



Review

# The Role of Structural Variants in the Genetic Architecture of Parkinson's Disease

Abigail Miano-Burkhardt <sup>1,2</sup>, Pilar Alvarez Jerez <sup>2</sup>, Kensuke Daida <sup>1,2</sup>, Sara Bandres Ciga <sup>2</sup>  
and Kimberley J. Billingsley <sup>1,2,\*</sup>

<sup>1</sup> Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA; abbylmiano@gmail.com (A.M.-B.); kensuke.daida@nih.gov (K.D.)

<sup>2</sup> Center for Alzheimer's and Related Dementias, National Institute on Aging, Bethesda, MD 20892, USA; pilar.alvarezjerez@nih.gov (P.A.J.); sara.bandresciga@nih.gov (S.B.C.)

\* Correspondence: kimberley.billingsley@nih.gov

**Abstract:** Parkinson's disease (PD) significantly impacts millions of individuals worldwide. Although our understanding of the genetic foundations of PD has advanced, a substantial portion of the genetic variation contributing to disease risk remains unknown. Current PD genetic studies have primarily focused on one form of genetic variation, single nucleotide variants (SNVs), while other important forms of genetic variation, such as structural variants (SVs), are mostly ignored due to the complexity of detecting these variants with traditional sequencing methods. Yet, these forms of genetic variation play crucial roles in gene expression and regulation in the human brain and are causative of numerous neurological disorders, including forms of PD. This review aims to provide a comprehensive overview of our current understanding of the involvement of coding and noncoding SVs in the genetic architecture of PD.

**Keywords:** Parkinson's disease; genetics; structural variants; long-read sequencing



**Citation:** Miano-Burkhardt, A.; Alvarez Jerez, P.; Daida, K.; Bandres Ciga, S.; Billingsley, K.J. The Role of Structural Variants in the Genetic Architecture of Parkinson's Disease. *Int. J. Mol. Sci.* **2024**, *25*, 4801. <https://doi.org/10.3390/ijms25094801>

Academic Editor: Cristoforo Comi

Received: 20 March 2024

Revised: 17 April 2024

Accepted: 22 April 2024

Published: 27 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Genetics of Parkinson's Disease

Parkinson's disease (PD) is the second most common neurodegenerative disorder and the most common movement disorder worldwide. It affected approximately 6.2 million individuals in 2015, with predictions estimating this number to double by 2040 [1]. PD has a known genetic component and can be both monogenic and apparently sporadic in nature. Mendelian forms of disease result from highly penetrant, disease-causing variants. Over twenty associated genes have been identified, although robust replication for the majority of these is missing [2]. On the other hand, sporadic PD is caused by common variants with moderate to low risk that increase disease susceptibility. To date, the majority of sporadic PD genetic studies have used large-scale genome-wide association studies (GWASs) focused on single nucleotide variants (SNVs) [3,4]. Through this work, over one hundred independent risk loci have been associated with apparently sporadic PD. However, these common variants explain around 16–30% of the heritable component of PD, leaving a large portion of the genetics of PD unknown [5,6].

### 1.2. Coding vs. Noncoding DNA

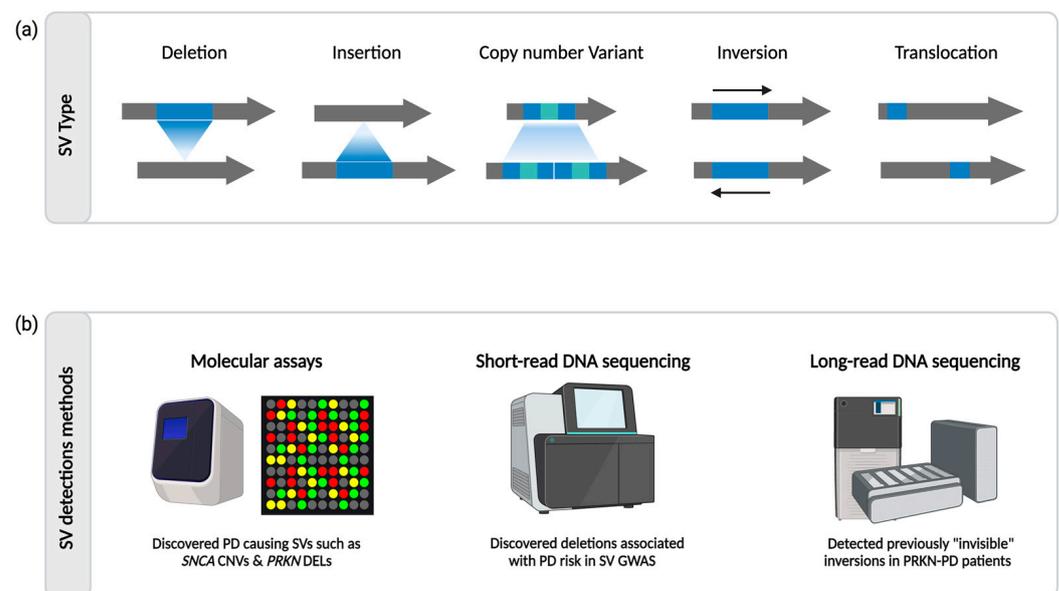
DNA can be categorized as either coding or noncoding. Coding DNA refers to regions of the genome that encode proteins, representing approximately 1.5% of the human genome [7]. In contrast, noncoding regions of the genome comprise the vast majority of DNA and include important regulatory elements and noncoding RNAs.

Causal variants identified in neurodegenerative diseases so far have tended to be located in coding regions of DNA, as these are more likely to have a substantial functional impact and are thus implicated in monogenic forms of disease. However, the majority

of common variants identified through GWASs are located in noncoding regions, which cumulatively contribute to disease risk. Despite being located in a noncoding region of the genome, these variants can still influence gene expression, protein function, and other molecular processes, thereby contributing to the risk of developing a disease or influencing a trait.

### 1.3. Structural Variants

The unaccounted “missing heritability” of PD is likely due to the fact that most existing genetic studies have focused on SNVs. Unlike SNVs that result in a single nucleotide change, structural variants (SVs) are any genomic rearrangement including fifty or more nucleotides, and these regions contribute to a more substantial amount of variation in the human genome compared to SNVs [8]. SVs are commonly classified as deletions, duplications, insertions, inversions, and translocations, each portraying distinctive combinations of gains, losses, or rearrangements within the DNA sequence [9,10]. Additionally, SVs include copy number variants (CNVs) and retrotransposon insertions such as *Alu*, LINE-1, and SINE-VNTR-*Alu* (SVA) elements, which can integrate themselves into the human genome. Finally, complex SVs constitute intricate combinations of these events. These genomic variations can considerably influence phenotypes by disrupting gene function, regulation, or altering gene dosage. Recent studies demonstrate that SVs drive functional changes not only across diverse populations but also within various cell and tissue types [10,11] (Figure 1a).

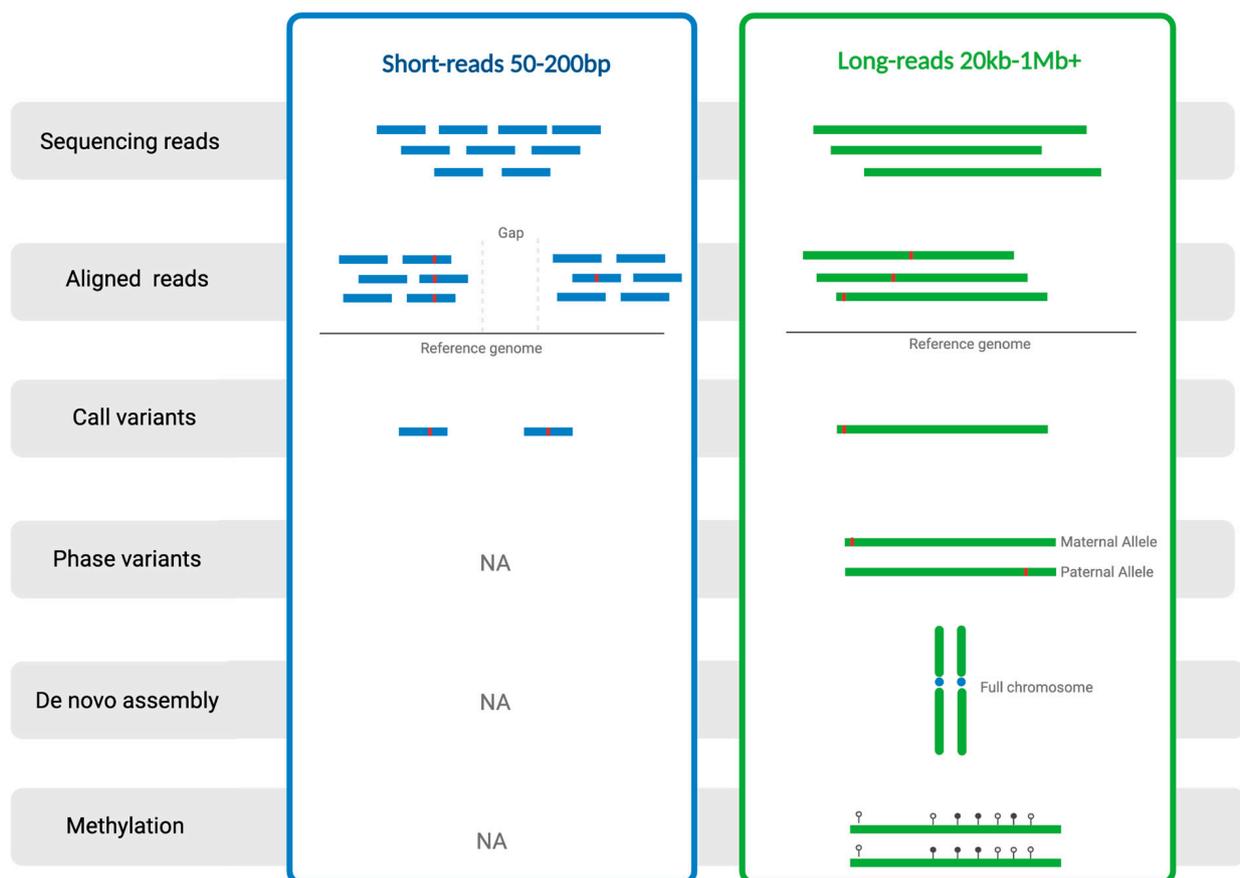


**Figure 1.** (a) Overview of different types of SVs. Each arrow represents a sequence of DNA. The blue portions of the arrows represent a variant. The first row of arrows represents the original DNA sequence while the second row of arrows represents the mutated DNA sequence. (b) Overview of the main technologies used to detect SVs. Created with [BioRender.com](https://www.biorender.com) (accessed on 16 January 2024).

### 1.4. Methods for Detecting Structural Variants

There are several methods for detecting SVs, including molecular assays, short-read DNA sequencing, and, more recently, long-read DNA sequencing (Figure 1b). Many of the first PD SV genetic studies used molecular techniques, such as quantitative polymerase chain reaction (qPCR), multiplex ligation-dependent amplification (MLPA), fluorescent in situ hybridization (FISH), and other microarrays to detect gene dosage. qPCR, in comparison to standard PCR, uses fluorescent labeling to measure the amount of PCR product in real time, which allows for more accurate quantification. MLPA is a variation of a multiplex PCR, which uses probes rather than primers to detect a target sequence and amplifies ligated probes, which can then be quantified using capillary electrophoresis. FISH

uses fluorescent probes to precisely localize a specific DNA sequence in an individual cell. However, these methods cannot provide sequence information. Short-read sequencing is a widely used technique in genomics that involves sequencing DNA fragments of relatively small length (typically ranging from 50 to 300 nucleotides). This approach, employed in technologies such as Illumina sequencing, is known for its high throughput and cost-effectiveness. In recent years, hundreds of tools have been developed for detecting SVs from short-read sequencing data, and using these tools, several groups have been successful in identifying SVs associated with disease risk [12–14]. However, short-read sequencing faces challenges in accurately calling SVs due to limitations in read length, difficulty in spanning large-scale variations, issues with repetitive regions, and potential chimeric read alignment, which often lead to spurious associations that lack replication. In comparison, “third generation” long-read sequencing technology makes it possible to sequence much larger stretches of DNA. Technologies such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) generate reads spanning tens of thousands to millions of base pairs. Because of this, long-read sequencing offers advantages in resolving complex genetic variation, including repetitive regions and SVs. This technology is particularly valuable for de novo genome assembly, accurately identifying large-scale SVs, and providing a more comprehensive view of the genome’s architecture. Figure 2 gives a detailed overview of the different capabilities of short- and long-read DNA sequencing technologies.



**Figure 2.** Comparison of the capabilities of short- and long-read sequencing technologies. The left panel displays short blue lines representing reads from short-read sequencing, and the right panel displays long green lines representing reads from long-read sequencing (except for the two vertical green lines which represent a homologous chromosome). The red lines on some of the reads are variants. Adapted from Kono and Arakawa [15]. Created with [BioRender.com](https://BioRender.com) (accessed on 16 January 2024).

### 1.5. Structural Variants in Neurological Conditions

SVs are well-established drivers of gene regulation in the human brain. In a recent study, Han and colleagues integrated SV calls from nearly fifteen thousand individuals with short-read bulk RNA-sequencing datasets from human post-mortem brain tissue. Through this work, they identified many coding SVs that overlapped with gene expression outliers [11]. Additionally, the authors of a large-scale study performed SV quantitative trait locus (QTL) analyses, and they reported that more than three thousand SVs impact histone modifications, mRNA expression, mRNA splicing, and protein abundance in brain-based short-read datasets [16]. Due to their potential to exert profound phenotypic influences in the brain, numerous mostly coding SVs have been linked with neurological conditions [13] and neurodevelopmental/neuropsychiatric disorders [17]. For example, Phelan-McDermid, or 22q13 deletion syndrome, which most commonly results in intellectual disability (ID), developmental delay, delayed or absent speech, and hypotonia, all of varying severity, can be caused by a deletion at the *SHANK3* gene [18]. Syndromic X-linked ID Lubs type is caused by a duplication of the gene *MECP2*, which almost exclusively occurs in males and is characterized by ID, developmental delay, recurrent infections, delayed speech and motor, hypotonia, seizures, and gastrointestinal dysfunction [19]. Further, Duchenne Muscular Dystrophy can be caused by deletions or duplications in the *DMD* gene [20], and deletions at the *SPAST* gene lead to hereditary spastic paraplegia [21]. In addition to coding SVs being causative of monogenic conditions, there are several SVs that significantly increase an individual's risk of disease. One example of this is a deletion at 1q21.1, which is associated with an increased risk of autism spectrum disorder (ASD), attention deficit hyperactivity disorder (ADHD), schizophrenia, ID, epilepsy [22], and parkinsonism [23].

In the context of SVs and their role in various forms of parkinsonism, it is noteworthy that a rare SV is associated with X-linked Dystonia Parkinsonism (XDP). XDP, also known as Lubag syndrome, is a rare genetic movement disorder characterized by progressive dystonia and parkinsonism, endemic to the island of Panay in the Philippines [24]. Specifically, XDP is associated with an insertion of an SVA retrotransposon that is inserted into intron 32 of the *TAF1* gene [25–27]. This SVA contains a hexameric DNA repeat domain (CCCTCT)<sub>n</sub> that is variable in length across patients, and the length of this expansion inversely correlates with age at onset (AAO). To understand the functional impact of the SVA, Aneichyk and colleagues demonstrated in cell models that this variant altered *TAF1* splicing and decreased expression of the full-length transcript [28]. Additionally, using induced pluripotent stem cell (iPSC) lines derived from XDP patients, Rakovic and colleagues showed that removal of the SVA using genome editing can successfully rescue *TAF1* expression [29], highlighting the causative nature of the SVA.

This review highlights key studies that have identified SVs associated with monogenic forms of PD (detailed in Supplementary Table S1), which are mostly in coding regions of the genome. We also outline SVs associated with the risk of apparently sporadic PD and finally detail ongoing research efforts to comprehensively catalog SVs in PD through large-scale long-read sequencing initiatives. Notably, while several studies have also reported repeat expansions that are associated with PD, these will not be discussed in the present review as our focus is on SVs.

## 2. Monogenic Parkinson's Disease

### 2.1. Autosomal Dominant Parkinson's Disease

#### SNCA/PARK1

The *SNCA* gene, Synuclein Alpha, is located on chromosome 4 (chr4:89724099–89837161) and encodes presynaptic neuronal protein alpha-synuclein [30–32]. A build-up of this protein in the brain is the major constituent of Lewy bodies, the main pathological hallmark of synucleinopathies, including PD. Although PD was originally believed to be caused only by environmental factors, in 1997, a missense SNV in *SNCA* was associated with PD [33]. Other point mutations at this locus have also been identified to cause familial PD [34–37]. However, the first time an SV was associated with *SNCA* was in 2003, when a

deleterious triplication of the entire gene was found in a family with autosomal dominant PD [38]. Following this, further studies have continued to identify *SNCA* multiplications (CNVs) as a rare cause of familial parkinsonism [39–46]. In addition to multiple *SNCA* triplication studies in individuals of European ancestry, other studies have reported this multiplication in other populations, such as the Japanese [44,46], Korean [39], South African, and Tunisian [31] populations.

Multiplications of *SNCA* include triplications and duplications. A heterozygous triplication has three copies of the *SNCA* gene on one allele and a single copy of the gene on the other allele, meaning a triplication carrier has a total of four copies of *SNCA*. A heterozygous duplication carrier has two copies of the *SNCA* gene on one allele and one copy on the other, for a total of three copies. Duplication carriers can also be homozygous, meaning they have two copies of *SNCA* on each allele for a total of four copies. In a large family study that comprised 264 families with typical autosomal dominant PD and 22 families with atypical autosomal dominant parkinsonism, about 1.5% of families with typical autosomal dominant PD had an *SNCA* duplication, while 4.5% of atypical families had a triplication [47]. In the typical PD population, duplications are more frequent than triplications. Most multiplications of this locus involve more than one gene and vary widely in length. Differences in multiplication length between families indicate that SV events occurred independently [31]. The rise of these de novo variations may be due to flanking repeat regions, which could cause genomic instability in this region and lead to gaps and breaks [41].

Along with being disease-causing, SVs at the *SNCA* gene are disease-modifying as the *SNCA* copy number determines the AAO and severity of symptoms, indicating a gene dosage effect. These multiplications are not only limited to *SNCA* but can include any number of neighboring genes, from two or three to dozens. However, neither the number of genes nor the size of the multiplication seems to have an effect on disease severity. The phenotype for duplication carriers includes symptoms similar to that of typical PD patients. These symptoms include mild to moderate common parkinsonism motor symptoms, such as postural instability and resting tremor, with little to no cognitive or autonomic dysfunction. However, some individuals have an *SNCA* duplication but no clinical presentation [44], while others show symptoms similar to that of atypical parkinsonian syndromes [39]. It is unclear why this phenotypic heterogeneity occurs, though it could be linked to other genetic modifiers. The phenotype for patients carrying the triplication, along with homozygous duplication carriers, however, includes atypical features reminiscent of dementia with Lewy bodies (DLB). *SNCA* triplication causes early AAO (younger than 50 years) with rapidly evolving, aggressive symptoms, both motor and non-motor, suggesting complete penetrance. Compared to heterozygous duplication carriers, triplication carriers' motor symptoms are those typically associated with PD, while non-motor symptoms include cognitive impairment and dysautonomia [40]. *SNCA* triplication can also cause levodopa-induced dyskinesia.

A number of studies have looked into the functional effects of these multiplications, particularly the triplication, in vitro and post-mortem. iPSCs derived from patients with an *SNCA* triplication recapitulate the accumulation of alpha-synuclein and doubling of gene expression [48] as seen in post-mortem blood and brain samples [49]. Various pathophysiological mechanisms have been reported to be triggered by overexpression of *SNCA* in mature iPSC-derived dopaminergic cells, such as elevated oxidative stress [50] and mitochondrial [51] and lysosomal dysfunction [52]. Studies have indicated that the *SNCA* locus triplication can impair neuronal differentiation and maturation in PD [53]. Severe neuronal degeneration in the substantia nigra, locus ceruleus, basal nucleus of Meynert, and hippocampal cornu ammonis has been seen in both triplication [40] and duplication carriers [54].

## 2.2. Autosomal Recessive Parkinson's Disease

### 2.2.1. PRKN/PARK2

A few years before the SNCA triplication was discovered, the Parkin RBR E3 ubiquitin-protein ligase (*PRKN*) gene was first characterized in 1998 when Kitada and colleagues [55] investigated a very large gene consisting of 1.3 Mb and 12 exons on chromosome 6 [chr6:161347417–162727766] [32]. This study identified a Japanese patient with autosomal recessive juvenile parkinsonism who was missing five exons (exons 3–7). This same study also identified an exon 4 deletion in four separate autosomal recessive PD families.

Parkin is an E3 ubiquitin ligase responsible for catalyzing the ubiquitination process and thus the degradation of aberrant proteins. Studies using fibroblasts from *PRKN* SV carriers implicate a loss of ubiquitin ligase function, ultimately leading to the pathogenesis of PD. Proteins are no longer effectively degraded by the parkin-dependent ubiquitin proteasome pathway, which leads to a protein build-up in substantia nigra pars compacta neurons [56]. *PRKN* SV patient fibroblasts [57,58] and mouse models [59–61] have been used to exhibit *PRKN* loss-of-function effects such as mitochondrial dysfunction and oxidative stress. Neuropathologically, *PRKN*-linked PD cases show significant substantia nigral degeneration and generally lack Lewy bodies [62–66]. Cases absent of Lewy bodies may be due to a younger AAO.

Causative SVs in *PRKN* have been reported in several other disorders [67], such as ADHD [68], cancer [69], and ASD [70,71]. *PRKN* variants are the most common cause of autosomal recessive PD, affecting 12.5% of individuals (who are either homozygous or compound heterozygous) in this population, and 42.2% of *PRKN* mutation carriers have an AAO less than 21 years [72]. SVs in this region account for 43.2% of *PRKN* mutations in autosomal recessive PD [73].

*PRKN* is located in FRA6E, a frequently observed common fragile site in the human genome, making it susceptible to gaps and breaks [74]. Deletions are the most common SVs among *PRKN* mutations. These deletions can encompass multiple exons or just one. The majority of exonic deletions span the noncoding region of *PRKN* between exons 2 and 4, and within this exonic region, patients most frequently carry deletions of exon 3 [73,75,76]. Patients with *PRKN* mutations on both alleles present with typical early-onset parkinsonism, including slow disease progression, positive levodopa responsiveness, absence of cognitive decline, resting tremor, rigidity, and levodopa-induced dyskinesia. Duplications, triplications, and occasionally inversions have also been reported as a cause of *PRKN* autosomal recessive PD [77].

There has been controversy among *PRKN* pathogenicity in heterozygous patients. Although homozygous and compound heterozygous carriers show a strong association with PD, researchers have also investigated heterozygous carriers with debate among the findings. In some studies, *PRKN* monoallelic mutations were indicated to be disease-causing [78]. However, evidence shows that these heterozygous participants were incorrectly genotyped due to researchers missing second *PRKN* mutations. In a pair of monozygotic twins presenting with early-onset PD, a very large, novel inversion spanning *PRKN* exons 1–11 was identified using long-read sequencing and went undetected using both short-read whole exome sequencing and MLPA [79]. After validating genotyping methods, heterozygous carriers were not found to be at higher risk for PD compared to those without a *PRKN* mutation [76,77]. Single *PRKN* mutations are relatively common in the general population at around 1.8%.

### 2.2.2. PARK7/DJ-1

Following the discovery of the deletion at *PRKN* was the identification of an SV at Parkinson's disease gene 7 (*PARK7*), formerly known as *DJ-1*, located on chromosome 1 [chr1:7961711–7985505] [32]. In 2003, *PARK7* was first characterized when a 14 kb homozygous deletion of exons 1–5 was identified in a Dutch family with early-onset parkinsonism [80]. Mutations in *PARK7*, encoding the deglycase protein DJ-1, are a rare cause of familial autosomal recessive PD, affecting less than 2% of cases [81,82]. Although the

majority of pathogenic variants in this gene are attributed to missense/nonsense mutations [73], exonic rearrangements have been identified, including a homozygous deletion of exon 5 [83] and another homozygous deletion of *PARK7* exons 1–5. This homozygous deletion of exons 1–5 was present with a second homozygous deletion of exons 1–6 in the neighboring gene *TNFRSF9* [84]. Individuals with homozygous SVs at *PARK7* tend to present with early-onset (20s and 30s) PD with typical features and a suitable levodopa response. A loss-of-function effect is observed in the same Dutch family with the homozygous 1–5 exonic deletion [80], but studies have yet to investigate the downstream effects of these large SVs in *PARK7*. The function of *PARK7* is hypothesized to play a protective role against oxidative stress and neuroinflammation, but the exact mechanism by which it performs this is still debated [85].

### 2.2.3. PINK1/PARK6

The next SV-harboring gene discovered in autosomal recessive PD was P-TEN-induced putative kinase 1 (*PINK1*). *PINK1* was first associated with PD in 2004 when Valente and colleagues [86] identified a nonsense and a missense variant located on chromosome 1 [chr1:20633458–20651511] in three consanguineous parkinsonian families [32]. Similar to *PARK7*, *PINK1* mutations are primarily missense [73,87], but SVs have additionally been identified at this locus. SVs include homozygous deletions of exons 6–8 [88], a compound heterozygous deletion of the entire *PINK1* gene [89], homozygous deletions of exons 4–8 [90], homozygous deletions of exons 2–4, homozygous deletions of exon 4, and homozygous deletions of exon 5 [83]. *PINK1* SV carriers tend to have a younger AAO but largely resemble typical idiopathic PD patients with features including a suitable levodopa response and a lack of dementia. *PINK1* phosphorylates parkin and ubiquitin to degrade damaged mitochondria [91]. Although mouse models have yet to accurately mirror the neurodegeneration seen in PD patient brains, more recent studies in monkeys with *PINK1* deletions show deterioration of the cortex, striatum, and substantia nigra [92,93].

## 3. Apparently Sporadic Parkinson's Disease

### 22Q11.2 Deletion

In addition to SVs being implicated as causative of monogenic forms of PD, numerous instances highlight SVs that are also genetic risk factors. One such region is located on chromosome 22q11.2, a segment in the human genome prone to meiotic chromosome rearrangements. These rearrangements can result in deletions, contributing to a spectrum of heterogeneous disorders. Among these, 22q11.2 deletion syndrome (22q11.2DS), historically known as DiGeorge syndrome, is the most common and well characterized of these disorders, occurring in ~1 in every 4000 live births [94–97]. The deletion can span 1.5–3 Mb and primarily arises de novo, although 5% of the events are inherited and the clinical phenotype varies widely. These include medical features such as congenital heart disease, autoimmune disease, immunodeficiency, palatal abnormalities, developmental delays, and learning difficulties [98].

Several studies have established a link between 22q11.2DS and increased risk of PD. Early independent case reports documented the co-occurrence of early-onset parkinsonism and 22q11.2DS [99,100]. Further, Butcher and colleagues reported that in a group of 68 young adults (18–34 years) with the syndrome, four had been diagnosed with early-onset PD [101]. The authors concluded that individuals with 22q11.2DS are estimated to have at least a twenty times increased risk of PD compared with estimates for the general population. Expanding on this association, in 2016, Mok and colleagues conducted an extensive analysis using SNV array datasets [102]. Their large-scale investigation of 22q11.2 deletions involved 9387 cases with apparently sporadic PD and 13,863 controls from European ancestry. Through this work, they identified that chromosome 22q11.2 deletions were enriched in PD cases (8 carriers in PD cases and none in controls). Moreover, PD patients carrying a deletion had an earlier AAO compared to non-carriers (mean 42.1 years vs. mean 61 years, respectively).

#### 4. Structural Variants as PD Genetic Risk Factors

Large-scale GWASs have enabled well-powered and unbiased analyses for identifying genetic risk factors. These genetic studies have played a fundamental role in discovering the PD risk loci known today. But, traditionally, due to the instability of accurately genotyping SVs at scale, these studies only focused on SNVs. Recent advances in short-read sequencing technologies and SV detection tools now allow for assessing the contribution of SVs to disease risk in large cohorts. Through this, large-scale SV GWASs have begun to identify SV genetic risk factors associated with various neurodegenerative diseases, including amyotrophic lateral sclerosis [103], Alzheimer's disease [104], non-Alzheimer's dementias (Lewy body dementia and frontotemporal dementia) [13], and PD [12].

In 2023, the first PD SV GWAS leveraged the SV detection tool GATK-SV [105] and called SVs using short-read sequencing whole-genome data from 2585 PD cases and 2779 controls from European ancestry. GATK-SV is a "gold standard" SV discovery pipeline developed by the Broad Institute that runs five different SV tools to capture all types of SVs. Using machine learning algorithms, it then merges together and filters the SV calls and outputs one final group vcf file that can be used for downstream genetic analyses [106]. Overall, after the implementation of this pipeline in the PD dataset, over 200,000 SVs were identified, and a mean of ~5600 SVs were detected per genome. Given that short-read sequencing data can still yield many false positives, especially in repetitive regions, it is crucial to validate all SVs of interest with additional molecular assays such as PCR or long-read DNA sequencing [107]. After generating high-quality ONT long-read sequencing from a subset of matched samples, each genome carried a mean of ~27,000 SVs overall. Through ONT validation, three deletions were confirmed that tagged known PD SNVs. Despite identifying these new SVs associated with PD risk, further benchmarking studies comparing the short- and long-read sequencing data demonstrated that ~30% of the SVs detected using short-read sequencing methods were false positives, and ~85% of the SVs detected by the long-read sequencing data were missed by the short-read caller [12]. This highlights that, although analysis of SVs using short-read sequencing datasets can yield valuable insights, it lacks the power to accurately detect most of the SVs in the human genome.

#### 5. Future Directions

To elucidate the genetic architecture of PD on a global scale, it is crucial to assess the impact of genetic variation across large, diverse populations. To date, however, the majority of PD GWASs have only been based on European samples. In fact, a recent systematic review quantifying the non-European representation in existing PD GWAS efforts revealed that, despite comprising only 16% of the global population, ~62% of all participants in PD genetic studies were European. This disparity was even more pronounced for neurodegenerative diseases as a whole, with only ~20% of participants in these GWASs being of non-European ancestry. Although there are several consortia-based efforts to improve diversity in PD research, much more is needed to better understand the risk of PD in other populations [108].

Gaining a comprehensive understanding of PD genetics, as outlined throughout this review, also requires the analysis of all forms of genetic variation, extending beyond just SNVs. Although most of the new efforts in the PD space are still focused on generating SNV datasets, these datasets can still be leveraged in conjunction with new SV reference panels to impute SVs into large datasets to gain valuable insights into the role of SVs. Additionally, there are several large-scale long-read sequencing efforts that aim to catalog SVs across diverse populations, including NIH-led initiatives such as the All of Us (AOU) initiative [109] and the Center of Alzheimer's Disease and Related Dementias (CARD) long-read sequencing initiative [110,111]. Most importantly, in terms of PD-specific initiatives, the Global Parkinson's Genetics Program (GP2) is conducting a large-scale ONT long-read initiative, which is split broadly into two sections [112,113]. The first section is part of the "Monogenic Network", whereby samples that have been prioritized based on AAO,

family structure, and negative results from short-read sequencing will undergo long-read sequencing to identify new causal variants. Additionally, ~1000 samples will be sequenced from a cohort of apparently sporadic PD patients and controls with the aim of identifying new genetic risk factors and resolving existing known risk loci.

Finally, another promising avenue for future research involves studying the impact of somatic SVs in PD. While the present review has focused on germline SVs, which are genetic variations that can be inherited from parent to offspring, somatic variation is a genetic variation that occurs post-zygotically and can consequently lead to mosaicism in the human brain. However, detecting somatic variation accurately is challenging and often requires high-depth sequencing to reliably identify variants present in a small subset of cells. Despite these difficulties, recent studies have started to address the role of somatic SVs in PD. Mokreter and colleagues detected somatic SVs in *SNCA* in multiple synucleinopathies, including PD and multiple system atrophy (MSA) [114], and using single-cell whole genome sequencing (WGS), Perez-Rodriguez and colleagues report that MSA patients have large CNVs in approximately 30% of all cells [115]. But, these studies have been limited by current technology, which has prevented them from accurately determining the breakpoints of somatic SVs, analyzing a large number of samples, and detecting CNVs smaller than 1 Mb in size. Moving forward, advancements in tools combining high-throughput long-read sequencing and single-cell methods to overcome previous limitations hold the potential to offer a more comprehensive understanding of the impact of somatic SVs in PD in the future.

### Conclusions

While many SVs are known to be associated with PD, our understanding of the overall role that SVs play in PD genetics remains very incomplete. This review highlights that conventional techniques frequently miss potential disease-relevant SVs, as exemplified by the recent identification of a large, rare *PRKN* inversion observed in Japanese PD patients. Long-read sequencing emerges as a powerful tool for resolving these complicated regions of the genome, making ongoing research initiatives that leverage long-read sequencing on a population scale particularly promising. The insights gained from these initiatives are anticipated to revolutionize our understanding of the genetic basis of PD, ultimately facilitating the development of more targeted and effective therapeutic interventions.

**Supplementary Materials:** The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25094801/s1>.

**Author Contributions:** Conceptualization, K.J.B.; writing—original draft preparation, A.M.-B. and K.J.B.; writing—review and editing, A.M.-B., P.A.J., K.D., S.B.C. and K.J.B. 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, National Institute of Neurological Disorders and Stroke project numbers: project numbers 1ZIA-NS003154, Z01-AG000949-02, and Z01-ES101986). K.D. reports receiving grants from the Japan Society for the Promotion of Science Research Fellowship for Japanese Biomedical and Behavioral Researchers at the National Institutes of Health.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

22q11.2DS	22q11.2 deletion syndrome
AAO	Age at onset
aCGH	Array comparative genomic hybridization
ADHD	Attention-deficit hyperactivity disorder
ASD	Autism spectrum disorder
AOU	All of Us
CARD	Center for Alzheimer’s and Related Dementias
CNV	Copy number variant
DaTSCAN	Dopamine transporter scan
DHPLC	Denaturing high performance liquid chromatography
DLB	Dementia with Lewy bodies
FISH	Fluorescent in situ hybridization
GP2	Global Parkinson’s Genetics Program
GWAS	Genome-wide association study
ID	Intellectual disability
iPSC	Induced pluripotent stem cell
MRI	Magnetic resonance imaging
MLPA	Multiplex ligation-dependent amplification
MSA	Multiple system atrophy
NGS	Next-generation sequencing
ONT	Oxford Nanopore Technologies
<i>PARK</i>	Parkinson’s disease gene
PD	Parkinson’s disease
PET	Positron emission tomography
<i>PINK1</i>	P-TEN-induced putative kinase 1
<i>PRKN</i>	Parkin RBR E3 ubiquitin-protein ligase
qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait locus
<i>SNCA</i>	Synuclein alpha
SPECT	Single-photon emission computed tomography
SNV	Single nucleotide variant
SV	Structural variant
SVA	SINE-VNTR- <i>Alu</i>
WGS	Whole genome sequencing
XDP	X-linked Dystonia Parkinsonism

## References

- Dorsey, E.R.; Bloem, B.R. The Parkinson Pandemic-A Call to Action. *JAMA Neurol.* **2018**, *75*, 9–10. [[CrossRef](#)] [[PubMed](#)]
- Blauwendraat, C.; Nalls, M.A.; Singleton, A.B. The genetic architecture of Parkinson’s disease. *Lancet Neurol.* **2020**, *19*, 170–178. [[CrossRef](#)] [[PubMed](#)]
- Chang, D.; Nalls, M.A.; Hallgrímsdóttir, I.B.; Hunkapiller, J.; Van Der Brug, M.; Cai, F.; International Parkinson’s Disease Genomics Consortium; 23andMe Research Team; Kerchner, G.A.; Ayalon, G.; 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](#)] [[PubMed](#)]
- 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](#)] [[PubMed](#)]
- Kim, J.J.; Vitale, D.; Otani, D.V.; Lian, M.M.; Heilbron, K.; Iwaki, H.; Lake, J.; Solsberg, C.W.; Leonard, H.; Makarious, M.B.; et al. Multi-ancestry genome-wide association meta-analysis of Parkinson’s disease. *Nat. Genet.* **2023**, *56*, 27–36. [[CrossRef](#)] [[PubMed](#)]
- 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](#)] [[PubMed](#)]
- Lander, E.S. Initial impact of the sequencing of the human genome. *Nature* **2011**, *470*, 187–197. [[CrossRef](#)]
- 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](#)] [[PubMed](#)]

9. Sedlazeck, F.J.; Lee, H.; Darby, C.A.; Schatz, M.C. Piercing the dark matter: Bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet.* **2018**, *19*, 329–346. [[CrossRef](#)]
10. Sudmant, P.H.; Rausch, T.; Gardner, E.J.; Handsaker, R.E.; Abyzov, A.; Huddleston, J.; Zhang, Y.; Ye, K.; Jun, G.; Fritz, M.H.-Y.; et al. An Integrated Map of Structural Variation in 2504 Human Genomes. *Nature* **2015**, *526*, 75–81. [[CrossRef](#)] [[PubMed](#)]
11. 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](#)]
12. Billingsley, K.J.; Ding, J.; Jerez, P.A.; Illarionova, A.; Levine, K.; Grenn, F.P.; Makarios, M.B.; Moore, A.; Vitale, D.; Reed, X.; et al. Genome-Wide Analysis of Structural Variants in Parkinson Disease. *Ann. Neurol.* **2023**, *93*, 1012–1022. [[CrossRef](#)] [[PubMed](#)]
13. Kaivola, K.; Chia, R.; Ding, J.; Rasheed, M.; Fujita, M.; Menon, V.; Walton, R.L.; Collins, R.L.; Billingsley, K.; Brand, H.; et al. Genome-wide structural variant analysis identifies risk loci for non-Alzheimer’s dementias. *Cell Genom.* **2023**, *3*, 100316. [[CrossRef](#)] [[PubMed](#)]
14. Li, Y.; Roberts, N.D.; Wala, J.A.; Shapira, O.; Schumacher, S.E.; Kumar, K.; Khurana, E.; Korbel, J.O.; Haber, J.E.; Imielinski, M.; et al. Patterns of somatic structural variation in human cancer genomes. *Nature* **2020**, *578*, 112–121. [[CrossRef](#)]
15. Kono, N.; Arakawa, K. Nanopore sequencing: Review of potential applications in functional genomics. *Dev. Growth Differ.* **2019**, *61*, 316–326. [[CrossRef](#)]
16. Vialle, R.A.; Lopes, K.d.P.; Bennett, D.A.; Crary, J.F.; Raj, T. Integrating whole-genome sequencing with multi-omic data reveals the impact of structural variants on gene regulation in the human brain. *Nat. Neurosci.* **2022**, *25*, 504–514. [[CrossRef](#)] [[PubMed](#)]
17. Mollon, J.; Almasy, L.; Jacquemont, S.; Glahn, D.C. The contribution of copy number variants to psychiatric symptoms and cognitive ability. *Mol. Psychiatry* **2023**, *28*, 1480–1493. [[CrossRef](#)] [[PubMed](#)]
18. Wilson, H.L.; Wong, A.C.; Shaw, S.R.; Tse, W.Y.; Stapleton, G.A.; Phelan, M.C.; Hu, S.; Marshall, J.; McDermid, H.E. Molecular Characterisation of the 22q13 Deletion Syndrome Supports the Role of Haploinsufficiency of SHANK3/PROSAP2 in the Major Neurological Symptoms. *J. Med. Genet.* **2003**, *40*, 575–584. [[CrossRef](#)]
19. Van Esch, H. MECP2 Duplication Syndrome. In *GeneReviews*<sup>®</sup> [Internet]; University of Washington: Seattle, WA, USA, 2020.
20. Aartsma-Rus, A.; Ginjaar, I.B.; Bushby, K. The importance of genetic diagnosis for Duchenne muscular dystrophy. *J. Med. Genet.* **2016**, *53*, 145–151. [[CrossRef](#)]
21. Méreaux, J.-L.; Banneau, G.; Papin, M.; Coarelli, G.; Valter, R.; Raymond, L.; Kol, B.; Ariste, O.; Parodi, L.; Tissier, L.; et al. Clinical and genetic spectra of 1550 index patients with hereditary spastic paraplegia. *Brain* **2022**, *145*, 1029–1037. [[CrossRef](#)] [[PubMed](#)]
22. Yoon, J.; Mao, Y. Dissecting Molecular Genetic Mechanisms of 1q21.1 CNV in Neuropsychiatric Disorders. *Int. J. Mol. Sci.* **2021**, *22*, 5811. [[CrossRef](#)] [[PubMed](#)]
23. Figura, M.; Geremek, M.; Milanowski, M.; Meisner-Kramarz, I.; Duszyńska-Was, K.; Szlufik, S.; Rózański, D.; Smyk, M.; Koziorowski, D. Movement disorders associated with chromosomal aberrations diagnosed in adult patients. *Neurol. Neurochir. Pol.* **2021**, *55*, 300–305. [[CrossRef](#)] [[PubMed](#)]
24. Lee, L.V.; Pascasio, F.M.; Fuentes, F.D.; Viterbo, G.H. Torsion Dystonia in Panay, Philippines. *Adv. Neurol.* **1976**, *14*, 137–151. [[PubMed](#)]
25. Domingo, A.; Westenberger, A.; Lee, L.V.; Brønne, I.; Liu, T.; Vater, I.; Rosales, R.; Jamora, R.D.; Pasco, P.M.; Paz, E.M.C.-D.; et al. New insights into the genetics of X-linked dystonia-parkinsonism (XDP, DYT3). *Eur. J. Hum. Genet.* **2015**, *23*, 1334–1340. [[CrossRef](#)] [[PubMed](#)]
26. 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](#)] [[PubMed](#)]
27. Nolte, D.; Niemann, S.; Müller, U. Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10347–10352. [[CrossRef](#)] [[PubMed](#)]
28. 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.e21. [[CrossRef](#)] [[PubMed](#)]
29. Rakovic, A.; Domingo, A.; Grütz, K.; Kulikovskaja, L.; Capetian, P.; Cowley, S.A.; Lenz, I.; Brüggemann, N.; Rosales, R.; Jamora, D.; et al. Genome Editing in Induced Pluripotent Stem Cells Rescues TAF1 Levels in X-Linked Dystonia-Parkinsonism. *Mov. Disord.* **2018**, *33*, 1108–1118. [[CrossRef](#)] [[PubMed](#)]
30. 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 Medica* **2017**, *108*, 438–447. [[CrossRef](#)]
31. Book, A.; Guella, I.; Candido, T.; Brice, A.; Hattori, N.; Jeon, B.; Farrer, M.J. SNCA Multiplication Investigators of the GEPD Consortium A Meta-Analysis of  $\alpha$ -Synuclein Multiplication in Familial Parkinsonism. *Front. Neurol.* **2018**, *9*, 1021. [[CrossRef](#)]
32. Kent, W.J.; Sugnet, C.W.; Furey, T.S.; Roskin, K.M.; Pringle, T.H.; Zahler, A.M.; Haussler, D. The Human Genome Browser at UCSC. *Genome Res.* **2002**, *12*, 996–1006. [[CrossRef](#)] [[PubMed](#)]
33. Polymeropoulos, M.H.; Lavedan, C.; Leroy, E.; Ide, S.E.; Dehejia, A.; Dutra, A.; Pike, B.; Root, H.; Rubenstein, J.; Boyer, R. Mutation in the  $\alpha$ -Synuclein Gene Identified in Families with Parkinson’s Disease. *Science* **1997**, *276*, 2045–2047. [[CrossRef](#)] [[PubMed](#)]

34. Appel-Cresswell, S.; Vilarino-Guell, C.; Encarnacion, M.; Sherman, H.; Yu, I.; Shah, B.; Weir, D.; Thompson, C.; Szu-Tu, C.; Trinh, J.; et al. Alpha-Synuclein p.H50Q, a Novel Pathogenic Mutation for Parkinson's Disease. *Mov. Disord.* **2013**, *28*, 811–813. [[CrossRef](#)] [[PubMed](#)]
35. Krüger, R.; Kuhn, W.; Müller, T.; Woitalla, D.; Graeber, M.; Kösel, S.; Przuntek, H.; Epplen, J.T.; Schöls, L.; Riess, O. AlaSOPro Mutation in the Gene Encoding  $\alpha$ -Synuclein in Parkinson's Disease. *Nat. Genet.* **1998**, *18*, 106–108. [[CrossRef](#)] [[PubMed](#)]
36. Lesage, S.; Anheim, M.; Letournel, F.; Bousset, L.; Honoré, A.; Rozas, N.; Pieri, L.; Mадiona, K.; Dürr, A.; Melki, R.; et al. G51D  $\alpha$ -Synuclein Mutation Causes a Novel Parkinsonian-Pyramidal Syndrome. *Ann. Neurol.* **2013**, *73*, 459–471. [[CrossRef](#)] [[PubMed](#)]
37. Zarranz, J.J.; Alegre, J.; Gómez-Esteban, J.C.; Lezcano, E.; Ros, R.; Ampuero, I.; Vidal, L.; Hoenicka, J.; Rodriguez, O.; Atarés, B.; et al. The New Mutation, E46K, of  $\alpha$ -Synuclein Causes Parkinson and Lewy Body Dementia. *Ann. Neurol.* **2004**, *55*, 164–173. [[CrossRef](#)] [[PubMed](#)]
38. 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](#)]
39. Ahn, T.B.; Kim, S.Y.; Kim, J.Y.; Park, S.S.; Lee, D.S.; Min, H.J.; Kim, Y.K.; Kim, S.E.; Kim, J.M.; Kim, H.J.; et al.  $\alpha$ -Synuclein Gene Duplication Is Present in Sporadic Parkinson Disease. *Neurology* **2008**, *70*, 43–49. [[CrossRef](#)] [[PubMed](#)]
40. 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](#)] [[PubMed](#)]
41. 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. *Parkinson's Dis.* **2015**, *2015*, 546462. [[CrossRef](#)] [[PubMed](#)]
42. Fuchs, J.; Nilsson, C.; Kachergus, J.; Munz, M.; Larsson, E.-M.; Schule, 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](#)] [[PubMed](#)]
43. Keyser, R.J.; Lombard, D.; Veikondis, R.; Carr, J.; Bardiën, S. Analysis of exon dosage using MLPA in South African Parkinson's disease patients. *Neurogenetics* **2009**, *11*, 305–312. [[CrossRef](#)] [[PubMed](#)]
44. 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](#)]
45. Olgiati, S.; Thomas, A.; Quadri, M.; Breedveld, G.J.; Graafland, J.; Eussen, H.; Douben, H.; de Klein, A.; Onofrij, M.; Bonifati, V. Early-onset parkinsonism caused by alpha-synuclein gene triplication: Clinical and genetic findings in a novel family. *Park. Relat. Disord.* **2015**, *21*, 981–986. [[CrossRef](#)]
46. 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](#)] [[PubMed](#)]
47. 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](#)]
48. Devine, M.J.; Ryten, M.; Vodicka, P.; Thomson, A.J.; Burdon, T.; Houlden, H.; Cavaleri, F.; Nagano, M.; Drummond, N.J.; Taanman, J.-W.; et al. Parkinson's disease induced pluripotent stem cells with triplication of the  $\alpha$ -synuclein locus. *Nat. Commun.* **2011**, *2*, 440. [[CrossRef](#)] [[PubMed](#)]
49. Miller, D.W.; Hague, S.M.; Clarimon, J.; Baptista, M.; Gwinn-Hardy, K.; Cookson, M.R.; Singleton, A.B.  $\alpha$ -Synuclein in Blood and Brain from Familial Parkinson Disease with SNCA Locus Triplication. *Neurology* **2004**, *62*, 1835–1838. [[CrossRef](#)] [[PubMed](#)]
50. 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  $\alpha$ -Synuclein and Are Susceptible to Oxidative Stress. *PLoS ONE* **2011**, *6*, e26159. [[CrossRef](#)]
51. Iannielli, A.; Luoni, M.; Giannelli, S.G.; Ferese, R.; Ordazzo, G.; Fossati, M.; Raimondi, A.; Opazo, F.; Corti, O.; Prehn, J.H.M.; et al. Modeling native and seeded Synuclein aggregation and related cellular dysfunctions in dopaminergic neurons derived by a new set of isogenic iPSC lines with SNCA multiplications. *Cell Death Dis.* **2022**, *13*, 881. [[CrossRef](#)] [[PubMed](#)]
52. Burbulla, L.F.; Song, P.; Mazzulli, J.R.; Zampese, E.; Wong, Y.C.; Jeon, S.; Santos, D.P.; Blanz, J.; Obermaier, C.D.; Strojny, C.; et al. Dopamine oxidation mediates mitochondrial and lysosomal dysfunction in Parkinson's disease. *Science* **2017**, *357*, 1255–1261. [[CrossRef](#)] [[PubMed](#)]
53. Oliveira, L.M.A.; Falomir-Lockhart, L.J.; Botelho, M.G.; Lin, K.-H.; Wales, P.; Koch, J.C.; Gerhardt, E.; Taschenberger, H.; Outeiro, T.F.; Lingor, P.; et al. Elevated  $\alpha$ -synuclein caused by SNCA gene triplication impairs neuronal differentiation and maturation in Parkinson's patient-derived induced pluripotent stem cells. *Cell Death Dis.* **2015**, *6*, e1994. [[CrossRef](#)] [[PubMed](#)]
54. Konno, T.; Ross, O.A.; Puschmann, A.; Dickson, D.W.; Wszolek, Z.K. Autosomal dominant Parkinson's disease caused by SNCA duplications. *Park. Relat. Disord.* **2015**, *22*, S1–S6. [[CrossRef](#)] [[PubMed](#)]
55. 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](#)] [[PubMed](#)]
56. Kalinderi, K.; Bostantjopoulou, S.; Fidani, L. The genetic background of Parkinson's disease: Current progress and future prospects. *Acta Neurol. Scand.* **2016**, *134*, 314–326. [[CrossRef](#)] [[PubMed](#)]

57. Grünewald, A.; Voges, L.; Rakovic, A.; Kasten, M.; Vandebona, H.; Hemmelmann, C.; Lohmann, K.; Orolicki, S.; Ramirez, A.; Schapira, A.H.V.; et al. Mutant Parkin Impairs Mitochondrial Function and Morphology in Human Fibroblasts. *PLoS ONE* **2010**, *5*, e12962. [[CrossRef](#)] [[PubMed](#)]
58. Müftüoğlu, M.; Elibol, B.; Dalmazrak, Ö.; Ercan, A.; Kulaksız, G.; Ögüs, H.; Dalkara, T.; Özer, N. Mitochondrial complex I and IV activities in leukocytes from patients with parkin mutations. *Mov. Disord.* **2003**, *19*, 544–548. [[CrossRef](#)] [[PubMed](#)]
59. Goldberg, M.S.; Fleming, S.M.; Palacino, J.J.; Cepeda, C.; Lam, H.A.; Bhatnagar, A.; Meloni, E.G.; Wu, N.; Ackerson, L.C.; Klapstein, G.J.; et al. Parkin-Deficient Mice Exhibit Nigrostriatal Deficits but Not Loss of Dopaminergic Neurons. *J. Biol. Chem.* **2003**, *278*, 43628–43635. [[CrossRef](#)] [[PubMed](#)]
60. Itier, J.-M.; Ibáñez, P.; Mena, M.A.; Abbas, N.; Cohen-Salmon, C.; Bohme, G.A.; Laville, M.; Pratt, J.; Corti, O.; Pradier, L.; et al. Parkin gene inactivation alters behaviour and dopamine neurotransmission in the mouse. *Hum. Mol. Genet.* **2003**, *12*, 2277–2291. [[CrossRef](#)] [[PubMed](#)]
61. Palacino, J.J.; Sagi, D.; Goldberg, M.S.; Krauss, S.; Motz, C.; Wacker, M.; Klose, J.; Shen, J. Mitochondrial Dysfunction and Oxidative Damage in Parkin-Deficient Mice. *J. Biol. Chem.* **2004**, *279*, 18614–18622. [[CrossRef](#)] [[PubMed](#)]
62. Hayashi, S.; Wakabayashi, K.; Ishikawa, A.; Nagai, H.; Saito, M.; Maruyama, M.; Takahashi, T.; Ozawa, T.; Tsuji, S.; Takahashi, H. An Autopsy Case of Autosomal-Recessive Juvenile Parkinsonism with a Homozygous Exon 4 Deletion in the Parkin Gene. *Mov. Disord.* **2000**, *15*, 884–888. [[CrossRef](#)]
63. Miyakawa, S.; Ogino, M.; Funabe, S.; Uchino, A.; Shimo, Y.; Hattori, N.; Ichinoe, M.; Mikami, T.; Saegusa, M.; Nishiyama, K.; et al. Lewy Body Pathology in a Patient with a Homozygous Parkin Deletion. *Mov. Disord.* **2013**, *28*, 388–391. [[CrossRef](#)] [[PubMed](#)]
64. Mori, H.; Kondo, T.; Yokochi, M.; Matsumine, H.; Nakagawa-Hattori, Y.; Miyake, T.; Suda, K.; Mizuno, Y. Pathologic and biochemical studies of juvenile parkinsonism linked to chromosome 6q. *Neurology* **1998**, *51*, 890–892. [[CrossRef](#)] [[PubMed](#)]
65. Sasaki, S.; Shirata, A.; Yamane, K.; Iwata, M. Parkin-Positive Autosomal Recessive Juvenile Parkinsonism with  $\alpha$ -Synuclein-Positive Inclusions. *Neurology* **2004**, *63*, 678–682. [[CrossRef](#)] [[PubMed](#)]
66. Yamamura, Y.; Kuzuhara, S.; Kondo, K.; Yanagi, T.; Uchida, M.; Matsumine, H.; Mizuno, Y. Clinical, pathologic and genetic studies on autosomal recessive early-onset parkinsonism with diurnal fluctuation. *Park. Relat. Disord.* **1998**, *4*, 65–72. [[CrossRef](#)] [[PubMed](#)]
67. Ahmad, A.; Nkosi, D.; Iqbal, M.A. PARK2 Microdeletion or Duplications Have Been Implicated in Different Neurological Disorders Including Early Onset Parkinson Disease. *Genes* **2023**, *14*, 600. [[CrossRef](#)] [[PubMed](#)]
68. Jarick, I.; Volckmar, A.L.; Pütter, C.; Pechlivanis, S.; Nguyen, T.T.; Dauvermann, M.R.; Beck, S.; Albayrak, Ö.; Scherag, S.; Gilsbach, S.; et al. Genome-Wide Analysis of Rare Copy Number Variations Reveals PARK2 as a Candidate Gene for Attention-Deficit/hyperactivity Disorder. *Mol. Psychiatry* **2014**, *19*, 115–121. [[CrossRef](#)]
69. Xu, L.; Lin, D.C.; Yin, D.; Koeffler, H.P. An Emerging Role of PARK2 in Cancer. *J. Mol. Med.* **2013**, *92*, 31–42. [[CrossRef](#)]
70. Glessner, J.T.; Wang, K.; Cai, G.; Korvatska, O.; Kim, C.E.; Wood, S.; Zhang, H.; Estes, A.; Brune, C.W.; Bradfield, J.P.; et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **2009**, *459*, 569–573. [[CrossRef](#)] [[PubMed](#)]
71. Yin, C.-L.; Chen, H.-I.; Li, L.-H.; Chien, Y.-L.; Liao, H.-M.; Chou, M.C.; Chou, W.-J.; Tsai, W.-C.; Chiu, Y.-N.; Wu, Y.-Y.; et al. Genome-wide analysis of copy number variations identifies PARK2 as a candidate gene for autism spectrum disorder. *Mol. Autism* **2016**, *7*, 23. [[CrossRef](#)] [[PubMed](#)]
72. Lesage, S.; Lunati, A.; Houot, M.; Ben Romdhan, S.; Clot, F.; Tesson, C.; Mangone, G.; Le Toullec, B.; Courtin, T.; Larcher, K.; et al. Characterization of Recessive Parkinson Disease in a Large Multicenter Study. *Ann. Neurol.* **2020**, *88*, 843–850. [[CrossRef](#)] [[PubMed](#)]
73. Kasten, M.; Hartmann, C.; Hampf, J.; Schaake, S.; Westenberger, A.; Vollstedt, E.J.; Balck, A.; Domingo, A.; Vulinovic, F.; Dulovic, M.; et al. Genotype-Phenotype Relations for the Parkinson’s Disease Genes Parkin, PINK1, DJ1: MDSGene Systematic Review. *Mov. Disord.* **2018**, *33*, 730–741. [[CrossRef](#)] [[PubMed](#)]
74. Denison, S.R.; Callahan, G.; Becker, N.A.; Phillips, L.A.; Smith, D.I. Characterization of FRA6E and its potential role in autosomal recessive juvenile parkinsonism and ovarian cancer. *Genes Chromosom. Cancer* **2003**, *38*, 40–52. [[CrossRef](#)] [[PubMed](#)]
75. Kilarski, L.L.; Pearson, J.P.; Newsway, V.; Majounie, E.; Knipe MD, W.; Misbahuddin, A.; Chinnery, P.F.; Burn, D.J.; Clarke, C.E.; Marion, M.-H.; et al. Systematic Review and UK-Based Study of PARK2 (parkin), PINK1, PARK7 (DJ-1) and LRRK2 in Early-Onset Parkinson’s Disease. *Mov. Disord.* **2012**, *27*, 1522–1529. [[CrossRef](#)] [[PubMed](#)]
76. Zhu, W.; Huang, X.; Yoon, E.; Bandres-Ciga, S.; Blauwendraat, C.; Billingsley, K.J.; Cade, J.H.; Wu, B.P.; Williams, V.H.; Schindler, A.B.; et al. Heterozygous PRKN Mutations Are Common but Do Not Increase the Risk of Parkinson’s Disease. *Brain* **2022**, *145*, 2077–2091. [[CrossRef](#)] [[PubMed](#)]
77. Kay, D.M.; Stevens, C.F.; Hamza, T.H.; Montimurro, J.S.; Zabetian, C.P.; Factor, S.A.; Samii, A.; Griffith, A.; Roberts, J.W.; Molho, E.S.; et al. A Comprehensive Analysis of Deletions, Multiplications, and Copy Number Variations in PARK2. *Neurology* **2010**, *75*, 1189. [[CrossRef](#)] [[PubMed](#)]
78. Huttenlocher, J.; Stefansson, H.; Steinberg, S.; Helgadóttir, H.T.; Sveinbjörnsdóttir, S.; Riess, O.; Bauer, P.; Stefansson, K. Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson’s disease. *Hum. Mol. Genet.* **2015**, *24*, 5637–5643. [[CrossRef](#)] [[PubMed](#)]
79. Daida, K.; Funayama, M.; Billingsley, K.J.; Malik, L.; Miano-Burkhardt, A.; Leonard, H.L.; Makarious, M.B.; Iwaki, H.; Ding, J.; Gibbs, J.R.; et al. Long-Read Sequencing Resolves a Complex Structural Variant in PRKN Parkinson’s Disease. *medRxiv* **2023**. [[CrossRef](#)] [[PubMed](#)]

80. Bonifati, V.; Rizzu, P.; van Baren, M.J.; Schaap, O.; Breedveld, G.J.; Krieger, E.; Dekker, M.C.J.; Squitieri, F.; Ibanez, P.; Joosse, M.; et al. Mutations in the *DJ-1* Gene Associated with Autosomal Recessive Early-Onset Parkinsonism. *Science* **2003**, *299*, 256–259. [[CrossRef](#)] [[PubMed](#)]
81. Abou-Sleiman, P.M.; Healy, D.G.; Quinn, N.; Lees, A.J.; Wood, N.W. The Role of Pathogenic DJ-1 Mutations in Parkinson's Disease. *Ann. Neurol.* **2003**, *54*, 283–286. [[CrossRef](#)] [[PubMed](#)]
82. Clark, L.N.; Afridi, S.; Mejia-Santana, H.; Harris, J.; Louis, E.D.; Cote, L.J.; Andrews, H.; Singleton, A.; Wavrant De-Vrieze, F.; Hardy, J.; et al. Analysis of an Early-Onset Parkinson's Disease Cohort for DJ-1 Mutations. *Mov. Disord.* **2004**, *19*, 796–800. [[CrossRef](#)] [[PubMed](#)]
83. Darvish, H.; Movafagh, A.; Omrani, M.D.; Firouzabadi, S.G.; Azargashb, E.; Jamshidi, J.; Khaligh, A.; Haghnejad, L.; Naeini, N.S.; Talebi, A.; et al. Detection of copy number changes in genes associated with Parkinson's disease in Iranian patients. *Neurosci. Lett.* **2013**, *551*, 75–78. [[CrossRef](#)] [[PubMed](#)]
84. Güler, S.; Gül, T.; Güler, Ş.; Haerle, M.C.; Başak, A.N. Early-Onset Parkinson's Disease: A Novel Deletion Comprising the DJ-1 and TNFRSF9 Genes. *Mov. Disord.* **2021**, *36*, 2973–2976. [[CrossRef](#)] [[PubMed](#)]
85. Ross, G.; Dalvi, S.P.; Dalvi, P.S. DJ-1 and Parkinson's Disease. *Brain Disord. Ther.* **2021**, *3*, 100020.
86. Valente, E.M.; Abou-Sleiman, P.M.; Caputo, V.; Muqit, M.M.K.; Harvey, K.; Gispert, S.; Ali, Z.; Del Turco, D.; Bentivoglio, A.R.; Healy, D.G.; et al. Hereditary Early-Onset Parkinson's Disease Caused by Mutations in PINK1. *Science* **2004**, *304*, 1158–1160. [[CrossRef](#)] [[PubMed](#)]
87. Nuytemans, K.; Theuns, J.; Cruys, M.; Van Broeckhoven, C. Genetic Etiology of Parkinson Disease Associated with Mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 Genes: A Mutation Update. *Hum. Mutat.* **2010**, *31*, 763–780. [[CrossRef](#)] [[PubMed](#)]
88. Li, Y.; Tomiyama, H.; Sato, K.; Hatano, Y.; Yoshino, H.; Atsumi, M.; Kitaguchi, M.; Sasaki, S.; Kawaguchi, S.; Miyajima, H.; et al. Clinicogenetic study of PINK1 mutations in autosomal recessive early-onset parkinsonism. *Neurology* **2005**, *64*, 1955–1957. [[CrossRef](#)] [[PubMed](#)]
89. Marongiu, R.; Brancati, F.; Antonini, A.; Ialongo, T.; Ceccarini, C.; Scarciolla, O.; Capalbo, A.; Benti, R.; Pezzoli, G.; Dallapiccola, B.; et al. Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum. *Hum. Mutat.* **2006**, *28*, 98. [[CrossRef](#)]
90. Cazeneuve, C.; Sãn, C.; Ibrahim, S.A.; Mukhtar, M.M.; Kheir, M.M.; LeGuern, E.; Brice, A.; Salih, M.A. A new complex homozygous large rearrangement of the PINK1 gene in a Sudanese family with early onset Parkinson's disease. *Neurogenetics* **2009**, *10*, 265–270. [[CrossRef](#)] [[PubMed](#)]
91. Vizziello, M.; Borellini, L.; Franco, G.; Ardolino, G. Disruption of Mitochondrial Homeostasis: The Role of PINK1 in Parkinson's Disease. *Cells* **2021**, *10*, 3022. [[CrossRef](#)]
92. Yang, W.; Li, S.; Li, X.-J. A CRISPR monkey model unravels a unique function of PINK1 in primate brains. *Mol. Neurodegener.* **2019**, *14*, 17. [[CrossRef](#)] [[PubMed](#)]
93. Yang, W.; Liu, Y.; Tu, Z.; Xiao, C.; Yan, S.; Ma, X.; Guo, X.; Chen, X.; Yin, P.; Yang, Z.; et al. CRISPR/Cas9-mediated PINK1 deletion leads to neurodegeneration in rhesus monkeys. *Cell Res.* **2019**, *29*, 334–336. [[CrossRef](#)] [[PubMed](#)]
94. Botto, L.D.; May, K.; Fernhoff, P.M.; Correa, A.; Coleman, K.; Rasmussen, S.A.; Merritt, R.K.; O'leary, L.A.; Wong, L.-Y.; Elixson, E.M.; et al. A Population-Based Study of the 22q11.2 Deletion: Phenotype, Incidence, and Contribution to Major Birth Defects in the Population. *Pediatrics* **2003**, *112*, 101–107. [[CrossRef](#)] [[PubMed](#)]
95. Devriendt, K.; Fryns, J.P.; Mortier, G.; van Thienen, M.N.; Keymolen, K. The annual incidence of DiGeorge/velocardiofacial syndrome. *J. Med. Genet.* **1998**, *35*, 789–790. [[CrossRef](#)] [[PubMed](#)]
96. Goodship, J.; Cross, I.; LiLing, J.; Wren, C. A population study of chromosome 22q11 deletions in infancy. *Arch. Dis. Child.* **1998**, *79*, 348–351. [[CrossRef](#)]
97. Scambler, P.J. The 22q11 Deletion Syndromes. *Hum. Mol. Genet.* **2000**, *9*, 2421–2426. [[CrossRef](#)] [[PubMed](#)]
98. Hacıhamdioğlu, B.; Hacıhamdioğlu, D.; Delil, K. 22q11 Deletion Syndrome: Current Perspective. *Appl. Clin. Genet.* **2015**, *8*, 123–132. [[CrossRef](#)] [[PubMed](#)]
99. McDonald-McGinn, D.M. (Ed.) *The Chromosome 22q11.2 Deletion Syndrome: A Multidisciplinary Approach to Diagnosis and Treatment*; Academic Press: Cambridge, MA, USA, 2022.
100. Ogaki, K.; Ross, O.A. Chromosome 22q11.2 Deletion May Contain a Locus for Recessive Early-Onset Parkinson's Disease. *Park. Relat. Disord.* **2014**, *20*, 945–946. [[CrossRef](#)] [[PubMed](#)]
101. 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](#)]
102. 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](#)] [[PubMed](#)]
103. Al Khleifat, A.; Iacoangeli, A.; van Vugt, J.J.F.A.; Bowles, H.; Moisse, M.; Zwamborn, R.A.J.; van der Spek, R.A.A.; Shatunov, A.; Cooper-Knock, J.; Topp, S.; et al. Structural variation analysis of 6500 whole genome sequences in amyotrophic lateral sclerosis. *npj Genom. Med.* **2022**, *7*, 8. [[CrossRef](#)] [[PubMed](#)]
104. Wang, H.; Dombroski, B.A.; Cheng, P.L.; Tucci, A.; Si, Y.Q.; Farrell, J.J.; Tzeng, J.-Y.; Leung, Y.Y.; Malamon, J.S.; Wang, L.-S.; et al. Structural Variation Detection and Association Analysis of Whole-Genome-Sequence Data from 16,905 Alzheimer's Diseases Sequencing Project Subjects. *medRxiv* **2023**. [[CrossRef](#)] [[PubMed](#)]

105. GitHub-Broadinstitute/gatk-Sv: A Structural Variation Pipeline for Short-Read Sequencing. n.d. GitHub. Available online: <https://github.com/broadinstitute/gatk-sv> (accessed on 21 January 2024).
106. 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](#)] [[PubMed](#)]
107. Zhao, X.; Collins, R.L.; Lee, W.-P.; Weber, A.M.; Jun, Y.; Zhu, Q.; Weisburd, B.; Huang, Y.; Audano, P.A.; Wang, H.; et al. Expectations and blind spots for structural variation detection from long-read assemblies and short-read genome sequencing technologies. *Am. J. Hum. Genet.* **2021**, *108*, 919–928. [[CrossRef](#)] [[PubMed](#)]
108. Leonard, H.; Jonson, C.; Levine, K.; Lake, J.; Hertslet, L.; Jones, L.; Patel, D.; Kim, J.; Bandres-Ciga, S.; Terry, N.; et al. Assessing the Lack of Diversity in Genetics Research across Neurodegenerative Diseases: A Systematic Review of the GWAS Catalog and Literature. *medRxiv* **2024**. [[CrossRef](#)] [[PubMed](#)]
109. Mahmoud, M.; Huang, Y.; Garimella, K.; Audano, P.A.; Wan, W.; Prasad, N.; Handsaker, R.E.; Hall, S.; Pionzio, A.; Schatz, M.C.; et al. Utility of Long-Read Sequencing for All of Us. *bioRxiv* **2023**. [[CrossRef](#)]
110. Long-Read Sequencing. n.d. Available online: <https://card.nih.gov/research-programs/long-read-sequencing> (accessed on 4 February 2024).
111. Website. n.d. Available online: <https://www.genomeweb.com/sequencing/nih-study-demonstrates-nanopore-sequencing-population-scale-genomics> (accessed on 21 January 2024).
112. Global Parkinson’s Genetics Program. GP2: The Global Parkinson’s Genetics Program. *Mov. Disord.* **2021**, *36*, 842–851. [[CrossRef](#)] [[PubMed](#)]
113. Blythe, L. 2020. Understanding the Genetic Architecture of Parkinson’s Disease. GP2. Available online: <https://gp2.org/> (accessed on 22 December 2023).
114. Mokretar, K.; Pease, D.; Taanman, J.-W.; Soenmez, A.; Ejaz, A.; Lashley, T.; Ling, H.; Gentleman, S.; Houlden, H.; Holton, J.L.; et al. Somatic copy number gains of  $\alpha$ -synuclein (SNCA) in Parkinson’s disease and multiple system atrophy brains. *Brain* **2018**, *141*, 2419–2431. [[CrossRef](#)] [[PubMed](#)]
115. Perez-Rodriguez, D.; Kalyva, M.; Leija-Salazar, M.; Lashley, T.; Tarabichi, M.; Chelban, V.; Gentleman, S.; Schottlaender, L.; Franklin, H.; Vasmatazis, G.; et al. Investigation of somatic CNVs in brains of synucleinopathy cases using targeted SNCA analysis and single cell sequencing. *Acta Neuropathol. Commun.* **2019**, *7*, 219. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.