Research Article ISSN 2689-1077

Genetics & Molecular Medicine

Challenges in Clinico-Genetic Correlations in Parkinson's disease (PD). The Role of Copy Number Variants (CNV)

Gatto Emilia Mabel MD, FAAN[1]*, Radrizzani Martin PhD[2], González Rojas Natalia MD[1], Cesarini Martin Emiliano MD[1], Etcheverry José Luis MD[1], Perandones Claudia MD, MSc, PhD[3]

[1]*Instituto de Neurociencias Buenos Aires. INEBA.*

[2]*School of Science and Technology, National University of San Martin. Buenos Aires, Argentina and Institute of Emerging Technologies and Applied Sciences - CONICET.*

[3]*National Administration of Laboratories and Institutes of Health, ANLIS, Dr. Carlos G. Malbrán. Buenos Aires, Argentina.*

***Correspondence:**
Emilia Mabel Gatto, Instituto de Neurociencias Buenos Aires. INEBA, Área de Enfermedad de Parkinson y Movimientos Anormales, Guardia Vieja 4435, CABA, Buenos Aires, Argentina, Tel: (+54) 1148677700.

Received: 17 July 2021; **Accepted:** 12 August 2021

Citation: Gatto EM, Radrizzani M, González Rojas N, et al. Challenges in Clinico-Genetic Correlations in Parkinson's disease (PD). The Role of Copy Number Variants (CNV). Genet Mol Med. 2021; 3(2): 1-10.

ABSTRACT

Parkinson's disease (PD) represents the second most common neurodegenerative disease and remains incurable. Mutations in multiple genes have been linked to monogenic PD (gPD); these monogenic forms, however, represent a small number of cases while in most instances PD appears as idiopathic (iPD). These findings raise the question of whether genetic and idiopathic parkinsonisms constitute the same disease. Nevertheless, monogenic-PD phenotypes and iPD both fulfill MDS criteria for PD, and show evidence of alpha-synuclein aggregates in both conditions.

Distinct genetic loci in rare Mendelian forms have been identified as causal mutations, others as possible disease-causing genes, and genome-wide association studies have reported several risk loci, many of them located in the genes associated with the dominant mutations.

Not only single-nucleotide polymorphisms (SNPs), but other kinds of DNA molecular defects as well have been spotted as significant disease-causing mutations, including large chromosomal structural rearrangements and copy number variations (CNVs). As their size varies, and detection methodologies have different sensitivity and resolution, CNVs pose a special challenge in genetic studies, and there currently is a debate on the pathogenetic or susceptibility impact of specific CNVs on PD.

In this review, through multiple instances of experimental evidence, we analyze the impact on histopathology of the different mutational mechanisms involved in the genesis and etiology of PD. We believe that increasing our knowledge about the changes and implications at tissue level produced by each of those mechanisms will allow to develop much more suitable and personalized potential therapeutic strategies, biomarker identification, as well as disease modeling, agreeing with the precision medicine concept.

Keywords

Parkinson Disease, Genetics, Copy Number Variant, Mosaicism.

Introduction

Genetics of Parkinson's disease

With over 7 million patients worldwide, Parkinson's disease (PD) is the most common neurodegenerative movement disorder and it is becoming more prevalent as life span and industrialization increase. Clinically, it is characterized by tremor, rigidity, bradykinesia/akinesia and postural instability, with a large number of non-motor symptoms (NMSs) involving multiple systems including behavioral, psychiatric, mood, and cognitive disorders, which make PD a systemic disease [1]. To date, the pathological hallmarks remain the dopaminergic neuronal loss and

intra-cytoplasmic alpha-synuclein aggregation [1]. However, in some monogenic familial PD (LRRK2), other proteins (MAPT) than alpha-synuclein protein aggregation may occur, raising the question of whether genetic parkinsonisms (gPD) constitute distinct diseases from idiopathic PD (iPD) [2].

The etiology of PD remains incompletely known. However, aging, genetic and environmental factors, and lifestyle play a complex interaction in the etiopathogenic mechanism of PD (Figure 1) [3]. The estimated prevalence in the global population over 65 years old is 1%, but this prevalence increases to 3% in patients older than 80 years old. Age at onset distribution has been estimated in 39% for late-onset, 51% for middle-onset and 10% for young-onset PD [4]. As epidemiologic studies have shown, aging is the most relevant risk factor for PD, but environmental factors like exposure to pesticides, manganese, and rural life contribute to increasing the risk. Moreover, some lifestyle aspects including cigarette smoking, coffee and yerba mate drinking, statins use, b2-adrenoreceptor agonists, and nonsteroidal anti-inflammatory drugs have been related with a decreased risk for PD development.

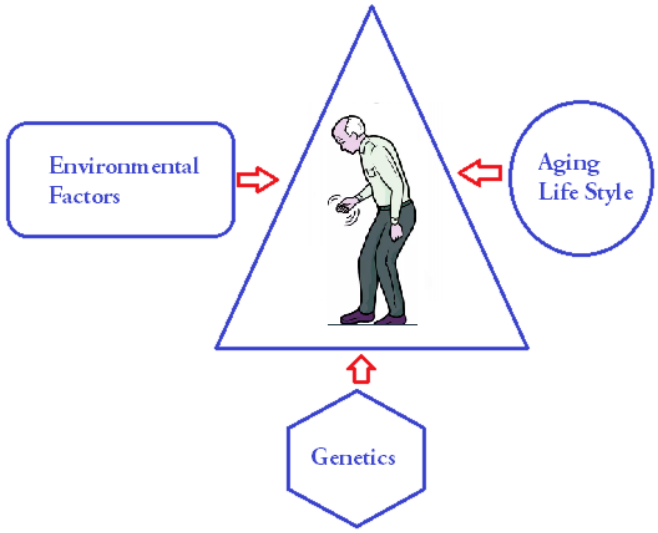

Figure 1: PD etiopathogenic mechanism. Modified by M J Armstrong 2020

Although genetic factors appear as one of the main relevant etiological mechanisms, inherited mutations have been identified in only 5% -10% of all PD cases [3].

After the description and identification of the first causal alpha–synuclein gene (SNCA), codifying α-synuclein protein in the Contursi family in 1997, the genetic contribution has never stopped to increase. At present, more than 19 loci have been described as causal mutations responsible for monogenic familial PD. Among them, autosomal dominant inheritance is related to SNCA (PARK 1/4), LRRK2 (PARK 8), VPS35 (PARK 17), UCHL1 (PARK 5) genes, respectively; while recessive inheritance is caused by mutations on parkin (PRKN-PARK 2), Pink1 (PARK 6), Dj-1 (PARK 7), ATP13A2 (PARK 9), PLA2G6 (PARK 14), FBXO7 (PARK 15) and DNAJC6 (PARK 19A), with a common early onset in general in this last group.

A special mention is required for GBA1 mutations, the gene codifying glucocerebrosidase (GBA) that in homozygosis causes Gaucher disease. However, carriers of GBA1 mutations constitute the population at the most relevant genetic risk for PD, particularly in the Ashkenazi population, with a prevalence of 5-25% of PD [3].

Over the last years, genome-wide association (GWA) studies have allowed to identify 90 genetic loci risk variants, with many of them harboring common risk variants including rare variants of SNCA, LRRK2, or GBA. Interestingly, these rare variants explained 16–36% of the heritable risk of PD [5-7].

In fact, the significant, massive development in genomic biotechnologies have led clinical molecular laboratories to detect novel pathogenic or susceptibility variants that have opened new questions. In these sense, it is required to determine the correlation between one phenotype and one or many genotypes, or the association of multiple phenotypes with a specific genotype.
In 2015, in order to understand the clinical significance of any variant, the American College of Medical Genetics and Genomics established standards and guidelines for variant classification into five categories: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign [8]. Despite this useful tool, the role of these genetic risk variants in the etiopathogenic and pathobiological mechanisms of PD is still to be elucidated.

To date, biotechnology has provided different approaches for genetic testing. Strengths and limitations of each one of them have to be individually evaluated. However, the main objective is to identify a positive causal genetic factor [9].

Among the currently available techniques, we could mention:
1. Single gene targeted variant testing.
2. Single gene sequencing (and/or deletion/duplication analysis).
3. Multigene targeted variant panel.
4. Multigene sequencing panel (and/or deletion/duplication analysis).
5. Whole-exome sequence (WES) or whole-genome sequence (WGS) testing.

In the setting of genetic research in PD, chromosomal abnormalities for dominant and recessive causes have to be included. Although to date single nucleotide variants (SNVs) have appeared as the main causal genetic factor for monogenic PD, copy number variants (CNVs-duplication, triplication or deletion) require particular consideration, even more so if we assume that deletions and duplications comprise 15% of all mutations in monogenic diseases [10,11].

A clear example of the role of the CNVs is duplications and triplication of the SNCA which prompt a dosage effect [3].

CNVs constitute a particular challenge in genetic studies, and the pathogenetic or susceptibility impact of specific CNVs on PD is currently under debate.

The objective of this article is to describe the available evidence about mosaicism and CNVs in PD.

Assessing the global genome-wide burden of large CNVs and elucidating the role of mosaicism of the CNVs on PD may reveal new candidate genes, and consequently improve diagnosis and counseling of mutations carriers.

CNVs: origin, classification, and clinical relevance

Copy number variants (CNVs) are germline or acquired (somatic) variations of the copy number of a given genomic region in comparison to its reference genome. The relevance of CNVs in genomic diversity is reportedly similar to that of single nucleotide polymorphisms (SNPs). Although not uniformly, there is a variety of evidence that definitively links CNVs with certain diseases or phenotypes [12,13].

They can be greater than the reference sequence (i.e. gain or duplication) or less than it (i.e. loss or deletion), and can partially or entirely encompass one or more contiguous genes, finally resulting in an altered DNA diploid status.

Classically, CNVs are DNA segments larger than 1kb by comparison with a referent genome; according to new detection methods, size has dropped to 50 bp, while smaller length fragments are called Indels.

CNVs are very common and occur in DNA regions very susceptible to rearrangements. Depending on whether the same rearrangement is identified in unrelated individuals, CNVs can be grouped as recurrent or non-recurrent events.

In a clinical setting, CNVs are categorized into five groups (according The American College of Medical Genetics and Genomics practice guidelines):
• Abnormal or pathogenic (e.g. well-established association with a disease).
• Likely pathogenic.
• VOUS (Variants of uncertain clinical significance – rare or private CNV).
• Likely benign.
• Benign (a polymorphic variant detected in a normal individual without clinical significance).

CNVs and SNVs can occur during cell division; they may be present in germline (germ cells within the gonads), somatic, or both cells post-zygotically, depending on mutation developmental timing [14].

Somatic and germline mosaicisms have been related with aging, neurodegenerative diseases, and cancer. Classically, germline mosaicism has been related with genetic diseases, while somatic mosaicism has been associated with cancer.

In fact, as we have mentioned before, germline mosaicism describes the presence of different genomes in germ cells of one individual. This concept is different from the germline variation that designates a DNA mutation transmitted through a parent's germline to all somatic and germ cells of a child [14].

Somatic mosaicism is uninheritable. However, discussion is still underway on whether somatic mosaicism correlates with disease recurrence risk factor and if it occurs during preimplantation stage it may be transmitted to the offspring [14,15].

Findings observed by our group in a family with early-onset Parkinson's disease allow to confirm transmission to offspring of SNCA duplication/ triplication mosaicisms. For further details, see Figure 2.

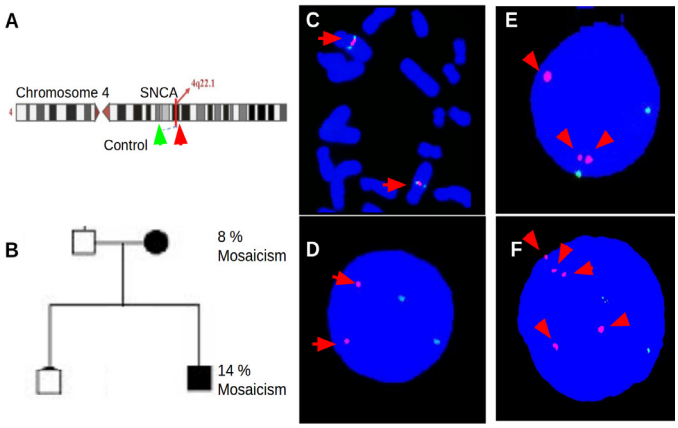

Figure 2: *SNCA familial mosaicism.*

Alpha-synuclein gene (SNCA) mosaicism in familial case: FISH analysis was developed on interphase or metaphase chromosomes with rhodamine-labeled SNCA (BAC RP11-614) or FITC-labeled control AFF1 (BAC RP11-711J3). Genomic DNA was counterstained with DAPI.
A) Location of the SNCA gene in chromosome 4, labeled in red, and the control gene (AFF1) in green.
B) Family tree showing the presence of the mosaicism in peripheral blood.
C) Fluorescence microscope image of chromosomes 22 carrying the AFF1 control gene (green) and the SNCA gene (red) in a normal metaphase of the patient (DAPI, in blue).
D) Image of the interphase showing the AFF1 control gene and the SNCA gene in a normal cell of the unaffected son.
E) Image of an interphase of the mother showing an SNCA duplication.
F) Image of an interphase of the affected son showing an SNCA duplication.

Mosaicism may be classified as general or confined if it is present in at least more than two cells lines, even before differentiation, while confined mosaicism is restricted to one organ [**16]**

In the past, investigations on CNVs have been restricted to the traditional single-gene mutation or gene-centric testing, but facing PD, a complex and heterogeneous molecular disease, this investigation has resulted inadequate. CNVs do not only contribute to monogenic diseases (e.g. SNCA triplication-PARK4), but can also control the phenotype affecting gene and protein expression

by dosage effect or regulatory changes in the transcription factors or via DNA methylation or histone acetylation. In fact, a variety of evidence supports an obligated role for somatic mosaicism, including in typical monogenic disorders [14]. Moreover, CNVs have been associated with non-disease modifications (e.g. genetic cerebral cortex architecture), aging-related disorders, and neurodegeneration. Altogether, determining the clinical significance of CNV alterations is very challenging.

CNVs have emerged as crucial actors involved in biological processes relevant for PD etiopathogenesis. In this scenario, it is necessary to explore underlying factors involved in PD, including new genomic regions (single gene or a contiguous set of genes) beyond the well-known familial PD-genes [17].

For CNVs identification, different strategies could be followed, for instance the whole genome (genome-wide level) or specific chromosome locations (locus-specific levels) [17]. In Table 1, we have included the available laboratory methods targeting specific locations on chromosomes (locus-specific levels), or the whole genome (genome-wide level) to identify CNVs.

Table 1: Laboratory methods to identify CNV.

	Lab. Methods	Resolution
Locus Specific Targeting	PCR-based approaches	~100bp
	FISH Assays	>50 kb
	RFLP- Southern Blot	>1kb
Whole-genome targeting	Karyotyping (G- bandage)	>10Mb
	aCGH	0.06 kb-
	NGS	>1kb

More recently, different tools combining NGS plus variant alleles for the true copy number status and genomic regions from WES data have been developed to improve detection (e.g. CVN Radar, "systems biology") [17,18].

However, these strategies still need post-experimental validations pending for a gold-standard analysis available.

CNVs in PD

Single-gene CNVs in familiar PD-genes

After the initial description of SNCA, not only SNVs, but also CNVs have been described on PRKN with an autosomal recessive pattern of inheritance and a phenotype characterized by early onset of disease (EAO 20 years earlier than iPD).

In patients from different populations (European and Latin American PD patients), a similar increased burden of CNVs overlapping SNCA and more extensively PRKN has been reported associated with PD susceptibility, with CNVs on PRKN in 5.6% of the Latino population [3,7,19]. It is interesting to mention that at the moment no CNVs have been reported for LRRK2 with a PD phenotype [3].

In this section, we will describe the CNVs in familial PD genes.

SNCA

In 1997, the identification of mutation in SNCA, the gene codifying alpha synuclein (aSyn), was the milestone in the knowledge about PD. The SNCA gene is located at 4q22.1. aSyn is a cytosolic,14 kDa protein expressed in the presynaptic terminals [20]. To date, five pathogenic disease-causing mutations have been identified (A30P, E46K, G51D, A53E, and A53T) [21]. The critical role of aSyn in PD pathogenesis has been largely provided, and currently aggregated aSyn is a key feature in the neurodegenerative pathway not only in PD but also in other synucleinopathies (MSA, LBD).

However, its aberrant accumulation is also observed as a *concomitant pathology* in tauopathies, TDP-43 proteinopathies, prionopathies, and amyloid-β (Aβ) deposition [20].

However, in synucleinopathies, aSyn accumulation occurs at different levels (neuronal or glial cytoplasm), with different points of aSyn phosphorylation, as well as different seeding capacity in recipient cells or tissues. This allows us to recognize conformational and functional differences among synucleinopathies [20,22-24].

On the other hand, triplications and duplications of SNCA are the best example of CNVs and their impact on mRNA expression and proteins causing a more severe phenotype in those individuals with triplication.

These multiplications are variable in extension, ranging from 41.2Mb to 0.2 Mb; however, size does not seem to significantly influence clinical manifestations [17].

Moreover, a recent population study analyzing the UK Biobank concluded that rare SNCA CNVs and mosaicism occur in the general population without PD symptoms. Nevertheless, some of these individuals (7/18) reported a history of blood-based cancer [21].

In this sense, a mosaicism condition (it will be discussed extensively later) is one of the main factors favoring SCNA rearrangements [17]. Perandones et al., (2014), highlighted two interesting PD cases with a parkinsonian phenotype in whom CNVs in peripheral blood were negative, while positive in oral mucosa cells [25].

The findings included in the abovementioned report suggest that, before excluding the involvement of SNCA rearrangements in early-onset PD with severe autonomic and early cognitive decline, the spectrum of evaluations should be extended to include more sensitive FISH analyses and immune-histochemical studies, as well as subjecting cells of ectodermal origin (such as those of the oral mucosa) to analysis (for further details, see Figure 3).

Mosaicism of SNCA rearrangements

Genetic variations arise and accumulate in normal cells of the human soma during development or ageing, constituting somatic mosaicism. Although mosaicism occurs in health conditions, a role in neurodevelopmental and neurodegenerative conditions is increasingly recognized (Huntington's disease, Parkinson's disease). Moreover, post-zygotic variation appears as a relevant confounder in genetic testing and opens new avenues for research [26].

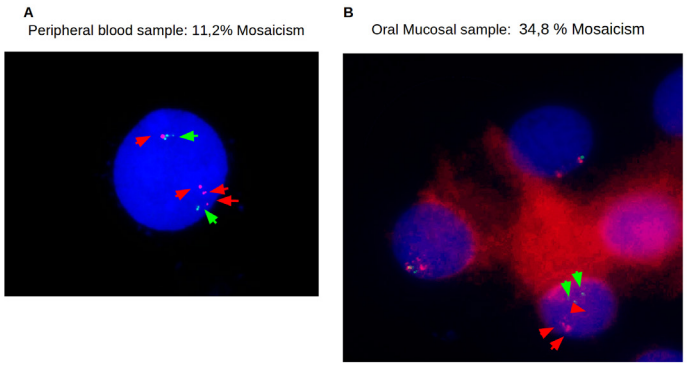

Figure 3: *Discrepancy between the findings in peripheral blood samples and buccal mucosa for SNCA mosaicisms.*
SNCA gene probes were used, labeled in red, and the control gene (AFF1) in green.
A) Fluorescence microscope image of the AFF1 control gene (green) and the SNCA gene (red) in an interphase with SNCA duplication (DAPI, in blue) from peripheral blood.
B) Image of the interphase showing the AFF1 control gene and the SNCA gene with amplifications in buccal mucosa cells.

Somatic mutations include CNVs, SNVs, structural variants, and transposable insertions. As we have mentioned before, somatic mutations may occur in synucleinopathies, such as duplications, and triplications in SNCA as well as in PRKN. Some of these mutations could be detected in neuroectodermal cells, but not in other cells or tissues.

These findings may help us explain why, in sporadic PD, SNCA and other gene mutations are rarely detected in peripheral blood cells.

Moreover, in synucleinopathies, mosaicism is observed in specific brain regions such as at the cingulate gyrus, and some CNVs correlated with age of death in PD patients.

Additionally, mosaicism due to SNCA CNVs in dopaminergic neurons at the substantia nigra was detected in MSA, supporting a role in aetiology and pathogenesis even in sporadic synucleinopathies [27].

Techniques for the evaluation of mosaicism of CNVs

Although CNVs are a more frequent genetic finding than SNVs, their association with clinical disease has been only recently highlighted, because of the limitations for detection of the initial techniques such as karyotyping. Currently, new techniques, with different precision and resolution, have increased detection resolution. These new techniques include fluorescence *in situ* hybridization (FISH) [28], polymerase chain reaction (PCR), Multiplex Ligation Dependent Probe Amplification (MLPA), comparative genomic hybridization, SNP arrays, DNA microarrays (CGH and SNP), and next-generation sequencing (NGS) based [29].

In Table 2**,** we have included the current available techniques for the identification of CNV mosaicisms [30,31].

PARK 2

Homozygous or compound heterozygous variants in the PRKN gene, harbored in chromosome 6 (6q25.2-q27), have been identified as the most common cause of early-onset parkinsonism (EOPD), with an estimated frequency of ~ 2.6% in individuals with PD onset before 50 years old affecting a number of ethnic populations [19] PRKN accounts for 50% of familial PD with early onset and recessive inheritance, and also explains ~15% of the sporadic EOPD cases [18].

The phenotype of the biallelic mutations is characterized by EOPD, levodopa response, dystonia (present in 65% of patients), slow progression; while psychiatric manifestations might be present, other non-motor symptoms such as olfactory dysfunction and cognitive impairment are rarely mentioned [32,33].

The PRKN gene encodes Parkin, a 465-residue E3 ubiquitin protein ligase protein, with various functional consequences, involved in the proteasomal degradation of misfolded or damaged proteins, with a pivotal role in mitochondrial quality control (autophagy/mitophagy). Mitochondrial control is mediated by the Pink1/Parkin pathway interaction on the machinery involved in mitochondrial fusion and fission, with signal amplification for mitophagy, and vesicular transport. On the other hand, PRKN possesses a regulatory role on transcription factors involved in mitochondrial biogenesis, for instance the peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1), a transcriptional coactivator of the mitochondrial biogenesis [32-37].

Although the clinical phenotype is indistinguishable from iPD except for the AAO, the neuropathological findings in biallelic PRKN variants have failed to identify a-Syn accumulation and Lewy bodies, raising again the question whether biallelic variants represent a distinct disease.

The PRKN genomic location, at the core of FRA6E site, is prone to rearrangements as a susceptible site of the genome. In this sense, to date more than 200 mutations have been reported, with a high rate of SNVs and CNVs [2,33].

To date, more than 200 homozygous or component heterozygous mutations of PRKN have been described, including duplications and deletions affecting the cloned 12 exons and the promoter region.

Additionally, alternative splicing and additional exons have been considered for modulation of *PARK2* [18].

Determination of mutational rearrangements is particularly challenging, and most of them are found in the region between *PARK2* exons 2 and 8.

Although the largest number of SNVs and CNVs has been described associated to PRKN, the role of these variants in heterozygous PRKN SNV and CNV remains controversial. A meta-analysis has shown an increased risk for PD in heterozygous carriers of CNVs

Table 2: Techniques for mosaicism detection.

Technique	Purpose	Principle	Benefits	Limitations	Ref
Karyotyping			To identify diseases whit numerical and structural balanced aberrations highly represented		**14**
Fluorescence *in situ* hybridization (FISH)	Localization of genes and specific genomic regions on target chromosomes in metaphase and interphase cells	Molecular cytogenetic technique using fluorescent probes.	Detection of multiple targets simultaneously in different colors.	Identifying genomic changes as small as ~2 kb but more typically those that are several hundred kb	**28, 14**
Copy Number Variant microarrays (maCNV)			Detection of mosaicism lower than 5-20% or balanced chromosomal aberrations	Do not provide information on the location of the structural variation	
Array based comparative genomic hybridization (aCGH)			to identify mosaicism when variant cells constitute >10% of the total cell population		**29,14**
Digital Droplet PCR(ddPCR)					
Multiplex Ligation Dependent Probe Amplification (MLPA)	To detect CNV changes using PCR	PCR	Simultaneous detection of ≥ 50 polymorphisms in a single reaction. Discrimination between two sequences differing by a single nucleotide. Characterization of the specific breakpoint site of gene deletions and duplications. To detect and characterize the methylation status of DNA and the presence of SNPs		**44**
Comparative genomic hybridization (CGH)	To detect the entire genome for copy number aberrations (CAN)	Molecular cytogenetic method, to compare fluorescent –labeled DNA to a reference DNA	To detect unbalanced chromosomal abnormalities.	No detection of single nucleotides changes reciprocal translocations, inversions or ring chromosomes	
Next-generation sequencing (NGS)	To produce millions or billions of short, high-confidence genomic reads	Sequencing reads are typically 100–150 bp and may be either single-read or paired-end sequenced.	NGS can discover all classes of mosaic variation across almost the entire genome.		**14**
DNA microarrays (CGH and SNP)	To analyze many, different and at different cycles cells simultaneously. Not culturing is require.		Medium- to large-sized mosaic CN changes at >5% aneuploidy can be detected by the balance of allele intensities across SNPs	To detect somatic mutations occurring at low variant allele frequency (e.g.,<5%) and small events due to limited probe coverage	**29,14**
Single-Cell Sequencing (SCS)		To amplify, and sequency DNA from single cells	Approximately 90% of the genome is accessible by SCS	Enables the discovery of somatic mutations within individual cells, cell types, and tissues.	**14**
Single-Molecule Sequencing (SMS)		Exceptionally long sequencing reads (>20 kb) from unamplified high-molecular-weight DNA.	SMS long reads permit sequencing through repetitive elements, improved variant phasing, and detection of epigenetic modifications	Higher per-base nucleotide error rates, large DNA input requirements (>5 µg), variable sequence read lengths, errors in low-sequence complexity regions, and prohibitive cost	**14**

for PRKN [38]. Some authors reported AAO 10 years early in Parkin heterozygotes with respect to the idiopathic PD patients, but these observations have not been replicated by others [33].

Facing the concept of precision medicine, it is crucial to identify these subpopulations, because they contribute to explain why some specific gene therapies could fail.

PINK1

PARK6 gene (*PINK1gene)* encodes the 581 amino acid protein phosphatase and tensin homolog (PTEN)-induced kinase 1 (PINK1), a mitochondrial serine/threonine protein kinase, and plays several important roles in mitochondrial pathways, including mitophagy, mitochondrial trafficking, and mitochondrial dynamics. The protein sequence reveals a predicted C-terminal kinase domain and a mitochondrial targeting sequence at the N-terminus suggesting that it is imported into the mitochondria, consistent with its mitochondrial localization in cells.

Some mutations in *PINK1* may decrease protein stability, whereas others significantly reduce the phosphorylation or kinase activity. Homozygous and compound heterozygous deletions, harboring

exons 4–8, have been reported in familial and sporadic EAOPD from different countries around the world, affecting neighboring *DDOST genes.*

The entire *PINK1* genetic region, two neighboring genes, and two highly similar AluJo repeat sequences are involved in the largest PINK1 heterozygous deletion [18].

Some heterozygous carriers have shown mild Parkinsonism, as in PRKN mutations in these cases disease may be caused by haplo insufficiency or a low-penetrance dominant mechanism. Moreover, carriers of heterozygous mutations in recessive genes could acts as a susceptibility factor or a disease modifier [38].

DJ1

PARK7 mutations is one of the genes involved in autosomal recessive EAOPD. This gene, mapped at chromosome 1p36, encodes a conserved multifunctional protein. DJ1 acts as a positive regulator of transcription, redox-sensitive chaperone, sensor for oxidative stress, and apparently protects neurons from ROS-induced apoptosis.

Heterozygous deletions and duplications in CNVs involving the exons of *DJ-1* gene have been reported, but these findings failed to explain the recessive pattern of the PD phenotype [18].

ATP13A2

ATP13A2 mutations are associated with autosomal recessive levodopa-responsive atypical Parkinsonism (Kufor-Rakeb syndrome, KRS) mapped at 1p36.13. This gene encodes a large protein belonging to the ATPase transmembrane transporters, and recently it has been identified as a potent modifier of the toxicity induced by alpha-synuclein. CNV, a homozygous deletion of exon2, has been reported in a family from Iran. The phenotype was characterized by moderate mental retardation, aggressive behavior, visual hallucinations, supranuclear vertical gaze paresis, and slow vertical saccades [18].

The 22q11.2 deletion

DiGeorge syndrome is caused by a deletion of a small segment of chromosome 22 (22q11.2 deletion). Syndrome, is a multi-systemic syndrome caused by the deletion of a small segment of chromosome 22 [39].

The classical syndrome is characterized by multiple system involvement, including cleft palate, dysmorphic facial features, cardiac defects, skeletal deformities, developmental delays, learning disabilities, and increased risk of developing schizophrenia and other mental disorders. In recent years, a link between 22q11.2 deletions and PD has been suspected after initially reported by Butcher et al. [40]. These authors have identified EOPD in 4/ 159 adult individuals with 22q11.2 deletion syndrome, who were later confirmed positive for typical Lewy bodies and Lewy neurite formations.

Mok et al. have demonstrated a statistically significant association between 22q deletion and the risk for PD [41]. Another case has been reported with EOPD lacking the classical 22qDs phenotype. A rationale for this association could be found in the fact that the deleted region contains candidate genes implicated in PD. Among these genes, we have to mention:

a) Catechol-O-Methyl transferase (COMT gene),

b) *SEPT5*gene, a vesicle- and membrane-associated protein inhibiting exocytosis, as well as a park in substrate,

c) *DGCR8*gene that encodes a complex subunit involved in the biogenesis of microRNAs, including miR-185 which is predicted to target *LRRK2* [18].

Interestingly, Perandones et al. have detected by FISH analysis a mosaicism of 22q deletion in blood cells from an Ashkenazi Jewish ethnic group with PD [42]

The abovementioned reported case does not only provide more evidence about the relationship between Parkinson's disease and the 22q11.2 deletion syndrome, but it also highlights the relevance of performing individual cell-by-cell tests like FISH analysis, at least until single-cell sequencing becomes optimized and generally available. The pathogenesis of early-onset PD in patients with 22qDS remains unknown but, if elucidated, it may contribute to understanding the etiology of PD and ultimately to developing prevention and treatment strategies. For further details, see Figure 4.

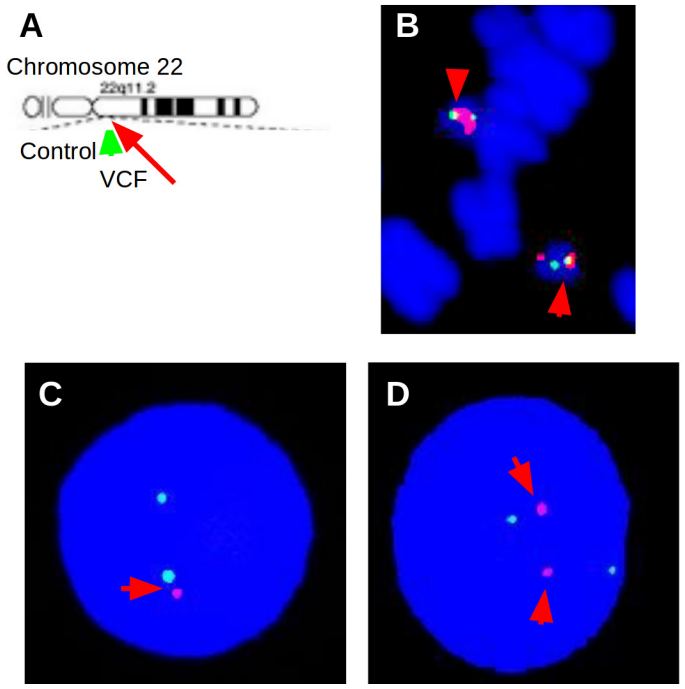

Figure 4: *22q11 deletion mosaicism.*

FISH (Fluorescent In-Situ Hybridization) was conducted using probes that had been labeled with fluorescent oligonucleotides by means of nick translation. BACs corresponding to the control telomeric region -RP11-976a21 were label with FITC or the RP5-882j5 deleted region 22q11.2 and RP5-925j7 labeled with rhodamine. FISH analysis confirmed a 22q deletion in 24% of the evaluated blood cells.

A) Graph with the location of the SNCA gene in chromosome 22, labeled in red, and the control gene (cep22) in green.

B) Fluorescence microscope image showing an interphase with the 22q11 microdeletion.

C) Fluorescence microscope image of chromosomes 22 carrying the control gene (green probe) and the VCF region (red probe; DAPI, in blue). D) Image of the interphase showing the Cep 22 control gene and the SNCA gene in a normal cell of the unaffected son.

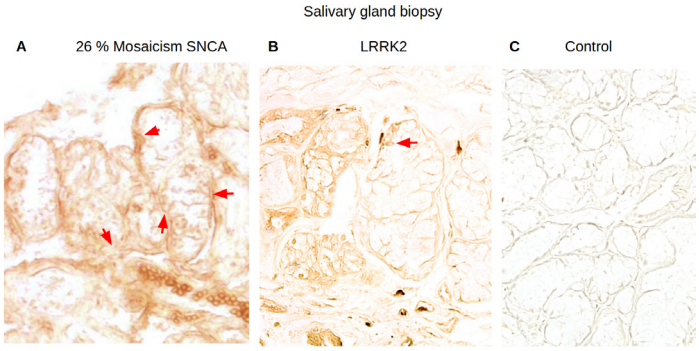

Figure 5: *a-syn staining in patients with SNCA mosaicism, LRRK2 mutations and control in salivary gland samples.*
Immunohistochemical studies of a-syn protein were performed on samples from the minor salivary glands using the rabbit anti-alpha synuclein antibody (sc-7011-r #-synuclein(C-20)-R Santa Cruz, CA, USA) and developed with secondary antibody (30 sc-2004 and g anti-rabbit IgG-HRP, Santa Cruz, Intl., CA, USA). The images were obtained in an optical microscope at 200 X.
A) Patient with 36% of alpha-synuclein duplication mosaicism.
B) Patient with **G2019S** LRRK2 mutation.
C) B patient control omitting the primary antibody.

LRRK2

The most common mutation in LRRK2 is the G2019S substitution in the kinase domain, which leads to an increase in the kinase activity. This mutation accounts for 5% of familial autosomal dominant PD and ~2% for sporadic PD. Interestingly, many PD patients with LRRK2 (G2019S) mutations exhibit α-synuclein-positive LBs, even though nearly half the LRRK2 (G2019S)-PD cases are LB-negative. This suggests that LRRK2 mutations could be implicated in α-synuclein-induced neurodegeneration in PD. For further details, see Figure 5.

The LRRK2 protein is widely expressed in brain, kidney, heart, lung and peripheral mononuclear cells. The protein function is not completely known; it has been implicated in autophagy, endocytosis, mitochondrial dysfunction, and cytoskeletal function, and it plays a role in neuro inflammation and in systemic immune pathways.

A variety of evidence suggests an inverse association between neurodegenerative diseases and cancer. However, more recent studies have suggested that overexpression of LRRK2 may activate proliferation of certain cancers (melanoma, papillary thyroid carcinoma) [45]. Nevertheless, a dual LRRK2 function has been described, on the one hand, a tumor-suppression function, mediated by p53 phosphorylation and p21 induction, JNK activation, and RCAN1 phosphorylation, while an oncogenic role appears related to the MET signaling activation.

Recently, Lopez et al. have investigated the prevalence of LRRK2 point mutations, indels or fusions, and copy-number variations (CNVs) in a large cohort of human malignancies, and the potential prognosis of this neoplasms [44]. The authors have demonstrated that LRRK2 genetic alterations account for 5.4% in the cancer genomic database, with MUTs in 76.9% of cases and CNVs in 22.1%. These alterations do not involve G2019S and R1441C mutations. LRRK2 CNVs, both amplifications and whole gene deletions, confer a poorer prognosis in terms of OS in comparison to unaltered cases, while LRRK2 somatic MUTs show no prognostic significance [44].

Conclusions

A variety of evidence suggests an extensive and complex genetic action of CNVs and CNV mosaicisms on PD etiopathogenesis. Thus far, unfortunately, only a small portion of the genetic variance has been identified; the remaining substantial components are still unknown. Assessing the global genome-wide burden of large CNVs and elucidating the role of *de novo* rare structural variants on PD may reveal new candidate genes, explain a portion of the "missing heritability," and consequently improve diagnosis and counseling of mutation carriers.

Despite extensive efforts and continuous progress, mosaic identification remains very challenging in clinical diagnostics and research laboratories.

Various biological materials may be tested in mosaicism investigation. Primarily, low invasive procedures are recommended in the sample collection process. Peripheral blood samples (PBSs), fibroblasts, and cells derived from other tissue biopsies, buccal tissue, and saliva are the most frequently used. Mosaicism detection is always correlated with the mutation identification *per se*. Targeted material, containing the causing mutation, would be optimal for testing and always better than the most accessible one. As already mentioned, frequently-dividing cells would be the examples of material to test or verify somatic mosaicism.

Since somatic mutations may occur at the same time in different tissues and in distinct body sites with different mosaic ratio, without extensive sampling the obtained results would be correct for the specific analyzed samples only.

In this era of constant progress in the technology used in clinical diagnostics and research, it is crucial to adjust the tested material to the detection technique.

Inconsistencies between mosaicism detection techniques and the variety in the downstream analyses may complicate the comparison of the clinical diagnostic or research outcomes. Mosaicism identification is usually a multi-step process, extensive, expensive and time-consuming. Usually more than one technique is used to recognize mosaicism. Then, an additional method, usually different from the first one used, is applied to confirm the feature.

Finally, with the clinical and diagnostic data, the genotype–phenotype correlation will be assessed to discuss and/or predict mosaicism clinical consequences. There is no gold standard test in mosaicism detection. While the improvement in mosaicism detection points to NGS technologies as the most promising method, there are some hesitations and exemptions. Moreover, it is worth pointing out that methods already established and functioning in the laboratory are less error-prone as compared to novel technologies to be implemented.

Also, while the NGS technologies in themselves are user-friendly, the subsequent bioinformatics analyses are too complex to be performed in a regular diagnostic laboratory.

Mosaicism can complicate clinical diagnosis and genetic counseling. Mosaic phenotypes may have incomplete syndromic features, which may stay unnoticed, especially in low-grade mosaicism.

Carriers of mosaic mutations may be at risk for abnormal pregnancy outcomes, including offspring with the phenotype caused by the mutation in the form of a dominant allele. Those findings have important implications for genetic counseling and for understanding patterns of recurrence in transmission genetics.

As NGS technologies constitute a promising methodological solution in mosaicism detection in the coming years, revisions in current diagnostic protocols are necessary to increase the detection rate of the unrevealed mosaicism events. In the NGS era, the apparently *de novo* mutations will be identified, in a percentage of cases, as a consequence of mosaicism occurred in the previous generation of the proband. Also, with various tissue samples tested per one individual, mosaics with even a low level of the mutation will be detected, and the findings will allow for better genotype–phenotype correlations and more precise clinical diagnosis.

Contribution to the field

In this review, through multiple instances of experimental evidence, we analyze the impact on histopathology of the different mutational mechanisms involved in the genesis and etiology of PD. We believe that increasing our knowledge about the changes and implications at tissue level produced by each of those mechanisms will allow to develop much more suitable and personalized potential therapeutic strategies, biomarker identification, as well as disease modeling, agreeing with the precision medicine concept.

Acknowledge

The authors would like to thank Daniela Converso and Daniel Nicolas Bruni for its technical support.

References

1. Balestrino R, Schapira AH. Parkinson disease. European journal of neurology. 2020; 27: 27-42.
2. Correia Guedes L, Mestre T, Outeiro TF, et al. Are genetic and idiopathic forms of Parkinson's disease the same disease? Journal of Neurochemistry. 2020; 152: 515-522.
3. Sarihan E.I, Pérez-Palma E, Niestroj L.M, et al. Genome-Wide Analysis of Copy Number Variation in Latin American Parkinson's Disease Patients. Movement Disorders. 2021; 36: 434-441
4. Armstrong MJ, Okun MS. Time for a New Image of Parkinson Disease. JAMA neurology. 2020; 77: 1345-1346.
5. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. The Lancet Neurology. 2019; 18: 1091-1102.
6. Bandres-Ciga S, Diez-Fairen M, Kim JJ, et al. Genetics of Parkinson's disease: An introspection of its journey towards precision medicine. Neurobiology of Disease. 2020; 137: 104782.
7. Kasten M, Hartmann C, Hampf J, et al. Genotype-Phenotype Relations for the Parkinson's Disease Genes Parkin, PINK1, DJ1: MDSGene Systematic Review. Mov Disord. 2018; 33: 730-741.
8. ichards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine. 2015; 17: 405-423.
9. Cook L, Schulze J, Kopil C, et al. Genetic testing for Parkinson disease: Are we ready?. Neurology: Clinical Practice. 2020; 11: 69-77.
10. Wirth T, Mariani LL, Bergant G, et al. Loss-of-Function Mutations in NR4A2 Cause Dopa-Responsive Dystonia Parkinsonism. Mov Disord. 2020; 35: 880-885.
11. Pfundt R, Del Rosario M, Vissers LELM, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. Genet Med. 2017; 19: 667- 675.
12. Gentile G, La Cognata V, Cavallaro S. The contribution of CNVs to the most common aging-related neurodegenerative diseases. Aging Clinical and Experimental Research. 2021; 33: 1187-1195.
13. Valsesia A, Macé A, Jacquemont S, et al. The growing importance of CNVs: new insights for detection and clinical interpretation. Frontiers in genetics. 2013; 4: 92.
14. Thorpe J, Osei-Owusu IA, Avigdor BE, et al. Mosaicism in human health and disease. Annual Review of Genetics. 2020; 54: 487-510.
15. Liu Q, Grochowski CM, Bi W, et al. Quantitative Assessment of Parental Somatic Mosaicism for Copy-Number Variant (CNV) Deletions. Current Protocols in Human Genetics. 2020; 106: e99.
16. Queremel Milani DA, Chauhan PR. Genetics, Mosaicism. Stat Pearls [Internet]. 2021.
17. La Cognata V, Morello G, D'Agata V, et al. Copy number

variability in Parkinson's disease: assembling the puzzle through a systems biology approach. Human genetics. 2017; 136: 13-37.

18. Soong D, Stratford J, Avet-Loiseau H, et al. CNV Radar: an improved method for somatic copy number alteration characterization in oncology. BMC bioinformatics. 2020; 21: 1-18.
19. Yu E, Rudakou U, Krohn L, et al. Analysis of heterozygous PRKN variants and copy number variations in Parkinson's disease. Mov Disord. 2021; 36: 178-187.
20. Visanji NP, Lang AE, Kovacs GG. Beyond the synucleinopathies: alpha synuclein as a driving force in neurodegenerative comorbidities. Translational neurodegeneration. 2019; 8: 28.
21. Blauwendraat C, Makarious MB, Leonard HL, et al. A population scale analysis of rare SNCA variation in the UK Biobank. Neurobiology of Disease.2021; 148: 105182.
22. De Luca CMG, Elia AE, Portaleone SM, et al. Efficient RT-QuIC seeding activity for α-synuclein in olfactory mucosa samples of patients with Parkinson's disease and multiple system atrophy. Translational Neurodegeneration. 2019; 8: 24.
23. Rossi M, Candelise N, Baiardi S, et al. Ultrasensitive RT-QuIC assay with high sensitivity and specificity for Lewy body-associated synucleinopathies. Acta neuropathologica. Erratum in: Acta Neuropathol. 2020; 140: 245.
24. Shahnawaz M, Mukherjee A, Pritzkow S, et al. Discriminating α-synuclein strains in Parkinson's disease and multiple system atrophy. Nature. 2020; 578: 273-277.
25. Perandones C, Giugni JC, Calvo DS, et al. Mosaicism of alpha-synuclein gene rearrangements: report of two unrelated cases of early-onset Parkinsonism. Parkinsonism & related disorders.2014; 20: 558-561.
26. Forsberg LA, Gisselsson D, Dumanski JP. Mosaicism in health and disease—clones picking up speed. Nature Reviews Genetics. 2017; 18: 128.
27. Perez-Rodriguez D, Kalyva M, Leija-Salazar M, et al. Investigation of somatic CNVs in brains of synucleinopathy cases using targeted SNCA analysis and single cell sequencing. Acta Neuropathologica Communications. 2019; 7: 1-22.
28. Garimberti E, Tosi S. Fluorescence in situ hybridization (FISH), basic principles and methodology. In Fluorescence in situ hybridization (FISH). Methods Mol Biol. 2010; 659:3-20.
29. Gorcenco S, Ilinca A, Almasoudi W, et al. Invited review-Highlights of 2019 congress new generation genetic testing entering the clinic. Parkinsonism & Related Disorders.2020; 73:72-84.
30. Campbell IM, Shaw CA, Stankiewicz P, et al. Somatic mosaicism: implications for disease and transmission genetics. Trends in Genetics. 2015; 31: 382-392.
31. Wang Z, Guo J, Guo Y, et al. Genome-Wide Detection of CNVs and Association With Body Weight in Sheep Based on 600K SNP Arrays. Frontiers in Genetics. 2020; 11: 558.
32. Borsche M, Pereira SL, Klein C, et al. Mitochondria and Parkinson's disease: Clinical, Molecular, and Translational Aspects. Journal of Parkinson's disease. 2021; 11: 45-60.
33. Wasner K, Grünewal A, Klein C. Parkin-linked Parkinson's disease: From clinical insights to pathogenic mechanisms and novel therapeutic approaches. Neuroscience Research. 2020; 159: 34-39.
34. Karimi-Moghadam A, Charsouei S, Bell B, et al. Parkinson disease from mendelian forms to genetic susceptibility: new molecular insights into the neurodegeneration process. Cellular and molecular neurobiology. 2018; 38: 1153-1178.
35. Tay SP, Lim GYG, Yeo CW, et al. Parkin and Parkinson's disease (pp. 47-70). InTech. Book: Etiology and Pathophysiology of Parkinson's disease. 2011.
36. Cookson MR, Bandmann O. Parkinson's disease: insights from pathways. Human molecular genetics. 2010; 19: R21-R27.
37. Sekine S, Youle RJ. PINK1 import regulation; a fine system to convey mitochondrial stress to the cytosol. BMC biology. 2018; 16: 2.
38. Reed X, Bandrés-Ciga S, Blauwendraat C, et al. The role of monogenic genes in idiopathic Parkinson's disease. Neurobiology of disease. 2019; 124: 230-239.
39. Boot E, Bassett AS, Marras C. 22q11. 2 Deletion Syndrome–Associated Parkinson's disease. Movement disorders clinical practice. 2019; 6: 11-16.
40. Butcher NJ, Kiehl TR, Hazrati LN, et al. 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 neurology. 2013; 70: 1359-1366.
41. Mok KY, Sheerin U, Simón-Sánchez J, et al. Deletions at 22q11. 2 in idiopathic Parkinson's disease: a combined analysis of genome-wide association data. The Lancet Neurology. 2016; 15: 585-596.
42. Perandones C, Farini VL, Pellene LA, et al. Parkinson's disease in a patient with 22q11. 2 deletion syndrome: the relevance of detecting mosaicisms by means of cell-by-cell evaluation techniques. Single Cell Biol. 2015; 4:4.
43. Eis PS, Huang N, Langston JW, et al. Loss-of-Function NUBPL Mutation May Link Parkinson's disease to Recessive Complex I Deficiency. Frontiers in Neurology. 2020; 11: 1332.
44. Lopez G, Lazzeri G, Rappa A, et al. Comprehensive genomic analysis reveals the prognostic role of LRRK2 copy-number variations in human malignancies. Genes. 2020; 11: 846.
45. Rojas NG, Cesarini M, Etcheverry JL, et al. Neurodegenerative diseases and cancer: sharing common mechanisms in complex interactions. Journal of integrative neuroscience. 2020; 19: 187-199.