C.A.; Collins, R.L.; Domingo, A.; Currall, B.; Stortchevoi, A.; et al. Dissecting the Causal Mechanism of X-Linked Dystonia-Parkinsonism by Integrating Genome and Transcriptome Assembly. *Cell* **2018**, *172*, 897–909. [[CrossRef](#)]
92. McDonald-McGinn, D.M.; Sullivan, K.E.; Marino, B.; Philip, N.; Swillen, A.; Vorstman, J.A.; Zackai, E.H.; Emanuel, B.S.; Vermeesch, J.R.; Morrow, B.E.; et al. 22q11.2 deletion syndrome. *Nat. Rev. Dis. Primers* **2015**, *1*, 15071. [[CrossRef](#)] [[PubMed](#)]

93. Krahn, L.E.; Maraganore, D.M.; Michels, V.V. Childhood-onset schizophrenia associated with parkinsonism in a patient with a microdeletion of chromosome 22. *Mayo Clin. Proc.* **1998**, *73*, 956–959. [[CrossRef](#)] [[PubMed](#)]
94. Booij, J.; van Amelsvoort, T.; Boot, E. Co-occurrence of early-onset Parkinson disease and 22q11.2 deletion syndrome: Potential role for dopamine transporter imaging. *Am. J. Med. Genet.* **2010**, *152A*, 2937–2938. [[CrossRef](#)] [[PubMed](#)]
95. Butcher, N.J.; Kiehl, T.R.; Hazrati, L.N.; Chow, E.W.; Rogaeva, E.; Lang, A.E.; Bassett, A.S. Association between early-onset Parkinson disease and 22q11.2 deletion syndrome: Identification of a novel genetic form of Parkinson disease and its clinical implications. *JAMA Neurol.* **2013**, *70*, 1359–1366. [[CrossRef](#)] [[PubMed](#)]
96. Mok, K.Y.; Sheerin, U.; Simón-Sánchez, J.; Salaka, A.; Chester, L.; Escott-Price, V.; Mantripragada, K.; Doherty, K.M.; Noyce, A.J.; Mencacci, N.E.; et al. Deletions at 22q11.2 in idiopathic Parkinson's disease: A combined analysis of genome-wide association data. *Lancet Neurol.* **2016**, *15*, 585–596. [[CrossRef](#)]
97. Sebat, J.; Lakshmi, B.; Malhotra, D.; Troge, J.; Lese-Martin, C.; Walsh, T.; Yamrom, B.; Yoon, S.; Krasnitz, A.; Kendall, J.; et al. Strong association of de novo copy number mutations with autism. *Science* **2007**, *316*, 445–449. [[CrossRef](#)]
98. Weiss, L.A.; Shen, Y.; Korn, J.M.; Arking, D.E.; Miller, D.T.; Fossdal, R.; Saemundsen, E.; Stefansson, H.; Ferreira, M.A.; Green, T.; et al. Association between microdeletion and microduplication at 16p11.2 and autism. *New Engl. J. Med.* **2008**, *358*, 667–675. [[CrossRef](#)]
99. Turner, T.N.; Coe, B.P.; Dickel, D.E.; Hoekzema, K.; Nelson, B.J.; Zody, M.C.; Kronenberg, Z.N.; Hormozdiari, F.; Raja, A.; Pennacchio, L.A.; et al. Genomic Patterns of De Novo Mutation in Simplex Autism. *Cell* **2017**, *171*, 710–722.e712. [[CrossRef](#)]
100. Walsh, T.; McClellan, J.M.; McCarthy, S.E.; Addington, A.M.; Pierce, S.B.; Cooper, G.M.; Nord, A.S.; Kusenda, M.; Malhotra, D.; Bhandari, A.; et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **2008**, *320*, 539–543. [[CrossRef](#)]
101. Consortium, I.S. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **2008**, *455*, 237–241. [[CrossRef](#)]
102. McCarthy, S.E.; Makarov, V.; Kirov, G.; Addington, A.M.; McClellan, J.; Yoon, S.; Perkins, D.O.; Dickel, D.E.; Kusenda, M.; Krastoshevsky, O.; et al. Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.* **2009**, *41*, 1223–1227. [[CrossRef](#)]
103. Marshall, C.R.; Howrigan, D.P.; Merico, D.; Thiruvahindrapuram, B.; Wu, W.; Greer, D.S.; Antaki, D.; Shetty, A.; Holmans, P.A.; Pinto, D.; et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **2017**, *49*, 27–35. [[CrossRef](#)] [[PubMed](#)]
104. Collins, R.L.; Brand, H.; Karczewski, K.J.; Zhao, X.; Alföldi, J.; Francioli, L.C.; Khera, A.V.; Lowther, C.; Gauthier, L.D.; Wang, H.; et al. A structural variation reference for medical and population genetics. *Nature* **2020**, *581*, 444–451. [[CrossRef](#)]
105. Mahmoud, M.; Gobet, N.; Cruz-Dávalos, D.I.; Mounier, N.; Dessimoz, C.; Sedlazeck, F.J. Structural variant calling: The long and the short of it. *Genome Biol.* **2019**, *1*, 246. [[CrossRef](#)] [[PubMed](#)]

