UNIVERSITY OF CATANIA

INTERNATIONAL PhD PROGRAM IN NEUROSCIENCE

XXIX Cycle

VALENTINA LA COGNATA

READING THROUGH THE *BUILDING BLOCKS* OF THE GENOME: EXONIC VARIATION IN PARKINSON'S DISEASE

PhD Thesis

Coordinator:

Prof. Salvatore Salomone

Supervisors:

Prof. Velia D'Agata

Dr. Sebastiano Cavallaro

January 2017

SUMMARY

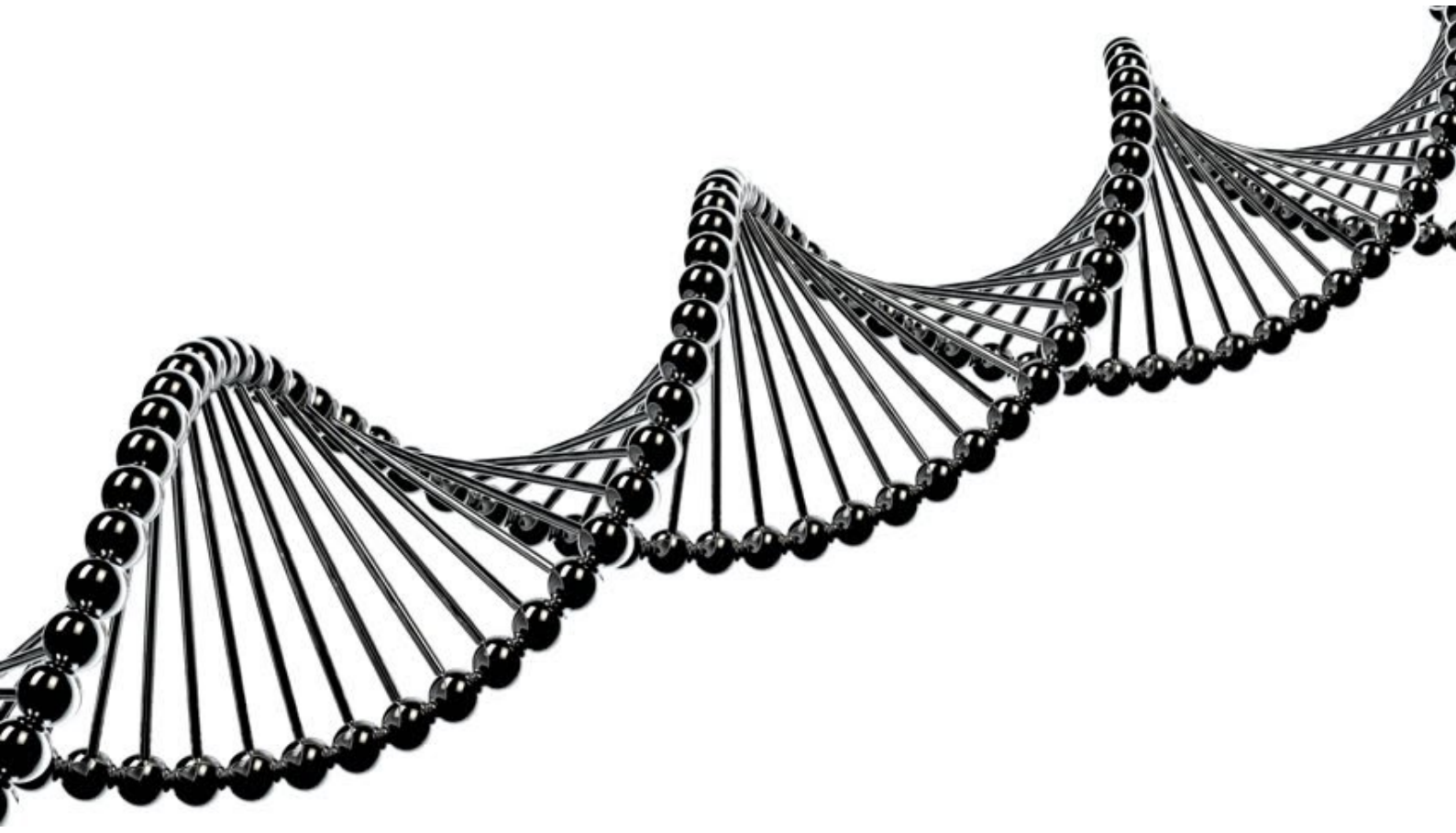
ABSTRACT

Parkinson's disease (PD) is one of the most common movement disorders worldwide, characterized by a profound and selective loss of dopaminergic neurons in the *substantia nigra pars compacta*. Treatments aimed at compensating dopamine deficit can alleviate the major motor symptoms and enhance the patients' quality of life, but finally are not able to halt or slow down disease progression. Therefore, there is an urgent need to better understand the molecular mechanisms underlying the physiopathology of PD and to identify new biomarkers and new therapeutic targets.

The hypothesis addressed in this PhD thesis aims to decipher the structural variability of exonic regions in PD-linked genes and in their relative mRNA transcripts, in order to investigate if these perturbations have some effects on PD pathogenesis. Two major cellular events able to trigger exonic variations in both DNA and mRNA molecules will be examined: copy number variations and alternative splicing. Both mechanisms are well known to play a crucial role in PD onset and can modulate disease severity. An improved comprehension of exonic variability at both genomic and transcriptomic level may prompt new insights to understand the "missing heritability" and the variety of phenotypic outcomes in PD patients.

CHAPTER 1

General Introduction & Aims



PARKINSON'S DISEASE

PD is the second most common progressive neurodegenerative disorder after Alzheimer's disease. It affects 1-2 % of all individuals above the age of 65 years old, increasing to 4-5% by the age of 85. Old age represents the greatest risk factor; indeed the onset is extremely rare before age 40. PD is a slowly progressive disorder, which begins insidiously, gradually worsens in severity and usually affects one side of the body before spreading to involve the other side. It is characterized by four cardinal symptoms: bradykinesia, resting tremor, rigidity, and postural instability. The early symptoms of PD are usually alleviated by the treatment with levodopa or dopamine agonists. As PD advances from year to year, late symptoms such as flexed posture, loss of postural reflexes and freezing phenomenon, do not respond to the treatment anymore. Surgical interventions such as deep brain stimulation of striatal output pathways have proven effective in some cases. While motor symptoms dominate PD clinical features, many patients show also non-motor symptoms. These include fatigue, depression, anxiety, sleep disturbances, constipation, decreased motivation, apathy and a decline in cognition that can progress to dementia.

PD is due to the relatively selective loss (70-90%) of dopaminergic neurons in the *Substantia Nigra pars compacta* (SNc), which leads to a profound reduction in striatal dopamine (DA). The loss of dopaminergic neurons is asymmetric, slow and progressive as the disease itself. With the progressive loss of dopaminergic neurons, there is a corresponding decrease of DA content in both the *Substantia Nigra* and the *striatum.* The loss of the nigrostriatal pathway can be detected during life using PET (positron emission tomography) and SPECT (single-photon emission computed tomography) scanning, showing a progressive reduction of fluoro-DOPA (FDOPA) and DA transporter ligand binding in the *striatum*. The neuronal loss is accompanied by an increase in glial cells and loss of neuromelanin, pigment normally contained in dopaminergic neurons. Lewy Bodies and dystrophic neuritis, called Lewy Neurites (LNs), are present in some of the remaining dopaminergic neurons and are the typical pathological hallmark of PD. LBs are round eosinophilic inclusions composed of a halo of radiating fibrils and a less defined core (Figure 1). Both LBs and LNs are composed by the accumulation of cytoplasmic aggregates containing a variety of proteins, of which α-synuclein is the major component.

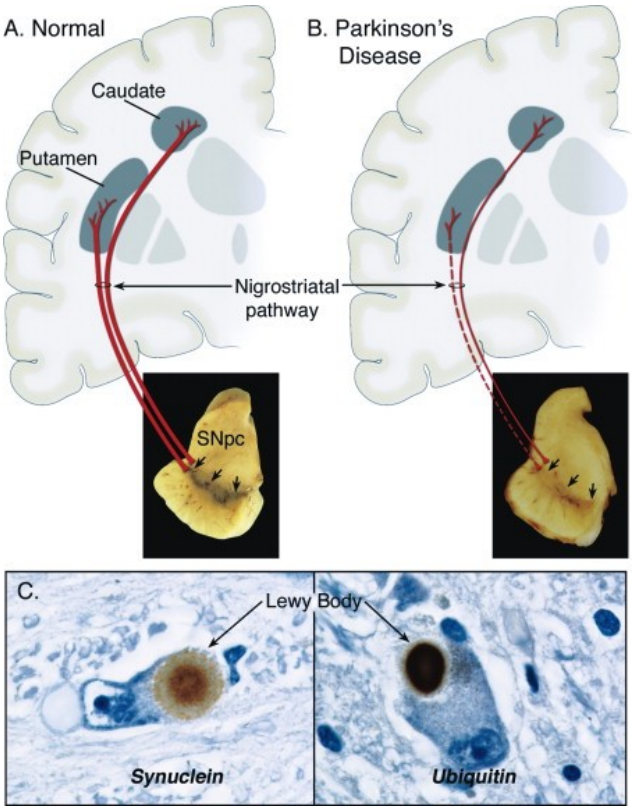


Figure 1. Neuropathology of Parkinson's disease

(A) Schematic representation of the normal nigrostriatal pathway (in red). It is composed of dopaminergic neurons whose cell bodies are located in the substantia nigra pars compacta (SNpc; see arrows). These neurons project (thick solid red lines) to the basal ganglia and synapse in the striatum (i.e., putamen and caudate nucleus). The photograph demonstrates the normal pigmentation of the SNpc, produced by neuromelanin within the dopaminergic neurons.

(B) Schematic representation of the diseased nigrostriatal pathway (in red). In Parkinson's disease, the nigrostriatal pathway degenerates. There is a marked loss of dopaminergic neurons that project to the putamen (dashed line) and a much more modest loss of those that project to the caudate (thin red solid line). The photograph demonstrates depigmentation (i.e., loss of dark-brown pigment neuromelanin; arrows) of the SNpc due to the marked loss of dopaminergic neurons.

(C) Immunohistochemical labeling of intraneuronal inclusions, termed Lewy bodies, in the SNpc dopaminergic neurons. Immunostaining with an antibody against α-synuclein reveals a Lewy body (black arrow) with an intensely immunoreactive central zone surrounded by a faintly immunoreactive peripheral zone (left photograph). Conversely, immunostaining with an antibody against ubiquitin yields more diffuse immunoreactivity within the Lewy body (right photograph).

From (Dauer and Przedborski 2003)

PD is a multifactorial disease caused by both genetic and environmental factors. The sporadic form of the disease has been suggested to spread from the interaction with chemicals in the environment. This may be due to the larger exposure to environmental toxins (like MPTP), herbicide (i.e. paraquat) or pesticides (i.e. rotenone). The familial cases are about the 10% of the total number of patients and are based on the genetic component of the disease. Patients with familial PD usually have an early start, greater consanguinity rate and greater frequency of a similar

disease in their parents. Highly penetrant mutations producing rare, monogenic forms of the disease have been discovered in singular genes such as *SNCA*, *LRRK2*, *Parkin*, *DJ1* and *PINK1* (Table 1). Moreover, a number of variants with incomplete penetrance have been shown to be strong risk factors for PD in certain populations. However, only a small portion of the genetic variance involved in PD has been identified; the remaining substantial components remain unknown and urgently need to be addressed.

Symbol	Gene locus	Gene	Inheritance	Disorder
PARK1	4q21-22	SNCA	AD	EOPD
PARK2	6q25.2–q27	Parkin	AR	EOPD
PARK3	2p13	Unknown	AD	Classical PD
PARK4	4q21–q23	SNCA	AD	EOPD
PARK5	4p13	UCHL1	AD	Classical PD
PARK6	1p35–p36	PINK1	AR	EOPD
PARK7	1p36	DJ-1	AR	EOPD
PARK8	12q12	LRRK2	AD	Classical PD
PARK9	1p36	ATP13A2	AR	Kufor-Rakeb syndrome; atypical PD with dementia, spasticity, and supranuclear gaze palsy
PARK10	1p32	Unknown	Risk factor	Classical PD
PARK11	2q36-37	Unknown	AD	Late-onset PD
PARK12	Xq21–q25	Unknown	Risk factor	Classical PD
PARK13	2p12	HTRA2	AD or risk factor	Classical PD
PARK14	22q13.1	PLA2G6	AR	Early-onset dystonia-parkinsonism
PARK15	22q12–q13	FBX07	AR	Early-onset parkinsonian-pyramidal syndrome
PARK16	1q32	Unknown	Risk factor	Classical PD
PARK17	16q11.2	VPS35	AD	Classical PD
PARK18	3q27.1	EIF4G1	AD	Classical PD
PARK19	1p31.3	DNAJC6	AR	Juvenile onset, atypical PD
PARK20	21q22.11	SYNJ1	AR	Juvenile onset, atypical PD
PARK21	3q22.1	DNAJC13	AD	Late-onset PD

Table 1. The table lists the set of Mendelian genes currently linked to PD onset. Abbreviations: AR (autosomal recessive), AD (autosomal dominant), EOPD (early-onset PD).
From (Kalinderi, Bostantjopoulou et al. 2016).

Despite we are still far from comprehensively understand the genetic basis of PD, investigating Mendelian forms has provided precious insights into the pathophysiological mechanisms that underline the more common idiopathic form. The most affected processes described until now include the ubiquitin-proteasomal pathway, synaptic transmission, endosomal trafficking, lysosomal autophagy, energy metabolism and mitophagy.

HAS EXONIC VARIATION A ROLE IN PD?

The human genome is a dynamic system, which constantly varies because of sequence and structural changes. Sequence variations comprise both single nucleotide polymorphisms (SNPs) and small insertions or deletions (indels), while structural variations include deletions, duplications, mobile-element insertions, inversions, balanced or unbalanced translocations, chromosomal aneuploidies and complex genomic rearrangements. Such kind of modifications constitute a natural phenomenon and represent major contributors to human phenotypes, leading to either benign or pathogenic consequences. Thanks to the massive advances in our ability to map the human genome at high resolution, it is now possible to characterize the distribution and the role of genomic variation in both physiological and pathological states.

Human genes are an integral part of the genome, and their assembly and regulation are the results of a slow adaptive process, which over time has become more refined and tidy. Each of our 20,000 genes may undergo changes in the coding areas (the exons), in the longer non-coding segments (the introns), or in one of the regions responsible for regulation (promoter, enhancer, silencer, splice sites). These genetic variants are able to influence the transcriptional expression and their proteins production in diverse ways according to size, type, or location.

The exons, the *building blocks* of genes, constitute the most fragile sites, harboring the higher fraction of pathogenic variants linked to human diseases. The scientific community has spent many efforts and time to find out the cause-effect relationship between mutations of exonic sequence and human pathologies. However, while a number of successes have been achieved for monogenic disorders, for complex multigenic pathologies (like PD) decoding the genetic contribution is not easy, and several controversial inconclusive findings have been reported.

The hypothesis addressed in this PhD work focuses on deciphering the structural variability of exonic regions in PD-related genes. Exonic variation in DNA molecules can arise after rearrangement events (deletions or duplications of genomic intervals) or can occur directly at transcriptomic level thanks to the alternative splicing process. The first mechanism (copy number changes) translates into a gene-dosage alteration of transcriptional regulation; the second one works in absence of genomic changes and is responsible for shuffling and assembling the cassette exons in

order to produce several protein-coding mRNA transcripts and multiple protein isoforms starting from a single-gene. Both copy number changes and alternative splicing, therefore, contribute to enhancing exonic variation and functional diversity by increasing the diversification of gene products (Jin, Kryukov et al. 2008).

The comprehension of exonic variability at both genomic and transcriptomic level may prompt new clues to understanding the "missing heritability" and the wide spectrum of phenotypic outcomes in PD patients. Indeed, it is a laborious task, which could be extended to other complex neurodegenerative multifactorial diseases. In the word of a metaphor, if we have so far investigated the small cracks that are able to alter the solid wall of our genome, now we are going to address a more complex challenge: deciphering the assembling of the bricks.

AIMS OF THE PhD WORK

The long-term goal of the present PhD thesis is to evaluate the role of exonic variations in genes and mRNA transcripts linked to PD, and to investigate if these alterations have some effects on PD pathogenesis. I will primarily focus on two different mechanisms able to trigger exonic variations in both DNA and mRNA elements: copy number variations and alternative splicing. The main aims discussed in next pages can be summarized as follows:

Aim 1)

To evaluate the global impact of exonic Copy Number Variations in PD and examine the effect of rare individual rearrangements observed in PD patient by a *system biology* approach. Our analysis revealed that disregarded individual CNVs functionally act in common deregulated biological processes relevant for PD pathogenesis and therefore, potentially account for a portion of the "missing heritability" underlying PD.

Aim 2)

To design a customized exon-centric Comparative Genomic Hybridization array tailored to detect single/multi-exon deletions and duplications in a large panel of PD-related genes. This *ad hoc* designed high-throughput platform provides a focused

evaluation of clinically relevant exonic regions at relatively low cost and enables the exploration of new potential genetic biomarkers underlying PD pathogenic mechanisms.

Aim 3)

To assess the potential relevance of alternative splicing mechanisms in PD pathogenesis. We characterize the alternative splicing regulation of PD-linked genes and discuss the globally splicing changes observed in PD patients through genome-wide approaches.

Aim 4)

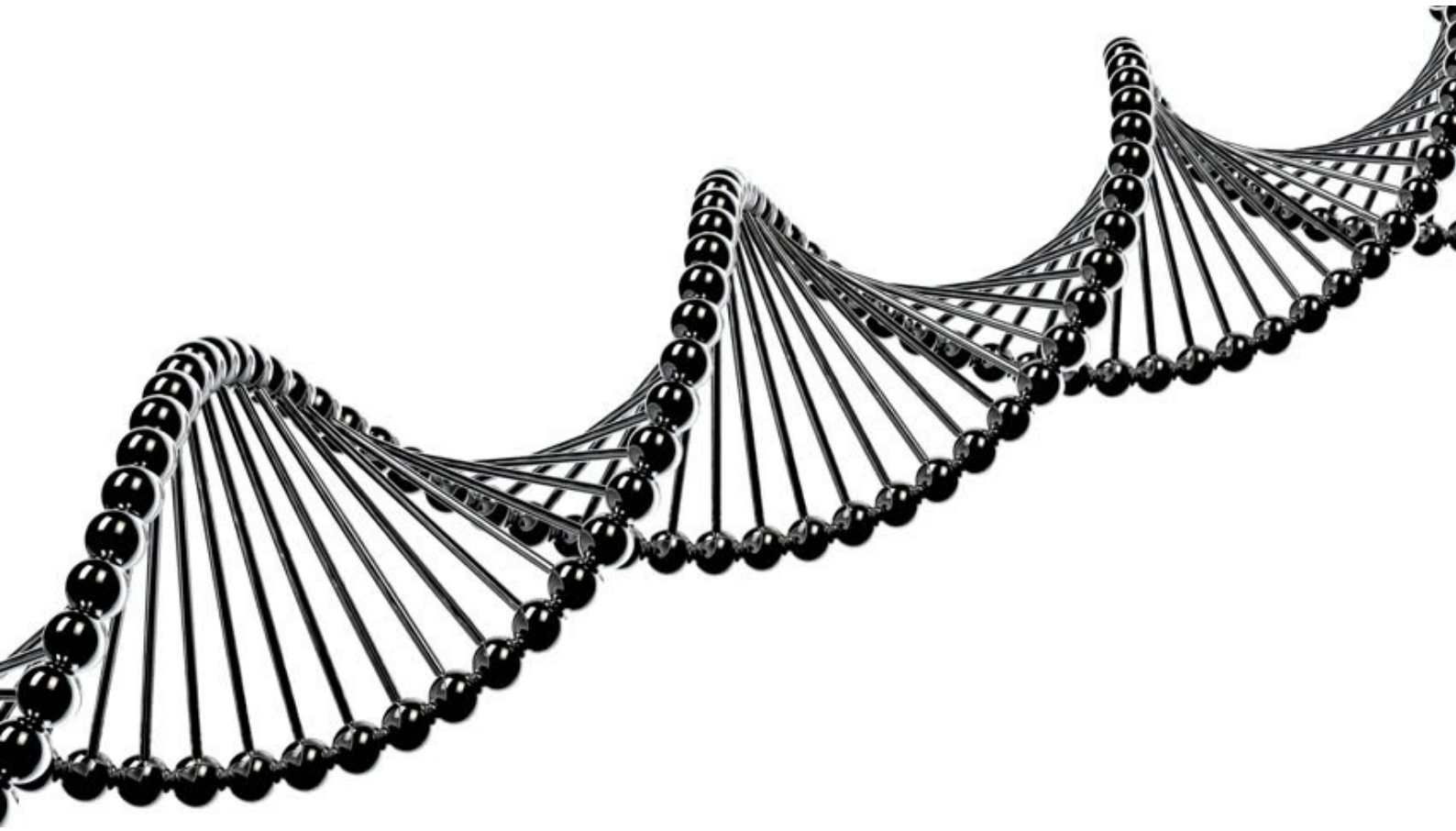
To better investigate the alternative spliced transcripts of a familiar autosomal recessive PD gene, *PARK2,* and correlate them to those in rat and mouse, two common animal models for studying human disease genes.

Aim 5)

To investigate the alternative splice protein isoforms of human, rat and mouse *PARK2* and analyze the diversified panel of commercially available antibodies recognizing different epitopes of parkin.

CHAPTER 2

Copy number variability in Parkinson's disease: assembling the puzzle through a systems biology approach



La Cognata V. et al. Human Genetics (2016); doi: 10.1007/s00439-016-1749-4

Copy number variability in Parkinson's disease: assembling the puzzle through a *systems biology* approach

Valentina La Cognata[1-2], Giovanna Morello[1], Velia D'Agata[2], Sebastiano Cavallaro[1*]

[1]*Institute of Neurological Sciences, National Research Council, Catania, Italy;*
[2]*Section of Human Anatomy and Histology, Department of Biomedical and Biotechnological Sciences, University of Catania, Catania, Italy.*

* Corresponding author: Sebastiano Cavallaro, M.D., Ph.D., Institute of Neurological Sciences (ISN), Italian National Research Council, Via Paolo Gaifami 18, 95125, Catania, Italy; Phone: +39 095 7338111; Fax: +39 095 7338 110; e-mail: sebastiano.cavallaro@cnr.it

ABSTRACT

Parkinson's disease (PD), the second most common progressive neurodegenerative disorder of aging, was long believed to be a non-genetic sporadic origin syndrome. The proof that several genetic loci are responsible for rare Mendelian forms has represented a revolutionary breakthrough, enabling to reveal molecular mechanisms underlying this debilitating still incurable condition.

While single nucleotide polymorphisms (SNPs) and small indels constitute the most commonly investigated DNA variations accounting for only a limited number of PD cases, larger genomic molecular rearrangements have emerged as significant PD-causing mutations, including submicroscopic Copy Number Variations (CNVs). CNVs constitute a prevalent source of genomic variations and substantially participate to each individual's genomic make-up and phenotypic outcome. However, the majority of genetic studies have focused their attention on single candidate-gene mutations or on common variants reaching a significant statistically level of acceptance. This gene-centric approach is insufficient to uncover the genetic background of polygenic multifactorial disorders like PD, and potentially masks rare individual CNVs that all together might contribute to disease development or progression.

In this review, we will discuss literature and bioinformatic data describing the involvement of CNVs on PD pathobiology. We will analyze the most frequent copy number changes in familiar PD genes and provide a "*systems biology"* overview of rare individual rearrangements that could functionally act on commonly deregulated molecular pathways. Assessing the global genome-wide burden of CNVs in PD patients may reveal new disease-related molecular mechanisms, and opens the window to a new possible genetic scenario in the unsolved PD puzzle.

Keywords: Parkinson's disease, Genetics, Genomics, Copy Number Variations, DNA rearrangements, *systems biology*

INTRODUCTION

Parkinson's disease (PD) is a progressive debilitating movement disorder, affecting approximately 1% of the population over 65 [1]. The characteristic major motor symptoms derive from the profound and selective loss of dopaminergic neurons from *substantia nigra pars compacta*, coupled with an accumulation of round cytoplasmic inclusions (Lewy bodies) and dystrophic neurites (Lewy neurites) in surviving neurons [1]. In more advanced stages, patients can also develop a range of non-motor symptoms, including rapid eye movement, sleep behavior disorder, constipation, depression and cognitive decline. Treatments aimed at compensating dopamine deficit (such as levodopa and deep brain stimulation) can alleviate the motor symptoms but finally are not effective to halt or slow down disease progression [2]. Despite the molecular mechanisms underlying PD are still far from being understood, the progressive deterioration of vulnerable dopaminergic neurons seems to arise from several cellular disturbances including protein misfolding and aggregation [3], synaptic damages, apoptosis, mitochondrial dysfunctions [4], oxidative stress [5], impairment of the Ubiquitin/Proteasome System (UPS) [6] and neuro-inflammation [7] (**Figure 1**).

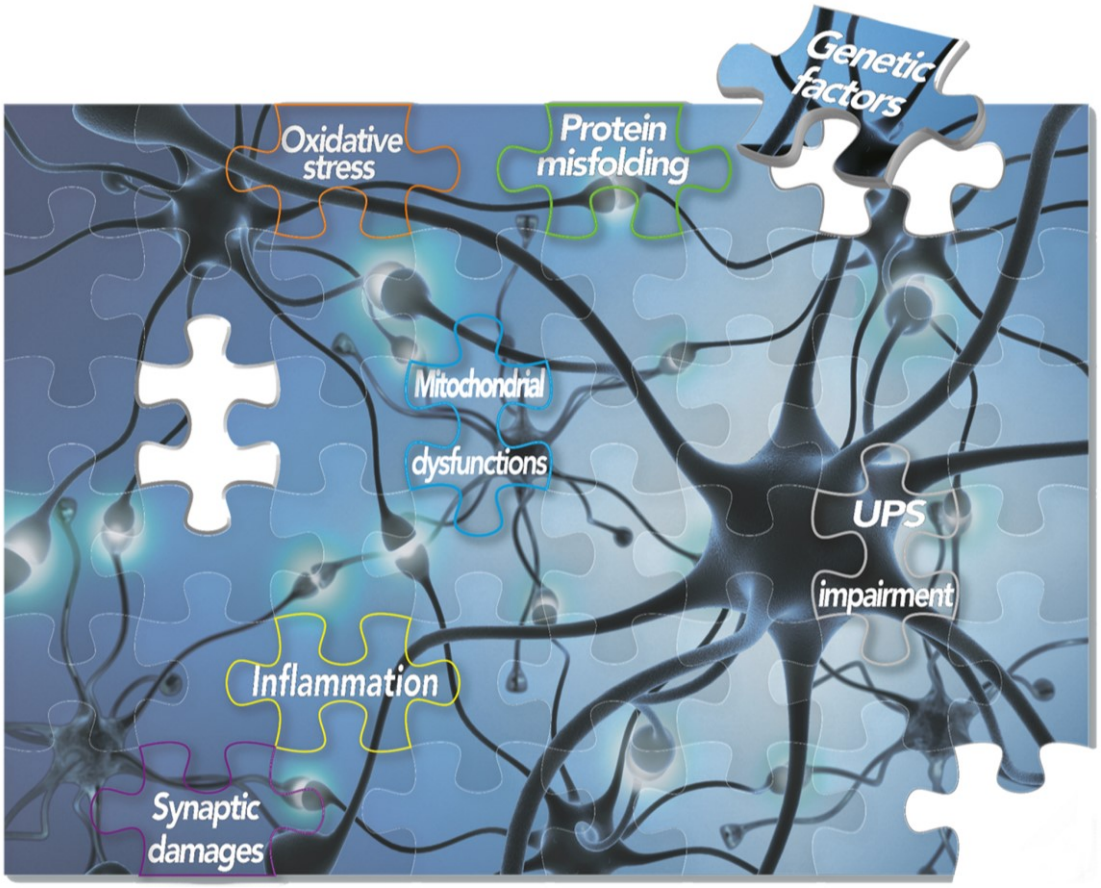


Figure 1. Schematic representation of molecular elements and common altered pathways underlying the complex PD puzzle.

PD was for a long time believed to be a typical non-genetic disorder. When in 1997 Polymeropoulos and colleagues reported the first *SNCA* pathogenic mutation in the Italian Contursi kindred [8], they revolutionized this view opening the way to new interesting perspectives about the genetic contribution to this still incurable condition [1]. From that moment, an increasing number of genetic loci and numerous risk factors have been discovered [9,10], starting from the familiar genes responsible of the Mendelian inherited forms, such as the autosomal dominant genes (*SNCA, LRRK2, VPS35, GBA*), the typical recessive (*PARK2, PINK1, PARK7*) and the atypical recessive ones (*ATP13A2, PLA2G6, FBXO7*) [11]. Despite the existence of these rare monogenic forms, it is now clear that PD is a genetically heterogeneous and most likely complex disorder, often complicated by incomplete penetrant traits and variable expressivity. The list of candidate genes is continuously updated [10,12,13], mainly thanks to the massive advancement in genomic biotechnologies that have allowed to detect hundreds of pathogenic or susceptibility variants at the single nucleotide

polymorphism (SNP) level. However, a lot of work still has to be done to identify additional sources of missing heritability or to assign a precise causal mechanism to the growing number of discovered loci [14].

While single nucleotide polymorphisms (SNPs) and small indels constitute the most commonly investigated DNA variations, submicroscopic chromosomal rearrangements, also known as Copy Number Variations (CNVs), are emerging as crucial players in the individual's genomic architecture and in modeling complex human diseases, including PD. However, the majority of CNVs association studies have been conducted by using the traditional candidate-gene approach that, although provides valuable information on common variants, is inadequate to completely dissect the genetic background of polygenic multifactorial disorders like PD. The search for single-gene mutations has to be changed, turning into the need to assess the collective effect of common and rare variants that together may converge on PD pathology. In this context, the "*systems biology*" approach represents a worthwhile instrument to analyze complex biological systems, moving beyond the conventional gene-centric scheme, and finally generating a more defined molecular picture of PD.

Herein, we will review the most common CNV-altered genes and detail the current knowledge about their pathogenic or susceptibility impact on PD pathobiology. Moreover, we will collect the set of rare individual CNVs reported so far in PD patients and analyze them by a "*systems biology*" approach. This new perspective reveals these private CNVs cluster in common deregulated biological processes that could contribute to disease onset or progression, and open the window to a new possible genetic scenario in the unsolved PD puzzle.

CNVs: A PREVALENT SOURCE OF GENOMIC VARIATIONS

The DNA sequence of human genome is constantly changing and this process allows humans to evolve and adapt. The scientific community has long been aware of genetic variations of extreme size (i.e. cytogenetically recognizable elements and SNPs) [15]. However, about 10 years ago, scientists began to recognize abundant variations of an intermediate size class known as structural variations. Within this class, Copy Number variations represent the largest component by far. CNVs are defined as genomic segments showing copy number variability among individuals compared to a

reference genome. The size of CNVs ranges from 50 bp to several Mb, with a significant drop of variant numbers in 50 bp to 1 kb range [16]. These structural variants can include either a single gene or a contiguous set of genes, encompassing more polymorphic base pairs than SNPs and finally resulting in an altered DNA diploid status (i.e. gain or loss of genomic region).

Depending on their size, CNVs can be measured by a multitude of laboratory testing methods, either targeting the whole genome (genome-wide level) or restricted to certain locations on chromosomes (locus-specific levels) **(Figure 2)** [17]. While targeted approaches such as FISH or quantitative PCR-based strategies have been long used in the past, the most advanced current screenings rely on whole-genome applications, such as array Comparative Genomic Hybridization or Next Generation Sequencing experiments. Both these biotechnologies have dramatically improved and catalyzed the detection and characterization of multiple CNVs, offering the simultaneous testing of thousands of loci with high reproducibility, high resolution, and scalability for complete mapping of imbalances [18-22]. However, these whole-genome strategies still need post-experimental validations and therefore, a gold-standard analysis has not been defined yet.

CNVs are very common and arise in presence of specific architectural genomic elements that render 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 [23]. The most common cause of recurrent genomic rearrangements is the non-allelic homologous recombination (NAHR), that occurs between two DNA blocks of high homology, like the region-specific low-copy repeats sequences (LCRs) (Figure 3, panel A). On the contrary, non-recurrent CNVs can result from non-homologous end joining (NHEJ) or fork stalling and template switching (FoSTeS) mechanisms. NHEJ represents the major cellular mechanism for double-strand break repair: upon a double-strand break, NHEJ reconnects chromosome ends leaving random nucleotides at the site of the breakage to facilitate the strands' alignment and ligation (Figure 3, panel B) [24]. FoSTeS occurs when the DNA replication machinery pauses, and the template is switched with another region in physical proximity to the original replication fork (Figure 3, panel C) [25]. Such template switching may occur several times before the replication process gets back to its original template, resulting in complex rearrangements [24].

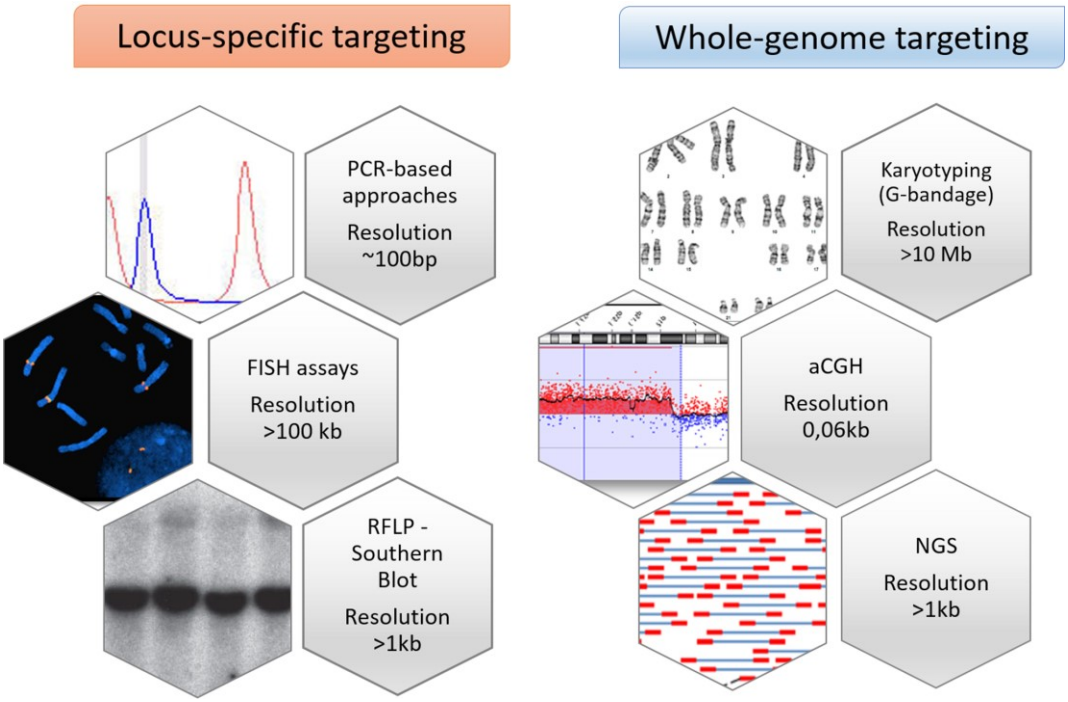


Figure 2. CNVs can be measured by a spectrum of laboratory-methods targeting specific locations on chromosomes (locus-specific levels), or the whole genome (genome-wide level). These numerous methodologies are characterized by different levels of resolutions. The locus-specific techniques encompass i) PCR-based strategies, such as quantitative real-time PCR (qPCR), Multiplex Ligand Probe Amplification (MLPA) or multiplex amplifiable probe hybridization (MAPH); ii) the Fluorescence in situ Hybridization (FISH) assays and iii) the RFLP (restriction fragment length polymorphism)– Southern blot analysis. The whole-genome methodologies include i) the classical chromosomal G-bandage (karyotyping); ii) the aCGH (Comparative Genomic Hybridization array) platforms and iii) the NGS (Next Generation sequencing) technology. These two latter are increasingly replacing both the classical detections methods and the locus-specific techniques.

CNVs can control phenotype in several ways: they can affect gene expression through the simple gene-dosage effects, or through more intricate mechanisms as, for example, insertions and deletions of regulatory regions and alterations of chromatin architecture [14]. To this regard, CNVs can interfere with a form of regulatory scaffold of the chromatin (the so-called Topologically Associating Domains or TADs) by disrupting or repositioning boundaries, and therefore, constraining the enhancer or silencer activity with their target genes [26]. Similarly, CNVs in other non-coding regions may alter the normal rate and tissue specific transcription pattern of the neighboring, otherwise intact, genes by changing, for example, the affinity for transcription factors. This *cis-acting* effect of non-coding variations has been recently demonstrated for a SNP in a distal enhancer element regulating the expression of *SNCA* [27]. Some representative pictures about the mechanisms of non-coding variants and their implication in human genetics are reported in a number of excellent reviews [28-30], whose the reader is referred for a better understanding.

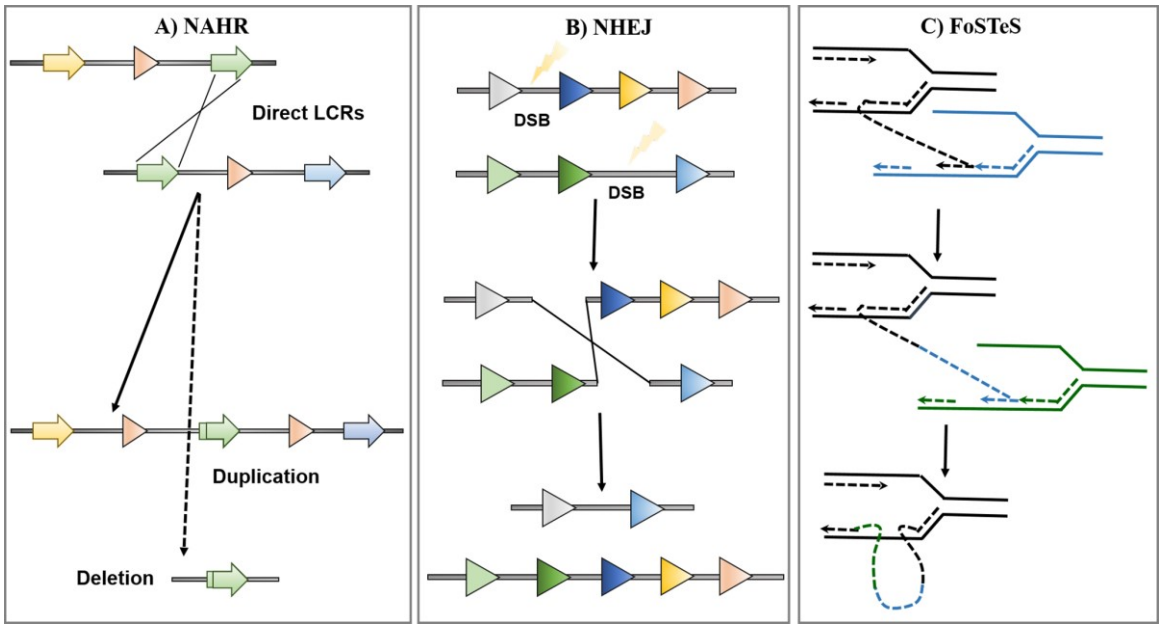


Figure 3. Schematic illustration of the three most common events causing genomic rearrangements. Panel A: NAHR generates CNVs when genomic segments with high sequence similarity (direct low-copy repeats sequences, green arrows) recombine. This recombination can generate a duplication of the similar locus (red arrow) on one chromosome, while removing the copy from the other. Panel B: Double stranded breaks (DBS) in DNA sequence recruit NHEJ associated proteins to repair and ligate DNA strands together. First, end-repair protein replaces lost nucleotides on the double strand break and DNA ligase associates broken DNA fragments together. If fragments from different chromosomes ligate together, duplications or deletions of sequence can occur. Panel C: After the original stalling of the replication fork (black lines), the lagging strand disengages and anneals to a second fork (blue lines), followed by extension of the now 'primed' second fork and DNA synthesis. After the fork disengages, the tethered original fork with its lagging strand (black and blue lines) could invade a third fork (green lines). Serial replication fork disengaging and lagging strand invasion could occur several times (e.g. FoSTeS x 2, FoSTeS x 3, etc.) before resumption of replication on the original template. It should be noted that the CNVs created through FoSTeS are difficult to be distinguished from those generated by micro-homology-mediated breakpoint-induced repair (MMBIR), a mechanism of end-joining that relies on small-scale homology of DNA sequence at the ends of DSBs.

All together, CNV alterations may account for adaptive or behavioral traits, may have no phenotypic effects or can underlie diseases. For this reason, determining the clinical significance of CNVs is very challenging and comprehensively relies on frequency information from healthy control cohorts, hereditability, size, gene content, type (copy number state) and location on chromosome (interstitial, centromeric or repeat-regions) [31].

Notwithstanding the difficulties in interpreting quantitative data, specific large CNVs and single-gene dosage alterations have emerged as critical elements for the development and maintenance of the nervous system [32] and have appeared to contribute to hereditable or sporadic neurological diseases, such as neuropathies,

epilepsy forms, autistic syndromes, psychiatric illnesses and also neurodegenerative diseases, including PD [23,33-37].

In the next paragraphs, we will focus on the current evidence about the occurrence of CNVs in familiar PD genes by highlighting strengths and weaknesses of interpretations for diagnosis and biomarkers usefulness. Moreover, we will collect from published literature the currently known set of rare CNVs observed in PD patients and analyze them through a *systems biology* point of view, in order to assess their biological role, their interactions and the possible functional impact on PD pathobiology.

COPY NUMBER VARIATIONS IN FAMILIAR PD GENES

SNCA

SNCA (alpha-synuclein) represents the most convincing locus causing both familiar and sporadic PD. This gene encodes a small natively unfolded presynaptic protein that aggregate in Lewy bodies and Lewy neurites, the pathological hallmark lesions of PD [38]. As we will discuss here below, *SNCA* is the best example of dosage-dependent toxicity: the more alpha-synuclein you have, the worse will be PD.

The first genomic triplication of *SNCA* was observed within the Spellman–Muenter family (better known as Iowa Kindred) a large family with autosomal dominant inheritance transmission of PD and dementia [39]. Later, several families with different ethnic background have been described, including members carrying four copies (triplication) or three copies (duplication) of *SNCA* (**Table 1**) [40-54,37,55-59]. In general, triplication generates very high expression of mRNA and protein molecules and influence the clinical manifestations of PD, causing severe forms of Parkinsonism similar to dementia with Lewy Body. In contrast, the clinical phenotype of patients with duplicated *SNCA* resembles idiopathic PD, mainly with late age at onset, good efficacy for levodopa therapy, slower disease progression and without early development of dementia.

An interesting familiar pedigree, the “Lister family”, present both duplicated and triplicated *SNCA* carriers within different branches of the pedigree (branches J and I), suggesting a primary duplication event followed later by another one and resulting in the triplication [60,61]. Similarly, the Ikeuchi family has both heterozygous and homozygous duplication carriers born from a consanguineous marriage (producing a

pseudo-triplication) [62]. The clinical features of individuals with the *SNCA* homozygous duplication showed severe parkinsonism similar to that of triplication carriers.

Along with the familiar forms, a good percentage of sporadic PD patients carry *de novo* duplication of *SNCA* (**Table 1**) [47,63-67]. Generally, their clinical course is similar to typical sporadic PD without severe progression or cognitive decline.

The breakpoint of *SNCA* multiplications is not the same in each patient. The largest multiplication detected so far is about 41.2 Mb, containing 150 genes and defined a partial trisomy 4q [63], while the smallest one counts about 0.2 Mb [48]. The size and gene make-up of each multiplicated region does not seem to severely influence the clinical presentation of the carriers.

Interesting insights derive from the mosaicism condition of *SNCA* rearrangements. To this regard, two interesting PD cases have been described, resulted negative to exon dosage test in peripheral blood, and positive for *SNCA* copy number changes on oral mucosa cells [68]. Both patients displayed a parkinsonian clinical phenotype of *SNCA* copy number carriers. Starting from this evidence, authors suggest to take into consideration the possibility to examine cells from both peripheral lymphocytes and other tissues in order to detect low-grade mosaicism.

CNVs in α-synuclein gene in PD						
	Size	**Ethnicity**	**Phenotype**	**Methodology**	**F-S-D**	**Reference**
Triplication (Spellman–Muenter family or Iowa Kindred)	1,61 – 2,04 Mb	Iowa	PD and dementia with LBs	qPCR, FISH	F	[39]
Duplication (Lister Family, branch J)	0,7987 – 0,9359 Mb	Sweden, United States	Late-onset parkinsonism and early dysautonomia	qPCR; Microsatellite markers analysis; Affymetrix 250K microarray	F	[60, 61]
Triplication (Lister Family, Swedish-America, Branch I)			Early-onset parkinsonism with dementia and dysautonomia			
Duplication (Ikeuchi family)	5 Mb	Japan	Progressive parkinsonism with dementia with LBs	Microsatellite markers analysis, qPCR	F	[62]
Homozygous duplication (Ikeuchi family - consanguineous marriage)						
Duplication (Uchiyama family)	0,5 – 1,6 Mb	Japan	Parkinsonism with dementia with LBs	qPCR	F	[46]
Duplication	n.a.	Korea	Early onset parkinsonism with rapidly progressive course, cognitive impairment, and dysautonomia (Ahn family)	Semi-quantitative multiplex PCR, FISH	F	[47]
			Typical PD		S	
Duplication	n.a.	Germany	Early-onset parkinsonism	qPCR, MLPA	D	[66]
Duplication	n.a.	European and North African	Early onset PD	MLPA , Microsatellite markers analysis	S	[65]
Duplication (Family A)	0,6 Mb	Japan	Parkinsonism with or without dementia	Microsatellite markers analysis, qPCR, FISH, aCGH (BACS and Affymetrix)	F	[48, 49, 56, 58]
Duplication (Family B)	0,4 Mb					
Duplication (Family C)	0,4 Mb					
Duplication (Family D)	0,4 Mb					
Duplication (Family E)	0,2 Mb					
Duplication (Family F)	0,6 Mb					
Duplication (Family G)	0,6 Mb					
Triplication (FPD-014) (Pat 011)	2,61 – 2,64 Mb	France. Italy	Atypical autosomal dominant parkinsonism	Semi-quantitative Multiplex PCR, Microsatellite analysis, Affymetrix GeneChip Human Mapping 250K microarray (just for P59	F	[50 -52, 57-59]
Duplication (FPD-131 o P59) (Pat 024-022-026)	4,928 Mb		Typical autosomal dominant PD			
Duplication (FPD-321) (Pat 021)	3,47 – 3,58 Mb					

Duplication (FPD-410) (Pat 001)	0,63 – 0,65 Mb			family: FISH, 44k CGH arrays Agilent)		
Duplication (FPD-437) (Pat 010-012)	0.42 - 0.43 Mb					
Duplication (Sironi family)	3,65 Mb	Italy	PD with progression to dementia	MLPA, Agilent 105A chip	F	[53]
Duplication	n.a.	Belgium	Parkinsonian syndrome	Multiplex amplicon quantification, qPCR	S	[67]
Triplication (Keyser family)	n.a.	South African (French-Italian origin)	PD with dementia	MLPA, qPCR	F	[40]
Duplication	n.a.	Korea	PD with cognitive dysfunction	Semi-quantitative multiplex PCR	S	[64]
Triplication	n.a.	Asian	Early-onset and severe clinical features of parkinsonism	qPCR, MLPA, microsatellite analysis	F	[41]
Duplication	3 Mb	n.a.	PD	Illumina370Duo arrays	F	[54]
Homozygous duplication	0,928 Mb	Pakistan	Young-Onset Parkinsonism	MLPA, Nimblegen 135 K array CGH	F	[42]
Duplication (Partial Trisomy 4q)	41,2 MB	Belgium	Young onset, dopa-responsive parkinsonism	Karyotype, aCGH, MLPA	D	[63]
Duplication	n.a.	Non-Hispanic Caucasian	Autosomal Dominant Early-onset PD	Customized 4 × 72 k format CGH microarrays by NimbleGen; Taqman qPCR	F	[37]
Duplication (family Elia A)	773 Kb	Northern Argentina	Early onset PD that was variably associated with nonmotor features, such as dysautonomia, cognitive deficits, and psychiatric disturbances	MLPA, qPCR, Affymetrix high-resolution single nucleotide polymorphism-array analysis	F	[55]
Duplication (Family Elia B)	4820 Kb	Italian	Early onset PD dementia with psychiatric disturbances to late onset PD with mild cognitive impairment			
Duplication	6,4 Mb	Caucasian English	Atypical clinical presentation strongly reminiscent of frontotemporal dementia and late-onset pallidopyramidal syndromes	MLPA, aCGH Agilent 8x60K	F	[58]
Duplication Triplication	n.a.	Iranian	PD typical clinical features PD with dementia	MLPA, qPCR	F-S	[43]

Duplication (mosaicism)	n.a.	American mitochondrial haplogroup and European autosomal markers	Early-onset Parkinsonism	MLPA (no dosage alteration in buccal swab), FISH (no rearrangements in peripheral leukocytes; duplication - triplication in oral mucosa)	F-S	[68]
Duplication	n.a.	American	Parkinsonism with LBs and Lewy neurites	n.a.	n.a.	[56]
Triplication	1,3 Mb	Italian	Early-onset parkinsonism combined with depression, behavior disturbances, sleep disorders, and cognitive decline	Genome-wide SNP microarrays, FISH, MLPA	F	[44]
Triplication	351 Kb	Italian	Severe parkinsonism featuring early onset dyskinesia, psychiatric symptoms, and cognitive deterioration	CGH-Array, MLPA, qPCR	F	[45]

Table 1 lists all the currently studies describing *SNCA* copy number changes in PD. The CNVs mutation type, the size of the mutation, the ethnicity of patients, the phenotype and the methodological approaches to measure quantitative genomic variations are reported. The column F-S-D reports if described cases are familial, sporadic or de novo.

PARK2

Although *SNCA* story suggests a gain of function, several early-onset forms of PD have demonstrated the role of loss of function genes in the etiology of the disease. The most common loss-of-function mutations belong to Parkin (or *PARK2*) gene, one of the largest in our genome harbored in the long arm of chromosome 6 (6q25.2-q27) and encoding an E3 ubiquitin ligase. Mutations of *PARK2* are particularly frequent in individuals with familiar recessive inheritance and account for 50% of the cases with autosomal recessive juvenile PD. Parkin mutations also explain ~15% of the sporadic cases with onset before 45 [69,70] and act as susceptibility alleles for late-onset forms of PD (2% of cases) [71].

PARK2 gene has a high mutation rate because it is located in the *core* of FRA6E site, one of the most mutation-susceptible common fragile site of human genome [24]. For this reason, more than 200 putative pathogenic mutations have been reported so far, affecting numerous ethnic populations [72-75,67,76,77,40,78-80,37,81-85]. The *PARK2* mutation spectrum includes homozygous or compound heterozygous missense and nonsense point mutations, as well as several exon rearrangements (both duplications and deletions) involving all the originally cloned 12 exons and the promoter region. Recently, our research group has outlined a complex alternative splicing mechanism regulating the expression of *PARK2* [86-88]. These data suggest that 5 additional exons exist, that however have never been considered for mutational or dosage screening. Overall, currently known Parkin CNVs are summarized in **Figure 4** and are collected in the Parkinson Disease Mutation database (http://www.molgen.vib-ua.be/PDMutDB), whose the reader is referred for more details.

CNV rearrangements involving *PARK2* exons accounts for 50–60% of all pathogenic anomalies, rendering gene-dosage assays essential in parkin mutational screening [89]. However, the *hot-spot* nature of this gene makes its quantitative analysis a particular challenge, and several issues need to be pointed out in this regard. Firstly, the determination of mutational phase of the rearrangements, meaning the assessment that amplified or deleted exons are really contiguous. Phase determination seems to be a fundamental requisite for *PARK2* molecular diagnosis: by phase determination, several patients with apparent contiguous multi-exon deletions were re-diagnosed as compound heterozygotes [89]. A second important point refers to

breakpoint mapping which can be useful to compare exon rearrangements between patients and families and to study the possible causing event mechanism [90]. Just a few papers have addressed this issue so far, but mostly report rearrangements into the region between *PARK2* exons 2 and 8 [90,24]. In the majority of mapped cases, micro-homologies at breakpoint junctions were present, thus supporting NHEJ and FoSTeS as the major mechanisms responsible for *PARK2* genomic rearrangements [24]. Moreover, some data underpin the possible effects of ancient common founder in minor ethnic groups [91]. For example, microsatellite markers analysis in four families from The Netherlands have shown that a common haplotype of 1.2 Mb could be distinguished for the exon 7 duplication and a common haplotype of 6.3 Mb for the deletion of exon 4, suggesting common founder effects for distinct large rearrangements in parkin [90].

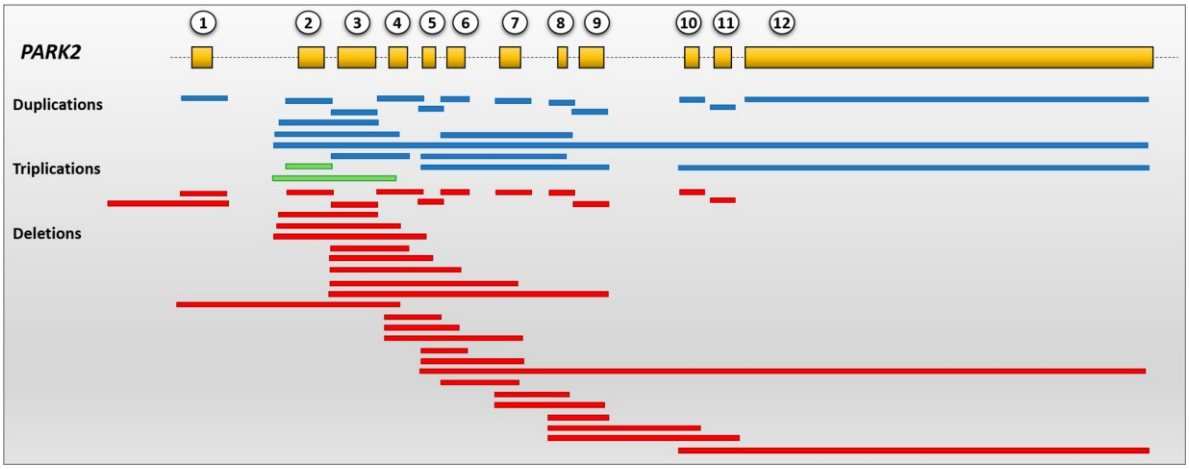


Figure 4. Schematic representation of PARK2 genetic structure and currently identified CNVs in PD patients. All the canonical PARK2 exons are involved in exons rearrangements. Red bars correspond to exons deletions, blue bars to duplications and green bars to triplications. All depicted CNVs can be found at the PDMutDB.

A relevant matter of ongoing debates is the pathogenic role of single heterozygous *PARK2* CNVs. Several studies have sought to address this issue, but the findings published so far are controversy and conflicting. Some reports indicate that CNVs heterozygous mutations in *PARK2* associate with increased PD risk [76,54,92], while others found no differences for association [77,37]. Also, examinations of family

pedigrees revealed heterozygous members with mild late-onset PD [93,94], or without typical clinical signs of the disease [37].

PINK1

Pathogenic mutations in *PINK1* (PTEN-induced kinase gene) are a less common cause of early-onset PD with a frequency variable from 1 to 9% depending on the ethnic background [95]. The encoded protein is a putative serine/threonine kinase of 581 amino acids involved in mitochondrial quality control and oxidative stress [96].

Homozygous and compound heterozygous mutations deletions involving different combinations of exons 4-8 have been described in both familial and sporadic early-onset cases coming from Japan, Brazil, Sudan and Iran (**Table 2**) [97-100,43]. A breakpoint analysis has been performed just in one of these patients, revealing a complex rearrangement involving the neighboring *DDOST* gene and maybe resulting from FoSTeS mechanism [100]. Moreover, single heterozygous cases have been described, albeit these mutations do not completely explain the recessive inheritance pattern. The largest heterozygous deletion known so far (56 kb) includes the entire *PINK1* genetic region, two neighboring genes, and two highly similar AluJo repeat sequences, which have been suggested as responsible of an unequal crossing-over [101]. Further heterozygous deletions involving exons 1, 3-8 and exon 7 have been described in familial or sporadic cases of early-onset PD (**Table 2**) [102,103,80].

CNVs overlapping PINK1 gene in PD						
CNV type	**Size**	**Ethnicity**	**Phenotype**	**Methodology**	**F-S**	**Reference**
Homoz. deletion	Exons 6-8	Japan	Early onset PD + dementia	n.a.	F	[98]
Homoz. deletion	~4600 bp	Japan	Early onset PD	n.a.	n.a.	[97]
Heteroz. deletion	~56kb	Italy	Definite PD	qPCR, FISH, Microsatellite markers analysis	S	[101]
Homoz. deletion	Exon 7	Brazil	Early onset PD	Sequencing, qPCR	S	[99]
Homoz. deletion	8,669 bp (Exons 4-8)	Sudan	Early onset PD	MLPA, sequencing	F	[100]
Heteroz. deletion	Exons 3-8	China	Early onset PD	qPCR	S	[102]
Heteroz. deletion	Exon 7	Spanish	Early onset PD with LBs	Sequencing, qPCR	F	[103]

Compound heteroz. deletion	Exon 2 + exons 2-4	Iran	Typical clinical features	MLPA, qPCR	S	
Homoz. deletion	Exon 5 and exon 4		Typical clinical features and PD with dementia		F	[43]
Heteroz. deletion	Exon 1	Brazil	Early onset PD	MLPA, qPCR	n.a.	[80]

Table 2 lists all the currently studies describing *PINK1* copy number changes in PD. The CNVs mutation type, the size of the mutation, the ethnicity of patients, the phenotype and the methodological approaches to measure quantitative genomic variations are reported. The column F-S reports if described cases are familial or sporadic.

PARK7

PARK7 was the third gene identified in 2001 as responsible of early-onset PD [104,105]. It encodes a conserved multifunctional protein belonging to the peptidase C56 family (also called DJ1) which acts as positive regulator of transcription, redox-sensitive chaperone, sensor for oxidative stress, and apparently protects neurons from ROS-induced apoptosis [106-108].

The proof that *PARK7* was a gene-causing disease came from a study on a Dutch family where members carried a 14 kb homozygous deletion involving the first five of seven exons [104]. Later, three siblings of Iranian origins born from consanguineous parents and carriers of a homozygous deletion of exon 5 have been reported (**Table 3**) [43]. Further heterozygous CNVs (both deletions and duplication) involving the exons of *DJ-1* gene have been published so far [102,109-111], although they do not completely explain the recessive pattern of the PD phenotype.

CNVs overlapping PARK7 gene in PD						
CNV type	**Size**	**Ethnicity**	**Phenotype**	**Methodology**	**F vs. S**	**Reference**
Homoz. Deletion	Exons 1-5 (14.082 bp)	Dutch	Autosomal Recessive Early-Onset PD	Microsatellite markers analysis, cloning, PCR, sequencing	F	[104]
Heteroz. Deletion	Exons 5-7	Caucasian (Tyrol, Austria)	Early-onset PD	Quantitative duplex PCR	S	[109]
Heteroz. Deletion	Exon 5	Serbian	Early-onset PD	qPCR	n.a.	[110]

Heteroz. Duplication	Exons 1-5	Dutch	Early-onset PD	MLPA, sequencing	S	[111]
Heteroz. Deletion	Exon 2	China	Early onset PD	qPCR	S	[102]
Homoz. Deletion	Exon 5	Iran	Typical clinical features	MLPA, qPCR	F	[43]

Table 3 lists all the currently studies describing *PARK7* copy number changes in PD. The CNVs mutation type, the size of the mutation, the ethnicity of patients, the phenotype and the methodological approaches to measure quantitative genomic variations are reported. The column F-S reports if described cases are familial or sporadic.

ATP13A2

ATP13A2 mutations are associated with Kufor-Rakeb syndrome (KRS), a form of recessively levodopa-responsive inherited atypical Parkinsonism [112]. It 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 [113]. To our knowledge, just one family from Iran with deletion of *ATP13A2* has been reported, including three affected siblings born from consanguineous parents and carriers of a homozygous deletion of exon 2 [43]. All three individuals presented moderate mental retardation, aggressive behaviors, visual hallucinations, supranuclear vertical gaze paresis, slow vertical saccades and dystonia. Cognitive function deteriorated rapidly, and all of them developed dementia by age 10. Further clinical and genetic follow-up of KRS patients will increase the knowledge of the natural history and clinical features of this syndrome.

THE 22q11.2 DELETION

A separate speech deserves the 22q11.2 deletion that lately is receiving more and more attention in PD field. Deletions at 22q11.2 are classically associated with a heterogeneous range of clinical syndromes, overall named 22q deletion syndrome (22qDS). The clinical phenotype of 22q deletion carriers varies widely, with 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. Despite the multiple system involvement, the association between 22q11.2 deletion and PD was not suspected

until the publication of independent case reports of co-occurrence of parkinsonism in patients with 22q11.2 deletion syndrome **(Table 4)** [114-116].

The interest in this possible link increased after Butcher and colleagues reported four patients with early-onset PD in their study of 159 adults with 22q11.2 deletion syndrome, founding that the use of antipsychotics in these patients delayed diagnosis of PD, and assessing after autopsy examination the presence of typical Lewy bodies and Lewy neurite formations too [117]. A couple of months ago, Mok et al. [118] performed the reverse experiment, namely pooling data from previous large PD case-control studies and assessing the frequency of 22q11.2 deletion carriers. Eight patients with PD and none of the controls had the deletion, providing a statistical significant association between the 22q deletion and an increased risk of developing the disease (**Table 4**). In according with this result, a single case-report from Virginia describes a 37-years-old early-onset PD patient carrying the 22q11.2 deletion but without any features of typical 22qDS [119]. All together, this evidence suggests 22q11.2 deletion might underlie early-onset PD, warning clinicians to take into consideration this genetic test as part of their evaluation for patients with early-onset PD.

The chromosome 22q11.2 region contains some excellent candidate genes for PD: *COMT* (or Catechol-O-Methyltransferase), a key regulator of synaptic dopamine levels and a target of inhibitory drugs for the treatment of wearing-off phenomena in PD patients [120]; *SEPT5*, a vescicle- and membrane-associated protein playing a significant role in inhibiting exocytosis, as well as a parkin substrate [121,122]; *DGCR8* that encodes a complex-subunit involved in the biogenesis of microRNAs, including miR-185 which is predicted to target *LRRK2* [123].

Interestingly, Perandones et al. [124] reported a case of mosaicism of a patient from the Ashkenazi Jewish ethnic group with a history of midline defects and PD onset at 46 years (**Table 4**). In this patient, FISH test detected a mosaicism of the 22q deletion in 24% of the analyzed blood cells, highlighting the relevance of performing individual cell-by-cell analysis.

CNVs involving the 22q11.2 region						
CNV type	**Size**	**Ethnicity**	**Phenotype**	**Methodology**	**Familial vs. Sporadic**	**Reference**
Hemiz. deletion	n.a.	n.a.	Childhood-Onset Schizophrenia Associated With Parkinsonism	FISH	No family history	[114]
Hemiz. deletion	3 Mb including COMT gene	n.a.	22qDS + Early onset PD	FISH	No family history	[115]
Hemiz. deletion	n.a.	n.a.	22qDS + PD (PD was considered a side effect of neuroleptic treatment or a clinical feature of early-onset PD)	FISH	n.a.	[116]
Hemiz. deletion	from 1,5 Mb to 3 Mb	Canada	22qDS + early-onset PD confirmed by neuropathological examination	FISH, quantitative real-time PCR	No family history	[117]
Mosaicism	n.a.	Ashkenazi Jewish	Midline defects and PD	FISH	Familial	[124]
Hemiz. deletion	2.88 Mb	n.a.	22q11.2DS + PD, dopa responsive	aCGH	Proband's father developed PD in later life	[119]
Deletion	Eight patients, whose deletion ranged around 3 Mb	n.a.	Both erly-onset and late-onset PD with typical motor signs and response to L-DOPA	Metanalysis of four previous indipendent studies and validation with CytoSure 15K-Array	n.a.	[118]

Table 4 lists all the currently studies describing 22q11.2 deletions in PD patients. The CNVs mutation type, the size of the mutation, the ethnicity of patients, the phenotype and the methodological approaches to measure quantitative genomic variations are reported. The column Familial vs. Sporadic reports if described cases are familial or sporadic PD.

HIGH-THROUGHPUT WHOLE-GENOME STUDIES TO MAP CNVs IN PD

The major reported PD-linked CNVs have actually been ascertained through single-gene investigations, and received most of the attention because of their already-known or hypothesized role in the disease. However, these mutations account only for a limited number of PD, and the vast majority of cases continue to remain without a valid explanation. Thanks to the rapid advancement of biotechnologies, scientists are now

able to scan entirely the human genome, producing high-quality ultra-dense genotypes and fast localization of genomic deletions and duplications. However, their applications in PD field are still not numerous, and only a few studies have investigated the overall contribution of global CNVs on PD etiology [125,54,126-128].

The first pilot analysis assessing the role of structural genetic variations in risk for PD was carried out in a population of 276 unique and unrelated Caucasian individual with PD by using two genome-wide SNP genotyping platforms and corrected metrics for CNVs interpretation [125]. In this study, along with several *PARK2* deletions and duplications confirmed by independent gene dosage experiments, a total of 182 genomic duplications and 161 heterozygous/homozygous deletions were measured, but no statistically significant regions associated with PD were identified. Among these CNVs, a subgroup (38 duplications and 44 deletions) was revealed only in patients and not in healthy controls or in DGV repository (http://dgv.tcag.ca/dgv/app/home), a web database collecting CNV alterations observed in the normal population (**Supplementary Table 1)**.

Some years later, Pankratz et al. [54] presented the results of a systematic CNV genome-wide analysis performed by using two CNV calling algorithms (PennCNV and QuantiSNP), two different association strategies (centric position and 400kb window) and multiple filters to improve the quality of CNVs calls. By intersection of results from all these criteria, they were able to replicate the association of PD susceptibility with *PARK2* CNVs, and then revealed two novel genes (*DOCK5* and *USP32*) associated with an increase in risk for PD at genome-wide significance (unfortunately not confirmed by independent molecular tests). Also in this study, a set of altered genetic regions were unique of PD patients (**Supplementary Table 1)**.

In order to identify novel CNVs and to evaluate their contribution to PD, Liu et al. [126] conducted a CNVs genome-wide scan in a case–control dataset (268 PD cases and 178 controls), focusing on a genetic isolate, the Ashkenazi Jewish population. Using high-confidence CNVs, they examined the global genome-wide burden of large and rare CNVs: this analysis did not reveal significant differences between cases and controls, but deletions were found 1.4 times more often in cases than controls. Interestingly, several rare genic CNVs were present in patients and absent in controls (**Supplementary Table 1)**. Among these, the duplication of *OVOS2* (ovostatin 2, a gene of unknown function) was classified as significant risk factors for PD. Other interesting PD-related CNVs alterations encompassed *NSF* and *WNT3* genes (later better

discussed), and *ATXN3, FBXW7, CHCHD3, HSF1, KLC1,* and *MBD3,* which participate in the PD disease pathways.

An unusual approach was performed by Pamphlett and colleagues [127], who investigated the existence of somatic candidate genetic CNVs missing in blood DNA in PD brains. A total of 45 PD -brain specific CNVs was found, some of which overlap with DGV regions. Candidate genes (not in controls nor in DGV) included *BCL2* involved in mitochondrial function and apoptosis (discussed in the following paragraphs), *NRSN1* implicated in cellular vesicle formation, and *RYR2* which participates in cellular calcium release (Supplementary Table 1). This study shows that specific-brain CNVs can be detected, and raises the possibility that brain-situated mutations could underlie some cases of PD.

A SYSTEMS BIOLOGY APPROACH FOR RARE AND SINGLETON CNVs

Altogether, genome-wide studies have revealed the existence of multiple genetic loci containing rare o singleton copy number changes in PD and not reported in control cohorts (**Supplementary Table 1)**. Although less frequent, these rare CNVs could represent potentially functional variants exerting small effects on PD pathogenesis, but not emphasized by single-gene investigations or association studies because do not reach a significant level of acceptance. These studies, in fact, are not the ideal approach for polygenic multifactorial diseases, where the pattern of allelic architecture could consist of hundreds of susceptibility loci acting together by modulating the disease itself. To overcome some of these limits, the *systems biology* perspective can be used in order to assess, in a comprehensive manner, the collective effect of these variants on PD outcome (**Figure 5**).

Interestingly, the Gene Ontologies (GO) enrichment of the total CNV-driven genes observed in PD patients until now reveals common deregulated biological processes (**Figure 6**) mainly related to nervous system functions and morphogenesis and including brain development (p-value= $2.137E^{-7}$), regulation of neurotransmission (p-value= $5.465E^{-7}$), neuronal signal transduction (p-value= $2.137E^{-6}$) and social behavior (p-value= $3.958E^{-7}$). Moreover, several potential relationships occur between rare CNV-affected genes and the currently known Mendelian PD genes (*SNCA*, *LRRK2*, *GBA*, *PARK2*, *PINK1*, *DJ1*, *VPS35, ATP13A2, PLA2G6, FBXO7, UCHL1, MAPT*). As shown in **Figure 7**, specific and meaningful associations

exist (i.e. proteins jointly contribute to a shared function, but this does not necessarily mean they are physically binding each other), and some rare CNV-affected genes could represent direct or indirect targets of Mendelian genes.

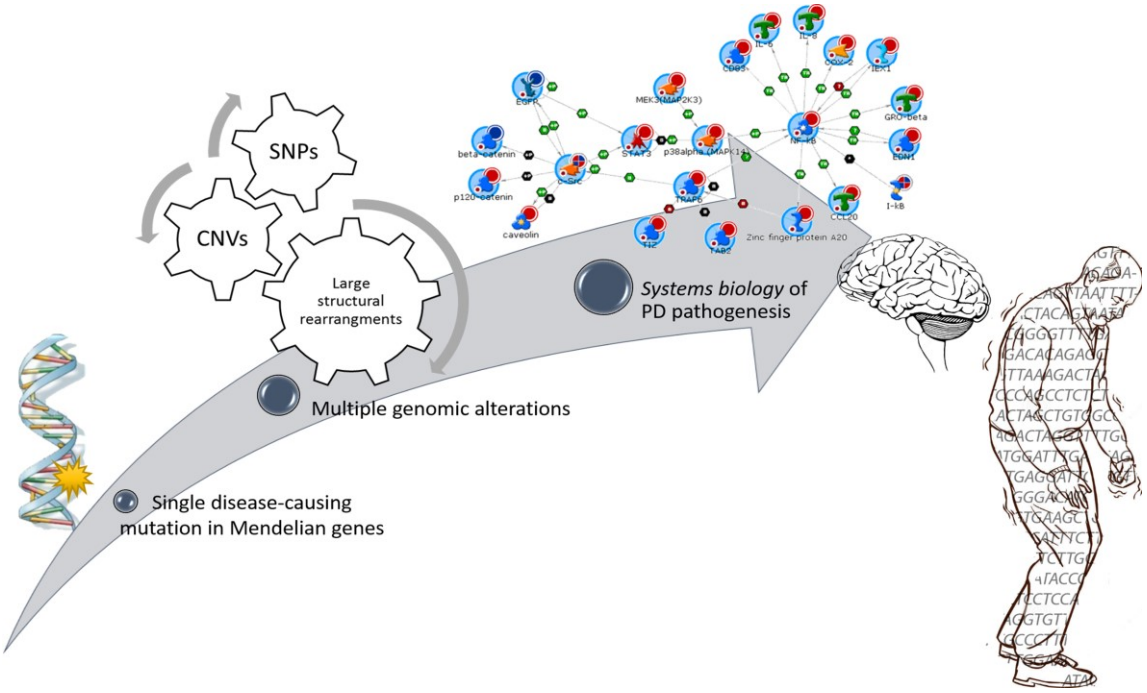


Figure 5. From a single-gene mutation perspective to a «systems biology» approach in order to dissect complex multifactorial diseases and improve the comprehension of the molecular basis underlying PD pathogenesis.

The global contribution of rare and singleton CNV-driven genes to nervous system pathophysiology and functions is also mirrored by the fact that, among those coinciding with the MIM MORBID/ORPHANET records, more than 50% are involved in syndromes with altered phenotypic nervous features, including ataxia conditions, neuropathies, dystrophies, learning and development disabilities and sensorineural disorders (**Supplementary Table 2**). All together, these findings support the evidence that uncommon individual CNVs may exert a susceptibility effect on PD, and strengthen the effectiveness of a *systems biology* approach to dissect complex multifactorial genetically heterogeneous diseases like PD.

Gene Ontology Processes	p-Value	FDR	Ratio
Protein complex assembly involved in synapse maturation	1.538e-7	4.755e-4	4/4
Hindbrain development	2.137e-7	4.755e-4	20/255
Multi-organism behavior	3.301e-7	4.755e-4	15/150
Social behavior	3.958e-7	4.755e-4	11/78
Regulation of neurotransmitter levels	5.465e-7	5.470e-4	22/321
Positive regulation of synaptic transmission, glutamatergic	8.174e-7	6.017e-4	8/39
Cerebellar granule cell differentiation or formation	9.016e-7	6.017e-4	6/18
Protein localization to synapse	2.029e-6	1.032e-3	7/31
Neuronal signal transduction	2.137e-6	1.032e-3	5/12
Guanylate kinase-associated protein clustering	2.234e-6	1.032e-3	4/6

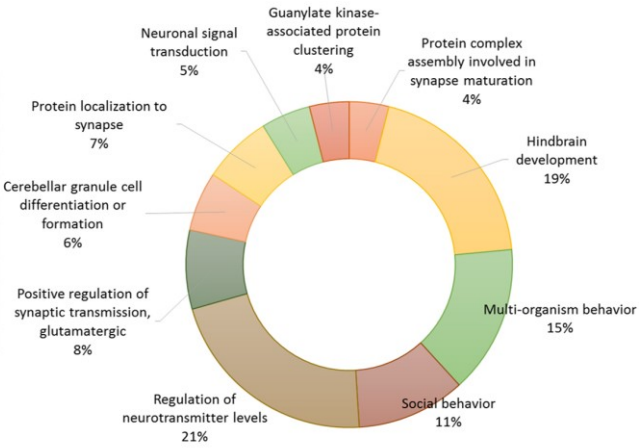


Figure 6. Gene Ontologies (GO) enrichment analysis of PD-specific CNV loci reveals biological processes relevant to PD pathogenesis. Panel A: Representation of the top 10 most significantly enriched (FDR < 0.05) canonical GO biological processes associated with candidate PD genes with copy number alterations (not reported in controls or DGV). The analysis was performed using MetaCore platform (GeneGo, Thompson Reuters). The list is arranged in descending order with the most significant biological process at the top. Detailed information about the entire list of CNVs and overlapping genes are reported in Supplementary Table 1. P-values have been obtained through hypergeometric analysis and corrected by FDR (false discovery rate) method. Panel B: Pie chart representing the percentage of genes with altered copy number in PD belonging to the top 10 enriched (P<0.05) GO Biological Processes.

Below, we will briefly discuss the GO-enriched rare genes by grouping them into the three main categories (Synaptic trafficking and neurotransmission, Brain development and cell fate differentiation, Cognitive impairment). GO-enriched genes are graphically illustrated in **Figure 8**.

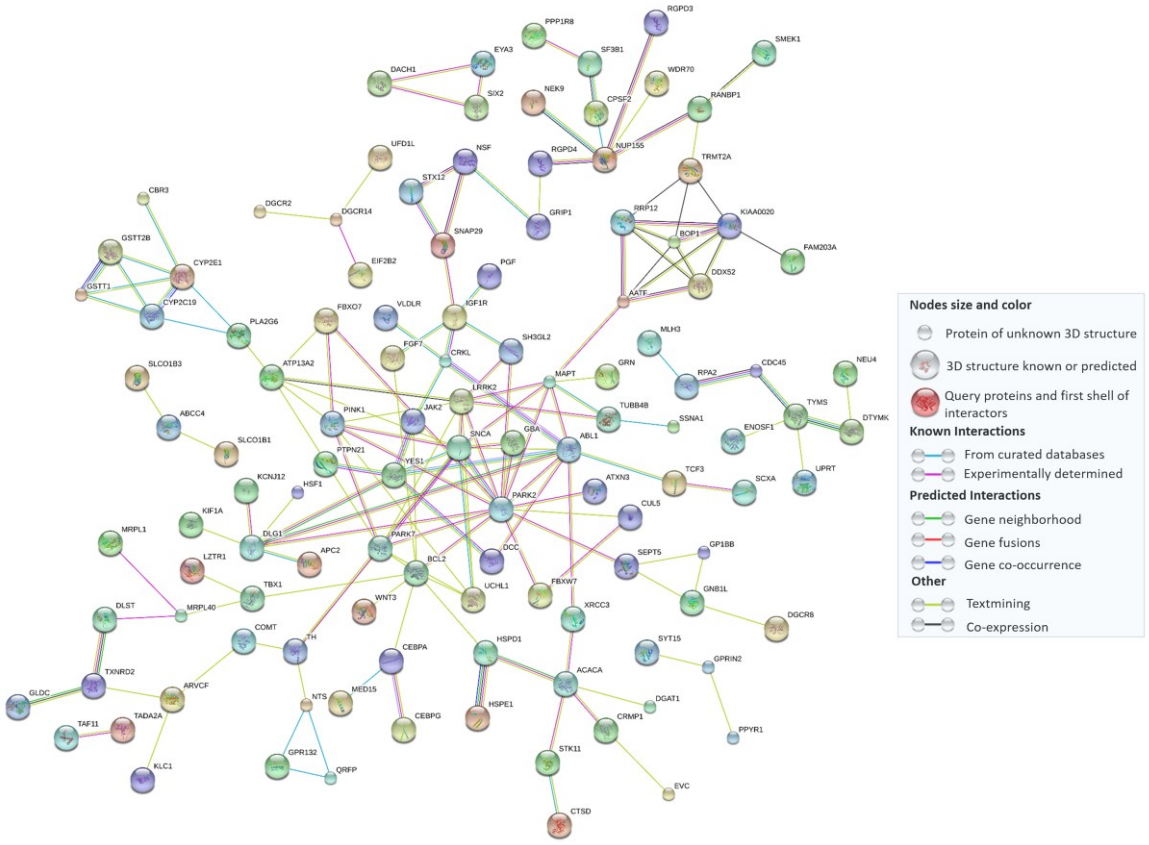


Figure 7. Potential protein-protein interactions among rare PD CNVs and the currently known Mendelian genes, as showed by STRING Software v.10 (http://string-db.org/) with high confidence settings (0.700 as minimum interaction score). The legend displays the meaning of nodes and edges.

Synaptic trafficking and neurotransmission. Synapses are specialized junctions of the central nervous system through which neurons connect each other to form extensive neural circuits. Synaptic functioning depends on a constant supply of energy and resources, essential for both neurotransmitters production and intracellular trafficking *via* repeated synaptic vescicle cycles. Alterations in synaptic stability result in a disruption of the neuronal networks, a common hallmark of several neurodegenerative conditions, including PD, Huntington's and Alzheimer's.

Along with the previously described *COMT*, some rare CNV-altered genes are involved in synaptic neurotransmission. One of the most interesting is the deletion of the entire *TH* (Tyrosine hydroxylase) gene, detected in a PD patient without evidence for dystonia but responsive to L-DOPA treatment, and in none of the controls (**Supplementary Table 1**) [128,129]. *TH* encodes a monooxygenase that catalyzes the conversion of L-tyrosine to L-dihydroxyphenylalanine (L-DOPA), the rate-limiting step in dopamine biosynthesis. Consistent with the essential role of TH in dopamine

homeostasis, missense mutations in *TH* have been previously investigated, providing links with severe Parkinsonism-related phenotypes, such as Segawa's syndrome, L-DOPA responsive infantile Parkinsonism, or L-DOPA-responsive dystonia (DRD) in the recessive form [129].

CNV-affected genes could alter not only dopamine metabolism but also other neurotransmitters signaling pathways. Convergent evidence agrees for early alterations in the cholinergic system in PD [130,131]. To this regard, CNVs in the *BChE* gene, a nonspecific cholinesterase enzyme that hydrolyses many different choline-based esters, and in the synaptic choline transporter *SLC5A7* have been reported in some PD patients (**Supplementary Table 1**). Moreover, glutamate plays a central role in basal ganglia circuitry, sometimes modulated by dopamine itself. Genetic variations in the glutamate receptors *GRID1,* observed in some PD patients, could conceivably affect either the risk of developing PD or the phenotype.

DLG1, observed deleted in a PD patient, encodes a multi-domain scaffolding protein acting in septate junction formation, signal transduction, cell proliferation and synaptogenesis. Interestingly, this gene was identified as differentially expressed in the blood of PD patients versus controls and was suggested at high confidence as candidates biomarker for PD [132]. Moreover, recent studies have highlighted the role of *DLG1* in the regulation of 5-HT2AR endocytosis and signaling [133].

Several genes working in the synaptic vesicular cargo trafficking are affected by CNV alterations, such as *SEPT5* and *SNAP29* (both overlapping the 22q11.2 deletion), *NSF, SYT15*, *PCLO,* and *KLC1*. A particular focus deserves *NSF* (also known as N-ethylmaleimide sensitive factor), previously identified as "top-hit" in a large GWAS meta-study (rs183211) [134]. While *NSF* functions in vesicular trafficking, membrane fusion, and synaptic neurotransmission are well documented, some recent studies also suggest a direct interaction between NSF and the Dopamine D1 receptor (D1R) [135]. This gene was also experimentally showed to be the directed target of miR-4519, a microRNA whose genetic variants are strongly associated with PD [136].

An interesting deletion concerns the *ADCYAP1* gene, encoding the pleiotropic bioactive peptide PACAP (or Pituitary adenylate cyclase activating polypeptide). PACAP is considered a potent neurotrophic and neuroprotective factor, playing an important role during the embryonic development of the nervous system, and protecting neurons against toxic insults and neurodegeneration [137,138]. In the specific case of PD pathology, PACAP has been demonstrated to safeguard *in vitro* PD

cell model against both salsolinol-induced and inflammatory-mediated toxicity [139,140], to protect rat dopaminergic neurons after injection of 6-OHDA into the *substantia nigra* [141], and to prevent Parkinson-like neuronal loss and motor deficits induced by prostaglandin J2 [142].

Further genes founded CNV-altered and involved in synaptic functions include ion channels (*KCNJ12*), biochemical enzymes (*NAT8L, PPP1R8*) and receptors (*RTN4R*).

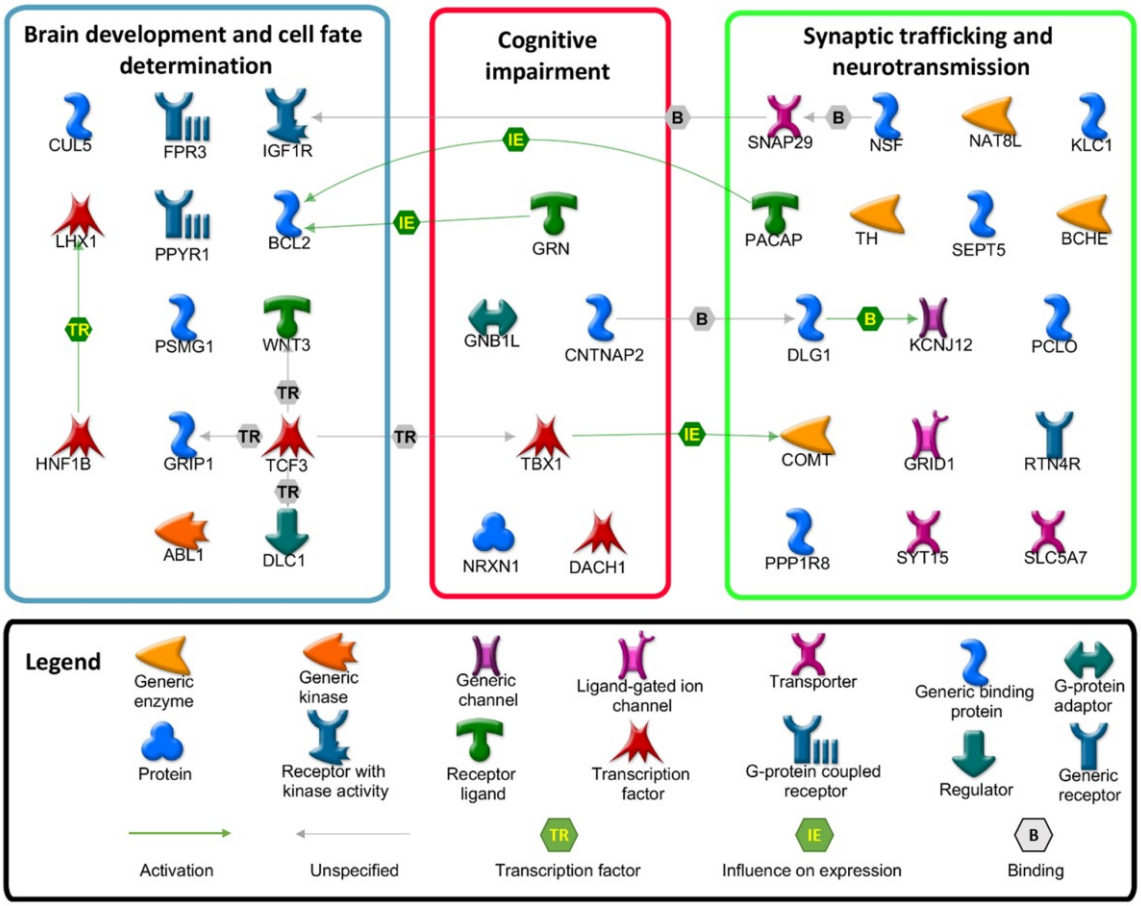


Figure 8. Interaction map representing CNV-altered genes enriched in GO classes and grouped on the basis of their main biological processes. The map was created using the MetaCore Pathway Map Creator tool (GeneGo, Thompson Reuters). Detailed information about type of CNV, chromosomal size and study references are reported in Supplementary Table 1.

Brain development and cell fate determination. Brain development is an orchestrated, tightly regulated, and genetically programmed process with influences from the environment. Alterations in genes regulating differentiation, elaboration, and maintenance of neuronal cells can compromise neural specification events and cellular homeostasis, turning into neurodevelopmental abnormalities and neurodegenerative diseases [143].

Despite dopamine neurons cell death is the leading event in PD pathology, it also physiologically occurs in developing brain during embryogenesis [144]. The proper development of dopaminergic system requires the action of BCL2 family members, responsible for dictating cell survival or commitment to apoptosis [144]. To this regard, a somatic deletion in *BCL2* has been observed in the brain of a PD patient in a homozygous state (**Supplementary Table 1**), likely producing deleterious effects on gene function.

Neuronal migration, differentiation, and death during brain development is also carefully tuned by a vast repertoire of growth and transcription factors, such as *LHX1* and *IGF1R,* observed duplicated in PD patients (**Supplementary Table 1**). *LHX1* encodes a transcription factor involved in axonal guidance and neurogenesis, and its overexpression has been demonstrated to inhibit the correct mesencephalic dopaminergic neurons differentiation [145]. *IGF1R* produces the tyrosine kinase receptor for the IGF1 (insulin-like growth factor) signaling pathway. This pathway was found dysregulated in a previous cross-sectional transcriptomic analysis performed on PD datasets [146], and IGF1 signaling inhibitors have been proposed as promising therapies for the treatment of various late-onset neurodegenerative disorders [147]. Moreover, the IGF1R reduction *per se* triggers protective effects in neurodegenerative conditions [148], suggesting that its duplication could have a deleterious impact on neuronal life.

Members of the WNT family are additional factors expressed and secreted in the midbrain, involved in regulation of cell fate and patterning during embryogenesis. Among WNT family members, a singleton deletion in *WNT3* has been described in PD (**Supplementary Table 1**). *WNT3* is located near *MAPT* locus and, according to previous genome-wide association studies, variations in its genetic regions can influence the risk of developing PD [149,134]. Moreover, the WNT/β-catenin signaling pathway is able to control the dopaminergic cell commitment in the midbrain, and is mediated by several transcription factors of TCF family [150], such as *TCF3* observed deleted in PD (**Supplementary Table 1**).

Another interesting rare CNV-affected gene is *ABL1* (**Supplementary Table 1**), which encodes the tyrosine kinase protein c-Abl controlling neurogenesis, neurite outgrowth, and neuronal plasticity. Several lines of evidence suggest that aberrant activation of c-Abl plays an important role in PD pathogenesis: c-Abl is upregulated in postmortem striatum of PD patients and its phosphorylation at Tyr412 is enhanced in

substantia nigra and striatum; (ii) c-Abl phosphorylates parkin and impairs its E3 ligase activity, leading to loss of dopaminergic neurons in SN; (iii) c-Abl regulates the clearance of α-syn, catalyzing its phosphorylation mainly at Tyr39 and to a lesser extent at Tyr125; (iv) the inhibition of c-Abl activity by several drugs (imatinib/Gleevec, nilotinib/Tasigna, bafetinib/INNO-406) protects against the loss of dopaminergic neurons in wild-type mice [151].

Additional genes founded CNV-altered and involved in brain development, neuronal migration, and synaptic plasticity include G-coupled proteins (*FPR3, PPYR1*), transcription factors and coactivator (*HNF1B, GRIP1*), elements of ubiquitin/proteasome system (*CUL5, PSMG1*) and a member of the rhoGAP family (*DLC1*).

Cognitive impairment. The core feature of cognitive decline associated with PD is represented by an impairment of executive functions. Deficits in planning, sequencing, and execution of complex-goal-directed behavior are usually reported; working memory, episodic memory, procedural learning and attention are compromised with presence of attentive fluctuations [152].

A set of rare CNV-altered genes are involved in learning, behavior and cognitive dysfunctions, including *NRXN1*, *CNTNAP2, GRN*, *TBX1*, *GNB1L* and *DACH1*.

Neurexin 1, encoded by *NRXN1*, is a presynaptic neuronal adhesion molecule that interacts with postsynaptic neuroligins in both glutamatergic and GABAergic synapses, and is important in synaptic specification and efficient neurotransmission. Deletions and point mutations in *NRXN1* are associated with a broad spectrum of neuropsychiatric and neurodevelopmental disorders, including autism, intellectual disability, epilepsy, developmental delay, and schizophrenia [153]. As *NRXN1,* also *CNTNAP2* belongs to neurexin superfamily and encodes a neuronal transmembrane protein involved in neural-glia interactions and clustering of potassium channels in myelinated axons. Variations in this gene have been involved in susceptibility to neurodevelopmental disorders and language impairment [154]. Moreover, downregulation of *CNTNAP2* has been associated with AD and PD conditions [154,155]. It has been suggested that neurexins-neuroligins level fluctuations sway the balance between excitatory and inhibitory neurotransmission, leading to damage of synapses and dendrites and maybe triggering protein aggregates in neurodegenerative conditions [156].

GRN encodes progranulin, a multifunction protein widely distributed throughout the CNS primarily in neurons and microglia and a potent autocrine neurotrophic factor and regulator of neuroinflammation [157]. Its loss-of-function mutations are known to be responsible for FTLDU-17 (ubiquitin-positive frontotemporal lobar degeneration linked to chromosome 17) and increase the risk for both Alzheimer's and PD, suggesting important roles of progranulin in neurodegenerative processes [158]. A deletion of *GRN* exons 1–11, resulting from a non-homologous recombination event, has been observed in a patient with typical GRN neuropathology, and in his sister presenting PD [159]. Moreover, it has been recently demonstrated that progranulin gene delivery protects dopaminergic neurons in a PD model, suggesting that GRN gene therapy may have beneficial effects in the treatment of PD [157].

Other CNV-driven genes include *TBX1* and *GNB1L,* both overlapping the 22q11.2 deletion, which have been associated with neuropsychiatric disorders as schizophrenia and autism [160-162].

CONCLUSIONS

A number of evidence suggests an extensive and complex genetic action of CNVs on PD etiopathogenesis. Thus far, unfortunately, only a small portion of the genetic variance has been identified; the remaining substantial components remain unknown and urgently need to be addressed. One way we can move on is using "*systems biology*", a worthwhile instrument to analyze complex biological processes and generate a more definite molecular picture of PD.

In this review, we showed that disregarded individual uncommon CNVs functionally act in common deregulated biological processes relevant for PD pathogenesis and therefore, potentially account for a portion of the "missing heritability" underlying PD. The comprehensive detection and functional characterization of rare CNVs in PD patients may be helpful to generate a more defined molecular picture of this complex disease, by revealing new candidate genes or disease-related molecular mechanisms, finally leading to improved diagnosis and counseling of mutation carriers. The forthcoming new era of genomics data promises to increase resolution and uncover new interesting clues.

ACKNOWLEDGEMENTS

This work was supported by the Italian Ministry of Education, Universities and Research through grant CTN01_00177_817708 and the international Ph.D. program in Neuroscience of the University of Catania. Authors gratefully acknowledge Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo and Francesco Marino for their administrative and technical support.

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there is no conflict of interest.

ABBREVIATIONS

ABL proto-oncogene 1, non-receptor tyrosine kinase (*ABL1);* adenylate cyclase activating polypeptide 1 (*ADCYAP1);* ATPase 13A2 (*ATP13A2);* ataxin 3 *(ATXN3);* butyrylcholinesterase (*BCHE*); apoptosis regulator *(BCL2);* coiled-coil-helix-coiled-coil-helix domain containing 3 *(CHCHD3);* contactin associated protein-like 2 (*CNTNAP2);* Copy Number Variations *(*CNVs); catechol-O-methyltransferase (*COMT*); cullin 5 (*CUL5);* dopamine receptor D1 (D1R); dachshund family transcription factor 1 (*DACH1);* dolichyl-diphosphooligosaccharide--protein glycosyltransferase non-catalytic subunit (*DDOST);* DGCR8 microprocessor complex subunit (*DGCR8);* DLC1 Rho GTPase activating protein (*DLC1);* discs large MAGUK scaffold protein 1 (*DLG1);* dedicator of cytokinesis 5 *(DOCK5);* F-box protein 7 (*FBXO7);* F-box and WD repeat domain containing 7 (*FBXW7);* Fluorescent in situ hybridization *(*FISH), fork stalling and template switching (FoSTeS); formyl peptide receptor 3 (*FPR3);* ubiquitin-positive frontotemporal lobar degeneration linked to chromosome 17 (FTLDU-17); cyclin G associated kinase- diacylglycerol kinase theta *(GAK-DGKQ);* glucosylceramidase beta (*GBA);* glucosylceramidase beta pseudogene 1 *(GBAP1),* G protein subunit beta 1 like (*GNB1L);* glutamate ionotropic receptor delta type subunit 1 (*GRID1);* glutamate receptor interacting protein 1 (*GRIP1);* granulin (*GRN);* HNF1 homeobox B (*HNF1B);* heat shock transcription factor 1 (*HSF1);* HtrA serine peptidase 2 (*HTRA2);* Insulin growth factor 1 *(*IGF1), insulin like growth factor 1 receptor (*IGF1R);* potassium voltage-gated channel subfamily J member 12 (*KCNJ12);* kinesin light chain 1 (*KLC1);* LIM homeobox 1 (*LHX1);* leucine rich repeat kinase 2 (*LRRK2);* microtubule associated protein tau (*MAPT);* methyl-CpG binding domain protein 3

(*MBD3);* non-allelic homologous recombination (NAHR); N-acetyltransferase 8 like (*NAT8L);* non-homologous end joining (NHEJ); neurensin 1 (*NRSN1);* neurexin 1 (*NRXN1);* N-ethylmaleimide sensitive factor, vesicle fusing ATPase (*NSF);* ovostatin 2 (*OVOS2);* Pituitary adenylate cyclase activating polypeptide (PACAP); Parkinson disease 10 (*PARK10);* Parkinson disease 12 (*PARK12);* Parkinson disease 16 (*PARK16);* parkin RBR E3 ubiquitin protein ligase (*PARK2);* Parkinson disease 3 (autosomal dominant, Lewy body) (*PARK3);* Parkinsonism associated deglycase (*PARK7);* piccolo presynaptic cytomatrix protein (*PCLO);* PTEN induced putative kinase 1 (*PINK1);* phospholipase A2 group VI *(PLA2G6);* protein phosphatase 1 regulatory subunit 8 (*PPP1R8);* neuropeptide Y receptor Y4 *(PPYR1*); proteasome assembly chaperone 1 (*PSMG1);* reticulon 4 receptor (*RTN4R);* ryanodine receptor 2 (*RYR2);* septin 5 (*SEPT5);* solute carrier family 5 member 7 (*SLC5A7);* synaptosome associated protein 29 (*SNAP29);* alpha-synuclein *(SNCA);* synaptojanin 1 (*SYNJ1);* synaptotagmin 15 (*SYT15);* T-box 1 (*TBX1);* transcription factor 3 (*TCF3);* tyrosine hydroxylase *(TH);* ubiquitin C-terminal hydrolase L1 (*UCHL1);* ubiquitin proteasome system (UPS); ubiquitin specific peptidase 32 (*USP32);* VPS35 retromer complex component *(VPS35);* Wnt family member 3 (*WNT3);* serotonin 2A receptor (5-HT2AR).

REFERENCES

1. Moore DJ, West AB, Dawson VL, Dawson TM (2005) Molecular pathophysiology of Parkinson's disease. Annual review of neuroscience 28:57-87. doi:10.1146/annurev.neuro.28.061604.135718

2. Toft M, Ross OA (2010) Copy number variation in Parkinson's disease. Genome medicine 2 (9):62. doi:10.1186/gm183

3. Michel PP, Hirsch EC, Hunot S (2016) Understanding Dopaminergic Cell Death Pathways in Parkinson Disease. Neuron 90 (4):675-691. doi:10.1016/j.neuron.2016.03.038

4. Winklhofer KF, Haass C (2010) Mitochondrial dysfunction in Parkinson's disease. Biochimica et biophysica acta 1802 (1):29-44. doi:10.1016/j.bbadis.2009.08.013

5. Dias V, Junn E, Mouradian MM (2013) The role of oxidative stress in Parkinson's disease. Journal of Parkinson's disease 3 (4):461-491. doi:10.3233/JPD-130230

6. Betarbet R, Sherer TB, Greenamyre JT (2005) Ubiquitin-proteasome system and Parkinson's diseases. Experimental neurology 191 Suppl 1:S17-27. doi:10.1016/j.expneurol.2004.08.021

7. Wang Q, Liu Y, Zhou J (2015) Neuroinflammation in Parkinson's disease and its potential as therapeutic target. Translational neurodegeneration 4:19. doi:10.1186/s40035-015-0042-0

8. Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science 276 (5321):2045-2047

9. Lewis PA, Cookson MR (2012) Gene expression in the Parkinson's disease brain. Brain Res Bull 88

(4):302-312. doi:10.1016/j.brainresbull.2011.11.016

10. Bonifati V (2014) Genetics of Parkinson's disease--state of the art, 2013. Parkinsonism & related disorders 20 Suppl 1:S23-28. doi:10.1016/S1353-8020(13)70009-9

11. Klein C, Westenberger A (2012) Genetics of Parkinson's disease. Cold Spring Harbor perspectives in medicine 2 (1):a008888. doi:10.1101/cshperspect.a008888

12. Lubbe S, Morris HR (2014) Recent advances in Parkinson's disease genetics. Journal of neurology 261 (2):259-266. doi:10.1007/s00415-013-7003-2

13. Trinh J, Farrer M (2013) Advances in the genetics of Parkinson disease. Nature reviews Neurology 9 (8):445-454. doi:10.1038/nrneurol.2013.132

14. Gamazon ER, Stranger BE (2015) The impact of human copy number variation on gene expression. Briefings in functional genomics 14 (5):352-357. doi:10.1093/bfgp/elv017

15. Zarrei M, MacDonald JR, Merico D, Scherer SW (2015) A copy number variation map of the human genome. Nature reviews Genetics 16 (3):172-183. doi:10.1038/nrg3871

16. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW (2014) The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic acids research 42 (Database issue):D986-992. doi:10.1093/nar/gkt958

17. Cantsilieris S, Baird PN, White SJ (2013) Molecular methods for genotyping complex copy number polymorphisms. Genomics 101 (2):86-93. doi:10.1016/j.ygeno.2012.10.004

18. Carter NP, Fiegler H, Piper J (2002) Comparative analysis of comparative genomic hybridization microarray technologies: report of a workshop sponsored by the Wellcome Trust. Cytometry 49 (2):43-48. doi:10.1002/cyto.10153

19. Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, Firth H, Sanlaville D, Winter R, Colleaux L, Bobrow M, Carter NP (2004) Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. Journal of medical genetics 41 (4):241-248

20. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C (2004) Detection of large-scale variation in the human genome. Nature genetics 36 (9):949-951. doi:10.1038/ng1416

21. Inazawa J, Inoue J, Imoto I (2004) Comparative genomic hybridization (CGH)-arrays pave the way for identification of novel cancer-related genes. Cancer science 95 (7):559-563

22. Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL (2004) A tiling resolution DNA microarray with complete coverage of the human genome. Nature genetics 36 (3):299-303. doi:10.1038/ng1307

23. Lee JA, Lupski JR (2006) Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 52 (1):103-121. doi:10.1016/j.neuron.2006.09.027

24. Ambroziak W, Koziorowski D, Duszyc K, Gorka-Skoczylas P, Potulska-Chromik A, Slawek J, Hoffman-Zacharska D (2015) Genomic instability in the PARK2 locus is associated with Parkinson's disease. Journal of applied genetics 56 (4):451-461. doi:10.1007/s13353-015-0282-9

25. Lee JA, Carvalho CM, Lupski JR (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. Cell 131 (7):1235-1247. doi:10.1016/j.cell.2007.11.037

26. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, Kraft K, Kempfer R, Jerkovic I, Chan WL, Spielmann M, Timmermann B, Wittler L, Kurth I, Cambiaso P, Zuffardi O, Houge G, Lambie L, Brancati F, Pombo A, Vingron M, Spitz F, Mundlos S (2016) Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature 538 (7624):265-269. doi:10.1038/nature19800

27. Soldner F, Stelzer Y, Shivalila CS, Abraham BJ, Latourelle JC, Barrasa MI, Goldmann J, Myers RH, Young RA, Jaenisch R (2016) Parkinson-associated risk variant in distal enhancer of alpha-synuclein

modulates target gene expression. Nature 533 (7601):95-99. doi:10.1038/nature17939

28. Spielmann M, Mundlos S (2016) Looking beyond the genes: the role of non-coding variants in human disease. Human molecular genetics 25 (R2):R157-R165. doi:10.1093/hmg/ddw205

29. Lupianez DG, Spielmann M, Mundlos S (2016) Breaking TADs: How Alterations of Chromatin Domains Result in Disease. Trends in genetics : TIG 32 (4):225-237. doi:10.1016/j.tig.2016.01.003

30. Spielmann M, Mundlos S (2013) Structural variations, the regulatory landscape of the genome and their alteration in human disease. BioEssays : news and reviews in molecular, cellular and developmental biology 35 (6):533-543. doi:10.1002/bies.201200178

31. Hehir-Kwa JY, Pfundt R, Veltman JA, de Leeuw N (2013) Pathogenic or not? Assessing the clinical relevance of copy number variants. Clinical genetics 84 (5):415-421. doi:10.1111/cge.12242

32. Gu W, Lupski JR (2008) CNV and nervous system diseases--what's new? Cytogenetic and genome research 123 (1-4):54-64. doi:10.1159/000184692

33. Kalman B, Vitale E (2009) Structural chromosomal variations in neurological diseases. The neurologist 15 (5):245-253. doi:10.1097/NRL.0b013e3181963cef

34. Hoyer H, Braathen GJ, Eek AK, Nordang GB, Skjelbred CF, Russell MB (2015) Copy number variations in a population-based study of Charcot-Marie-Tooth disease. BioMed research international 2015:960404. doi:10.1155/2015/960404

35. Olson H, Shen Y, Avallone J, Sheidley BR, Pinsky R, Bergin AM, Berry GT, Duffy FH, Eksioglu Y, Harris DJ, Hisama FM, Ho E, Irons M, Jacobsen CM, James P, Kothare S, Khwaja O, Lipton J, Loddenkemper T, Markowitz J, Maski K, Megerian JT, Neilan E, Raffalli PC, Robbins M, Roberts A, Roe E, Rollins C, Sahin M, Sarco D, Schonwald A, Smith SE, Soul J, Stoler JM, Takeoka M, Tan WH, Torres AR, Tsai P, Urion DK, Weissman L, Wolff R, Wu BL, Miller DT, Poduri A (2014) Copy number variation plays an important role in clinical epilepsy. Annals of neurology 75 (6):943-958. doi:10.1002/ana.24178

36. Grayton HM, Fernandes C, Rujescu D, Collier DA (2012) Copy number variations in neurodevelopmental disorders. Progress in neurobiology 99 (1):81-91. doi:10.1016/j.pneurobio.2012.07.005

37. Wang L, Nuytemans K, Bademci G, Jauregui C, Martin ER, Scott WK, Vance JM, Zuchner S (2013) High-resolution survey in familial Parkinson disease genes reveals multiple independent copy number variation events in PARK2. Human mutation 34 (8):1071-1074. doi:10.1002/humu.22344

38. Stefanis L (2012) alpha-Synuclein in Parkinson's disease. Cold Spring Harbor perspectives in medicine 2 (2):a009399. doi:10.1101/cshperspect.a009399

39. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, Lincoln S, Crawley A, Hanson M, Maraganore D, Adler C, Cookson MR, Muenter M, Baptista M, Miller D, Blancato J, Hardy J, Gwinn-Hardy K (2003) alpha-Synuclein locus triplication causes Parkinson's disease. Science 302 (5646):841. doi:10.1126/science.1090278

40. Keyser RJ, Lombard D, Veikondis R, Carr J, Bardien S (2010) Analysis of exon dosage using MLPA in South African Parkinson's disease patients. Neurogenetics 11 (3):305-312. doi:10.1007/s10048-009-0229-6

41. Sekine T, Kagaya H, Funayama M, Li Y, Yoshino H, Tomiyama H, Hattori N (2010) Clinical course of the first Asian family with Parkinsonism related to SNCA triplication. Movement disorders : official journal of the Movement Disorder Society 25 (16):2871-2875. doi:10.1002/mds.23313

42. Kojovic M, Sheerin UM, Rubio-Agusti I, Saha A, Bras J, Gibbons V, Palmer R, Houlden H, Hardy J, Wood NW, Bhatia KP (2012) Young-onset parkinsonism due to homozygous duplication of alpha-synuclein in a consanguineous family. Movement disorders : official journal of the Movement Disorder Society 27 (14):1827-1829. doi:10.1002/mds.25199

43. Darvish H, Movafagh A, Omrani MD, Firouzabadi SG, Azargashb E, Jamshidi J, Khaligh A, Haghnejad L, Naeini NS, Talebi A, Heidari-Rostami HR, Noorollahi-Moghaddam H, Karkheiran S,

Shahidi GA, Paknejad SM, Ashrafian H, Abdi S, Kayyal M, Akbari M, Pedram N, Emamalizadeh B (2013) Detection of copy number changes in genes associated with Parkinson's disease in Iranian patients. Neuroscience letters 551:75-78. doi:10.1016/j.neulet.2013.07.013

44. Olgiati S, Thomas A, Quadri M, Breedveld GJ, Graafland J, Eussen H, Douben H, de Klein A, Onofrj M, Bonifati V (2015) Early-onset parkinsonism caused by alpha-synuclein gene triplication: Clinical and genetic findings in a novel family. Parkinsonism & related disorders 21 (8):981-986. doi:10.1016/j.parkreldis.2015.06.005

45. Ferese R, Modugno N, Campopiano R, Santilli M, Zampatti S, Giardina E, Nardone A, Postorivo D, Fornai F, Novelli G, Romoli E, Ruggieri S, Gambardella S (2015) Four Copies ofSNCAResponsible for Autosomal Dominant Parkinson's Disease in Two Italian Siblings. Parkinson's Disease 2015:1-6. doi:10.1155/2015/546462

46. Uchiyama T, Ikeuchi T, Ouchi Y, Sakamoto M, Kasuga K, Shiga A, Suzuki M, Ito M, Atsumi T, Shimizu T, Ohashi T (2008) Prominent psychiatric symptoms and glucose hypometabolism in a family with a SNCA duplication. Neurology 71 (16):1289-1291. doi:10.1212/01.wnl.0000327607.28928.e6

47. Ahn TB, Kim SY, Kim JY, Park SS, Lee DS, Min HJ, Kim YK, Kim SE, Kim JM, Kim HJ, Cho J, Jeon BS (2008) alpha-Synuclein gene duplication is present in sporadic Parkinson disease. Neurology 70 (1):43-49. doi:10.1212/01.wnl.0000271080.53272.c7

48. Nishioka K, Ross OA, Ishii K, Kachergus JM, Ishiwata K, Kitagawa M, Kono S, Obi T, Mizoguchi K, Inoue Y, Imai H, Takanashi M, Mizuno Y, Farrer MJ, Hattori N (2009) Expanding the clinical phenotype of SNCA duplication carriers. Movement disorders : official journal of the Movement Disorder Society 24 (12):1811-1819. doi:10.1002/mds.22682

49. Nishioka K, Hayashi S, Farrer MJ, Singleton AB, Yoshino H, Imai H, Kitami T, Sato K, Kuroda R, Tomiyama H, Mizoguchi K, Murata M, Toda T, Imoto I, Inazawa J, Mizuno Y, Hattori N (2006) Clinical heterogeneity of alpha-synuclein gene duplication in Parkinson's disease. Annals of neurology 59 (2):298-309. doi:10.1002/ana.20753

50. Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S, Levecque C, Larvor L, Andrieux J, Hulihan M, Waucquier N, Defebvre L, Amouyel P, Farrer M, Destee A (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. Lancet 364 (9440):1167-1169. doi:10.1016/S0140-6736(04)17103-1

51. Ibanez P, Bonnet AM, Debarges B, Lohmann E, Tison F, Pollak P, Agid Y, Durr A, Brice A (2004) Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. Lancet 364 (9440):1169-1171. doi:10.1016/S0140-6736(04)17104-3

52. Ibanez P, Lesage S, Janin S, Lohmann E, Durif F, Destee A, Bonnet AM, Brefel-Courbon C, Heath S, Zelenika D, Agid Y, Durr A, Brice A, French Parkinson's Disease Genetics Study G (2009) Alpha-synuclein gene rearrangements in dominantly inherited parkinsonism: frequency, phenotype, and mechanisms. Archives of neurology 66 (1):102-108. doi:10.1001/archneurol.2008.555

53. Sironi F, Trotta L, Antonini A, Zini M, Ciccone R, Della Mina E, Meucci N, Sacilotto G, Primignani P, Brambilla T, Coviello DA, Pezzoli G, Goldwurm S (2010) alpha-Synuclein multiplication analysis in Italian familial Parkinson disease. Parkinsonism & related disorders 16 (3):228-231. doi:10.1016/j.parkreldis.2009.09.008

54. Pankratz N, Dumitriu A, Hetrick KN, Sun M, Latourelle JC, Wilk JB, Halter C, Doheny KF, Gusella JF, Nichols WC, Myers RH, Foroud T, DeStefano AL, Psg P, GenePd Investigators C, Molecular Genetic L (2011) Copy number variation in familial Parkinson disease. PloS one 6 (8):e20988. doi:10.1371/journal.pone.0020988

55. Elia AE, Petrucci S, Fasano A, Guidi M, Valbonesi S, Bernardini L, Consoli F, Ferraris A, Albanese A, Valente EM (2013) Alpha-synuclein gene duplication: marked intrafamilial variability in two novel pedigrees. Movement disorders : official journal of the Movement Disorder Society 28 (6):813-817. doi:10.1002/mds.25518

56. Konno T, Ross OA, Puschmann A, Dickson DW, Wszolek ZK (2016) Autosomal dominant Parkinson's disease caused by SNCA duplications. Parkinsonism & related disorders 22 Suppl 1:S1-6. doi:10.1016/j.parkreldis.2015.09.007

57. Ross OA, Braithwaite AT, Skipper LM, Kachergus J, Hulihan MM, Middleton FA, Nishioka K, Fuchs J, Gasser T, Maraganore DM, Adler CH, Larvor L, Chartier-Harlin MC, Nilsson C, Langston JW, Gwinn K, Hattori N, Farrer MJ (2008) Genomic investigation of alpha-synuclein multiplication and parkinsonism. Annals of neurology 63 (6):743-750. doi:10.1002/ana.21380

58. Kara E, Kiely AP, Proukakis C, Giffin N, Love S, Hehir J, Rantell K, Pandraud A, Hernandez DG, Nacheva E, Pittman AM, Nalls MA, Singleton AB, Revesz T, Bhatia KP, Quinn N, Hardy J, Holton JL, Houlden H (2014) A 6.4 Mb duplication of the alpha-synuclein locus causing frontotemporal dementia and Parkinsonism: phenotype-genotype correlations. JAMA neurology 71 (9):1162-1171. doi:10.1001/jamaneurol.2014.994

59. Mutez E, Lepretre F, Le Rhun E, Larvor L, Duflot A, Mouroux V, Kerckaert JP, Figeac M, Dujardin K, Destee A, Chartier-Harlin MC (2011) SNCA locus duplication carriers: from genetics to Parkinson disease phenotypes. Human mutation 32 (4):E2079-2090. doi:10.1002/humu.21459

60. Farrer M, Kachergus J, Forno L, Lincoln S, Wang DS, Hulihan M, Maraganore D, Gwinn-Hardy K, Wszolek Z, Dickson D, Langston JW (2004) Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. Annals of neurology 55 (2):174-179. doi:10.1002/ana.10846

61. Fuchs J, Nilsson C, Kachergus J, Munz M, Larsson EM, Schule B, Langston JW, Middleton FA, Ross OA, Hulihan M, Gasser T, Farrer MJ (2007) Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication. Neurology 68 (12):916-922. doi:10.1212/01.wnl.0000254458.17630.c5

62. Ikeuchi T, Kakita A, Shiga A, Kasuga K, Kaneko H, Tan CF, Idezuka J, Wakabayashi K, Onodera O, Iwatsubo T, Nishizawa M, Takahashi H, Ishikawa A (2008) Patients homozygous and heterozygous for SNCA duplication in a family with parkinsonism and dementia. Archives of neurology 65 (4):514-519. doi:10.1001/archneur.65.4.514

63. Garraux G, Caberg JH, Vanbellinghen JF, Jamar M, Bours V, Moonen G, Dive D (2012) Partial trisomy 4q associated with young-onset dopa-responsive parkinsonism. Archives of neurology 69 (3):398-400. doi:10.1001/archneurol.2011.802

64. Shin CW, Kim HJ, Park SS, Kim SY, Kim JY, Jeon BS (2010) Two Parkinson's disease patients with alpha-synuclein gene duplication and rapid cognitive decline. Movement disorders : official journal of the Movement Disorder Society 25 (7):957-959. doi:10.1002/mds.23043

65. Troiano AR, Cazeneuve C, Le Ber I, Bonnet AM, Lesage S, Brice A (2008) Re: Alpha-synuclein gene duplication is present in sporadic Parkinson disease. Neurology 71 (16):1295; author reply 1295. doi:10.1212/01.wnl.0000338435.78120.0f

66. Brueggemann N, Odin P, Gruenewald A, Tadic V, Hagenah J, Seidel G, Lohmann K, Klein C, Djarmati A (2008) Re: Alpha-synuclein gene duplication is present in sporadic Parkinson disease. Neurology 71 (16):1294; author reply 1294. doi:10.1212/01.wnl.0000338439.00992.c7

67. Nuytemans K, Meeus B, Crosiers D, Brouwers N, Goossens D, Engelborghs S, Pals P, Pickut B, Van den Broeck M, Corsmit E, Cras P, De Deyn PP, Del-Favero J, Van Broeckhoven C, Theuns J (2009) Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population. Human mutation 30 (7):1054-1061. doi:10.1002/humu.21007

68. Perandones C, Giugni JC, Calvo DS, Raina GB, De Jorge Lopez L, Volpini V, Zabetian CP, Mata IF, Caputo M, Corach D, Radrizzani M, Micheli FE (2014) Mosaicism of alpha-synuclein gene rearrangements: report of two unrelated cases of early-onset parkinsonism. Parkinsonism & related disorders 20 (5):558-561. doi:10.1016/j.parkreldis.2013.11.014

69. Bonifati V (2012) Autosomal recessive parkinsonism. Parkinsonism & related disorders 18 Suppl 1:S4-6. doi:10.1016/S1353-8020(11)70004-9

70. Lucking CB, Durr A, Bonifati V, Vaughan J, De Michele G, Gasser T, Harhangi BS, Meco G, Denefle P, Wood NW, Agid Y, Brice A, French Parkinson's Disease Genetics Study G, European Consortium on Genetic Susceptibility in Parkinson's D (2000) Association between early-onset Parkinson's disease and mutations in the parkin gene. N Engl J Med 342 (21):1560-1567. doi:10.1056/NEJM200005253422103

71. Oliveira SA, Scott WK, Martin ER, Nance MA, Watts RL, Hubble JP, Koller WC, Pahwa R, Stern MB, Hiner BC, Ondo WG, Allen FH, Jr., Scott BL, Goetz CG, Small GW, Mastaglia F, Stajich JM, Zhang F, Booze MW, Winn MP, Middleton LT, Haines JL, Pericak-Vance MA, Vance JM (2003) Parkin mutations and susceptibility alleles in late-onset Parkinson's disease. Ann Neurol 53 (5):624-629. doi:10.1002/ana.10524

72. Chaudhary S, Behari M, Dihana M, Swaminath PV, Govindappa ST, Jayaram S, Goyal V, Maitra A, Muthane UB, Juyal RC, Thelma BK (2006) Parkin mutations in familial and sporadic Parkinson's disease among Indians. Parkinsonism & related disorders 12 (4):239-245. doi:10.1016/j.parkreldis.2005.12.004

73. Klein C, Djarmati A, Hedrich K, Schafer N, Scaglione C, Marchese R, Kock N, Schule B, Hiller A, Lohnau T, Winkler S, Wiegers K, Hering R, Bauer P, Riess O, Abbruzzese G, Martinelli P, Pramstaller PP (2005) PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism. European journal of human genetics : EJHG 13 (9):1086-1093. doi:10.1038/sj.ejhg.5201455

74. Shadrina MI, Semenova EV, Slominsky PA, Bagyeva GH, Illarioshkin SN, Ivanova S, II, Limborska SA (2007) Effective quantitative real-time polymerase chain reaction analysis of the parkin gene (PARK2) exon 1-12 dosage. BMC medical genetics 8:6. doi:10.1186/1471-2350-8-6

75. Choi JM, Woo MS, Ma HI, Kang SY, Sung YH, Yong SW, Chung SJ, Kim JS, Shin HW, Lyoo CH, Lee PH, Baik JS, Kim SJ, Park MY, Sohn YH, Kim JH, Kim JW, Lee MS, Lee MC, Kim DH, Kim YJ (2008) Analysis of PARK genes in a Korean cohort of early-onset Parkinson disease. Neurogenetics 9 (4):263-269. doi:10.1007/s10048-008-0138-0

76. Pankratz N, Kissell DK, Pauciulo MW, Halter CA, Rudolph A, Pfeiffer RF, Marder KS, Foroud T, Nichols WC, Parkinson Study Group PI (2009) Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations. Neurology 73 (4):279-286. doi:10.1212/WNL.0b013e3181af7a33

77. Kay DM, Stevens CF, Hamza TH, Montimurro JS, Zabetian CP, Factor SA, Samii A, Griffith A, Roberts JW, Molho ES, Higgins DS, Gancher S, Moses L, Zareparsi S, Poorkaj P, Bird T, Nutt J, Schellenberg GD, Payami H (2010) A comprehensive analysis of deletions, multiplications, and copy number variations in PARK2. Neurology 75 (13):1189-1194. doi:10.1212/WNL.0b013e3181f4d832

78. Wang C, Ma H, Feng X, Xie S, Chan P (2010) Parkin dosage mutations in patients with early-onset sporadic and familial Parkinson's disease in Chinese: an independent pathogenic role. Brain research 1358:30-38. doi:10.1016/j.brainres.2010.08.060

79. Guerrero Camacho JL, Monroy Jaramillo N, Yescas Gomez P, Rodriguez Violante M, Boll Woehrlen C, Alonso Vilatela ME, Lopez Lopez M (2012) High frequency of Parkin exon rearrangements in Mexican-mestizo patients with early-onset Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 27 (8):1047-1051. doi:10.1002/mds.25030

80. Moura KC, Junior MC, de Rosso AL, Nicaretta DH, Pereira JS, Jose Silva D, Santos-Reboucas CB, Pimentel MM (2012) Exon dosage variations in Brazilian patients with Parkinson's disease: analysis of SNCA, PARKIN, PINK1 and DJ-1 genes. Disease markers 32 (3):173-178. doi:10.3233/DMA-2011-0873

81. Yonova-Doing E, Atadzhanov M, Quadri M, Kelly P, Shawa N, Musonda ST, Simons EJ, Breedveld GJ, Oostra BA, Bonifati V (2012) Analysis of LRRK2, SNCA, Parkin, PINK1, and DJ-1 in Zambian patients with Parkinson's disease. Parkinsonism & related disorders 18 (5):567-571. doi:10.1016/j.parkreldis.2012.02.018

82. Chu MK, Kim WC, Choi JM, Hong JH, Kang SY, Ma HI, Kim YJ (2014) Analysis of Dosage Mutation in PARK2 among Korean Patients with Early-Onset or Familial Parkinson's Disease. Journal of clinical neurology 10 (3):244-248. doi:10.3988/jcn.2014.10.3.244

83. Al-Mubarak BR, Bohlega SA, Alkhairallah TS, Magrashi AI, AlTurki MI, Khalil DS, AlAbdulaziz BS, Abou Al-Shaar H, Mustafa AE, Alyemni EA, Alsaffar BA, Tahir AI, Al Tassan NA (2015) Parkinson's Disease in Saudi Patients: A Genetic Study. PloS one 10 (8):e0135950. doi:10.1371/journal.pone.0135950

84. Guo JF, Dong XL, Xu Q, Li N, Yan XX, Xia K, Tang BS (2015) Exon dosage analysis of parkin gene in Chinese sporadic Parkinson's disease. Neuroscience letters 604:47-51. doi:10.1016/j.neulet.2015.07.046

85. Mata IF, Alvarez V, Coto E, Blazquez M, Guisasola LM, Salvador C, Kachergus JM, Lincoln SJ, Farrer M (2005) Homozygous partial genomic triplication of the parkin gene in early-onset parkinsonism. Neuroscience letters 380 (3):257-259. doi:10.1016/j.neulet.2005.01.051

86. La Cognata V, Iemmolo R, D'Agata V, Scuderi S, Drago F, Zappia M, Cavallaro S (2014) Increasing the Coding Potential of Genomes Through Alternative Splicing: The Case of PARK2 Gene. Current genomics 15 (3):203-216. doi:10.2174/1389202915666140426003342

87. Scuderi S, La Cognata V, Drago F, Cavallaro S, D'Agata V (2014) Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain. BioMed research international 2014:690796. doi:10.1155/2014/690796

88. La Cognata V, D Agata V., Cavalcanti F., Cavallaro S. (2016) Parkin Alternative Splicing: Not Only Parkinsonism. ARC Journal of Neuroscience 1 (1):9-18

89. Kim SY, Seong MW, Jeon BS, Kim SY, Ko HS, Kim JY, Park SS (2012) Phase analysis identifies compound heterozygous deletions of the PARK2 gene in patients with early-onset Parkinson disease. Clinical genetics 82 (1):77-82. doi:10.1111/j.1399-0004.2011.01693.x

90. Elfferich P, Verleun-Mooijman MC, Maat-Kievit JA, van de Warrenburg BP, Abdo WF, Eshuis SA, Leenders KL, Hovestadt A, Zijlmans JC, Stroy JP, van Swieten JC, Boon AJ, van Engelen K, Verschuuren-Bemelmans CC, Lesnik-Oberstein SA, Tassorelli C, Lopiano L, Bonifati V, Dooijes D, van Minkelen R (2011) Breakpoint mapping of 13 large parkin deletions/duplications reveals an exon 4 deletion and an exon 7 duplication as founder mutations. Neurogenetics 12 (4):263-271. doi:10.1007/s10048-011-0302-9

91. Periquet M, Lucking C, Vaughan J, Bonifati V, Durr A, De Michele G, Horstink M, Farrer M, Illarioshkin SN, Pollak P, Borg M, Brefel-Courbon C, Denefle P, Meco G, Gasser T, Breteler MM, Wood N, Agid Y, Brice A, French Parkinson's Disease Genetics Study Group. The European Consortium on Genetic Susceptibility in Parkinson's D (2001) Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. American journal of human genetics 68 (3):617-626

92. Huttenlocher J, Stefansson H, Steinberg S, Helgadottir HT, Sveinbjornsdottir S, Riess O, Bauer P, Stefansson K (2015) Heterozygote carriers for CNVs in PARK2 are at increased risk of Parkinson's disease. Human molecular genetics 24 (19):5637-5643. doi:10.1093/hmg/ddv277

93. Klein C, Pramstaller PP, Kis B, Page CC, Kann M, Leung J, Woodward H, Castellan CC, Scherer M, Vieregge P, Breakefield XO, Kramer PL, Ozelius LJ (2000) Parkin deletions in a family with adult-onset, tremor-dominant parkinsonism: expanding the phenotype. Annals of neurology 48 (1):65-71

94. Farrer M, Chan P, Chen R, Tan L, Lincoln S, Hernandez D, Forno L, Gwinn-Hardy K, Petrucelli L, Hussey J, Singleton A, Tanner C, Hardy J, Langston JW (2001) Lewy bodies and parkinsonism in families with parkin mutations. Annals of neurology 50 (3):293-300

95. Pogson JH, Ivatt RM, Whitworth AJ (2011) Molecular mechanisms of PINK1-related neurodegeneration. Current neurology and neuroscience reports 11 (3):283-290. doi:10.1007/s11910-011-0187-x

96. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S, Ali Z, Del Turco D, Bentivoglio AR, Healy DG, Albanese A, Nussbaum R, Gonzalez-Maldonado R, Deller T, Salvi S, Cortelli P, Gilks WP, Latchman DS, Harvey RJ, Dallapiccola B, Auburger G, Wood NW (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304 (5674):1158-1160. doi:10.1126/science.1096284

97. Atsumi M, Li Y, Tomiyama H, Sato K, Hattori N (2006) [A 62-year-old woman with early-onset Parkinson's disease associated with the PINKi gene deletion]. Rinsho shinkeigaku = Clinical neurology 46 (3):199-202

98. Li Y, Tomiyama H, Sato K, Hatano Y, Yoshino H, Atsumi M, Kitaguchi M, Sasaki S, Kawaguchi S, Miyajima H, Toda T, Mizuno Y, Hattori N (2005) Clinicogenetic study of PINK1 mutations in autosomal recessive early-onset parkinsonism. Neurology 64 (11):1955-1957. doi:10.1212/01.WNL.0000164009.36740.4E

99. Camargos ST, Dornas LO, Momeni P, Lees A, Hardy J, Singleton A, Cardoso F (2009) Familial Parkinsonism and early onset Parkinson's disease in a Brazilian movement disorders clinic: phenotypic characterization and frequency of SNCA, PRKN, PINK1, and LRRK2 mutations. Movement disorders : official journal of the Movement Disorder Society 24 (5):662-666. doi:10.1002/mds.22365

100. Cazeneuve C, San C, Ibrahim SA, Mukhtar MM, Kheir MM, Leguern E, Brice A, Salih MA (2009) A new complex homozygous large rearrangement of the PINK1 gene in a Sudanese family with early onset Parkinson's disease. Neurogenetics 10 (3):265-270. doi:10.1007/s10048-009-0174-4

101. Marongiu R, Brancati F, Antonini A, Ialongo T, Ceccarini C, Scarciolla O, Capalbo A, Benti R, Pezzoli G, Dallapiccola B, Goldwurm S, Valente EM (2007) Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum. Human mutation 28 (1):98. doi:10.1002/humu.9472

102. Guo JF, Zhang XW, Nie LL, Zhang HN, Liao B, Li J, Wang L, Yan XX, Tang BS (2010) Mutation analysis of Parkin, PINK1 and DJ-1 genes in Chinese patients with sporadic early onset parkinsonism. Journal of neurology 257 (7):1170-1175. doi:10.1007/s00415-010-5485-8

103. Samaranch L, Lorenzo-Betancor O, Arbelo JM, Ferrer I, Lorenzo E, Irigoyen J, Pastor MA, Marrero C, Isla C, Herrera-Henriquez J, Pastor P (2010) PINK1-linked parkinsonism is associated with Lewy body pathology. Brain : a journal of neurology 133 (Pt 4):1128-1142. doi:10.1093/brain/awq051

104. Bonifati V, Rizzu P, Squitieri F, Krieger E, Vanacore N, van Swieten JC, Brice A, van Duijn CM, Oostra B, Meco G, Heutink P (2003) DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism. Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology 24 (3):159-160. doi:10.1007/s10072-003-0108-0

105. van Duijn CM, Dekker MC, Bonifati V, Galjaard RJ, Houwing-Duistermaat JJ, Snijders PJ, Testers L, Breedveld GJ, Horstink M, Sandkuijl LA, van Swieten JC, Oostra BA, Heutink P (2001) Park7, a novel locus for autosomal recessive early-onset parkinsonism, on chromosome 1p36. Am J Hum Genet 69 (3):629-634. doi:10.1086/322996

106. Lev N, Roncevic D, Ickowicz D, Melamed E, Offen D (2006) Role of DJ-1 in Parkinson's disease. Journal of molecular neuroscience : MN 29 (3):215-225

107. Ariga H, Takahashi-Niki K, Kato I, Maita H, Niki T, Iguchi-Ariga SM (2013) Neuroprotective function of DJ-1 in Parkinson's disease. Oxidative medicine and cellular longevity 2013:683920. doi:10.1155/2013/683920

108. Xu J, Zhong N, Wang H, Elias JE, Kim CY, Woldman I, Pifl C, Gygi SP, Geula C, Yankner BA (2005) The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. Human molecular genetics 14 (9):1231-1241. doi:10.1093/hmg/ddi134

109. Hedrich K, Djarmati A, Schafer N, Hering R, Wellenbrock C, Weiss PH, Hilker R, Vieregge P, Ozelius LJ, Heutink P, Bonifati V, Schwinger E, Lang AE, Noth J, Bressman SB, Pramstaller PP, Riess O, Klein C (2004) DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. Neurology 62 (3):389-394

110. Djarmati A, Hedrich K, Svetel M, Schafer N, Juric V, Vukosavic S, Hering R, Riess O, Romac S, Klein C, Kostic V (2004) Detection of Parkin (PARK2) and DJ1 (PARK7) mutations in early-onset Parkinson disease: Parkin mutation frequency depends on ethnic origin of patients. Human mutation 23 (5):525. doi:10.1002/humu.9240

111. Macedo MG, Verbaan D, Fang Y, van Rooden SM, Visser M, Anar B, Uras A, Groen JL, Rizzu P, van Hilten JJ, Heutink P (2009) Genotypic and phenotypic characteristics of Dutch patients with early onset Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 24 (2):196-203. doi:10.1002/mds.22287

112. Vilarino-Guell C, Soto AI, Lincoln SJ, Ben Yahmed S, Kefi M, Heckman MG, Hulihan MM, Chai H, Diehl NN, Amouri R, Rajput A, Mash DC, Dickson DW, Middleton LT, Gibson RA, Hentati F, Farrer MJ (2009) ATP13A2 variability in Parkinson disease. Human mutation 30 (3):406-410. doi:10.1002/humu.20877

113. Murphy KE, Cottle L, Gysbers AM, Cooper AA, Halliday GM (2013) ATP13A2 (PARK9) protein levels are reduced in brain tissue of cases with Lewy bodies. Acta Neuropathol Commun 1 (1):11. doi:10.1186/2051-5960-1-11

114. Krahn LE, Maraganore DM, Michels VV (1998) Childhood-onset schizophrenia associated with parkinsonism in a patient with a microdeletion of chromosome 22. Mayo Clinic proceedings 73 (10):956-959. doi:10.4065/73.10.956

115. Zaleski C, Bassett AS, Tam K, Shugar AL, Chow EW, McPherson E (2009) The co-occurrence of early onset Parkinson disease and 22q11.2 deletion syndrome. American journal of medical genetics Part A 149A (3):525-528. doi:10.1002/ajmg.a.32650

116. Booij J, van Amelsvoort T, Boot E (2010) Co-occurrence of early-onset Parkinson disease and 22q11.2 deletion syndrome: Potential role for dopamine transporter imaging. American journal of medical genetics Part A 152A (11):2937-2938. doi:10.1002/ajmg.a.33665

117. Butcher NJ, Kiehl TR, Hazrati LN, Chow EW, Rogaeva E, Lang AE, Bassett AS (2013) 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 70 (11):1359-1366. doi:10.1001/jamaneurol.2013.3646

118. Mok KY, Sheerin U, Simon-Sanchez J, Salaka A, Chester L, Escott-Price V, Mantripragada K, Doherty KM, Noyce AJ, Mencacci NE, Lubbe SJ, International Parkinson's Disease Genomics C, Williams-Gray CH, Barker RA, van Dijk KD, Berendse HW, Heutink P, Corvol JC, Cormier F, Lesage S, Brice A, Brockmann K, Schulte C, Gasser T, Foltynie T, Limousin P, Morrison KE, Clarke CE, Sawcer S, Warner TT, Lees AJ, Morris HR, Nalls MA, Singleton AB, Hardy J, Abramov AY, Plagnol V, Williams NM, Wood NW (2016) Deletions at 22q11.2 in idiopathic Parkinson's disease: a combined analysis of genome-wide association data. The Lancet Neurology. doi:10.1016/S1474-4422(16)00071-5

119. Rehman AF, Dhamija R, Williams ES, Barrett MJ (2015) 22q11.2 deletion syndrome presenting with early-onset Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 30 (9):1289-1290. doi:10.1002/mds.26305

120. Muller T (2015) Catechol-O-methyltransferase inhibitors in Parkinson's disease. Drugs 75 (2):157-174. doi:10.1007/s40265-014-0343-0

121. Son JH, Kawamata H, Yoo MS, Kim DJ, Lee YK, Kim S, Dawson TM, Zhang H, Sulzer D, Yang L, Beal MF, Degiorgio LA, Chun HS, Baker H, Peng C (2005) Neurotoxicity and behavioral deficits associated with Septin 5 accumulation in dopaminergic neurons. Journal of neurochemistry 94 (4):1040-1053. doi:10.1111/j.1471-4159.2005.03257.x

122. Marttinen M, Kurkinen KM, Soininen H, Haapasalo A, Hiltunen M (2015) Synaptic dysfunction and septin protein family members in neurodegenerative diseases. Molecular neurodegeneration 10:16. doi:10.1186/s13024-015-0013-z

123. Ogaki K, Ross OA (2014) Chromosome 22q11.2 deletion may contain a locus for recessive early-onset Parkinson's disease. Parkinsonism & related disorders 20 (9):945-946. doi:10.1016/j.parkreldis.2014.06.020

124. C P, Vl F, La P (2015) 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 Biology 04 (04). doi:10.4172/2168-9431.1000123

125. Simon-Sanchez J, Scholz S, Matarin Mdel M, Fung HC, Hernandez D, Gibbs JR, Britton A, Hardy J, Singleton A (2008) Genomewide SNP assay reveals mutations underlying Parkinson disease. Human mutation 29 (2):315-322. doi:10.1002/humu.20626

126. Liu X, Cheng R, Ye X, Verbitsky M, Kisselev S, Mejia-Santana H, Louis E, Cote L, Andrews H, Waters C, Ford B, Fahn S, Marder K, Lee J, Clark L (2013) Increased Rate of Sporadic and Recurrent

Rare Genic Copy Number Variants in Parkinson's Disease Among Ashkenazi Jews. Molecular genetics & genomic medicine 1 (3):142-154. doi:10.1002/mgg3.18

127. Pamphlett R, Morahan JM, Luquin N, Yu B (2012) An approach to finding brain-situated mutations in sporadic Parkinson's disease. Parkinsonism & related disorders 18 (1):82-85. doi:10.1016/j.parkreldis.2011.08.024

128. Bademci G, Edwards TL, Torres AL, Scott WK, Zuchner S, Martin ER, Vance JM, Wang L (2010) A rare novel deletion of the tyrosine hydroxylase gene in Parkinson disease. Human mutation 31 (10):E1767-1771. doi:10.1002/humu.21351

129. Bademci G, Vance JM, Wang L (2012) Tyrosine hydroxylase gene: another piece of the genetic puzzle of Parkinson's disease. CNS & neurological disorders drug targets 11 (4):469-481

130. Bohnen NI, Albin RL (2011) The cholinergic system and Parkinson disease. Behavioural brain research 221 (2):564-573. doi:10.1016/j.bbr.2009.12.048

131. Muller ML, Bohnen NI (2013) Cholinergic dysfunction in Parkinson's disease. Current neurology and neuroscience reports 13 (9):377. doi:10.1007/s11910-013-0377-9

132. Sun AG, Wang J, Shan YZ, Yu WJ, Li X, Cong CH, Wang X (2014) Identifying distinct candidate genes for early Parkinson's disease by analysis of gene expression in whole blood. Neuro endocrinology letters 35 (5):398-404

133. Dunn HA, Walther C, Yuan GY, Caetano FA, Godin CM, Ferguson SS (2014) Role of SAP97 in the regulation of 5-HT2AR endocytosis and signaling. Molecular pharmacology 86 (3):275-283. doi:10.1124/mol.114.093476

134. Liu X, Cheng R, Verbitsky M, Kisselev S, Browne A, Mejia-Sanatana H, Louis ED, Cote LJ, Andrews H, Waters C, Ford B, Frucht S, Fahn S, Marder K, Clark LN, Lee JH (2011) Genome-wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. BMC medical genetics 12:104. doi:10.1186/1471-2350-12-104

135. Chen S, Liu F (2010) Interaction of dopamine D1 receptor with N-ethylmaleimide-sensitive factor is important for the membrane localization of the receptor. Journal of neuroscience research 88 (11):2504-2512. doi:10.1002/jnr.22401

136. Ghanbari M, Darweesh SK, de Looper HW, van Luijn MM, Hofman A, Ikram MA, Franco OH, Erkeland SJ, Dehghan A (2016) Genetic Variants in MicroRNAs and Their Binding Sites Are Associated with the Risk of Parkinson Disease. Human mutation 37 (3):292-300. doi:10.1002/humu.22943

137. Reglodi D, Kiss P, Lubics A, Tamas A (2011) Review on the protective effects of PACAP in models of neurodegenerative diseases in vitro and in vivo. Current pharmaceutical design 17 (10):962-972

138. Lee EH, Seo SR (2014) Neuroprotective roles of pituitary adenylate cyclase-activating polypeptide in neurodegenerative diseases. BMB reports 47 (7):369-375

139. Brown D, Tamas A, Reglodi D, Tizabi Y (2013) PACAP protects against salsolinol-induced toxicity in dopaminergic SH-SY5Y cells: implication for Parkinson's disease. Journal of molecular neuroscience : MN 50 (3):600-607. doi:10.1007/s12031-013-0015-7

140. Brown D, Tamas A, Reglodi D, Tizabi Y (2014) PACAP protects against inflammatory-mediated toxicity in dopaminergic SH-SY5Y cells: implication for Parkinson's disease. Neurotoxicity research 26 (3):230-239. doi:10.1007/s12640-014-9468-x

141. Reglodi D, Lubics A, Tamas A, Szalontay L, Lengvari I (2004) Pituitary adenylate cyclase activating polypeptide protects dopaminergic neurons and improves behavioral deficits in a rat model of Parkinson's disease. Behavioural brain research 151 (1-2):303-312. doi:10.1016/j.bbr.2003.09.007

142. Shivers KY, Nikolopoulou A, Machlovi SI, Vallabhajosula S, Figueiredo-Pereira ME (2014) PACAP27 prevents Parkinson-like neuronal loss and motor deficits but not microglia activation induced by prostaglandin J2. Biochimica et biophysica acta 1842 (9):1707-1719. doi:10.1016/j.bbadis.2014.06.020

143. Mehler MF, Gokhan S (2000) Mechanisms underlying neural cell death in neurodegenerative diseases: alterations of a developmentally-

mediated cellular rheostat. Trends in neurosciences 23 (12):599-605

144. van der Heide LP, Smidt MP (2013) The BCL2 code to dopaminergic development and Parkinson's disease. Trends in molecular medicine 19 (4):211-216. doi:10.1016/j.molmed.2013.02.003

145. Nakatani T, Kumai M, Mizuhara E, Minaki Y, Ono Y (2010) Lmx1a and Lmx1b cooperate with Foxa2 to coordinate the specification of dopaminergic neurons and control of floor plate cell differentiation in the developing mesencephalon. Developmental biology 339 (1):101-113. doi:10.1016/j.ydbio.2009.12.017

146. Sutherland GT, Matigian NA, Chalk AM, Anderson MJ, Silburn PA, Mackay-Sim A, Wells CA, Mellick GD (2009) A cross-study transcriptional analysis of Parkinson's disease. PloS one 4 (3):e4955. doi:10.1371/journal.pone.0004955

147. Cohen E (2011) Countering neurodegeneration by reducing the activity of the insulin/IGF signaling pathway: current knowledge and future prospects. Experimental gerontology 46 (2-3):124-128. doi:10.1016/j.exger.2010.08.032

148. Biondi O, Branchu J, Ben Salah A, Houdebine L, Bertin L, Chali F, Desseille C, Weill L, Sanchez G, Lancelin C, Aid S, Lopes P, Pariset C, Lecolle S, Cote J, Holzenberger M, Chanoine C, Massaad C, Charbonnier F (2015) IGF-1R Reduction Triggers Neuroprotective Signaling Pathways in Spinal Muscular Atrophy Mice. The Journal of neuroscience : the official journal of the Society for Neuroscience 35 (34):12063-12079. doi:10.1523/JNEUROSCI.0608-15.2015

149. Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, Paisan-Ruiz C, Lichtner P, Scholz SW, Hernandez DG, Kruger R, Federoff M, Klein C, Goate A, Perlmutter J, Bonin M, Nalls MA, Illig T, Gieger C, Houlden H, Steffens M, Okun MS, Racette BA, Cookson MR, Foote KD, Fernandez HH, Traynor BJ, Schreiber S, Arepalli S, Zonozi R, Gwinn K, van der Brug M, Lopez G, Chanock SJ, Schatzkin A, Park Y, Hollenbeck A, Gao J, Huang X, Wood NW, Lorenz D, Deuschl G, Chen H, Riess O, Hardy JA, Singleton AB, Gasser T (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nature genetics 41 (12):1308-1312. doi:10.1038/ng.487

150. Chen L-W (2013) Roles of Wnt/β-Catenin Signaling in Controlling the Dopaminergic Neuronal Cell Commitment of Midbrain and Therapeutic Application for Parkinson’s Disease. doi:10.5772/53282

151. Mahul-Mellier AL, Fauvet B, Gysbers A, Dikiy I, Oueslati A, Georgeon S, Lamontanara AJ, Bisquertt A, Eliezer D, Masliah E, Halliday G, Hantschel O, Lashuel HA (2014) c-Abl phosphorylates alpha-synuclein and regulates its degradation: implication for alpha-synuclein clearance and contribution to the pathogenesis of Parkinson's disease. Human molecular genetics 23 (11):2858-2879. doi:10.1093/hmg/ddt674

152. Calabresi P, Picconi B, Parnetti L, Di Filippo M (2006) A convergent model for cognitive dysfunctions in Parkinson's disease: the critical dopamine-acetylcholine synaptic balance. The Lancet Neurology 5 (11):974-983. doi:10.1016/S1474-4422(06)70600-7

153. Jenkins A, Apud JA, Zhang F, Decot H, Weinberger DR, Law AJ (2014) Identification of candidate single-nucleotide polymorphisms in NRXN1 related to antipsychotic treatment response in patients with schizophrenia. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 39 (9):2170-2178. doi:10.1038/npp.2014.65

154. Infante J, Prieto C, Sierra M, Sanchez-Juan P, Gonzalez-Aramburu I, Sanchez-Quintana C, Berciano J, Combarros O, Sainz J (2015) Identification of candidate genes for Parkinson's disease through blood transcriptome analysis in LRRK2-G2019S carriers, idiopathic cases, and controls. Neurobiology of aging 36 (2):1105-1109. doi:10.1016/j.neurobiolaging.2014.10.039

155. van Abel D, Michel O, Veerhuis R, Jacobs M, van Dijk M, Oudejans CB (2012) Direct downregulation of CNTNAP2 by STOX1A is associated with Alzheimer's disease. Journal of Alzheimer's disease : JAD 31 (4):793-800. doi:10.3233/JAD-2012-120472

156. Sindi IA, Tannenberg RK, Dodd PR (2014) Role for the neurexin-neuroligin complex in Alzheimer's

disease. Neurobiology of aging 35 (4):746-756. doi:10.1016/j.neurobiolaging.2013.09.032

157. Van Kampen JM, Baranowski D, Kay DG (2014) Progranulin gene delivery protects dopaminergic neurons in a mouse model of Parkinson's disease. PloS one 9 (5):e97032. doi:10.1371/journal.pone.0097032

158. Chen Y, Li S, Su L, Sheng J, Lv W, Chen G, Xu Z (2015) Association of progranulin polymorphism rs5848 with neurodegenerative diseases: a meta-analysis. Journal of neurology 262 (4):814-822. doi:10.1007/s00415-014-7630-2

159. Rovelet-Lecrux A, Deramecourt V, Legallic S, Maurage CA, Le Ber I, Brice A, Lambert JC, Frebourg T, Hannequin D, Pasquier F, Campion D (2008) Deletion of the progranulin gene in patients with frontotemporal lobar degeneration or Parkinson disease. Neurobiology of disease 31 (1):41-45. doi:10.1016/j.nbd.2008.03.004

160. Chen YZ, Matsushita M, Girirajan S, Lisowski M, Sun E, Sul Y, Bernier R, Estes A, Dawson G, Minshew N, Shellenberg GD, Eichler EE, Rieder MJ, Nickerson DA, Tsuang DW, Tsuang MT, Wijsman EM, Raskind WH, Brkanac Z (2012) Evidence for involvement of GNB1L in autism. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics 159B (1):61-71. doi:10.1002/ajmg.b.32002

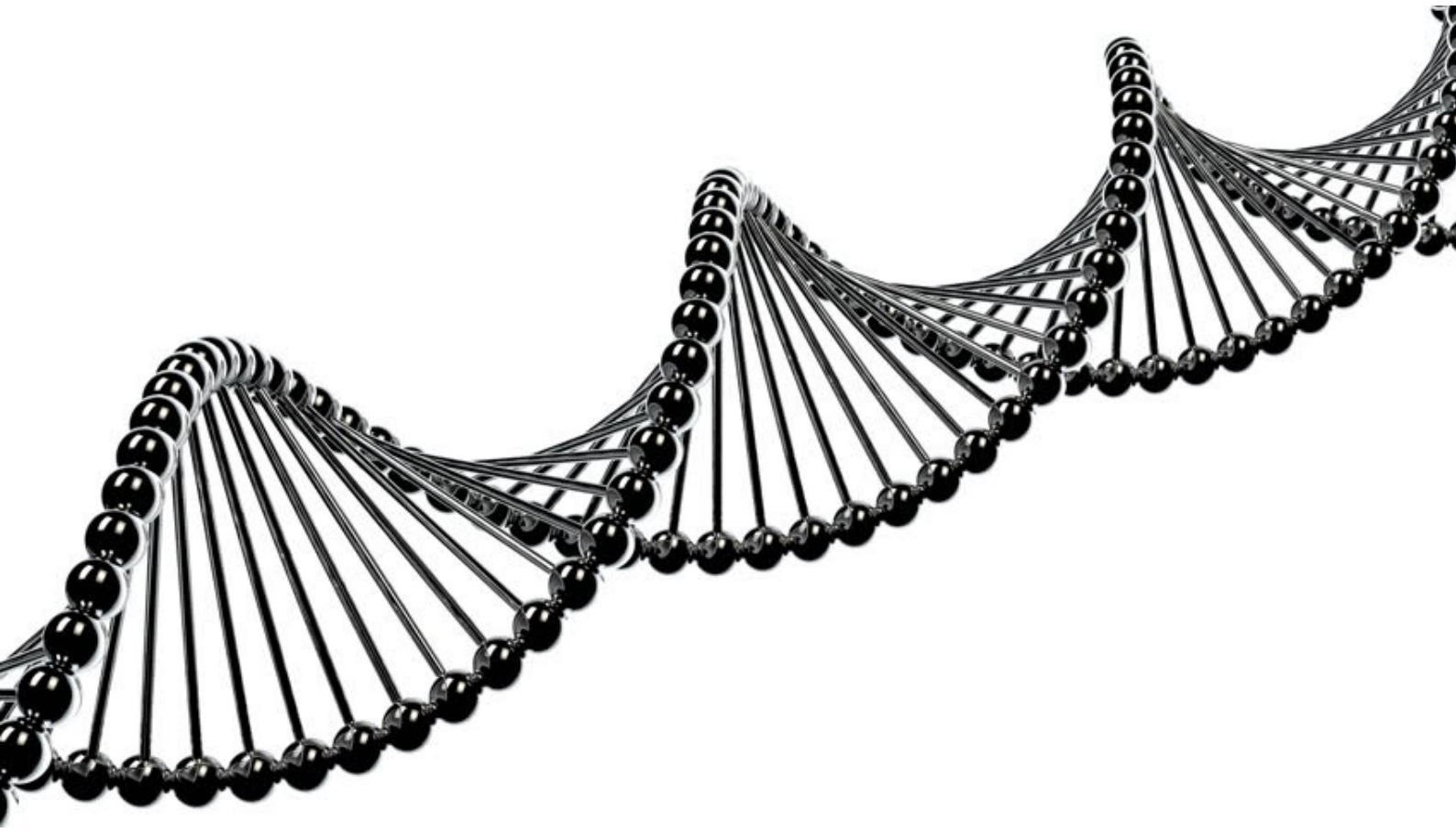
161. Ishiguro H, Koga M, Horiuchi Y, Noguchi E, Morikawa M, Suzuki Y, Arai M, Niizato K, Iritani S, Itokawa M, Inada T, Iwata N, Ozaki N, Ujike H, Kunugi H, Sasaki T, Takahashi M, Watanabe Y, Someya T, Kakita A, Takahashi H, Nawa H, Arinami T (2010) Supportive evidence for reduced expression of GNB1L in schizophrenia. Schizophrenia bulletin 36 (4):756-765. doi:10.1093/schbul/sbn160

162. Hiramoto T, Kang G, Suzuki G, Satoh Y, Kucherlapati R, Watanabe Y, Hiroi N (2011) Tbx1: identification of a 22q11.2 gene as a risk factor for autism spectrum disorder in a mouse model. Human molecular genetics 20 (24):4775-4785. doi:10.1093/hmg/ddr404

SUPPLEMENTARY DATA ARE AVAILABLE IN EXCEL FILE FORMAT IN THE ONLINE VERSION.

CHAPTER 3

A customized high-resolution array-comparative genomic hybridization to explore copy number variations in Parkinson's disease



La Cognata V. et al., Neurogenetics (2016). doi:10.1007/s10048-016-0494-0

A customized high-resolution array-comparative genomic hybridization to explore copy number variations in Parkinson's disease

Valentina La Cognata[1-2], Giovanna Morello[1], Giulia Gentile[1], Velia D'Agata[2], Chiara Criscuolo[3], Francesca Cavalcanti[4], Sebastiano Cavallaro*[1,4]

[1]Institute of Neurological Sciences, National Research Council, Catania, Italy; [2]Dept. of Biomedical and Biotechnological Sciences, Section of Human Anatomy and Histology, University of Catania, Italy; [3]Dept. of Neurosciences, Reproductive and Odontostomatological Sciences, Federico II University, Naples, Italy; [4]Institute of Neurological Sciences, National Research Council, Mangone (CS), Italy.

**Correspondence to: Sebastiano Cavallaro, Institute of Neurological Sciences, National Research Council, Via Paolo Gaifami 18, 95125, Catania, Italy; Tel/Fax: +39 0957338128, E-mail: sebastiano.cavallaro@cnr.it*

ABSTRACT

Parkinson's disease (PD), the second most common progressive neurodegenerative disorder, was long believed to be a non-genetic sporadic syndrome. Today, only a small percentage of PD cases with genetic inheritance patterns are known, often complicated by reduced penetrance and variable expressivity. The few well-characterized Mendelian genes, together with a number of risk factors, contribute to the major sporadic forms of the disease, thus delineating an intricate genetic profile at the basis of this debilitating and incurable condition. Along with single nucleotide changes, gene-dosage abnormalities and Copy Number Variations (CNVs) have emerged as significant disease-causing mutations in PD. However, due to their size variability and to the quantitative nature of the assay, CNV genotyping is particularly challenging. For this reason, innovative high-throughput platforms and bioinformatics algorithms are increasingly replacing classical CNV detection methods.

Here, we report the design strategy, development, validation, and implementation of *NeuroArray*, a customized exon-centric high-resolution array-based comparative genomic hybridization (aCGH) tailored to detect single/multi-exon deletions and duplications in a large panel of PD-related genes. This targeted design allows for a focused evaluation of structural imbalances in clinically relevant PD genes, combining exon-level resolution with genome-wide coverage. The *NeuroArray* platform may offer new insights in elucidating inherited potential or *de novo* structural alterations in PD patients and investigating new candidate genes.

Keywords: aCGH, CNVs, Parkinson's disease, Neurological Diseases, Genes.

INTRODUCTION

Parkinson's disease (PD) is a progressive debilitating movement disorder that affects approximately 1% of the population older than 65 years of age worldwide [1]. Clinically, most patients present resting tremor, bradykinesia, stiffness of movement and postural instability. These major symptoms derive from the profound and selective loss of dopaminergic neurons in the *substantia nigra pars compacta* (SNc), coupled with the accumulation of eosinophilic intracytoplasmic aggregates termed Lewy bodies (LBs) [1]. Like other complex diseases, PD is believed to be a multifactorial syndrome, resulting from an elaborate interplay of numerous elements (genes, susceptibility alleles, environmental exposures, and gene-environment interactions), and its molecular aetiology remains incompletely understood [2].

In recent years, the intensive efforts of the scientific community and the significant and rapid advancement of biotechnologies have fuelled several steps towards the elucidation of the genetic components of PD. Genome-wide linkage scans and exome sequencing of well-characterized PD families have been successful in discovering disease-causing mutations in dominant (*SNCA*, *LRRK2*, *VPS35* and the recent *TMEM230*), recessive (*PARK2*, *PINK1*, *DJ1, DNAJC6*) [2-4] and X-linked (*RAB39B*) PD genes [5,6]. Other genes, such as *CHCHD2* and *EIF4G1,* are associated with familial PD inheritance but still require independent confirmations [7,8]. Moreover, a set of genes related to atypical parkinsonian forms is known and includes *ATP13A2*, whose mutations cause the Kufor-Rakeb syndrome (PARK9) [9]. Despite

the existence of these rare Mendelian monogenic forms, it is now clear that PD is a genetically heterogeneous and most likely complex disorder. This complexity is underlined by the notion that we are currently aware of dozens of loci, genes and risk factors that seem to contribute to PD [2,10]. These genes are involved in numerous cellular pathways, such as the ubiquitin-proteasome system, synaptic transmission, autophagy, lysosomal autophagy, endosomal trafficking, mitochondrial metabolism, apoptosis and inflammatory mechanisms, all of which are generally implicated in neuronal cell death [11].

While the major pathogenic mutations are single nucleotide polymorphisms (SNPs) in the coding regions of PD-linked genes, the contribution of other types of DNA molecular defects (e.g., structural chromosome abnormalities such as Copy Number Variations – CNVs) to the genomic architecture is less emphasized but equally significant [12,13]. CNVs are unbalanced rearrangements larger than 50 bp and arise from genomic instability [12]. They are recognized as critical elements for the development and maintenance of the nervous system and appear to contribute to hereditable or sporadic neurological diseases, including neuropathies, epilepsy, autistic syndromes, psychiatric illnesses and neurodegenerative diseases, such as PD [14-16]. In this regard, several CNVs have been reported in PD patients, including specific pathogenic anomalies mapped in PD loci or involving candidate PD-related genes [17]. To mention the most recurrent, *SNCA* copy-number gains have been proven to play a major role in the disease severity of PARK1, while *PARK2* homozygous or compound heterozygous exon copy number changes are very common among the early-onset cases, rendering the gene-dosage assay essential in mutational screening.

Currently, the detection of CNVs and gene dosage imbalances mainly relies on traditional methodological approaches (karyotyping and PCR-based approaches such as quantitative PCR and Multiple Ligation Probe Analysis). However, these methodologies bear objective limits: they are time-consuming and labour-intensive, require multiple phase steps and severe equipment costs and, above all, do not provide a complete genomic overview of structural imbalances at sufficiently high resolution. The development of the aCGH technology has dramatically improved and catalysed the detection and characterization of multiple CNVs, offering high reproducibility, high resolution, and scalability for complete genome-wide mapping

of imbalances [18]. The aCGH technique has been refined to the most advanced aCGH plus SNPs edition, a widely used array able to simultaneously perform SNPs genotyping and CNV detection. This methodology shows higher sensitivity for the detection of low-level mosaic aneuploidies and chimerism and offers the ability to detect loss of heterozygosity, but it has a limited ability to detect single-exon CNVs due to the distribution of SNPs across the genome. For this reason, several customized aCGH suitably designed to focus on specific clinically relevant chromosomal locations have been developed and are already applied to different human diseases, including neuromuscular diseases, cancer, autism, epilepsy, multiple sclerosis, mitochondrial and metabolic disorders [19-24].

In this study, we developed a customized exon-centric aCGH (hereafter called *NeuroArray*), tailored to detect single/multi-exon deletions and duplications in a large panel of PD-related genes. We will first report the design strategy and the applied analysis methods. Then, we will show two representative PD cases tested on *NeuroArray*. Our findings show the advantages of the *NeuroArray* platform in terms of results, time, and costs, as well as for the discovery of new potential genetic biomarkers underlying the pathogenic mechanisms of PD and commonly shared genetic signatures with other neurological diseases.

MATERIALS AND METHODS

GENE SELECTION AND aCGH DESIGN STRATEGY

To build the customized *NeuroArray* aCGH platform, we aimed to obtain a high-density probe coverage in the coding region of clinically relevant genes associated with PD. Gene selection relied on our expertise in the clinic, genetics and literature data, and has been extended to the entire currently known sets of genes collected in PDgene (http://www.pdgene.org/) [25]. The list of selected genes embraces disease-causing genes, known and putative risk factors, and other genetic regions affected by different types of mutations. To perform a differential diagnosis, we also included genes related to other neurological conditions (see **Supplementary information and Supplementary Tables**).

The array design was carried out by using the web-based Agilent SureDesign Software (Agilent Technologies, Santa Clara, CA), a web application that allows one to

define regions of interest and select the "best-performing" probes from the High-Density (HD) Agilent probe library. Candidate probes were scored and filtered using bioinformatics prediction criteria for probe sensitivity, specificity, and responsiveness under appropriate conditions. We also selected a limited number of probes by genomic tiling to cover regions inadequately represented in the Agilent database. All probes had similar characteristics: isothermal probes, with melting temperature (Tm) of 80°C and probe length of approximately 60-mers, in accordance with the manufacturer's specifications. Further details about the design method, the number of genes, exons, the median probe spacing and other characteristics of *NeuroArray* are summarized in **Table 1, Supplementary information and Supplementary Table 1**.

CLINICAL SAMPLES SELECTION

To validate the *NeuroArray*, we selected DNA samples from individuals suffering from PD or other neurological disorders and previously subjected to gene dosage through MLPA, qPCR or other commercially available whole-genome aCGH. Moreover, DNA samples of patients with PD phenotypes but an incomplete molecular diagnosis were referred for *NeuroArray* molecular cytogenetic testing. Informed consent was obtained for the use of DNA samples and for the access to medical records for research purposes.

MICROARRAY EXPERIMENT AND DATA ANALYSIS

Genomic DNA was extracted from peripheral blood lymphocytes using the EZ1 DNA Blood extraction kit (Qiagen, Hilden, Germany) by the Biorobot EZ1 following the manufacturer's recommendations (Qiagen, Hilden, Germany). Highly concentrated DNA was checked for quality using the NanoDrop spectrophotometer (Thermo Scientific, Wilmington, Delaware, USA). Array experiments were performed as recommended by the manufacturer (Agilent Technologies, Santa Clara, CA), and data were extracted using Feature Extraction software (Agilent Technologies, Santa Clara, CA). After the quality control check, data visualization and analysis were performed with CytoGenomics software v. 3.0.6.6. (Agilent Technologies, Santa Clara, CA) using both ADM-2 and ADM-1 algorithms. Moreover, we took into account a single-probe analysis to include putative exonic variants. Significant single exonic probe signals were clustered for pathologies according to their location on causative

or susceptibility genes through a homemade script on R-platform [26]. Full details on microarray experiments and data analysis are available in the **Supplementary information**.

VALIDATION

Ad hoc quantitative real-time polymerase chain reaction (qPCR) assays were performed to validate genomic imbalances detected by the *NeuroArray* as previously described [27]. Primers flanking the putative exonic imbalances were designed using the PrimerBlast tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Each qPCR assay was performed in triplicate using the LightCycler 1.5 (Roche Diagnostics, Germany). The relative quantification was measured using the ΔΔCt method, which requires a healthy control sample (diploid) as a calibrator in all amplifications [28]. As a calibrator control, we used the same DNA reference hybridized in the *NeuroArray* experiments. A control gene, checked as normal double-copies on *NeuroArray*, was used as a reference for normalization. We considered a ΔΔCt value ≤0.6 as a loss, included from 0.8 to 1.2 as normal diploid, and ≥1.4 as a gain. PCR products were visualized by agarose gel electrophoresis.

RESULTS

aCGH DESIGN ON A TARGETED PD GENE PANEL

To perform a comprehensive analysis of CNVs in PD-related genes, we developed a focused customized oligonucleotide aCGH design targeting 505 genes and 6826 exonic regions linked to PD. Overall, 11,161 probes with a median probe spacing of 391 bp were enriched in the coding regions of these genes (**Table 1**). The majority of targeted genes map on chromosome 1, while lower numbers are distributed among the other chromosomes (**Figure 1**).

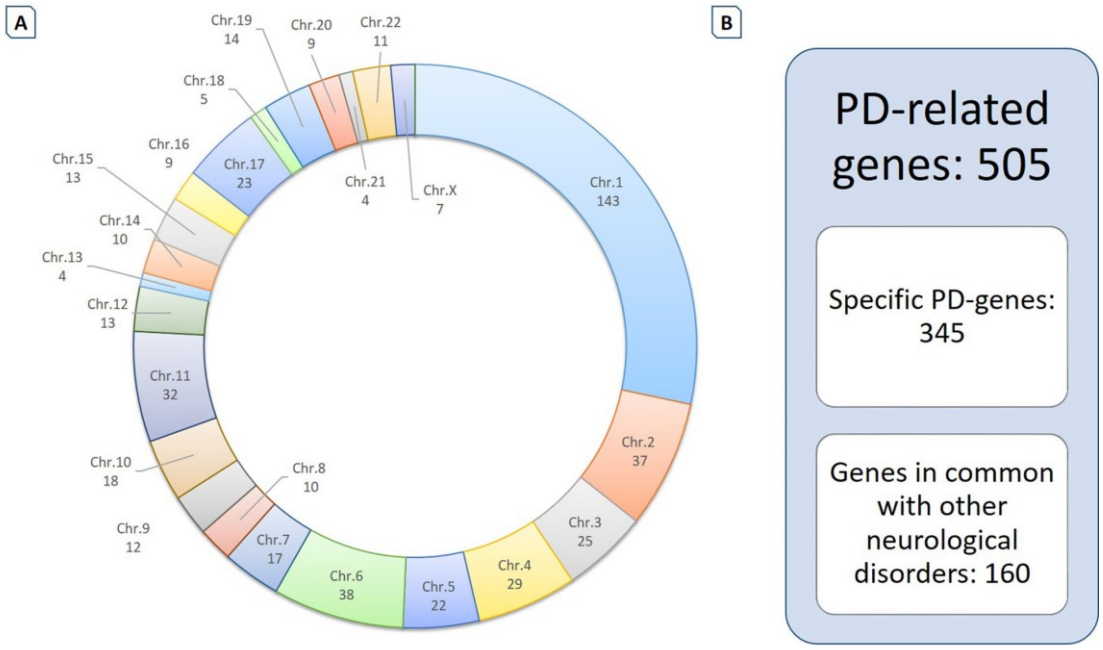


Figure 1. Distribution of selected PD genes on the human genome and overlap with other neurological diseases.

A: Graphical representation showing the number of clinically relevant genes for chromosomes included in the customized PD panel. The total number of selected genes is 505, mostly enclosed in chromosome 1. Chromosome Y does not include PD-related genes. B: The PD panel globally targets 505 PD-related genes. Of these, 345 are specific for PD, while 160 are in common with other neurological diseases. These latter ones can be useful to study the potential overlapping genetic signatures among different neurological conditions and to better define the genotype/phenotype correlations.

The tightly restricted criteria used for the array customization have allowed a higher exonic probe enrichment on selected gene panels, overcoming the resolution of commercially available genome-wide CGH array platforms. Overall, 94% of the total exon targets are covered by at least one probe in the *NeuroArray* design (**Table 1**), while other commercially available aCGH platforms provide a lower probe coverage of the same selected exonic regions. For example, the Agilent SurePrint G3

Human CGH Microarray 8x60K slide format covers our selected regions by 8.2%, while the highest resolution 1x1M array provides 25% of our target coverage. A representative illustration is reported in **Figure 2** and focuses on *PINK1* (RefSeq Acc. Num. *NM_032409.2*).

Customized PD-panel design	
Total genes	505
Total exonic targets	6826
Target coverage	94 %
Total target/exon size	1,935 Mbp
Total probes (1-2 probes per exon)	11161
Total unique probes from HD database	10411
Total unique probes by genomic tiling	750
Median probe spacing	391 bp
Mean target size	323 bp
Uncovered targets	431

The table lists the total number of selected genes and exon targets, the mean exon size, the number of probes, the median probe spacing and the total coverage of the customized design for CNV detection in PD. The array design was performed through the Agilent SureDesign software (https://earray.chem.agilent.com/suredesign). The majority of probes have been scored and filtered from the High-Density (HD) Agilent probe library. A limited number of probes has been designed with the Genomic Tiling option to cover regions inadequately represented in the Agilent database. All probes have been chosen with similar characteristics: isothermal probes, with melting temperature (Tm) of 80° C and probe length of 60-mers.

To perform an accurate differential analysis between PD patients and other neurological phenotypes, we also included genes related to other neurological disorders (S**upplementary information and Supplementary Tables).** Specifically, 160 of the 505 PD-related genes were linked to other neurological conditions (**Figure 1**).

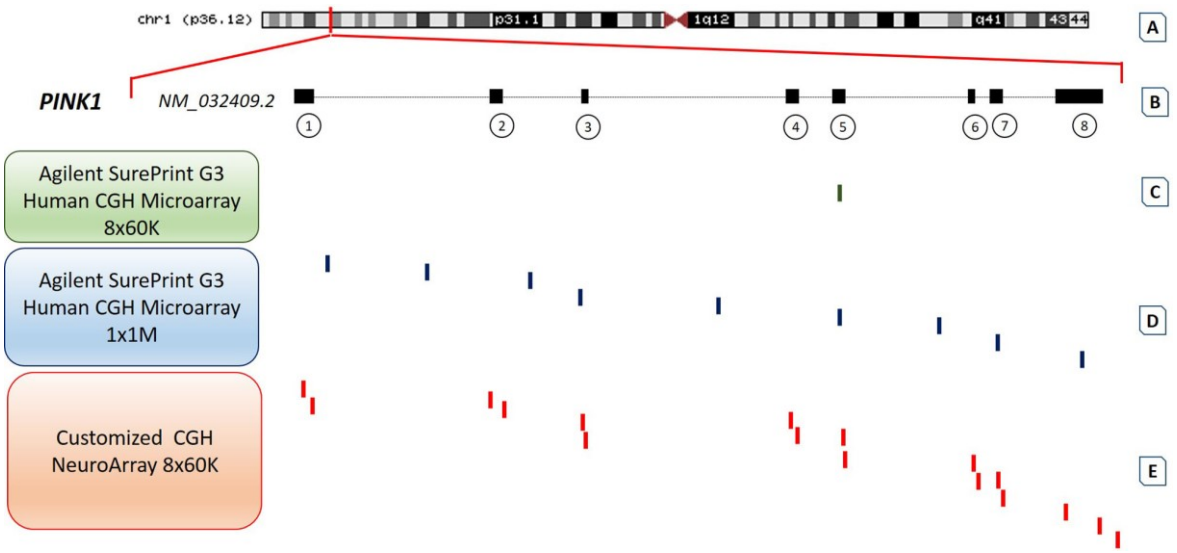


Figure 2. Oligonucleotide probe distribution on *PINK1* in different commercially available whole-genome aCGH platforms and *NeuroArray*.

A: The human *PINK1* gene is located on chromosome 1 (cytoband p36.12), spanning approximately 18 kb of genomic DNA. **B:** This gene produces an mRNA transcript encompassing 8 exonic regions (NCBI Accession Number NM_032409.2). Exons are represented in the figure by black boxes and are numbered consecutively. The dashed line represents intronic regions. **C:** Distribution of oligonucleotide probes (green bars) in the commercially available whole-genome Agilent SurePrint G3 Human CGH Microarray 8x60K. As evidenced in the figure, this platform has just one probe overlapping *PINK1* exon 5, proving low-resolution coverage. D: Distribution of oligonucleotide probes (blue bars) in the whole-genome Agilent SurePrint G3 Human CGH Microarray 1x1M slide format. The highest-resolution 1x1M array CGH reveals the *PINK1* genetic region with a greater number of oligonucleotide probes; however, it is five times more expensive per sample than the Agilent 8x60K slide format and leaves uncovered some exonic traits (for example, exons 1 or 2). E: Distribution of oligonucleotide probes (red bars) in the entire exonic regions of the *PINK1* gene in the customized *NeuroArray* design. The *NeuroArray* design allows high-density probe enrichment in the entire exonic regions of *PINK1*, enabling a focused evaluation of structural imbalances at a single-exon resolution with costs comparable to an 8x60K slide format.

CNVs OF PD-RELATED GENES DETECTED THROUGH THE NEUROARRAY PLATFORM

NeuroArray was able to confirm copy number changes previously characterized by other methodological strategies and revealed new interesting genomic imbalances. In the following sections, we will show two representative examples of *NeuroArray* tests obtained by using genomic DNA samples of PD patients. Further CNVs were observed in other neurological disease-related panels and were validated by qPCR (data not shown).

APPLICATION OF AN INTEGRATED ADM-1 AND ADM-2 ALGORITHM-BASED DATA ANALYSIS TO IMPROVE CNV CALLING

The DNA sample of patient #1 was referred to our laboratory for molecular testing of *PARK2*, *PINK1* and *DJ1*, to confirm the clinical diagnosis of familial recessive early-onset PD. Mutation analysis showed a heterozygous C1305T single nucleotide substitution in the coding region of *PARK2*. *NeuroArray* (with the ADM-1 algorithm) revealed 10 different CNVs, overall composed of 6 gains and 4 losses (**Table 2**). Four of them included genes previously linked to PD [29-32], while the others overlapped with genes related to other neurological conditions [33-42] (**Table 2**).

The most interesting findings regarded two principal dosage anomalies: i) the gain of a 1442-kb region on chromosome 1, which encompasses *PARK7*, and ii) the loss of the *NSF* (N-ethylmaleimide sensitive factor) gene on chromosome 17 (**Figure 3, panels A and B**). Mutations in *PARK7* comprehensively account for ~1% of the early-onset familial cases [1], and its copy number changes have been previously observed in PD patients [43,44]. *NSF* is involved in vesicular trafficking, membrane fusion, and synaptic neurotransmission, and its genetic alterations (both SNPs and deletion) have been previously reported in PD patients [45,31]. Validations of these genomic rearrangements were performed with qPCR assays, suitably designed to target *PARK7* exon 1 and *NSF* exon 11. Both assays confirmed the CNVs with 100% concordance and confirmed the heterozygous gain/loss (**Figure 3, panel C**). Primer sequences and PCR conditions are available upon request.

It should be highlighted that the default analysis with the ADM-2 algorithm revealed the loss of only the *NSF* gene. If this method were the only one applied, other relevant real CNVs (like the *PARK7* gain, later confirmed by qPCR) would have been lost. On the other hand, the analysis with ADM-2 allowed for the filtering of possible false positive CNVs within the ADM-1 analysis. It appears important, therefore, to integrate data from both CNV calling algorithms in order to provide a more accurate data analysis and, consequently, ensure a more effective quality assessment and experimental validation.

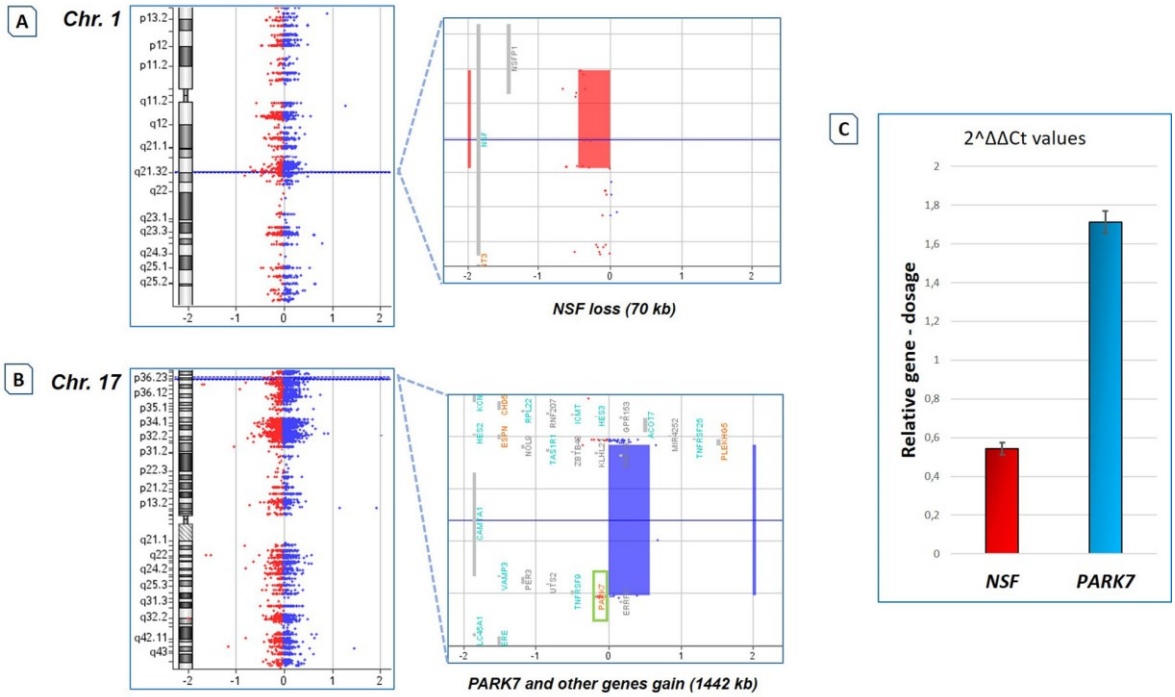


Figure 2. A representative example of CNV detection involving PD-related genes in a patient with early-onset PD.

The *NeuroArray* platform detected several CNVs in a female patient with early-onset PD and a mild phenotype (the reader is also referred to Table 2). **A:** Visualization of the *NSF* deletion detected by *NeuroArray* as shown by CytoGenomics software. The left panel shows the entire chromosome 1, while the right panel is a zoom in of the deleted region (indicated by the red area). Red and blue dots represent the log2 ratios for the relative hybridization intensities of each spotted probe. **B:** Visualization of the *PARK7* amplification detected by *NeuroArray* as shown by CytoGenomics software. The left panel shows the entire chromosome 17, while the right panel is a zoom in of the amplified region (indicate by the blue area). For red and blue dots, see panel A. Dots with an average log2 ratio of approximately +0.58 indicate a heterozygous amplification. **C:** Validation of both CNVs of *NSF* ad *PARK7* by qPCR. Relative gene dosage levels of *NSF* and *PARK7* genes are based on delta Ct calculation. Ct values of both genes were normalized to the Ct value of a normal diploid gene. The relative level of each gene of interest is presented as the mean of 2-ΔΔCt, as described in the methods section. Error bars indicate standard deviations from the mean.

CNV type	Chr	Start - Stop (bp)	Size (kb)	Cytoband	#Probes	Annotations	Disease-linked genes	Ref.	Previously described in DGV?	ClinGen nomenclature (Described pathogenic variants)
Gain	1	6579851-8021801	1,442	p36.31 - p36.23	3	PLEKHG5, NOL9, TAS1R1, ZBTB48, THAP3, DNAJC11, KLHL21, PHF13, CAMTA1, VAMP3, PER3, UTS2, TNFRSF9, **PARK7**	PLEKHG5 (Neuropathies and ALS), PARK7 (**PD** and ALS)	[2, 39-41]	Overlaps nsv1004598	Almost completely contained within nssv1602095, nssv578510, nssv578509
Gain	1	54704828-54747170	42	p32.3	12	**SSBP3**	**PD**	[29]	No	Completely contained within nssv578523 and nssv578522
Gain	2	179541928-179542606	1	q31.2	3	TTN	Limb-Girdle Muscular Dystrophy	[38]	No	Not reported
Loss	3	142216000-142222244	6	q23	6	ATR	Epilepsy	[42]	Overlaps nsv528954	Completely contained within nssv583804, nssv577921, nssv577927, nssv583018, nssv1602712, nssv3395109
Loss	3	178922270-178927435	5	q26.32	3	PIK3CA	Epilepsy	[37]	No	Not reported
Gain	5	70307077-70308602	2	q13.2	4	NAIP	ALS	[36]	Yes (completely contained in several genomic structural variants)	Completely contained within nssv1602328
Gain	9	131012433-131314975	303	q34.11	6	DNM1, GOLGA2, C9orf119, TRUB2, COQ4, SLC27A4, URM1, CERCAM, ODF2, GLE1, SPTAN1	DNM1 (Epilepsy), SPTAN1 (Epilepsy)	[34, 35]	Overlaps nsv1051081	Completely contained within nssv579112, nssv579118, nssv579123, nssv579147, nssv579149, nssv576650, nssv584344, nssv584434, nssv579136, nssv579139, nssv579121, nssv579124, nssv579127, nssv579138, nssv579140, nssv1602335, nssv1494937, nssv1415412, nssv1603388, nssv3397050, nssv3397066, nssv3397108

Loss	9	139903473-139904058	1	q34.3	4	ABCA2	AD	[33]	Yes (completely contained in several genomic structural variants)	Completely contained within nssv579112, nssv579118, nssv579123, nssv579147, nssv579149, nssv576610, nssv576650, nssv584344, nssv584430, nssv584434, nssv579136, nssv579139, nssv1603388, nssv1604614, nssv3397050, nssv3397066, nssv3397108, nssv579121, nssv579124, nssv579127, nssv579138, nssv579145, nssv707149, nssv582161, nssv1602986, nssv1494937, nssv1415412, nssv1602335, nssv1494929
Gain	12	56615423-56615694	0,27	q13.3	3	**RNF41**	**PD**	[30]	Yes (completely contained within nsv1051961)	Completely contained with nssv1603959
Loss	17	44701610-44771900	70	q21.31	14	**NSF**, NSFP1	**PD**	[45]	Yes (completely contained in several genomic structural variants)	Not reported

The *NeuroArray* aCGH (ADM-1 analysis method) revealed 10 different CNVs overall composed of 6 gains and 4 losses. Four of them overlapped with genes belonging to the PD panel (in bold type), while the others mainly included ALS or epilepsy-related genes. Overlapping structural variants previously described and deposited in DGV (http://dgv.tcag.ca/) or ClinGen (https://www.clinicalgenome.org/) databases are also indicated. Start and stop coordinates refer to UCSC Genome Browser Human Feb. 2009 Assembly (GRCh37/hg19).

DETECTION OF SINGLE-EXON COPY NUMBER CHANGES BY NEUROARRAY

Although some authors have outlined the evidence that a significant proportion of single probe intervals represents real events [46], in aCGH studies, it is often recommended to report only intervals detected by three or more consecutive probes. Due to this approach, deletions or duplications below certain size cut-offs are usually ignored in the aCGH reports and not reported. However, these genomic alterations (detected by less than three probes) have been demonstrated to be definitively crucial for particular clinical diagnoses [47]. Along this line, we applied a single probe analysis to reveal short genomic imbalances in the exonic regions of strongly linked causative genes. The utility of this approach on *NeuroArray* data analysis is shown in the following case.

Patient #2 was a sporadic PD patient, carrying a heterozygous deletion of two adjacent exons (4 and 5) of the *PARK2* gene. This deletion was previously revealed by an MLPA assay (SALSA MLPA Kit P051/P052 Parkinson; MRC-Holland). The *NeuroArray* test was able to detect and confirm the deletion of exon 5 through two consecutive probes (**Figure 4**) but was not able to detect the exon 4 deletion because during the phase of array design, this exon skipped the optimum parameters for probe coverage. The total concordance with the MLPA test was 91%. Despite this limit, the one-probe analysis was essential to detect the exon 5 *PARK2* deletion, which otherwise would not have been properly outlined using the analysis of three consecutive probes. However, this approach may result in a great number of false positives. Therefore, it is advisable to use it as a validation strategy for previously known exonic imbalances, i.e., NGS targeted panels, or to investigate copy number changes in a small set of strongly causative genes.

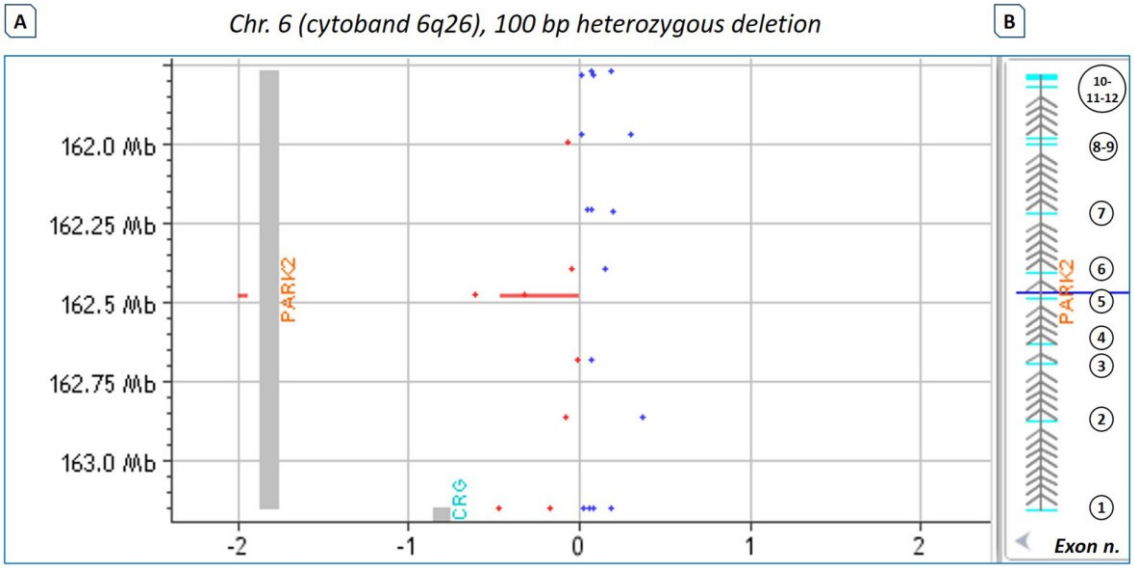


Figure 3. Detection of intragenic *PARK2* deletion (exon 5) in a patient with Autosomal Juvenile Parkinson's disease.

Heterozygous deletion of exon 5 of the *PARK2* gene detected by *NeuroArray* in a patient with Juvenile Parkinson's disease (PD) and previously revealed by an MLPA assay. **A:** *NeuroArray* aCGH data visualization and analysis as shown by CytoGenomics software. The red area represents the deleted region. The top of the panel shows the size of the deletion and the chromosomal locus. Red and blue dots represent the log2 ratios for the relative hybridization intensities of each spotted probe. The dots with an average log2 ratio around −1 indicate a heterozygous deletion. B: The panel displays the *PARK2* gene as annotated in the UCSC Genome Browser Feb. 2009 GRCh37/hg19 (https://genome-euro.ucsc.edu). Blue boxes represent exons and are numbered consecutively, whereas grey arrows are the intronic regions.

DISCUSSION

In recent years, several studies have highlighted the key role of CNVs in the development of hereditable or sporadic neurological diseases, including PD [14-16]. Many gene-dosage anomalies have been previously mapped in PD patients, including familiar genes (*SNCA, PARK2, PINK1, PARK7, ATP13A2*) [48,49], as well as several rare CNVs in candidate regions [45]. The aCGH biotechnology currently represents a useful tool for the detection of unbalanced chromosomal changes across the human genome, and its applications to screen common benign and rare pathogenetic CNVs are extensively growing [19-23]. The classical methodologic approaches are a gold-standard test when applied to monogenic disorders, but when applied to multigenic complex pathologies (such as PD), they require higher equipment costs, time, steps and personnel [50]. Conversely, targeted aCGH is rapid, relatively inexpensive, highly sensitive and an accurate method to simultaneously detect single- and multi-exon CNVs in numerous genes on a unique common platform. For this reason, several

whole-genome and exon-targeted aCGH platforms have already been implemented in human diseases [19-24], and their utility has been demonstrated in patients with various clinical complex phenotypes [51-53].

In this study, we have designed and validated a targeted exon-centric aCGH platform (*NeuroArray*) as a molecular testing tool to simultaneously screen CNV imbalances in a large set of clinically relevant genes for PD and other complex neurological diseases. This customized design offers some considerable advantages: it allows an exon-focused evaluation of structural imbalances in clinically relevant regions at a higher resolution than whole-genome commercially available platforms and lowers the costs of an "exon by exon" analysis through PCR-based approaches, simultaneously providing an extensive window of further potentially involved genetic alterations.

In addition to the customized design, we also applied several approaches for data analysis. The first interesting result was the need to integrate data from both the ADM-1 and ADM-2 algorithms for CNV calling aberrations in order to reduce the number of false positives and to bring out relevant CNVs that otherwise would have been lost. We have also employed a one-probe analysis to reveal small imbalances at the single-exon level. Although this approach has the potential to detect crucial genetic variations ignored by multi-probe analysis, it largely increases the quantity of false-positive probe signals. Therefore, the single-probe analysis would be a useful validation strategy for NGS experiments or to investigate exon copy number changes in a smaller set of causative genes (as we performed with the script in the R-platform).

The use of dedicated high-throughput genotyping platforms like our *NeuroArray* could offer new opportunities for the PD genomic research field, mainly for familiar PD cases with an incomplete molecular diagnosis or sporadic cases without any detected genetic anomalies. The large-scale screening of genes that are involved in nervous system dysfunctions could allow for differential diagnosis with other common neurological disorders, refine the genotype-phenotype correlations and explore the potential genetic overlapping signatures among different neurological conditions [54]. Specifically, the PD panel shares a good number of genes with other neurological diseases (**Figure 1**). Given the existence of PD patients with combined clinical and pathological features [55-57], this strategy could be useful to investigate common genetic anomalies underlying very complex phenotypes.

Similarly to other aCGH-based technology, *NeuroArray* has some limitations, such as the inability to detect mosaicism poorly represented, balanced structural chromosomal abnormalities, nucleotide repeat expansions (e.g., in *C9orf72* or *ATXN2* genes), and mutations included in regions not covered by probes. To overcome some of these limits and reduce the number of false positive signals, we are developing a second version of the *NeuroArray* design with the aim of improving probe coverage in non-targeted genomic regions, including (where necessary) the intronic flanking regions and the alternatively spliced cassette exons of relevant PD genes [58-60].

CONCLUSIONS

Our *NeuroArray* platform represents a powerful and reliable tool for the analysis of genomic imbalances associated with PD and other neurological diseases. Compared to PCR-based approaches applied to multigene analysis or to whole-genome commercially available CGH arrays, it provides a focused higher resolution at a lower cost, enabling a more detailed analysis of clinically relevant exonic regions and offering a better cost/benefit ratio. In future years, the use of this platform may offer new insights into the investigation of new genetic molecular anomalies contributing to PD, as well as a more precise definition of genotype-phenotype relationships. It may also offer novel clues in the elucidation of potential genetic overlapping among different neurological conditions.

LIST OF ABBREVIATIONS

aCGH: Comparative Genomic Hybridization array; **AD:** Alzheimer's disease; **ALS:** Amyotrophic Lateral Sclerosis; **BMD:** Becker Muscular Dystrophy; **CNVs:** Copy Number Variations; **DMD:** Duchenne Muscular Dystrophy; **HSP:** Hereditary Spastic Paraplegia; **LGMD:** Limb-Girdle Muscular Dystrophy; **MLPA**: Multiplex Ligation-dependent Probe Amplification; **NF**: Neurofibromatosis; **NGS**: Next Generation Sequencing; **PD:** Parkinson's disease; **PN**: Peripheral Neuropathy; **qPCR:** Quantitative Polymerase Chain Reaction; **RTT:** Rett Syndrome; **SCA:** Spinocerebellar Ataxia; **TSC:** Tuberous sclerosis.

ACKNOWLEDGEMENTS

This work was supported by the Italian Ministry of Education, Universities and Research through grant CTN01_00177_817708 and the international Ph.D. program in Neuroscience of the University of Catania. The authors gratefully acknowledge Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo, and Francesco Marino for their administrative and technical support.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

ETHICAL APPROVAL

Experiments involving human participants have been approved by an ethical committee for medical research and have been performed in accordance with ethical standards.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

REFERENCES

1. Moore DJ, West AB, Dawson VL, Dawson TM (2005) Molecular pathophysiology of Parkinson's disease. Annual review of neuroscience 28:57-87. doi:10.1146/annurev.neuro.28.061604.135718

2. Klein C, Westenberger A (2012) Genetics of Parkinson's disease. Cold Spring Harbor perspectives in medicine 2 (1):a008888. doi:10.1101/cshperspect.a008888

3. Deng HX, Shi Y, Yang Y, Ahmeti KB, Miller N, Huang C, Cheng L, Zhai H, Deng S, Nuytemans K, Corbett NJ, Kim MJ, Deng H, Tang B, Yang Z, Xu Y, Chan P, Huang B, Gao XP, Song Z, Liu Z, Fecto F, Siddique N, Foroud T, Jankovic J, Ghetti B, Nicholson DA, Krainc D, Melen O, Vance JM, Pericak-Vance MA, Ma YC, Rajput AH, Siddique T (2016) Identification of TMEM230 mutations in familial Parkinson's disease. Nature genetics 48 (7):733-739. doi:10.1038/ng.3589

4. Olgiati S, Quadri M, Fang M, Rood JP, Saute JA, Chien HF, Bouwkamp CG, Graafland J, Minneboo M, Breedveld GJ, Zhang J, International Parkinsonism Genetics N, Verheijen FW, Boon AJ, Kievit AJ, Jardim LB, Mandemakers W, Barbosa ER, Rieder CR, Leenders KL, Wang J, Bonifati V (2016) DNAJC6 Mutations Associated With Early-Onset Parkinson's Disease. Annals of neurology 79 (2):244-256. doi:10.1002/ana.24553

5. Wilson GR, Sim JC, McLean C, Giannandrea M, Galea CA, Riseley JR, Stephenson SE, Fitzpatrick E, Haas SA, Pope K, Hogan KJ, Gregg RG, Bromhead CJ, Wargowski DS, Lawrence CH, James PA, Churchyard A, Gao Y, Phelan DG, Gillies G, Salce N, Stanford L, Marsh AP, Mignogna ML, Hayflick SJ, Leventer RJ, Delatycki MB, Mellick GD, Kalscheuer VM, D'Adamo P, Bahlo M, Amor DJ, Lockhart PJ (2014) Mutations in RAB39B cause X-linked intellectual disability and early-onset Parkinson disease with alpha-synuclein pathology. American journal of human genetics 95 (6):729-735. doi:10.1016/j.ajhg.2014.10.015

6. Mata IF, Jang Y, Kim CH, Hanna DS, Dorschner MO, Samii A, Agarwal P, Roberts JW, Klepitskaya O, Shprecher DR, Chung KA, Factor SA, Espay AJ, Revilla FJ, Higgins DS, Litvan I, Leverenz JB, Yearout D, Inca-Martinez M, Martinez E, Thompson TR, Cholerton BA, Hu SC, Edwards KL, Kim KS, Zabetian CP (2015) The RAB39B p.G192R

mutation causes X-linked dominant Parkinson's disease. Molecular neurodegeneration 10:50. doi:10.1186/s13024-015-0045-4

7. Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, Li Y, Ogaki K, Ando M, Yoshino H, Tomiyama H, Nishioka K, Hasegawa K, Saiki H, Satake W, Mogushi K, Sasaki R, Kokubo Y, Kuzuhara S, Toda T, Mizuno Y, Uchiyama Y, Ohno K, Hattori N (2015) CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. The Lancet Neurology 14 (3):274-282. doi:10.1016/S1474-4422(14)70266-2

8. Chartier-Harlin MC, Dachsel JC, Vilarino-Guell C, Lincoln SJ, Lepretre F, Hulihan MM, Kachergus J, Milnerwood AJ, Tapia L, Song MS, Le Rhun E, Mutez E, Larvor L, Duflot A, Vanbesien-Mailliot C, Kreisler A, Ross OA, Nishioka K, Soto-Ortolaza AI, Cobb SA, Melrose HL, Behrouz B, Keeling BH, Bacon JA, Hentati E, Williams L, Yanagiya A, Sonenberg N, Lockhart PJ, Zubair AC, Uitti RJ, Aasly JO, Krygowska-Wajs A, Opala G, Wszolek ZK, Frigerio R, Maraganore DM, Gosal D, Lynch T, Hutchinson M, Bentivoglio AR, Valente EM, Nichols WC, Pankratz N, Foroud T, Gibson RA, Hentati F, Dickson DW, Destee A, Farrer MJ (2011) Translation initiator EIF4G1 mutations in familial Parkinson disease. American journal of human genetics 89 (3):398-406. doi:10.1016/j.ajhg.2011.08.009

9. Vilarino-Guell C, Soto AI, Lincoln SJ, Ben Yahmed S, Kefi M, Heckman MG, Hulihan MM, Chai H, Diehl NN, Amouri R, Rajput A, Mash DC, Dickson DW, Middleton LT, Gibson RA, Hentati F, Farrer MJ (2009) ATP13A2 variability in Parkinson disease. Human mutation 30 (3):406-410. doi:10.1002/humu.20877

10. Lesage S, Brice A (2012) Role of Mendelian genes in “sporadic” Parkinson's disease. Parkinsonism & Related Disorders 18, Supplement 1:S66-S70. doi:http://dx.doi.org/10.1016/S1353-8020(11)70022-0

11. Lin MK, Farrer MJ (2014) Genetics and genomics of Parkinson's disease. Genome medicine 6 (6):48. doi:10.1186/gm566

12. Zarrei M, MacDonald JR, Merico D, Scherer SW (2015) A copy number variation map of the human genome. Nature reviews Genetics 16 (3):172-183. doi:10.1038/nrg3871

13. Carvalho CM, Lupski JR (2016) Mechanisms underlying structural variant formation in genomic disorders. Nature reviews Genetics. doi:10.1038/nrg.2015.25

14. Lee JA, Lupski JR (2006) Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 52 (1):103-121. doi:10.1016/j.neuron.2006.09.027

15. Kalman B, Vitale E (2009) Structural chromosomal variations in neurological diseases. The neurologist 15 (5):245-253. doi:10.1097/NRL.0b013e3181963cef

16. Wang L, Nuytemans K, Bademci G, Jauregui C, Martin ER, Scott WK, Vance JM, Zuchner S (2013) High-resolution survey in familial Parkinson disease genes reveals multiple independent copy number variation events in PARK2. Human mutation 34 (8):1071-1074. doi:10.1002/humu.22344

17. Pankratz N, Dumitriu A, Hetrick KN, Sun M, Latourelle JC, Wilk JB, Halter C, Doheny KF, Gusella JF, Nichols WC, Myers RH, Foroud T, DeStefano AL, Psg P, GenePd Investigators C, Molecular Genetic L (2011) Copy number variation in familial Parkinson disease. PloS one 6 (8):e20988. doi:10.1371/journal.pone.0020988

18. Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, Firth H, Sanlaville D, Winter R, Colleaux L, Bobrow M, Carter NP (2004) Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. Journal of medical genetics 41 (4):241-248

19. Ferlini A, Bovolenta M, Neri M, Gualandi F, Balboni A, Yuryev A, Salvi F, Gemmati D, Liboni A, Zamboni P (2010) Custom CGH array profiling of copy number variations (CNVs) on chromosome 6p21.32 (HLA locus) in patients with venous malformations associated with multiple sclerosis. BMC medical genetics 11:64. doi:10.1186/1471-2350-11-64

20. Piluso G, Dionisi M, Del Vecchio Blanco F, Torella A, Aurino S, Savarese M, Giugliano T, Bertini E, Terracciano A, Vainzof M, Criscuolo C, Politano L, Casali C, Santorelli FM, Nigro V (2011) Motor chip: a comparative genomic hybridization microarray for copy-number mutations in 245 neuromuscular disorders. Clinical chemistry 57 (11):1584-1596. doi:10.1373/clinchem.2011.168898

21. Inazawa J, Inoue J, Imoto I (2004) Comparative genomic hybridization (CGH)-arrays pave the way for identification of novel cancer-related genes. Cancer science 95 (7):559-563

22. Wisniowiecka-Kowalnik B, Kastory-Bronowska M, Bartnik M, Derwinska K, Dymczak-Domini W, Szumbarska D, Ziemka E, Szczaluba K, Sykulski M, Gambin T, Gambin A, Shaw CA, Mazurczak T, Obersztyn E, Bocian E, Stankiewicz P (2013) Application of custom-designed oligonucleotide array CGH in 145 patients with autistic spectrum disorders. European journal of human genetics : EJHG 21 (6):620-625. doi:10.1038/ejhg.2012.219

23. Mefford HC, Muhle H, Ostertag P, von Spiczak S, Buysse K, Baker C, Franke A, Malafosse A, Genton P, Thomas P, Gurnett CA, Schreiber S, Bassuk AG, Guipponi M, Stephani U, Helbig I, Eichler EE (2010) Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS genetics 6 (5):e1000962. doi:10.1371/journal.pgen.1000962

24. Wang J, Rakhade M (2012) Utility of array CGH in molecular diagnosis of mitochondrial disorders. Methods in molecular biology 837:301-312. doi:10.1007/978-1-61779-504-6_20

25. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BM, Schjeide LM, Meissner E, Zauft U, Allen NC, Liu T, Schilling M, Anderson KJ, Beecham G, Berg D, Biernacka JM, Brice A, DeStefano AL, Do CB, Eriksson N, Factor SA, Farrer MJ, Foroud T, Gasser T, Hamza T, Hardy JA, Heutink P, Hill-Burns EM, Klein C, Latourelle JC, Maraganore DM, Martin ER, Martinez M, Myers RH, Nalls MA, Pankratz N, Payami H, Satake W, Scott WK, Sharma M, Singleton AB, Stefansson K, Toda T, Tung JY, Vance J, Wood NW, Zabetian CP, andMe Genetic Epidemiology of Parkinson's Disease C, International Parkinson's Disease Genomics C, Parkinson's Disease GC, Wellcome Trust Case Control C, Young P, Tanzi RE, Khoury MJ, Zipp F, Lehrach H, Ioannidis JP, Bertram L (2012) Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. PLoS genetics 8 (3):e1002548. doi:10.1371/journal.pgen.1002548

26. Team RC (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0,

27. D'haene B, Vandesompele J, Hellemans J (2010) Accurate and objective copy number profiling using real-time quantitative PCR. Methods 50 (4):262-270. doi:http://dx.doi.org/10.1016/j.ymeth.2009.12.007

28. Patitucci A, Muglia M, Magariello A, Gabriele AL, Peluso G, Sprovieri T, Conforti FL, Mazzei R, Ungaro C, Condino F, Valentino P, Bono F, Rodolico C, Mazzeo A, Toscano A, Vita G, Quattrone A (2005) Comparison of different techniques for detecting 17p12 duplication in CMT1A. Neuromuscular Disorders 15 (7):488-492. doi:http://dx.doi.org/10.1016/j.nmd.2005.04.006

29. Haugarvoll K, Toft M, Skipper L, Heckman MG, Crook JE, Soto A, Ross OA, Hulihan MM, Kachergus JM, Sando SB, White LR, Lynch T, Gibson JM, Uitti RJ, Wszolek ZK, Aasly JO, Farrer MJ (2009) Fine-mapping and candidate gene investigation within the PARK10 locus. European journal of human genetics : EJHG 17 (3):336-343. doi:10.1038/ejhg.2008.187

30. Mo X, Liu D, Li W, Hu Z, Hu Y, Li J, Guo J, Tang B, Zhang Z, Bai Y, Xia K (2010) Genetic screening for mutations in the Nrdp1 gene in Parkinson disease patients in a Chinese population. Parkinsonism & related disorders 16 (3):222-224. doi:10.1016/j.parkreldis.2009.09.001

31. Liu X, Cheng R, Verbitsky M, Kisselev S, Browne A, Mejia-Sanatana H, Louis ED, Cote LJ, Andrews H, Waters C, Ford B, Frucht S, Fahn S, Marder K, Clark LN, Lee JH (2011) Genome-wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. BMC medical genetics 12:104. doi:10.1186/1471-2350-12-104

32. Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, Dekker MC, Squitieri F, Ibanez P, Joosse M, van Dongen JW, Vanacore N, van Swieten JC, Brice A, Meco G, van Duijn CM, Oostra BA, Heutink P (2003) Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science 299 (5604):256-259. doi:10.1126/science.1077209

33. Mace S, Cousin E, Ricard S, Genin E, Spanakis E, Lafargue-Soubigou C, Genin B, Fournel R, Roche S, Haussy G, Massey F, Soubigou S, Brefort G, Benoit P, Brice A, Campion D, Hollis M, Pradier L, Benavides J, Deleuze JF (2005) ABCA2 is a strong genetic risk factor for early-onset Alzheimer's disease. Neurobiology of disease 18 (1):119-125. doi:10.1016/j.nbd.2004.09.011

34. Euro E-RESC, Epilepsy Phenome/Genome P, Epi KC (2014) De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. American journal of human genetics 95 (4):360-370. doi:10.1016/j.ajhg.2014.08.013

35. Tohyama J, Nakashima M, Nabatame S, Gaik-Siew C, Miyata R, Rener-Primec Z, Kato M, Matsumoto N, Saitsu H (2015) SPTAN1 encephalopathy: distinct phenotypes and genotypes. Journal of human genetics 60 (4):167-173. doi:10.1038/jhg.2015.5

36. Jackson M, Morrison KE, Al-Chalabi A, Bakker M, Leigh PN (1996) Analysis of chromosome 5q13 genes in amyotrophic lateral sclerosis: homozygous NAIP deletion in a sporadic case. Annals of neurology 39 (6):796-800. doi:10.1002/ana.410390616

37. Lee JH, Huynh M, Silhavy JL, Kim S, Dixon-Salazar T, Heiberg A, Scott E, Bafna V, Hill KJ, Collazo A, Funari V, Russ C, Gabriel SB, Mathern GW, Gleeson JG (2012) De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. Nature genetics 44 (8):941-945. doi:10.1038/ng.2329

38. Zheng W, Chen H, Deng X, Yuan L, Yang Y, Song Z, Yang Z, Wu Y, Deng H (2015) Identification of a Novel Mutation in the Titin Gene in a Chinese Family with Limb-Girdle Muscular Dystrophy 2J. Molecular neurobiology. doi:10.1007/s12035-015-9439-0

39. Hanagasi HA, Giri A, Guven G, Bilgic B, Hauser AK, Emre M, Heutink P, Basak N, Gasser T, Simon-Sanchez J, Lohmann E (2016) A novel homozygous DJ1 mutation causes parkinsonism and ALS in a Turkish family. Parkinsonism & related disorders. doi:10.1016/j.parkreldis.2016.03.001

40. Kim HJ, Hong YB, Park JM, Choi YR, Kim YJ, Yoon BR, Koo H, Yoo JH, Kim SB, Park M, Chung KW, Choi BO (2013) Mutations in the PLEKHG5 gene is relevant with autosomal recessive intermediate Charcot-Marie-Tooth disease. Orphanet journal of rare diseases 8:104. doi:10.1186/1750-1172-8-104

41. Ozoguz A, Uyan O, Birdal G, Iskender C, Kartal E, Lahut S, Omur O, Agim ZS, Eken AG, Sen NE, Kavak P, Saygi C, Sapp PC, Keagle P, Parman Y, Tan E, Koc F, Deymeer F, Oflazer P, Hanagasi H, Gurvit H, Bilgic B, Durmus H, Ertas M, Kotan D, Akalin MA, Gulluoglu H, Zarifoglu M, Aysal F, Dosoglu N, Bilguvar K, Gunel M, Keskin O, Akgun T, Ozcelik H, Landers JE, Brown RH, Basak AN (2015) The distinct genetic pattern of ALS in Turkey and novel mutations. Neurobiology of aging 36 (4):1764 e1769-1718. doi:10.1016/j.neurobiolaging.2014.12.032

42. Lemke JR, Riesch E, Scheurenbrand T, Schubach M, Wilhelm C, Steiner I, Hansen J, Courage C, Gallati S, Burki S, Strozzi S, Simonetti BG, Grunt S, Steinlin M, Alber M, Wolff M, Klopstock T, Prott EC, Lorenz R, Spaich C, Rona S, Lakshminarasimhan M, Kroll J, Dorn T, Kramer G, Synofzik M, Becker F, Weber YG, Lerche H, Bohm D, Biskup S (2012) Targeted next generation sequencing as a diagnostic tool in epileptic disorders. Epilepsia 53 (8):1387-1398. doi:10.1111/j.1528-1167.2012.03516.x

43. Guo JF, Zhang XW, Nie LL, Zhang HN, Liao B, Li J, Wang L, Yan XX, Tang BS (2010) Mutation analysis of Parkin, PINK1 and DJ-1 genes in Chinese patients with sporadic early onset parkinsonism. Journal of neurology 257 (7):1170-1175. doi:10.1007/s00415-010-5485-8

44. Hedrich K, Djarmati A, Schafer N, Hering R, Wellenbrock C, Weiss PH, Hilker R, Vieregge P, Ozelius LJ, Heutink P, Bonifati V, Schwinger E, Lang AE, Noth J, Bressman SB, Pramstaller PP, Riess O, Klein C (2004) DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. Neurology 62 (3):389-394

45. Liu X, Cheng R, Ye X, Verbitsky M, Kisselev S, Mejia-Santana H, Louis E, Cote L, Andrews H, Waters C, Ford B, Fahn S, Marder K, Lee J, Clark L (2013) Increased Rate of Sporadic and Recurrent Rare Genic Copy Number Variants in Parkinson's Disease Among Ashkenazi Jews. Molecular genetics & genomic medicine 1 (3):142-154. doi:10.1002/mgg3.18

46. de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, Tsang P, Ben-Dor A, Yakhini Z, Ellis RJ, Bruhn L, Laderman S, Froguel P, Blakemore AI (2007) Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. Hum Mol Genet 16 (23):2783-2794. doi:10.1093/hmg/ddm208

47. Mendelsohn BA, Pronold M, Long R, Smaoui N, Slavotinek AM (2014) Advanced bone age in a girl with Wiedemann-Steiner syndrome and an exonic deletion in KMT2A (MLL). American journal of medical genetics Part A 164A (8):2079-2083. doi:10.1002/ajmg.a.36590

48. Darvish H, Movafagh A, Omrani MD, Firouzabadi SG, Azargashb E, Jamshidi J, Khaligh A, Haghnejad L, Naeini NS, Talebi A, Heidari-Rostami HR, Noorollahi-Moghaddam H, Karkheiran S, Shahidi GA, Paknejad SM, Ashrafian H, Abdi S, Kayyal M, Akbari M, Pedram N, Emamalizadeh B (2013) Detection of copy number changes in genes associated with Parkinson's disease in Iranian patients. Neuroscience letters 551:75-78. doi:10.1016/j.neulet.2013.07.013

49. Toft M, Ross OA (2010) Copy number variation in Parkinson's disease. Genome medicine 2 (9):62. doi:10.1186/gm183

50. Tayeh MK, Chin EL, Miller VR, Bean LJ, Coffee B, Hegde M (2009) Targeted comparative genomic hybridization array for the detection of single- and multiexon gene deletions and duplications. Genetics in medicine : official journal of the American College of Medical Genetics 11 (4):232-240. doi:10.1097/GIM.0b013e318195e191

51. Boone PM, Bacino CA, Shaw CA, Eng PA, Hixson PM, Pursley AN, Kang SH, Yang Y, Wiszniewska J, Nowakowska BA, del Gaudio D, Xia Z, Simpson-Patel G, Immken LL, Gibson JB, Tsai AC, Bowers JA, Reimschisel TE, Schaaf CP, Potocki L, Scaglia F, Gambin T, Sykulski M, Bartnik M, Derwinska K, Wisniowiecka-Kowalnik B, Lalani SR, Probst FJ, Bi W, Beaudet AL, Patel A, Lupski JR, Cheung SW, Stankiewicz P (2010) Detection of clinically relevant exonic copy-number changes by array CGH. Human mutation 31 (12):1326-1342. doi:10.1002/humu.21360

52. Tucker T, Zahir FR, Griffith M, Delaney A, Chai D, Tsang E, Lemyre E, Dobrzeniecka S, Marra M, Eydoux P, Langlois S, Hamdan FF, Michaud JL, Friedman JM (2014) Single exon-resolution targeted chromosomal microarray analysis of known and candidate intellectual disability genes. European journal of human genetics : EJHG 22 (6):792-800. doi:10.1038/ejhg.2013.248

53. Aradhya S, Lewis R, Bonaga T, Nwokekeh N, Stafford A, Boggs B, Hruska K, Smaoui N, Compton JG, Richard G, Suchy S (2012) Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders. Genetics in medicine : official journal of the American College of Medical Genetics 14 (6):594-603. doi:10.1038/gim.2011.65

54. Desikan RS, Schork AJ, Wang Y, Witoelar A, Sharma M, McEvoy LK, Holland D, Brewer JB, Chen CH, Thompson WK, Harold D, Williams J, Owen MJ, O'Donovan MC, Pericak-Vance MA, Mayeux R, Haines JL, Farrer LA, Schellenberg GD, Heutink P, Singleton AB, Brice A, Wood NW, Hardy J, Martinez M, Choi SH, DeStefano A, Ikram MA, Bis JC, Smith A, Fitzpatrick AL, Launer L, van Duijn C, Seshadri S, Ulstein ID, Aarsland D, Fladby T, Djurovic S, Hyman BT, Snaedal J, Stefansson H, Stefansson K, Gasser T, Andreassen OA, Dale AM (2015) Genetic overlap between Alzheimer's disease and Parkinson's disease at the MAPT locus. Molecular psychiatry. doi:10.1038/mp.2015.6

55. Zaleski C, Bassett AS, Tam K, Shugar AL, Chow EW, McPherson E (2009) The co-occurrence of early onset Parkinson disease and 22q11.2 deletion syndrome. American journal of medical genetics Part A 149A (3):525-528. doi:10.1002/ajmg.a.32650

56. Annesi G, Savettieri G, Pugliese P, D'Amelio M, Tarantino P, Ragonese P, La Bella V, Piccoli T, Civitelli D, Annesi F, Fierro B, Piccoli F, Arabia G, Caracciolo M, Ciro Candiano IC, Quattrone A (2005) DJ-1 mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex. Annals of neurology 58 (5):803-807. doi:10.1002/ana.20666

57. Chung SJ, Jung Y, Hong M, Kim MJ, You S, Kim YJ, Kim J, Song K (2013) Alzheimer's disease and Parkinson's disease genome-wide association study top hits and risk of Parkinson's disease in Korean population. Neurobiology of aging 34 (11):2695 e2691-2697. doi:10.1016/j.neurobiolaging.2013.05.022

58. La Cognata V, D'Agata V, Cavalcanti F, Cavallaro S (2015) Splicing: is there an alternative contribution to Parkinson's disease? Neurogenetics. doi:10.1007/s10048-015-0449-x

59. La Cognata V, Iemmolo R, D'Agata V, Scuderi S, Drago F, Zappia M, Cavallaro S (2014) Increasing the Coding Potential of Genomes Through Alternative Splicing: The Case of PARK2 Gene. Current genomics 15 (3):203-216. doi:10.2174/1389202915666140426003342

60. Scuderi S, La Cognata V, Drago F, Cavallaro S, D'Agata V (2014) Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain. BioMed research international 2014:690796. doi:10.1155/2014/690796

SUPPLEMENTARY INFORMATIONS

Gene selection and aCGH design strategy. 505 PD-related genes were selected from PDgene (http://www.pdgene.org/), a comprehensive, unbiased and regularly updated database of genetic association studies performed in PD (Release February 2013). In order to perform a differential diagnosis with other neurological disorders, we also included genes related to other neurological conditions, such as Amyotrophic Lateral Sclerosis (ALS), Epilepsies, Rett Syndrome (RTT), Autosomal dominant and recessive Limb-Girdle Muscular Dystrophy (LGMD), Muscular Duchenne (DMD)/Becker Dystrophy (BMD), Hereditary Spastic Paraplegia (HSP), Spinocerebellar Ataxia (SCA), Neurofibromatosis (NF), Tuberous sclerosis (TSC), Peripheral Neuropathy (PN) and Stroke. The genes related to these disorders were selected from literature or screened from other databases available online (ALZgene - http://www.alzgene.org/ and ALSgene - http://www.alsgene.org/).

The array design was carried out by using the web-based Agilent SureDesign Software (Advanced Design Wizard option), version 1.2.1.15 (Agilent Technologies, Santa Clara, CA). This web application allows to define regions of interest and select the "best-performing" probes from the High-Density (HD) Agilent probe library. Chromosomal coordinates of all RefSeq genes were extrapolated from open-source databases, Biomart (http://www.biomart.org/) and UCSC Genome Browser according to Human Feb. 2009 Assembly (GRCh37/hg19) (http://genome.ucsc.edu). Exon coordinates were selected and formatted using a homemade R script [1] and then uploaded on SureDesign. Candidate probes were scored and filtered using bioinformatics prediction criteria for probe sensitivity, specificity, and responsiveness under appropriate conditions. We also selected a limited number of probes with the SureDesign Genomic Tiling option to cover regions inadequately represented in the Agilent database. All probes had similar characteristics: isothermal probes, with melting temperature (Tm) of 80° C and probe length of about 60-mers, in accordance to the manufacturer's specifications. Biological probes were randomly distributed in an 8x60K array format that allows to process simultaneously eight samples in a single experiment. The routinely used Agilent Human CGH Normalization Probe Group (1262 features) and the Agilent Human CGH Replicate Probe Groups (5000 features) were also included in the array design. Microarray slides were produced using Agilent's Sure-Print Inkjet technology (Agilent Technologies, Santa Clara, CA).

Microarray experiment. Array experiments were performed as recommended by the manufacturer (Agilent Technologies, Santa Clara, CA). DNA test and a reference of the same sex (Euro Reference, Agilent Technologies, Santa Clara, CA), both at the concentration of 500 ng, were double digested with RsaI and AluI for 2 hours at 37°C. After heat inactivation of the enzymes at 65°C for 20 min, each digested sample was labeled by random priming by using the genomic DNA Enzymatic Labelling Kit (Agilent Technologies, Santa Clara, CA) for 2 hours using Cy5-dUTP for patient DNAs and Cy3-dUTP for reference DNAs. Labeled products were column purified by using the SureTag DNA Labeling Kit Purification Columns (Agilent Technologies, Santa Clara, CA). After probe denaturation and pre-annealing with Cot-1 DNA, hybridization was performed at 65°C with rotation for 24 hr. After two washing steps, the array was scanned by SureScan scanner (Agilent Technologies, Santa Clara, CA) at 3 microns. Array data were extracted from scanned images using Feature Extraction software (Agilent Technologies, Santa Clara, CA) and underwent a quality control step in order to

check signal intensities and background noise. In particular, the following evaluation metrics were used to pass our quality control test: Derivative Log Ratio Spread ≤ 0.30, signal intensity ≥200, background noise of both channels ≤ 25, and signal to noise ≥30.

Data analysis. A first data analysis was carried out by using the Aberration Detection Method 2 (ADM-2) algorithm with default normalization parameters. This statistical algorithm permits to identify all aberrant intervals in a given sample with consistently high or low log ratios based on a statistical score. The statistical score is calculated on the average log ratios of the probes and the number of probes per region. To make a positive call, our threshold settings for the default analysis method were 6.0 for sensitivity, 0.25 for minimum absolute average log ratio and 3 as the minimum number of probes in the region. The ADM-2 algorithm is the most stringent since incorporates quality information about each probe measurement.
A second data analysis was performed with the Aberration Detection Method 1 (ADM-1) algorithm with the threshold settings described above. ADM-1 algorithm uses the same iterative procedure as ADM-2 to find all genomic intervals with the score above a user specified threshold, but does not take into account the Quality-Weighted Interval Score, and therefore results less stringent.

Supplementary Table 1. Main characteristics of the global aCGH *NeuroArray* design

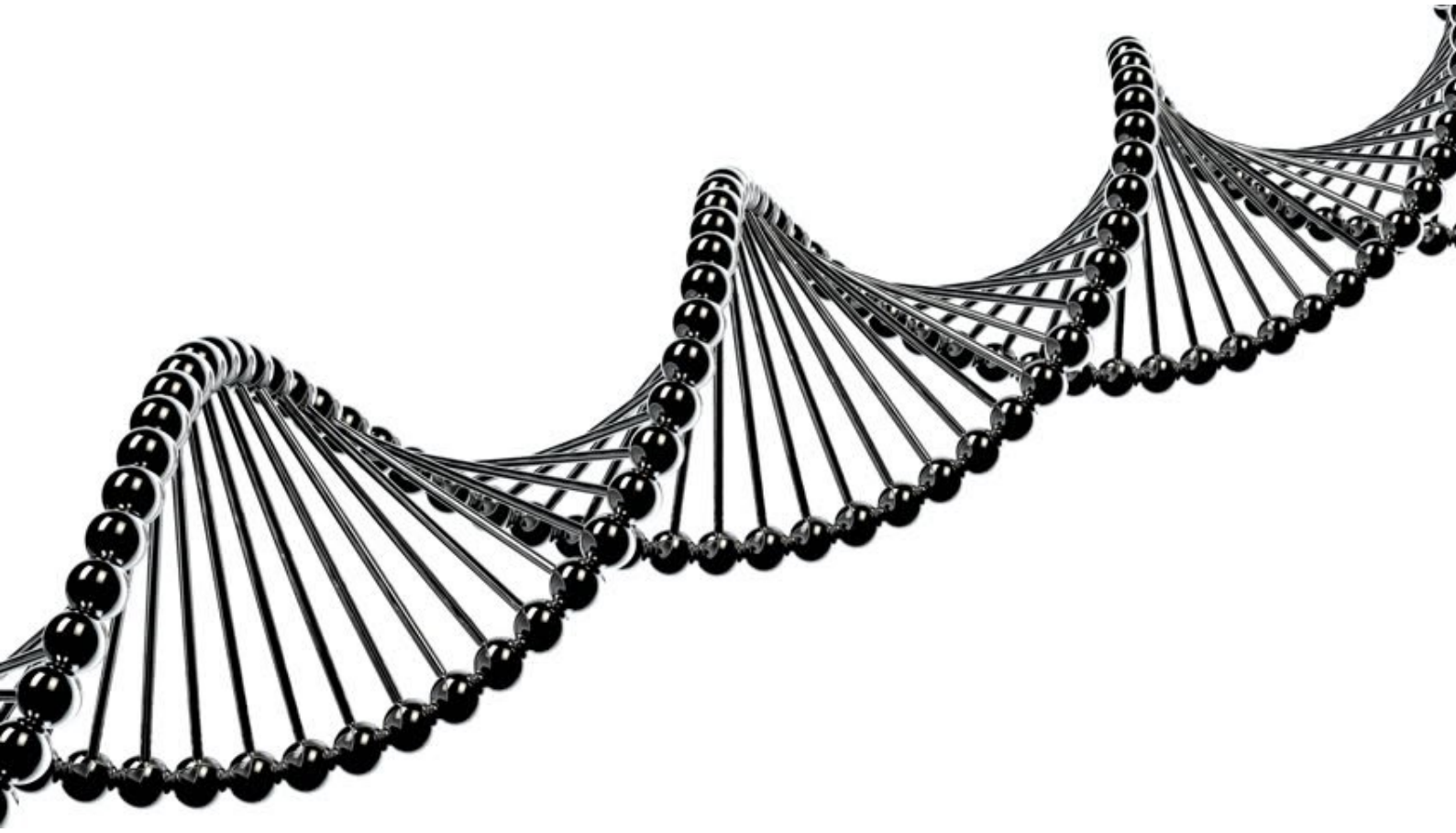
NeuroArray design	
Total genes	1632
Total exonic targets	24929
Target coverage	94 %
Total target/exon size	6,5 Mb
Total probes (1-2 probes per exon)	40973
Total unique probes from HD Database	38636
Total unique probes by genomic tiling	2337
Median probe spacing	355 bp
Mean target size	318,3 bp
Uncovered targets	1520

The table lists the total number of selected genes and exon targets, the mean exon size, the number of probes, the median probe spacing and the total coverage of the customized design. The array design was performed through the Agilent SureDesign software (https:/earray.chem.agilent.com/suredesign). The majority of probes have been scored and filtered from the High-Density (HD) Agilent probe library. A limited number of probes has been designed with the Genomic Tiling option to cover regions inadequately represented in the Agilent database. All probes have been chosen with similar characteristics: isothermal probes, with melting temperature (Tm) of 80° C and probe length of 60-mers.

Supplementary Table 2 is available in Excel file format or in the online version.

CHAPTER 4

Splicing: is there an alternative contribution to Parkinson's disease?



La Cognata, V., et al. Neurogenetics (2015) 16: 245. doi:10.1007/s10048-015-0449-x

Splicing: is there an alternative contribution to Parkinson's disease?

Valentina La Cognata[1,2], Velia D'Agata[2], Francesca Cavalcanti[1] and Sebastiano Cavallaro[1]*

[1]*Institute of Neurological Sciences, Italian National Research Council, Via Paolo Gaifami 18, 95125, Catania and 87050 Piano Lago di Mangone, Italy;* [2]*Department of Biomedical and Biotechnological Sciences, Section of Human Anatomy and Histology, University of Catania, Italy*

* *Corresponding author*: Institute of Neurological Sciences, Italian National Research Council, Via Paolo Gaifami 18, 95125, Catania, Italy; Tel: +39 095 7436634; Fax: + 39 095 7122426; E-mail address: sebastiano.cavallaro@cnr.it

ABSTRACT

Alternative splicing is a crucial mechanism of gene expression regulation that enormously increases the coding potential of our genome and represents an intermediate step between mRNA transcription and protein post-translational modifications. Alternative splicing occupies a central position in the development and functions of the nervous system. Therefore, its deregulation frequently leads to several neurological human disorders. In the present review, we provide an updated overview on the impact of alternative splicing in Parkinson's disease (PD), the second most common neurodegenerative disorder worldwide. We will describe the alternative splicing of major PD-linked genes by collecting the current evidences about this intricate not carefully explored aspect. Assessing the role of this mechanism on PD pathobiology may represent a central step toward an improved understanding of this complex disease.

KEYWORDS: Parkinson's disease; Alternative splicing; PD genes; mRNA splice transcripts; protein isoforms.

INTRODUCTION

The flow of genetic information from DNA to RNA to protein has traditionally been considered the central dogma of molecular biology. Additional steps of regulation are currently well-known, greatly expanding this simplistic framework and revealing the complex network that controls gene expression [1]. One of these steps is represented by alternative splicing (AS), whereby a single gene gives rise to multiple mRNA transcripts and protein isoforms with different functional properties [1]. It is estimated that 94% of human protein-coding genes are alternatively spliced [2,3], and the main site of alternative splicing events is the central nervous system [4,5].

The alternative splicing process consists in the removal of the intronic regions from the RNA primary transcript and simultaneous assembly of the exonic regions in different combinations to form a mature mRNA, which is then polyadenilated, exported to cytoplasm and translated into protein. The accuracy and efficiency of pre-mRNA splicing process depend on a range of constitutive DNA sequence motifs: the donor and the acceptor splice sites, the lariat branch point, the polypyrimidine tract and splicing enhancers and silencers (Fig. 1, Panel A). These motifs are recognized by a large macromolecular splicing machinery (called the spliceosome), which models the pre-mRNA while RNA polymerase II synthesizes it in the nucleus. The splicing machinery includes five spliceosomal uridine-rich small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4, U5 and U6) and several non-snRNP protein splicing factors, such as the Serine/Arginine (SR) rich protein family and hnRNP proteins [6,7]. The splicing reaction relies on two transesterification steps that occur within the highly dynamic splicing machine. The stepwise molecular mechanisms of the splicing reaction are detailed in Figure 1 (Panel A).

Alternative splicing works as an on–off switch in gene expression. It affects the expression levels, stability, half-life (*via* the Nonsense-Mediated mRNA Decay – NMD) and localization of the RNA messengers. It has also the potential to generate several protein isoforms with different biological properties, protein-protein interactions, subcellular localization, signaling pathway or catalytic ability. During the last years, great efforts have been made to decipher the intricate alternative splicing code. Five major alternative splicing events (i.e. cassette exons, use of alternative acceptor and/or donor sites, intron retention, and mutually exclusive exons) have been

described up to now and are detailed in Figure 1 (Panel B) [2,8]. However, how the spliceosome recognizes alternative exons and decides which exons to include remains not fully understood. Undoubtedly, there is more diversity in splice transcript variants than in protein isoforms. Although this is still not clear, different variants encode the same protein, but probably translate it with different efficiency [9].

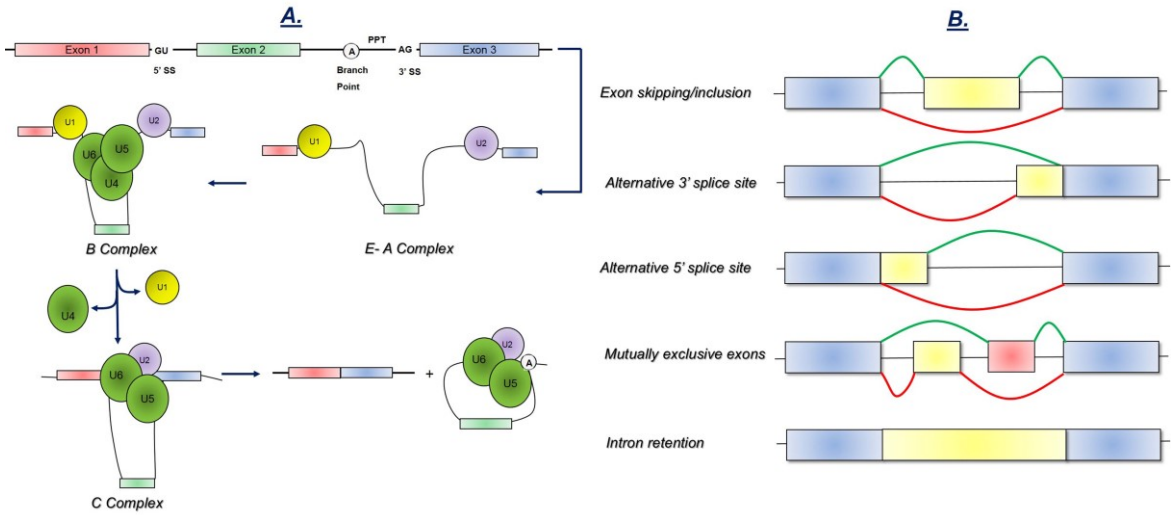


Figure 1. The alternative splicing mechanism

PANEL A. Four main conserved DNA sequence motifs allow the splicing mechanism: the donor splice site GU (5' SS), the acceptor splice site AG (3' SS), the lariat branch point (A) located upstream of the acceptor site and the polypyrimidine tract (PPT) placed between the acceptor site and the branch point. The splicing machinery includes mainly five spliceosomal uridine-rich small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4, U5, and U6), and further auxiliary RNA binding proteins. During the first step of spliceosome assembly, U1 snRNP base-pairs with the 5'-splice site of the pre-mRNA (E complex), whereas U2 base-pairs with the branch-point (A complex). Then, the tri-snRNP complex U4, U5 and U6 associates with the forming spliceosome (B complex), and both U1 and U4 are ejected. This allows U6 to replace U1 at the 5' splice site (C complex) and leads to a U6–U2 interaction that gets close together the 5'-splice site and the branch point, allowing for a transesterification step. At the end, U5 brings near the two exons, joining them through a second transesterification reaction.

PANEL B. Five major alternative splicing events are currently known: exon skipping/inclusion, use of alternative 3' splice site, use of alternative 5' splice site, mutually exclusive exons, and intron retention. In blue, are represented the constitutive exons. In yellow and red are represented the alternatively spliced exons. The splicing events rely on the interplay between the constitutive splicing motifs, the splicing regulatory sequences, the RNA secondary structures, the components of the spliceosome and further auxiliary RNA-binding proteins. However, how the spliceosome decides which exons to include remains currently not clear.

The finely tuned splicing regulatory network can easily undergo alterations. An aberrant alternative splicing may arise from changes in regulatory sequences required for correct pre-mRNA processing (so-called *cis*-acting mutations), as well as from

mutations that affect components necessary for splicing regulation (*trans*-acting mutations). *Cis*- and *trans*- splicing aberrations represent direct causative agents of disease or more subtle contributions to the determinants of disease susceptibility or modulators of disease severity. An extensive range of neurological diseases has been already associated to both splicing defects, including Alzheimer disease, Retinitis Pigmentosa, Spinal Muscular Atrophy, Muscular Dystrophy, Neurofibromatosis and Fragile-X-associated Tremor/Ataxia Syndrome [10-12,1,13]. In this broad neurological disorders scenario, the relevance of alternative splicing in Parkinson's disease is not still clear, and the splicing mechanisms that regulate PD-related genes remain mostly unknown.

Here, we provide an updated overview of the current knowledge about the impact of alternative splicing on Parkinson's disease. Firstly, we will take into account the most common PD-related genes "one by one", by analyzing their alternative transcripts currently known and their involvement in this disease. Then, we will describe the few studies that have globally analyzed the changes of splice variants expression in PD patients through genome-wide RNA expression approaches. Finally, we will briefly describe the current evidences about the alternative splicing modulation in PD through non-coding RNAs (miRNA and lcnRNA).

GENETICS OF PARKINSON'S DISEASE

PD is the second most common neurodegenerative disorder worldwide, characterized by resting tremor, bradykinesia, stiffness of movement and postural instability. These symptoms derive from the progressive loss of neurons from *substantia nigra pars compacta*, coupled with an accumulation of intraneuronal aggregates called Lewy bodies.

Despite significant progresses in the understanding of PD pathogenesis, the exact etiology of PD remains unknown. Over the 15 past years, an even more detailed knowledge of the genetic factors that contribute to PD has emerged through different research strategies [14,15]. Linkage mapping analysis, genome-wide association studies (GWAS) and next generation sequencing technologies are revealing an increasing number of locus and genes strongly linked to either autosomal dominant (*SNCA-PARK1, LRRK2-PARK8, VPS35-PARK17, GBA),* or typical recessive (*PARKIN*-

PARK2, *PINK1-PARK6*, *DJ1-PARK7*) and atypical recessive (*ATP13A2-PARK9, PLA2G6-PARK14*, *FBXO7-PARK15*) or X-linked (*ATP6A2*, *TAF1*) forms of disease. For the sake of completeness, we mention here further monogenic loci, not confirmed genes or risk factors genes (i.e. *PARK3, UCHL1*, *PARK10*, *GIGYF2*, *PARK12*, *HTRA2*, *PARK16*, *EIF4G1, DNAJ, HLA-DR, GAK-DGKQ, SYNJ1, GBAP1*) [15-17]. Furthermore, large-scale meta-analysis of genome-wide association data are revealing a wide range of additional loci having genome-wide significant association [18]. However, we will overlook their discussion because of the few data in the literature regarding their splicing regulation in pathological conditions.

In the next paragraphs, we will describe the alternative spliced mRNA variants of PD genes, and the current scientific data demonstrating their involvement in PD pathogenesis. For a more complete picture, we have also added some further implicated genes (*SRRM2, MAO-B, SNCAIP, MAPT, GBA),* indicated as other PD-related genes, which are not directly causative genes, but whose splicing regulation seems to be altered in PD states.

AUTOSOMAL DOMINANT PD GENES

SNCA

Alpha-synuclein, encoded by *SNCA* gene, is a small, natively unfolded presynaptic protein linked to PD [19]. Aggregates of alpha-synuclein protein represent the neuropathological hallmark lesions of PD and constitute the major components of Lewy bodies. Genetically, mutations in *SNCA* gene were the first to be associated with PD family inheritance. Missense mutations in coding regions (Ala53Thr, Ala30Pro, Glu46Lys), single nucleotide substitution in 3' UTR and dose-dependent genomic multiplications (duplications or triplications) of the gene cause both monogenic and sporadic forms of PD [20,19,21]. Some point mutations in splice donor sites have also been reported (IVS2+9A>C) [22].

SNCA gene maps to chromosome 4q22.1 and contains six exons spanning about 114 kb [21]. The set of mRNAs produced by *SNCA* gene includes the *full-length* transcripts, commonly known as SNCA140 from the amino acidic length of the encoded protein, and corresponding to SNCA-001, SNCA-002, SNCA-003, SNCA-006

and SNCA-008 mRNAs from Ensembl library (Table 1 and Fig. 2). Further additional splicing variants, known as SNCA126, SNCA112 and SNCA98 and correponding to i) SNCA-004, SNCA-203, SNCA-201, ii) SNCA-005, SNCA-202 and iii) SNCA-010 respectively are generated by in-frame excision of exons 3, 5, or both (Table 1 and Fig. 2). Two additional splice variants (SNCA-009 and SNCA-007) are generated from an inner transcription start and encode proteins of 115 and 97 amino acids respectively (Table 1 and Fig. 2). SNCA140, 126 and 112 are expressed in a broad spectrum of human tissues, while SNCA98 seems to be a brain-specific splice variant with varying expression levels in different areas of fetal and adult brain [23].

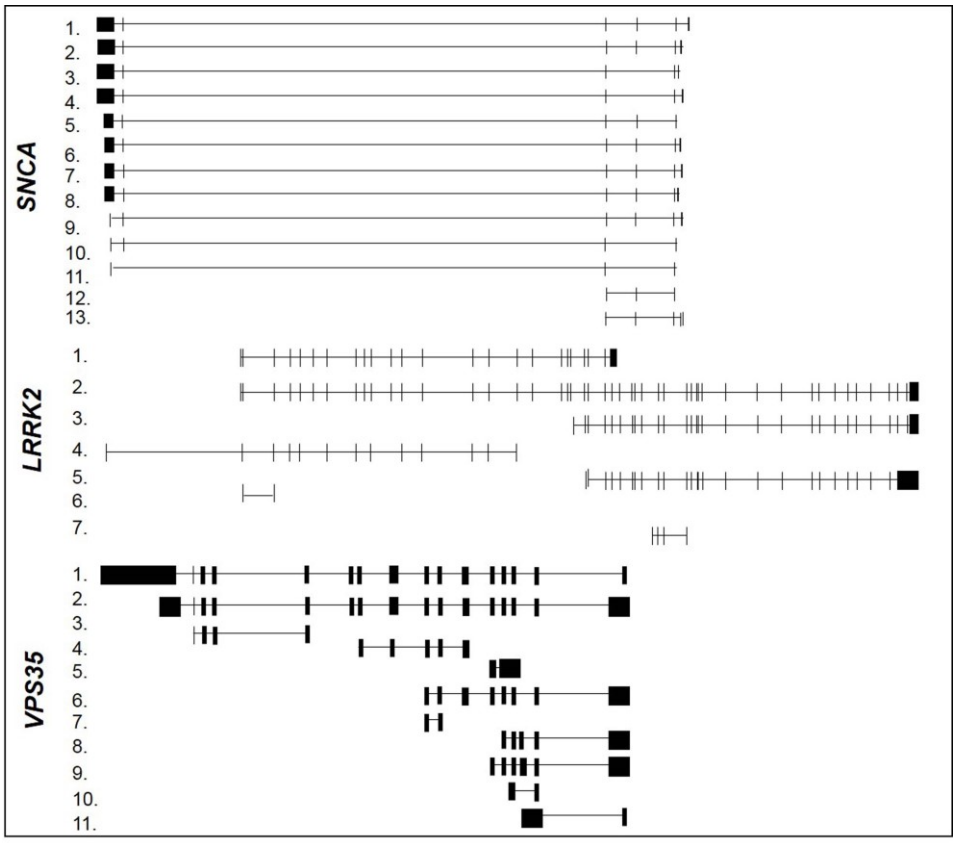


Figure 2. Structures of the alternative splicing variants of human dominant PD genes.

Structures of the described mRNA splicing variants are represented in figure as reported in Ensembl library (http://www.ensembl.org/index.html). On the left, each variant is indicated with a number corresponding to that indicated in Table 1. *LRRK2* gene is illustrated in 5'-3' sense, while *SNCA* and *VPS35* genes are illustrated in antisense corresponding to their 3'-5' sense transcription.

The expression profile of SNCA140, SNCA126, SNCA112 and SNCA98 splice variants is different in the various brain areas under normal and pathological states. Compared to healthy controls, in PD *frontal cortex* all these four transcripts are over-

expressed, with significant upregulation of SNCA126 [24]. In PD *substantia nigra*, only the three shorter transcripts have been observed significantly over-expressed [25,26], while higher SNCA112 and SNCA98 levels are also present in the *cerebellum* [25]. Different expression profiles of *SNCA* variants also occur in other forms of neurodegenerative disorders. Both SNCA140 and SNCA126 downregulation and SNCA98 overexpression have been reported in dementia with Lewy bodies and Alzheimer disease, while SNCA112 is upregulated in dementia with Lewy bodies and downregulated in Alzheimer disease [27,28,24].

Some interesting data emerge on SNCA112 variant. An association between PD risk-associated SNPs (Single Nucleotide Polymorphisms) within the 3' region of *SNCA* gene and higher SNCA112 ratio level has been observed in about one hundred of *frontal cortex* samples. These data reveal the *cis*-regulatory effect of these mutations on splicing mechanism [29]. The expression of SNCA112 is also abundantly induced by some parkinsonism mimetics (MPP+, rotenone) and related oxidants [30]. However, the reason for these effects remains unclear.

In addition to splice variants, specific RNA transcript isoforms of *SNCA* with an extended 3' untranslated region have been described and appear selectively linked to pathological processes [31]. However, this review is focusing only on the mRNA splice variants, thus their discussion will be omitted.

The 140 amino acids isoform is a small protein with a molecular weight of 14.5 kDa. It is composed of three distinct regions: (1) an amino terminus containing amphipathic helices conferring the propensity to bind membranes, (2) a central hydrophobic region, so-called NAC (non-Ab component), which confers the b-sheet potential, and (3) an acid glutamatergic carboxyl terminus that is highly negatively charged and prone to be unstructured [19]. Structural changes in the shorter splicing isoforms can be predicted as result of exon skipping events. SNCA126 predicted isoform shows interruption of the N-terminal protein-membrane interaction domain [32], SNCA112 is significant shorter in the unstructured C-terminal [32], while SNCA98 isoform results in a truncated protein consisting almost only of the central region containing NAC [23]. Recently, it has been demonstrated *in vitro* a lower aggregation propensity of the shorter isoforms [33]. In addition, morphology studies by using electron microscopy have shown straight fibrils for SNCA140, shorter fibrils

mostly arranged in parallel arrays for SNCA126 and circular structures for SNCA98 [33]. These data open new insight regard the formation of Lewy bodies induced by alpha-synuclein.

Numerous functions of alpha-synuclein have been proposed, counting molecular chaperone, regulator of dopamine uptake and homeostasis, inhibitor of phospholipase D2, down regulator of p53 pathway [32] and promoter of the SNARE-complex assembling [34]. Unfortunately, nothing is known about the specific pathophysiological roles of each alpha-synuclein isoform and their relative post-translational modifications (i.e. phosphorylations, nitration, sumoylation, oxidation, glycosylation, cleavage and ubiquitination), which are known to play a key role in SNCA functions and regulation [32].

LRRK2

LRRK2 encodes for leucine-rich repeat kinase 2 (or dardarin), which is a large 2,527 amino-acid multi-domain protein. The protein consists of multiple conserved well-defined domains including a small GTPase-like domain (Ras of complex proteins or ROC), a domain of unknown function termed the C-terminal of ROC (COR), a kinase domain, as well as several protein interaction domains (e.g. the leucine-rich repeat - LRR, the WD40 domain, the ankyrin repeat domain and the armadillo repeat region). The precise physiological function of LRRK2 is unknown. However, LRRK2 seems implicated in different cellular functions as neurite outgrowth, cytoskeletal maintenance, vescicle trafficking, and autophagic protein degradation [35].

The *LRRK2* gene spans a genomic region of 144 kb, with 51 exons, and harbors the most common mutations linked to both autosomal dominant inherited late-onset and sporadic PD. The missense mutations known so far are spread over the whole *LRRK2* gene and affect all functional domains. Some mutations have much higher frequencies than others, such as Gly2019Ser and mutations altering codon Arg1441, respectively in the kinase and ROC domains. In addition, several unclear pathogenic mutations affecting splice sites have been observed (IVS19+5_8delGTAA, IVS25-8delT, IVS27-9C>T, IVS30-6C>T, IVS31+3A>G, IVS32+14G>A, IVS33+6T>A, IVS37-9A>G, IVS38+7C>T, IVS46-14T>A, IVS46-8delT) [36,22,37-43].

In addition to the *full-length* transcript (LRRK2-004), further *LRRK2* shorter transcripts are deposited in Ensembl library (Table 1 and Fig. 2). Despite the existence of these transcripts, there are currently no data analyzing the splicing profile of this gene in PD states. Recently, a gene expression and splicing analysis of the *LRRK2* locus has been carried on [44]. Both exon array and RT-PCR methods confirm the existence of an isoform with spliced out exons 32–33 in *substantia nigra* and an isoform with exon 32 alone spliced out in *occipital cortex*, *medulla* and *cerebellum* of healthy humans [44].

Further evidences on *LRRK2* splicing have been observed by Giesert and collaborators [45], who have conducted a study in various brain regions and organs from adult mice. To this regard, it should be considered that *LRRK2* is highly conserved in human and mouse and that several transgenic animal models have been created. Giesert et al. [45] have identified two *LRRK2* splice-variants: one with skipped exon 5, primarily expressed in astrocytes, and another truncated variant terminating with an alternative exon 42a, barely detectable in microglia but highly expressed in neurons and astrocytes. Protein-structure predictions reveal that the loss of exon 5 may generate a smaller protein with changed affinity of binding partners, while the alternative exon 42a may lead to changes of its enzymatic activity. In addition, the protein-interaction domain WD40 would also be absent in such truncation. Interestingly, the deletion of this domain in the Zebrafish LRRK2 ortholog (zLRRK2) causes Parkinsonism-like phenotype, including loss of dopaminergic neurons in diencephalon and locomotion defects [46]. Further studies will need to assess the involvement of *LRRK2* alternative splice variants in PD.

VPS35

In 2011 two groups reported the identification of the same missense mutation (p.Asp620Asn) in the vacuolar protein sorting 35 (*VPS35*) gene as a novel cause of autosomal dominant PD [47,48]. *VPS35* was the first PD gene found by a direct whole exome sequencing in large families of Austrian and Swiss origins. An in-depth sequence analyses of all coding, non-coding and exon-intron boundaries *VPS35* genetic regions has been performed in a large well-characterized cohort of Lewy body disorders, including PD patients, PD with dementia and dementia with Lewy bodies

[49]. In addition to three novel missense mutations, a silent and an intronic variations, predicted to activate cryptic splice sites, have been observed in patient's group, but not in controls. However, pathogenicity of these mutations was not completely conclusive since they were not supported by segregation analysis in family relatives [49].

Various spliced transcript variants of this gene are reported in Ensembl library (Table 1 and Fig. 2), but the majority of them are processed for degradation and do not encode proteins.

Gene Name	Transcript Number	Ensembl Name	GenBank Accession Number	Protein Length
SNCA	1.	SNCA-003	NM_001146055	140 aa
	2.	SNCA-202	NM_007308	112 aa
	3.	SNCA-203	-	126 aa
	4.	SNCA-201	-	126 aa
	5.	SNCA-005	-	112 aa
	6.	SNCA-001	NM_001146054	140 aa
	7.	SNCA-002	NM_000345	140 aa
	8.	SNCA-008	-	140 aa
	9.	SNCA-006	-	140 aa
	10.	SNCA-004	-	126 aa
	11.	SNCA-010	-	98 aa
	12.	SNCA-009	-	115 aa
	13.	SNCA-007	-	67 aa
LRRK2	1.	LRRK2-002	-	1271 aa
	2.	LRRK2-004	NM_198578	2527 aa
	3.	LRRK2-005	-	207 aa
	4.	LRRK2-001	-	521 aa
	5.	LRRK2-003	-	No protein
	6.	LRRK2-006	-	No protein
	7.	LRRK2-007	-	No protein
VPS35	1.	VPS35-001	NM_018206	796 aa
	2.	VPS35-002	-	48 aa
	3.	VPS35-011	-	No protein
	4.	VPS35-012	-	No protein
	5.	VPS35-006	-	No protein
	6.	VPS35-003	-	No protein
	7.	VPS35-010	-	No protein
	8.	VPS35-005	-	47 aa
	9.	VPS35-008	-	41 aa
	10.	VPS35-007	-	No protein

11.	VPS35-004	-	No protein

Gene name, Ensembl transcript names, GenBank accession numbers and relative encoded amino acidic protein length of splice variants are reported in the table. Number in the column "Transcript number" identifies the transcript in Figure 2.

AUTOSOMAL RECESSIVE PD GENES

EARLY- ONSET TYPICAL PD GENES

PARK2

Mutations in *PARK2* gene (also known as *PARK2 parkin RBR E3 ubiquitin-protein ligase*) are the most common cause (50% of cases) of autosomal recessive juvenile Parkinsonism (AR-JP), a form of early-onset Parkinsonism characterized by good and prolonged response to levodopa and a benign, slow course. *PARK2* mutations also explain ~15% of the sporadic cases with onset before 45 [50,51] and act as susceptibility alleles for late-onset forms of Parkinson disease (2% of cases) [52]. Along with the about 200 mutations currently identified in *PARK2* coding region, several point mutations in splice acceptor or donor sites (introns 1, 6, 7, 10, 12, 13 and 16) have been identified in PD patients [53-57,22,58,59].

PARK2 gene spans more than 1.38 Mb of genomic DNA in the long arm of chromosome 6 (6q25.2-q27) and contains 12 exons, which are alternatively spliced to produce at least 11 different splicing variants (Table 2 and Fig. 3) [59]. The *full length PARK2* transcript (PARK2-004) encodes a protein of 465 amino acids (parkin) [60,61,59] acting in numerous molecular pathways (protein turnover, stress response, mitochondrial homeostasis, mitophagy, mitochondrial DNA stability, metabolism, cell growth and survival) [62]. Multiple parkin isoforms likely arising from *PARK2* splicing variants have been observed in different brain areas through western blot studies [9].

The extensive alternative splicing of *PARK2* is differently regulated both at transcript and protein level in tissues and cells [63,64,24,65-68]. Distinct expression patterns of *PARK2* splice variants emerge in human brain regions [69] and leukocytes [68], in rat brain, neuronal and glial cells [67], and in a wide variety of mouse tissues (brain, heart, lung, liver, skeletal muscle, kidney, and testis) [64]. At the protein level, *PARK2* protein isoforms show a differential distribution in human leukocytes [70] and

aged brain [71], as well as in different rat and mouse nervous system areas (cerebral cortex/diencephalons, hippocampus, cerebellum, brainstem, striatum, spinal cord, *substantia nigra*), peripheral tissues (heart, liver, spleen, pancreas, kidney), and developmental stages [72-76].

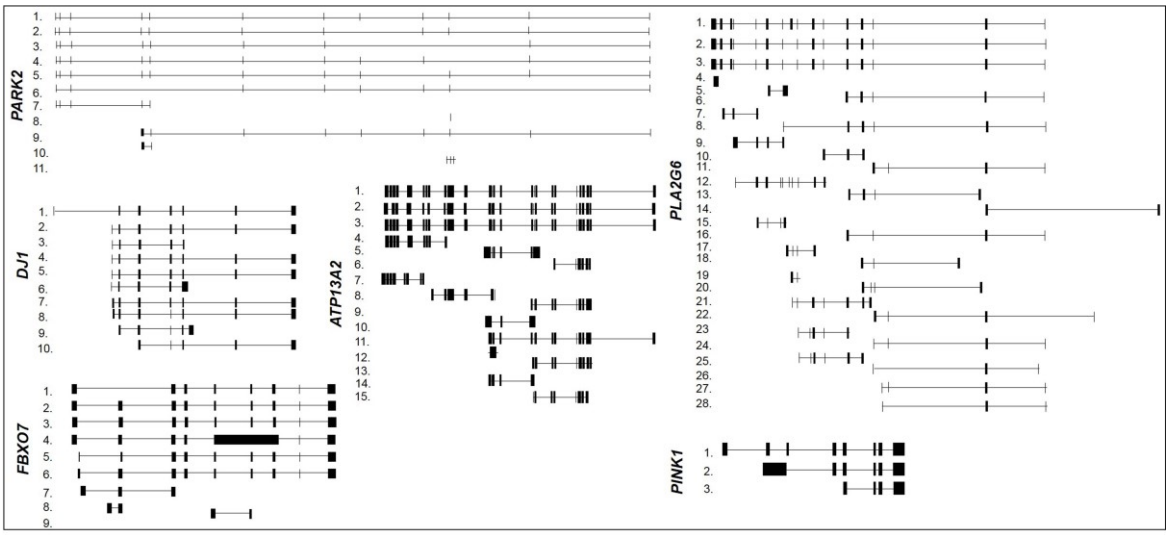


Figure 3. Structures of the alternative splicing variants of human recessive PD genes

Structures of the described mRNA splicing variants are represented in figure as reported in Ensembl library (http://www.ensembl.org/index.html). On the left, each variant is indicated with a number corresponding to that indicated in Table 2. All transcripts are illustrated in 5'-3' sense, except *PARK2*, *ATP13A2* and *PLA2G*6 genes, which are illustrated in antisense corresponding to their 3'-5' sense transcription.

Emerging evidences support the importance of *PARK2* splice variants expression changes in disease development. Differential expression of *PARK2* transcripts have been identified in the *frontal cortex* of Parkinson's disease, pure dementia with Lewy bodies, common Lewy body disease, and Alzheimer's disease patients, compared to controls [65,24]. Particularly, two *PARK2* splicing variants are significantly overexpressed in PD [65]. Another study reports both an increase in the expression level of a parkin splice variant and a decrease of the wild type between PD patients and healthy controls [66]. The differential and disease-specific expression profiles of *PARK2* alternative splice variants suggest a role for splicing deregulation in the development of neurodegenerative disorders.

PINK1

Homozygous or compound heterozygous loss-of-function mutations in *PINK1* (*PTEN-Induced putative kinase 1*) are the second most frequent cause of autosomal

recessive early-onset Parkinsonism. Mutation frequency varies geographically from 1 to 9% depending on ethnic background [77]. The *PINK1* mutations spectrum involves nonsense and missense mutations, insertions or deletions, and whole gene or single/multiple exon copy number variants located across the entire gene [78].

PINK1 gene maps in the short arm of chromosome 1 (1p36.12), encompassing ~18 kb of genomic DNA. Its coding sequence is spread over eight exons. In addition to the *full-length* (PINK1-001), two shorter variants exist but do not produce proteins. (Table 2 and Fig. 3).

Some interesting findings emerge regarding the splicing regulation of exon 7. A 23bp deletion disrupting the splice acceptor site of exon 7 has been detected in a sporadic parkinsonian patient, producing several aberrant mRNAs [79]. Moreover, whole exon 7 deletion and a novel U1-dependent 5' splice-site mutation in exon 7 have been found in a large Spanish family with PD members [80].

The PINK1 protein is a putative serine/threonine kinase of 581 amino acids involved in mitochondrial response to cellular and oxidative stress [81]. It has been demonstrated that in human brain are expressed at least two isoforms: a full-length protein of ~63 kDa and an N-terminally truncated isoform of 52 kDa [77,82-84]. An additional isoform of approximately 45 kDa has been suggested, although it has not extensively studied [85]. The 52 kDa isoform seems to originate by enzymatic cleavage of PARL [86]; however, the exact nature of the isoforms, the precise reason for the cleavage and the functional roles of these three different isoforms require further studies.

DJ1

Mutations in the *DJ1* (also known as *PARK7*) gene are the less common cause of autosomal recessive Parkinsonism (~1% of early-onset PD) [87,88]. A large homozygous deletion and a missense mutation (L166P) in *DJ-1* gene were first identified in both Italian and Dutch consanguineous families [89,90]. Additional mutations have been collected in other PD families and include missense mutations in coding and promoter regions, frame-shifts, copy number variations [91,88] and splice sites alterations [92,93].

DJ-1 gene maps to chromosome 1 (1p36.23) and includes seven exons. Several spliced transcript variants have been identified encoding the same protein (Table 2 and Fig. 3). Two shorter transcripts (the first lacking exon 4, and the second starting in an inner transcription point) encode for smaller proteins (Table 2 and Fig. 3).

The product of *DJ-1* gene is a highly conserved protein of 189 amino acids belonging to the peptidase C56 family [94]. It is a multifunctional protein, acting as positive regulator of transcription, redox-sensitive chaperone, sensor for oxidative stress, and apparently protects neurons from ROS-induced apoptosis [95,96]. In human brain and peripheral blood, several *DJ-1* isoforms exist and differ on their isoelectric point (pI) [97-100]. The relative abundance of these different *DJ-1* isoforms appears to be altered in PD, and, therefore, blood *DJ-1* isoforms have been proposed as potential biomarkers for Parkinson's disease [101]. The different pI of each variant is believed to result from post-translational modifications that alter the intrinsic charge of the protein [101]. Interestingly, it has been demonstrated that one of the major binding partners of DJ-1 in dopaminergic neuronal cells is the splicing factor proline/glutamine-rich (SFPQ protein) [96,102]. SFPQ, originally identified as a polypyrimidine tract binding protein, is part of the spliceosome C complex and is required for *in vitro* splicing of pre-mRNA [96,102]. DJ-1 binding to SFPQ modulates its transcriptional activity and, therefore, tunes its effect on splicing regulation. DJ-1 mutations could reverberate on its downstream targets, including the splicing factor SFPQ and altering the splicing control.

Gene Name	Transcript Number	Ensembl Name	GenBank Accession Number	Protein Length
PARK2	1.	PARK2-004	NM_004562	465 aa
	2.	PARK2-005	NM_013987	437 aa
	3.	PARK2-006	NM_013988	316 aa
	4.	PARK2-001	-	274 aa
	5.	PARK2-003	-	274 aa
	6.	PARK2-007	-	218 aa
	7.	PARK2-201	-	176 aa
	8.	PARK2-204	-	87 aa
	9.	PARK2-002	-	368 aa
	10.	PARK2-202	-	74 aa
	11.	PARK2-203	-	201 aa
PINK1	1.	PINK1-001	NM_032409	581 aa
	2.	PINK1-002	-	No protein
	3.	PINK1-003	-	No protein

DJ1	1.	PARK7-004	-	189 aa
	2.	PARK7-002	NM_001123377; NM_007262	189 aa
	3.	PARK7-007	-	No protein
	4.	PARK7-001	-	189 aa
	5.	PARK7-003	-	169 aa
	6.	PARK7-008	-	No protein
	7.	PARK7-005	-	189 aa
	8.	PARK7-006	-	189 aa
	9.	PARK7-009	-	No protein
	10.	PARK7-010	-	160 aa

Gene name, Ensembl transcript names, GenBank accession numbers and relative encoded amino acidic protein length of splice variants are reported in the table. Number in the column "Transcript number" identifies the transcript in Figure 3.

JUVENILE ATYPICAL PD GENES

ATP13A2

ATP13A2 mutations are associated with Kufor-Rakeb syndrome, a form of recessively levodopa-responsive inherited atypical Parkinsonism [103]. It 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 [104].

ATP13A2 is composed of 29 exons and lies on chromosome 1 covering about 26 kb of genomic DNA. One of the first identified disease-causing mutations was a guanine-to-adenine transition in the donor splice site of exon 13, leading to the skipping of exon 13 and resulting in an deletion of part of the third transmembrane domain [105].

According to data repositories, at least 15 alternatively spliced transcripts are expressed in humans (Table 2 and Fig. 3). The longest transcripts are ATP13A2-001, ATP13A2-002 and ATP13A2-005. Transcript variants ATP13A2-001 and ATP13A2-005 differ only in a nucleotide segment on exon 5, while transcript variant ATP13A2-002 lacks exons 22 and 28. The ATP13A2 mRNA is highly expressed in the brain, particularly in the *substantia nigra* of patients with classical late-onset PD [91]. However, nothing is known about the splicing expression profiles of this gene in PD and healthy subjects.

The products of these transcripts have been studied at the protein level. The isoform 1 encoded by ATP13A2-001 is a protein of 1180 amino acids with 10 transmembrane domains. Isoform 2 encoded by ATP13A2-005 contains a five amino acid deletion near the N-terminus, while isoform 3, encoded by ATP13A2-002, contains two deletions, generating a highly diverged C-terminus [105]. Functional studies have shown that the isoform 1 is located in the lysosome membrane, whereas the isoform 3 protein is retained in the endoplasmic reticulum and rapidly degraded by the proteasome. In addition, both isoform 1 and 3 are eliminated via the endoplasmic-reticulum-associated degradation pathway [105].

PLA2G6

Recessive mutations in the *PLA2G6* (*phospholipase A2 group VI*) gene have been initially described as the cause of infantile neuroaxonal dystrophy and neurodegeneration associated with brain iron accumulation. Recently, this gene has also been associated with a particular parkinsonian phenotype, consisting of levodopa-responsive dystonia, pyramidal signs, and cognitive/psychiatric features, with onset in early adulthood [106]. Among *PLA2G6* identified mutations, the c.1077G>A mutation at the last nucleotide of exon 7 (apparently a synonymous mutation) stands out as a cause of abnormal mRNA splicing. This single nucleotide substitution causes the activation of a cryptic splice site, producing a 4-bp deleted transcript with altered frame-shift in leukocytes [106].

PLA2G6 gene maps on chromosome 22 (q13.1), covering 70 kb of genomic DNA. Several transcript variants encoding multiple isoforms have been described up to now (Table 2 and Fig. 3). The longest *PLA2G6* mRNA PLA2G6-001 includes 17 exonic regions and encodes the 85/88 kDa calcium-independent phospholipase, known as A2 isoform a. The other two long transcripts (PLA2G6-002, PLA2G6-201) differ in the start point, both lack of exon 9 and encode the same protein, called isoform b. The expression profile of this gene in healthy and disease states remains unknown.

FBXO7

Mutations in the *FBXO7* (*F-box only protein 7*) gene cause Parkinsonian-Pyramidal Disease (PPD or PARK15-associated parkinsonism), an autosomal recessive neurodegenerative disease with juvenile onset, severe levodopa-response, and additional pyramidal signs. Some pathogenic mutations have been identified (R378G,

R498X, T22M), including a compound heterozygous mutation (IVS7 + 1G/T), that removes the invariable splice donor of intron 7 and may disrupt *FBXO7* messenger RNA splicing [107-109].

The *FBXO7* gene, mapped on chromosome 22q12.3, contains nine exons spanning about 24.1 kb. It encodes a 522 amino acids protein consisting of several domains [108], which target proteins for ubiquitination [108]. Alternatively spliced transcript variants of this gene have been identified, (Table 2 and Fig. 3) [107]. FBXO7-001 is the longest and more abundant transcript ubiquitously expressed [110], particularly in skin fibroblasts [111]. FBXO7-002 arises from an inner alternative exon 1, differ in the start codon and produce a shorter isoform. Both these encoded protein isoforms have been detected in cells [111].

Gene Name	Transcript Number	Ensembl Name	GenBank Accession Number	Protein Length
ATP13A2	1.	ATP13A2-001	NM_022089	1180 aa
	2.	ATP13A2-002	NM_001141974	1158 aa
	3.	ATP13A2-005	NM_001141973	1175 aa
	4.	ATP13A2-004	-	No protein
	5.	ATP13A2-003	-	No protein
	6.	ATP13A2-010	-	191 aa
	7.	ATP13A2-007	-	398 aa
	8.	ATP13A2-014	-	258 aa
	9.	ATP13A2-009	-	321 aa
	10.	ATP13A2-006	-	No protein
	11.	ATP13A2-201	-	228 aa
	12.	ATP13A2-013	-	No protein
	13.	ATP13A2-011	-	190 aa
	14.	ATP13A2-012	-	191 aa
	15.	ATP13A2-008	-	188 aa
PLA2G6	1.	PLA2G6-001	NM_003560	806 aa
	2.	PLA2G6-201	NM_001004426	752 aa
	3.	PLA2G6-002	NM_001199562	752 aa
	4.	PLA2G6-025	-	No protein
	5.	PLA2G6-021	-	No protein
	6.	PLA2G6-014	-	166 aa
	7.	PLA2G6-024	-	No protein
	8.	PLA2G6-013	-	No protein
	9.	PLA2G6-026	-	168 aa
	10.	PLA2G6-015	-	120 aa

	11.	PLA2G6-010	-	99 aa
	12.	PLA2G6-023	-	226 aa
	13.	PLA2G6-009	-	No protein
	14.	PLA2G6-027	-	51 aa
	15.	PLA2G6-022	-	No protein
	16.	PLA2G6-012	-	151 aa
	17.	PLA2G6-019	-	124 aa
	18.	PLA2G6-011	-	No protein
	19.	PLA2G6-020	-	No protein
	20.	PLA2G6-008	-	No protein
	21.	PLA2G6-016	-	229 aa
	22.	PLA2G6-005	-	99 aa
	23.	PLA2G6-018	-	157 aa
	24.	PLA2G6-003	-	99 aa
	25.	PLA2G6-017	-	197 aa
	26.	PLA2G6-007	-	80 aa
	27.	PLA2G6-004	-	No protein
	28.	PLA2G6-006	-	No protein
FBXO7	1.	FBXO7-003	-	41 aa
	2.	FBXO7-001	NM_012179	522 aa
	3.	FBXO7-004	-	49 aa
	4.	FBXO7-005	-	No protein
	5.	FBXO7-006	-	129 aa
	6.	FBXO7-002	NM_001033024; NM_001257990	408 aa
	7.	FBXO7-007	-	54 aa
	8.	FBXO7-008	-	No protein
	9.	FBXO7-010	-	No protein

X-LINKED PARKINSONISM

X-linked dystonia Parkinsonism (XDP) is an X-linked recessive adult onset movement disorder characterized by both dystonia and Parkinsonism. *TAF1* gene (*TATA-box binding protein-associated factor 1*), located in the disease locus Xq13.1, has been reported as the first related XPD gene, harboring disease-specific single-nucleotide changes and a small deletion within the multiple transcript panel [112]. This gene is part of a complex region of DNA (the TAF1/DYT3 multiple transcript systems), which encompasses the exonic regions of TAF1 gene and further additional downstream exons [112,113]. This system includes multiple different transcription start sites and encodes multiple spliced transcripts and isoforms (Table 3 and Fig. 4) [112].

Recently, the ATP6AP2 gene (table 3 and fig. 4) has been proposed as a novel gene for X-linked parkinsonism with spasticity (xpds) by exome sequencing analysis [114]. A silent mutation (p.s115s) in the ATP6AP2 gene has been identified in one affected individual, resulting in the aberrant splicing of ATP6AP2 mrna and the overexpression of a minor splice isoform [114]. Noteworthy, the ATP6AP2 is an essential accessory component of the vacuolar atpase required for lysosomal degradative functions and autophagy, a pathway frequently affected in PD.

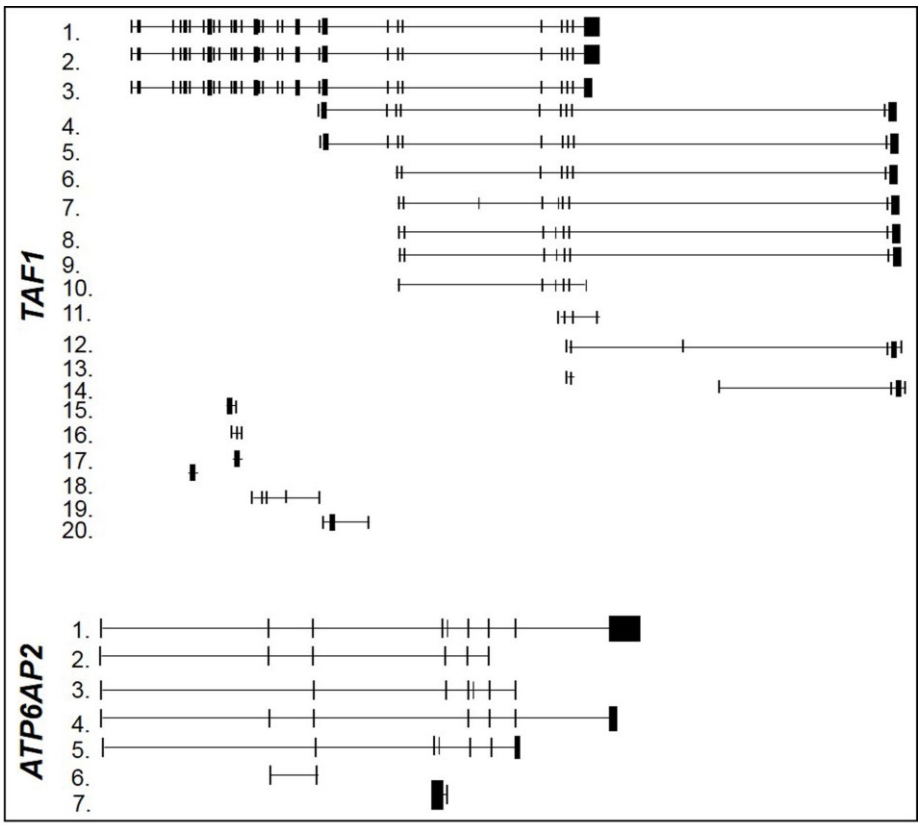


Figure 4. **Structures of the alternative splicing variants of human X-linked PD genes**

Structures of the described mRNA splicing variants are represented in figure as reported in Ensembl library (http://www.ensembl.org/index.html). On the left, each variant is indicated with a number corresponding to that indicated in Table 3. All transcripts are illustrated in 5'-3' sense.

Gene Name	Transcript Number	Ensembl Name	GenBank Accession Number	Protein Length
TAF1	1.	TAF1-201	NM_001286074	1895 aa
	2.	TAF1-009	NM_138923	1872 aa
	3.	TAF1-008	NM_004606	1893 aa

	4.	TAF1-014	-	458 aa
	5.	TAF1-010	-	490 aa
	6.	TAF1-021	-	No protein
	7.	TAF1-011	-	No protein
	8.	TAF1-013	-	No protein
	9.	TAF1-012	-	No protein
	10.	TAF1-015	-	No protein
	11.	TAF1-018	-	No protein
	12.	TAF1-016	-	No protein
	13.	TAF1-022	-	No protein
	14.	TAF1-023	-	No protein
	15.	TAF1-005	-	No protein
	16.	TAF1-006	-	No protein
	17.	TAF1-020	-	No protein
	18.	TAF1-019	-	No protein
	19.	TAF1-017	-	279 aa
	20.	TAF1-007	-	150 aa
ATP6AP2	1.	ATP6AP2-004	NM_005765	350 aa
	2.	ATP6AP2-007	-	203 aa
	3.	ATP6AP2-005	-	No protein
	4.	ATP6AP2-006	-	243 aa
	5.	ATP6AP2-001	-	259 aa
	6.	ATP6AP2-003	-	No protein
	7.	ATP6AP2-002	-	No protein

OTHER PD-RELATED GENES

SNCAIP

Synphilin-1, encoded by *SNCAIP* gene, is a presynaptic protein containing several protein-protein interaction motifs, including ankyrin-like repeats, a coiled-coil domain, and an ATP/GTP-binding domain [115]. It interacts strongly with alpha-synuclein in neuronal tissue and may play a role in the formation of Lewy bodies during neurodegeneration. It is also implicated in Parkinsonism as one of the parkin substrates. In addition, some studies have identified *SNCAIP* sequence variants in PD patients and have suggested it as a candidate PD gene [116,117].

SNCAIP gene maps on chromosome 5 (5q23.2) and spans about 152 kb of genomic DNA. Although the database Gene reports *SNCAIP* composed of 11 exons, additional exonic regions emerge by aligning the sequence of the gene with each transcript. To date, at least 22 alternative spliced transcript variants have been

identified (Table 4 and Fig. 5), but the most studied are synphilin-1 and 1A. Synphilin-1 (SNCAIP-001) is the *full-length* transcript, while synphilin-1A variant is a shorter form (SNCAIP-201). This latter lacks exons 4 and 5, and contains an extra exon located between exons 10 and 11. Synphilin-1A isoform is thought to be involved in the pathogenesis of PD and may play an important role in the formation of Lewy bodies [118-120]. Interestingly, synphilin-1A protein shows enhanced aggregation properties, which cause neuronal toxicity [118-120].

The mRNA expression levels of synphilin 1, 1A, and other two additional synphilin variants have been simultaneously investigated in the *frontal cortex* of PD patients. Their overall overexpression has been demonstrated when compared to healthy controls [24,65].

MAPT

MAPT gene encodes the microtubule-associated protein tau, a protein involved in microtubule assembly and stability [121]. It is located on chromosome 17q21 and contains 15 exons. It gives rise to multiple splice transcripts (Table 4 and Fig. 5) which are differentially expressed in human tissues [11]. In the adult human central nervous system, *MAPT* splicing generates six tau isoforms composed of either three or four microtubule-binding repeat motifs in the C-terminal (3R and 4R-tau).

A number of mutations within and around *MAPT* exon 10, disrupt exonic and intronic splicing elements as well as the formation of an RNA stem-loop structure at the 5'splice site (which normally functions to restrict spliceosome assembly). This event results in an altered ratio of 3R/4R isoforms [10,13]. The disruption of the balance between them results in hyperphosphorylation and aggregation of tau proteins into neurofibrillary tangles, causing the frontotemporal dementia with Parkinsonism linked to chromosome 17 (FTDP-17) [10,13]. These data support a direct relationship between aberrant alternative splicing of tau and neuropathology.

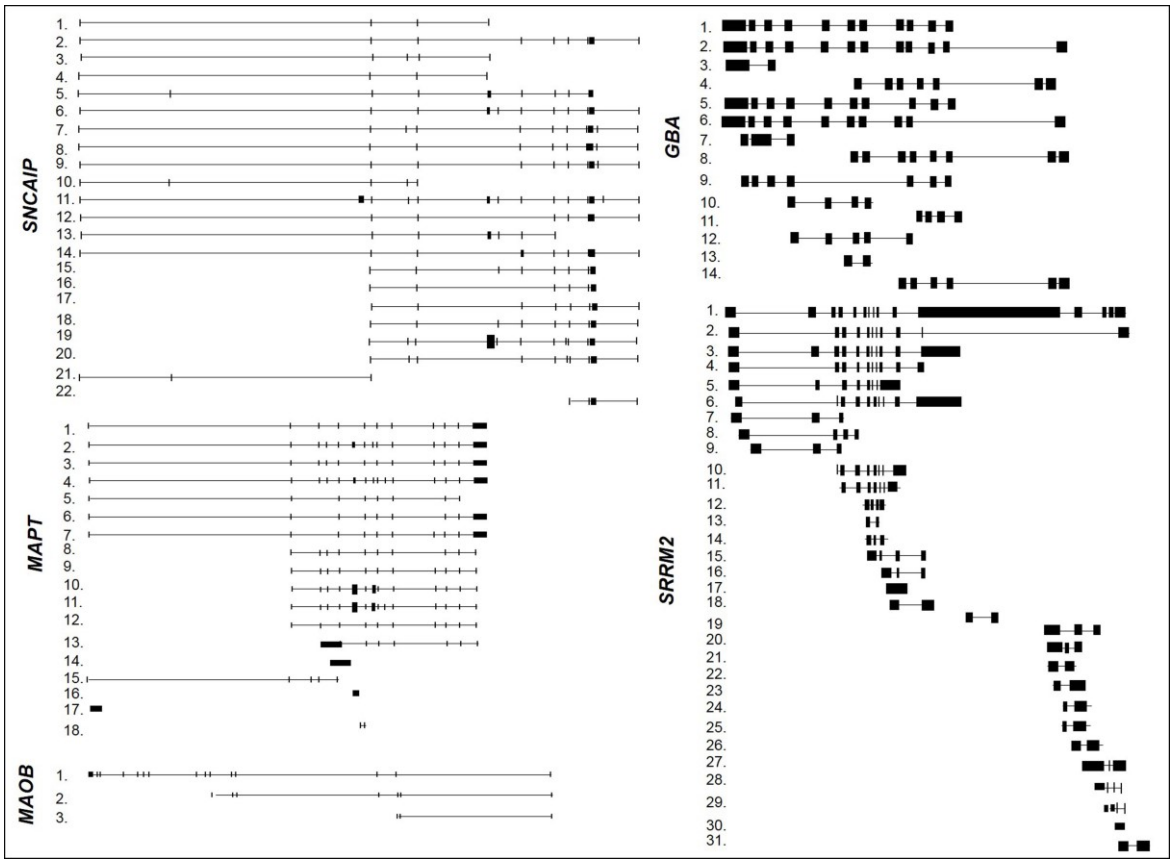


Figure 4. **Structures of the alternative splicing variants of the other human PD-related genes**

Structures of the described mRNA splicing variants are represented in figure as reported in Ensembl library (http://www.ensembl.org/index.html). On the left, each variant is indicated with a number corresponding to that indicated in Table 4. All transcripts are illustrated in 5'-3' sense, except *MAO-B* and *GBA* genes, which are illustrated in antisense corresponding to their 3'-5' sense transcription.

GBA

Mutations in *GBA* gene cause Gaucher disease, a lysosomal storage disease characterized by an accumulation of glucocerebrosides. Some studies have identified *GBA* genetic variants as significant risk factors for the development of PD [122,15].

GBA gene is located in a gene-rich region on chromosome 1q21. It spans 10.4 kb and contains 12 exons. Currently, there are four annotated alternative transcripts encoding proteins (GBA-001, GBA-002, GBA-015, GBA-016) (Table 4 and Fig. 5). Two of them originate from an alternative promoter located 2.6 kb upstream of the first ATG [123]. All transcripts share the same start codon, with the exception of GBA-012, whose open reading frame starts upon exon 4 and produces a shorter protein isoform. Further transcripts are produced, but they do not encode proteins. The *GBA* splicing

profile has not still been studied, and it is unknown if its alternative splicing is involved in PD.

MAO-B

MAO-B gene is located on chromosome X and includes 15 exons (Table 4 and Fig. 5). Although it is not a confirmed susceptibility gene [18], increased levels of Monoamine Oxidase B (MAO) mRNA and enzymatic activity have been reported in platelets from patients with both Parkinson's and Alzheimer's diseases [124]. Furthermore, it is well established that MAO-B inhibitors delay progression of both pathologies [125,126].

Several DNA polymorphisms in the *MAO-B* gene have been described in populations with distinct ethnic backgrounds [124]. A SNP common in all ethnic groups and associated with two-fold risk of PD is the G/A dimorphism in intron 13 sequence [127-129]. This SNP does not change the coding sequence and does not affect the consensus acceptor and donor sites. However, it has been demonstrated the G/A dimorphism in intron 13 sequence creates a splicing enhancer that stimulates intron 13 removal, spliceosomal complex assembly and alter splicing factors binding site efficiency [124].

SRRM2

Along with *cis-acting* elements, alternative splicing regulation relies on *trans*-splicing factors, including the serine/arginine (SR) proteins. One of these proteins, the RNA splicing factor SRRM2 (or serine/arginine repetitive matrix 2), has been identified as the only gene that stood out as differentially expressed in multiple gene expression PD datasets [130].

SRRM2 gene generates two main alternative splicing transcripts different at their 3' end (Table 4 and Fig. 5). The *full-length SRRM2* isoform SRRM2-001 contains 15 exons, while the shorter isoform SRRM2-003 contains 11 exons and lacks exons 12–15. These two isoforms are differentially expressed in postmortem PD brain regions [130]. The shorter transcript was upregulated in *substantia nigra* but unchanged in the *amygdala* of PD patients versus healthy controls. On the contrary, the longer transcript was downregulated in both *substantia nigra* and *amygdala* of PDs as compared to controls [130]. Furthermore, in the peripheral blood of patients with

PD, SRRM2 short isoform is overexpressed, while the expression of longest isoform is reduced [130].

Gene Name	Transcript Number	Ensembl Name	GenBank Accession Number	Protein Length
SNCAIP	1.	SNCAIP-019	-	135 aa
	2.	SNCAIP-003	-	66 aa
	3.	SNCAIP-016	-	161 aa
	4.	SNCAIP-017	-	98 aa
	5.	SNCAIP-010	-	858 aa
	6.	SNCAIP-001	NM_005460	919 aa
	7.	SNCAIP-204	-	113 aa
	8.	SNCAIP-201	NM_001242935	603 aa
	9.	SNCAIP-004	-	66 aa
	10.	SNCAIP-018	-	68 aa
	11.	SNCAIP-002	-	1016 aa
	12.	SNCAIP-006	-	62 aa
	13.	SNCAIP-005	-	No protein
	14.	SNCAIP-007	-	66 aa
	15.	SNCAIP-203	-	88 aa
	16.	SNCAIP-202	-	62 aa
	17.	SNCAIP-012	-	66 aa
	18.	SNCAIP-011	-	88 aa
	19.	SNCAIP-009	-	588 aa
	20.	SNCAIP-008	-	113 aa
	21.	SNCAIP-015	-	14 aa
	22.	SNCAIP-013	-	No protein
MAO-B	1.	MAOB-001	NM_000898	520 aa
	2.	MAOB-002	-	No protein
	3.	MAOB-004	-	No protein
GBA	1.	GBA-011	-	No protein
	2.	GBA-001	NM_000157	536 aa
	3.	GBA-002	NM_001005741; NM_001005742	536 aa
	4.	GBA-003	-	No protein
	5.	GBA-009	-	No protein
	6.	GBA-015	NM_001171812	487 aa
	7.	GBA-016	NM_001171811	449 aa
	8.	GBA-005	-	No protein
	9.	GBA-012	-	No protein
	10.	GBA-006	-	No protein
	11.	GBA-010	-	No protein
	12.	GBA-014	-	No protein
	13.	GBA-007	-	No protein
	14.	GBA-013	-	No protein
MAPT	1.	MAPT-204	NM_005910	441 aa
	2.	MAPT-202	NM_001123067	412 aa
	3.	MAPT-201	NM_016835	758 aa
	4.	MAPT-205	NM_001203251; NM_001203252	410 aa
	5.	MAPT-203	NM_001123066	776 aa
	6.	MAPT-013	-	No protein
	7.	MAPT-001	NM_016841	352 aa
	8.	MAPT-002	NM_016834	383 aa
	9.	MAPT-006	-	410 aa
	10.	MAPT-007	-	441 aa
	11.	MAPT-008	-	758 aa

	12.	MAPT-004	-	776 aa
	13.	MAPT-003	-	412 aa
	14.	MAPT-009	-	341 aa
	15.	MAPT-014	-	No protein
	16.	MAPT-011	-	59 aa
	17.	MAPT-010	-	No protein
	18.	MAPT-012	-	No protein
	1.	SRRM2-001	NM_016333	2752 aa
	2.	SRRM2-201	-	311aa
	3.	SRRM2-003	-	1018 aa
	4.	SRRM2-006	-	297 aa
	5.	SRRM2-004	-	No protein
	6.	SRRM2-007	-	895 aa
	7.	SRRM2-011	-	No protein
	8.	SRRM2-028	-	94 aa
	9.	SRRM2-012	-	115 aa
	10.	SRRM2-013	-	251 aa
	11.	SRRM2-029	-	No protein
	12.	SRRM2-014	-	No protein
	13.	SRRM2-030	-	No protein
	14.	SRRM2-015	-	No protein
	15.	SRRM2-016	-	No protein
SRRM2	16.	SRRM2-017	-	No protein
	17.	SRRM2-009	-	No protein
	18.	SRRM2-018	-	No protein
	19.	SRRM2-019	-	No protein
	20.	SRRM2-022	-	No protein
	21.	SRRM2-020	-	184 aa
	22.	SRRM2-021	-	No protein
	23.	SRRM2-023	-	No protein
	24.	SRRM2-024	-	No protein
	25.	SRRM2-025	-	No protein
	26.	SRRM2-026	-	No protein
	27.	SRRM2-010	-	No protein
	28.	SRRM2-027	-	No protein
	29.	SRRM2-031	-	78 aa
	30	SRRM2-032	-	41 aa
	31.	SRRM2-033	-	No protein

GENOME-WIDE RNA EXPRESSION ANALYSIS REVEALS GLOBAL ALTERNATIVE SPLICING CHANGES IN PD

Although a “gene by gene” approach may simplify splicing analysis, global alternative splicing changes in PD have to be considered. The majority of the whole gene-expression array studies in PD brain regions have unfortunately looked at a single transcript per gene, ignoring the multiple transcripts generated by alternative splicing [14]. Nonetheless, mRNA splicing has been identified as a mechanism significantly altered in cortical neurons of PD patients [131].

In order to investigate the splicing expression changes, some studies have used exon arrays. This kind of approach, enabling better monitoring and detection of the

alternative splicing events, has allowed to observe significant changes in overall gene splicing in PD blood cells compared to healthy controls [130,132]. Another exon array study has been conducted in blood of advanced PD patients prior to and following deep brain stimulation neurosurgery, a technique that efficiently improves the motor symptoms of PD [133]. This analysis has showed preliminary results suggesting brain electrical stimulation may correlate with significant profile changes in Nonsense Mediated mRNA Decay (an mRNA surveillance process that detects, and selectively degrades splice transcripts harboring premature termination codons) in blood cell transcripts [133]. Potashkin et al.(2012) has also used specific splice variant microarrays in PD patients in order to identify mRNAs splice transcripts as molecular biomarkers for an early PD diagnosis [134]. Through this approach, they identified 13 splice variants with an altered expression in early-stage PD patients versus healthy controls [134,135].

A recent technology to better study splicing defects is deep sequencing of RNA (RNAseq) [14]. The advantages of RNAseq are that it is theoretically feasible to measure both RNA expression levels and modifications such as splicing. In addition, RNAseq gives the possibility of discovering novel transcripts. Whole transcriptome RNAseq data have been obtained from blood leukocytes of PD patients' pre- and post-deep brain stimulation treatment [136]. This approach has enabled to discover novel human exons and junctions in protein-coding RNA molecules, as well as a large range of differential splicing events pre and post treatment compared to healthy controls [136]. Although this is the first study using in-depth PD transcriptome sequencing, RNAseq represents a promising technique to better study PD alternative splicing.

THE ROLE of miRNA AND lncRNA in PD ALTERNATIVE SPLICING MODULATION

A large number of alternative exon regions have been predicted as binding sites of microRNAs (miRNAs). These latter are a class of small non-coding RNA molecules, which mainly act as post-transcriptional modulators of multiple target genes by partial sequence complementarity. Through this mechanism, they may also influence splicing process.

The interplay between miRNA differential expression and alternative splicing modification in PD has been recently investigated [137]. Parallel changes in miRNA profiles and their spliced targets have been observed in PD leukocytes and PD-relevant brain regions (including the *substantia nigra* as well as the *frontal lobe*). This study was conducted through coupled analysis of small RNA sequencing data, splice junction arrays and exon arrays [137].

Another novel fascinating class of RNAs with unknown functions is long non-coding RNAs (lncRNAs), defined as transcripts of over 200 nucleotides. The GENCODE non-coding RNA set collects all lncRNAs known so far, including several spliced transcript shorter than 200 bp. LncRNA profiling has been recently assessed in PD leukocytes pre- and post-deep brain stimulation via RNAseq [136]. This survey allowed to identify some lncRNAs overexpressed in PD and inversely decreased following deep brain stimulation [136]. Differentially expressed lncRNA include the spliceosome component U1, supporting the hypothesis of disease-involved splicing modulations [136].

The identification of existing networks between non-coding mRNAs and alternative splicing modifications represents an important step forward the road to understanding the molecular basis of PD.

CONCLUSIONS

Alternative splicing is a highly harmonized process, based on a combination of DNA sequence motifs, intronic and exonic elements, regulatory factors and temporal and spatial signaling pathways. Mutations that disrupt any of these critical features may alter the finely tuned splicing processes, upsetting the production or functions of the encoded proteins, and finally causing human diseases. Assessing the alternative splicing modulation of PD-related genes represents an important point to understand PD molecular etiology. Future studies, both with the standard or the new currently available large-scale techniques, will offer a complete data pool of the alternative splicing events in PD and will provide new possible insights in order to develop strategies for PD therapy and diagnosis.

ACKNOWLEDGMENTS

This study was supported by the international Ph.D. program in Neuroscience, University of Catania. We also gratefully acknowledge Cristina Calì, Alfia Corsino, Maria Patrizia D'Angelo and Francesco Marino for their administrative and technical support.

REFERENCES

1. Ward AJ, Cooper TA (2010) The pathobiology of splicing. J Pathol 220 (2):152-163. doi:10.1002/path.2649

2. Yap K, Makeyev EV (2013) Regulation of gene expression in mammalian nervous system through alternative pre-mRNA splicing coupled with RNA quality control mechanisms. Mol Cell Neurosci 56:420-428. doi:10.1016/j.mcn.2013.01.003

3. Calarco JA, Zhen M, Blencowe BJ (2011) Networking in a global world: establishing functional connections between neural splicing regulators and their target transcripts. RNA 17 (5):775-791. doi:10.1261/rna.2603911

4. Li Q, Lee JA, Black DL (2007) Neuronal regulation of alternative pre-mRNA splicing. Nat Rev Neurosci 8 (11):819-831. doi:10.1038/nrn2237

5. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature genetics 40 (12):1413-1415. doi:10.1038/ng.259

6. Behzadnia N, Golas MM, Hartmuth K, Sander B, Kastner B, Deckert J, Dube P, Will CL, Urlaub H, Stark H, Luhrmann R (2007) Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. The EMBO journal 26 (6):1737-1748. doi:7601631 [pii]

10.1038/sj.emboj.7601631

7. Shefer K, Sperling J, Sperling R (2014) The Supraspliceosome - A Multi-Task Machine for Regulated Pre-mRNA Processing in the Cell Nucleus. Computational and structural biotechnology journal 11 (19):113-122. doi:10.1016/j.csbj.2014.09.008

8. Gamazon ER, Stranger BE (2014) Genomics of alternative splicing: evolution, development and pathophysiology. Human genetics 133 (6):679-687. doi:10.1007/s00439-013-1411-3

9. Scuderi S, La Cognata V, Drago F, Cavallaro S, D'Agata V (2014) Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain. Biomed Res Int 2014:690796. doi:10.1155/2014/690796

10. Licatalosi DD, Darnell RB (2006) Splicing regulation in neurologic disease. Neuron 52 (1):93-101. doi:10.1016/j.neuron.2006.09.017

11. Tazi J, Bakkour N, Stamm S (2009) Alternative splicing and disease. Biochim Biophys Acta 1792 (1):14-26. doi:10.1016/j.bbadis.2008.09.017

12. Singh RK, Cooper TA (2012) Pre-mRNA splicing in disease and therapeutics. Trends Mol Med 18 (8):472-482. doi:10.1016/j.molmed.2012.06.006

13. Cooper TA, Wan L, Dreyfuss G (2009) RNA and disease. Cell 136 (4):777-793. doi:10.1016/j.cell.2009.02.011

14. Lewis PA, Cookson MR (2012) Gene expression in the Parkinson's disease brain. Brain Res Bull 88 (4):302-312. doi:10.1016/j.brainresbull.2011.11.016

15. Bonifati V (2014) Genetics of Parkinson's disease--state of the art, 2013. Parkinsonism & related disorders 20 Suppl 1:S23-28. doi:10.1016/S1353-8020(13)70009-9

16. Lubbe S, Morris HR (2014) Recent advances in Parkinson's disease genetics. Journal of neurology 261 (2):259-266. doi:10.1007/s00415-013-7003-2

17. Trinh J, Farrer M (2013) Advances in the genetics of Parkinson disease. Nature reviews

Neurology 9 (8):445-454. doi:10.1038/nrneurol.2013.132

18. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, DeStefano AL, Kara E, Bras J, Sharma M, Schulte C, Keller MF, Arepalli S, Letson C, Edsall C, Stefansson H, Liu X, Pliner H, Lee JH, Cheng R, International Parkinson's Disease Genomics C, Parkinson's Study Group Parkinson's Research: The Organized GI, andMe, GenePd, NeuroGenetics Research C, Hussman Institute of Human G, Ashkenazi Jewish Dataset I, Cohorts for H, Aging Research in Genetic E, North American Brain Expression C, United Kingdom Brain Expression C, Greek Parkinson's Disease C, Alzheimer Genetic Analysis G, Ikram MA, Ioannidis JP, Hadjigeorgiou GM, Bis JC, Martinez M, Perlmutter JS, Goate A, Marder K, Fiske B, Sutherland M, Xiromerisiou G, Myers RH, Clark LN, Stefansson K, Hardy JA, Heutink P, Chen H, Wood NW, Houlden H, Payami H, Brice A, Scott WK, Gasser T, Bertram L, Eriksson N, Foroud T, Singleton AB (2014) Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nature genetics 46 (9):989-993. doi:10.1038/ng.3043

19. Stefanis L (2012) alpha-Synuclein in Parkinson's disease. Cold Spring Harb Perspect Med 2 (2):a009399. doi:10.1101/cshperspect.a009399

20. Pihlstrom L, Toft M (2011) Genetic variability in SNCA and Parkinson's disease. Neurogenetics 12 (4):283-293. doi:10.1007/s10048-011-0292-7

21. Deng H, Yuan L (2014) Genetic variants and animal models in SNCA and Parkinson disease. Ageing Res Rev 15C:161-176. doi:10.1016/j.arr.2014.04.002

22. Nuytemans K, Meeus B, Crosiers D, Brouwers N, Goossens D, Engelborghs S, Pals P, Pickut B, Van den Broeck M, Corsmit E, Cras P, De Deyn PP, Del-Favero J, Van Broeckhoven C, Theuns J (2009) Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population. Hum Mutat 30 (7):1054-1061. doi:10.1002/humu.21007

23. Beyer K, Domingo-Sabat M, Lao JI, Carrato C, Ferrer I, Ariza A (2008) Identification and characterization of a new alpha-synuclein isoform and its role in Lewy body diseases. Neurogenetics 9 (1):15-23. doi:10.1007/s10048-007-0106-0

24. Beyer K, Domingo-Sabat M, Humbert J, Carrato C, Ferrer I, Ariza A (2008) Differential expression of alpha-synuclein, parkin, and synphilin-1 isoforms in Lewy body disease. Neurogenetics 9 (3):163-172. doi:10.1007/s10048-008-0124-6

25. McLean JR, Hallett PJ, Cooper O, Stanley M, Isacson O (2012) Transcript expression levels of full-length alpha-synuclein and its three alternatively spliced variants in Parkinson's disease brain regions and in a transgenic mouse model of alpha-synuclein overexpression. Mol Cell Neurosci 49 (2):230-239. doi:10.1016/j.mcn.2011.11.006

26. Cardo LF, Coto E, de Mena L, Ribacoba R, Mata IF, Menendez M, Moris G, Alvarez V (2014) Alpha-synuclein transcript isoforms in three different brain regions from Parkinson's disease and healthy subjects in relation to the SNCA rs356165/rs11931074 polymorphisms. Neurosci Lett 562:45-49. doi:10.1016/j.neulet.2014.01.009

27. Beyer K, Lao JI, Carrato C, Mate JL, Lopez D, Ferrer I, Ariza A (2004) Differential expression of alpha-synuclein isoforms in dementia with Lewy bodies. Neuropathol Appl Neurobiol 30 (6):601-607. doi:10.1111/j.1365-2990.2004.00572.x

28. Beyer K, Humbert J, Ferrer A, Lao JI, Carrato C, Lopez D, Ferrer I, Ariza A (2006) Low alpha-synuclein 126 mRNA levels in dementia with Lewy bodies and Alzheimer disease. Neuroreport 17 (12):1327-1330. doi:10.1097/01.wnr.0000224773.66904.e7

29. McCarthy JJ, Linnertz C, Saucier L, Burke JR, Hulette CM, Welsh-Bohmer KA, Chiba-Falek O (2011) The effect of SNCA 3' region on the levels of SNCA-112 splicing variant. Neurogenetics 12 (1):59-64. doi:10.1007/s10048-010-0263-4

30. Kalivendi SV, Yedlapudi D, Hillard CJ, Kalyanaraman B (2010) Oxidants induce alternative splicing of alpha-synuclein: Implications for Parkinson's disease. Free Radic Biol Med 48 (3):377-383. doi:10.1016/j.freeradbiomed.2009.10.045

31. Rhinn H, Qiang L, Yamashita T, Rhee D, Zolin A, Vanti W, Abeliovich A (2012) Alternative alpha-synuclein transcript usage as a convergent

mechanism in Parkinson's disease pathology. Nature communications 3:1084. doi:10.1038/ncomms2032

32. Beyer K (2006) Alpha-synuclein structure, posttranslational modification and alternative splicing as aggregation enhancers. Acta Neuropathol 112 (3):237-251. doi:10.1007/s00401-006-0104-6

33. Bungeroth M, Appenzeller S, Regulin A, Volker W, Lorenzen I, Grotzinger J, Pendziwiat M, Kuhlenbaumer G (2014) Differential aggregation properties of alpha-synuclein isoforms. Neurobiology of aging 35 (8):1913-1919. doi:10.1016/j.neurobiolaging.2014.02.009

34. Burre J, Sharma M, Tsetsenis T, Buchman V, Etherton MR, Sudhof TC (2010) Alpha-synuclein promotes SNARE-complex assembly in vivo and in vitro. Science 329 (5999):1663-1667. doi:10.1126/science.1195227

35. Rideout HJ, Stefanis L (2014) The neurobiology of LRRK2 and its role in the pathogenesis of Parkinson's disease. Neurochem Res 39 (3):576-592. doi:10.1007/s11064-013-1073-5

36. Johnson J, Paisan-Ruiz C, Lopez G, Crews C, Britton A, Malkani R, Evans EW, McInerney-Leo A, Jain S, Nussbaum RL, Foote KD, Mandel RJ, Crawley A, Reimsnider S, Fernandez HH, Okun MS, Gwinn-Hardy K, Singleton AB (2007) Comprehensive screening of a North American Parkinson's disease cohort for LRRK2 mutation. Neurodegener Dis 4 (5):386-391. doi:10.1159/000105160

37. Skipper L, Shen H, Chua E, Bonnard C, Kolatkar P, Tan LC, Jamora RD, Puvan K, Puong KY, Zhao Y, Pavanni R, Wong MC, Yuen Y, Farrer M, Liu JJ, Tan EK (2005) Analysis of LRRK2 functional domains in nondominant Parkinson disease. Neurology 65 (8):1319-1321. doi:10.1212/01.wnl.0000180517.70572.37

38. Zabetian CP, Samii A, Mosley AD, Roberts JW, Leis BC, Yearout D, Raskind WH, Griffith A (2005) A clinic-based study of the LRRK2 gene in Parkinson disease yields new mutations. Neurology 65 (5):741-744. doi:10.1212/01.wnl.0000172630.22804.73

39. Shojaee S, Sina F, Farboodi N, Fazlali Z, Ghazavi F, Ghorashi SA, Parsa K, Sadeghi H, Shahidi GA, Ronaghi M, Elahi E (2009) A clinic-based screening of mutations in exons 31, 34, 35, 41, and 48 of LRRK2 in Iranian Parkinson's disease patients. Mov Disord 24 (7):1023-1027. doi:10.1002/mds.22503

40. Di Fonzo A, Tassorelli C, De Mari M, Chien HF, Ferreira J, Rohe CF, Riboldazzi G, Antonini A, Albani G, Mauro A, Marconi R, Abbruzzese G, Lopiano L, Fincati E, Guidi M, Marini P, Stocchi F, Onofrj M, Toni V, Tinazzi M, Fabbrini G, Lamberti P, Vanacore N, Meco G, Leitner P, Uitti RJ, Wszolek ZK, Gasser T, Simons EJ, Breedveld GJ, Goldwurm S, Pezzoli G, Sampaio C, Barbosa E, Martignoni E, Oostra BA, Bonifati V, Italian Parkinson's Genetics N (2006) Comprehensive analysis of the LRRK2 gene in sixty families with Parkinson's disease. Eur J Hum Genet 14 (3):322-331. doi:10.1038/sj.ejhg.5201539

41. Grimes DA, Racacho L, Han F, Panisset M, Bulman DE (2007) LRRK2 screening in a Canadian Parkinson's disease cohort. Can J Neurol Sci 34 (3):336-338

42. Paisan-Ruiz C, Nath P, Washecka N, Gibbs JR, Singleton AB (2008) Comprehensive analysis of LRRK2 in publicly available Parkinson's disease cases and neurologically normal controls. Hum Mutat 29 (4):485-490. doi:10.1002/humu.20668

43. Lesage S, Condroyer C, Lannuzel A, Lohmann E, Troiano A, Tison F, Damier P, Thobois S, Ouvrard-Hernandez AM, Rivaud-Pechoux S, Brefel-Courbon C, Destee A, Tranchant C, Romana M, Leclere L, Durr A, Brice A, French Parkinson's Disease Genetics Study G (2009) Molecular analyses of the LRRK2 gene in European and North African autosomal dominant Parkinson's disease. J Med Genet 46 (7):458-464. doi:10.1136/jmg.2008.062612

44. Trabzuni D, Ryten M, Emmett W, Ramasamy A, Lackner KJ, Zeller T, Walker R, Smith C, Lewis PA, Mamais A, de Silva R, Vandrovcova J, International Parkinson Disease Genomics C, Hernandez D, Nalls MA, Sharma M, Garnier S, Lesage S, Simon-Sanchez J, Gasser T, Heutink P, Brice A, Singleton A, Cai H, Schadt E, Wood NW, Bandopadhyay R, Weale ME, Hardy J, Plagnol V (2013) Fine-mapping, gene expression and splicing analysis of the disease associated LRRK2 locus. PLoS One 8 (8):e70724. doi:10.1371/journal.pone.0070724

45. Giesert F, Hofmann A, Burger A, Zerle J, Kloos K, Hafen U, Ernst L, Zhang J, Vogt-Weisenhorn DM, Wurst W (2013) Expression analysis of Lrrk1, Lrrk2 and Lrrk2 splice variants in mice. PLoS One 8 (5):e63778. doi:10.1371/journal.pone.0063778

46. Sheng D, Qu D, Kwok KH, Ng SS, Lim AY, Aw SS, Lee CW, Sung WK, Tan EK, Lufkin T, Jesuthasan S, Sinnakaruppan M, Liu J (2010) Deletion of the WD40 domain of LRRK2 in Zebrafish causes Parkinsonism-like loss of neurons and locomotive defect. PLoS Genet 6 (4):e1000914. doi:10.1371/journal.pgen.1000914

47. Vilarino-Guell C, Wider C, Ross OA, Dachsel JC, Kachergus JM, Lincoln SJ, Soto-Ortolaza AI, Cobb SA, Wilhoite GJ, Bacon JA, Behrouz B, Melrose HL, Hentati E, Puschmann A, Evans DM, Conibear E, Wasserman WW, Aasly JO, Burkhard PR, Djaldetti R, Ghika J, Hentati F, Krygowska-Wajs A, Lynch T, Melamed E, Rajput A, Rajput AH, Solida A, Wu RM, Uitti RJ, Wszolek ZK, Vingerhoets F, Farrer MJ (2011) VPS35 mutations in Parkinson disease. American journal of human genetics 89 (1):162-167. doi:10.1016/j.ajhg.2011.06.001

48. Zimprich A, Benet-Pages A, Struhal W, Graf E, Eck SH, Offman MN, Haubenberger D, Spielberger S, Schulte EC, Lichtner P, Rossle SC, Klopp N, Wolf E, Seppi K, Pirker W, Presslauer S, Mollenhauer B, Katzenschlager R, Foki T, Hotzy C, Reinthaler E, Harutyunyan A, Kralovics R, Peters A, Zimprich F, Brucke T, Poewe W, Auff E, Trenkwalder C, Rost B, Ransmayr G, Winkelmann J, Meitinger T, Strom TM (2011) A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. American journal of human genetics 89 (1):168-175. doi:10.1016/j.ajhg.2011.06.008

49. Verstraeten A, Wauters E, Crosiers D, Meeus B, Corsmit E, Elinck E, Mattheijssens M, Peeters K, Cras P, Pickut B, Vandenberghe R, Engelborghs S, De Deyn PP, Van Broeckhoven C, Theuns J (2012) Contribution of VPS35 genetic variability to LBD in the Flanders-Belgian population. Neurobiology of aging 33 (8):1844 e1811-1843. doi:10.1016/j.neurobiolaging.2012.01.006

50. Bonifati V (2012) Autosomal recessive parkinsonism. Parkinsonism & related disorders 18 Suppl 1:S4-6. doi:10.1016/S1353-8020(11)70004-9

51. Lucking CB, Durr A, Bonifati V, Vaughan J, De Michele G, Gasser T, Harhangi BS, Meco G, Denefle P, Wood NW, Agid Y, Brice A, French Parkinson's Disease Genetics Study G, European Consortium on Genetic Susceptibility in Parkinson's D (2000) Association between early-onset Parkinson's disease and mutations in the parkin gene. N Engl J Med 342 (21):1560-1567. doi:10.1056/NEJM200005253422103

52. Oliveira SA, Scott WK, Martin ER, Nance MA, Watts RL, Hubble JP, Koller WC, Pahwa R, Stern MB, Hiner BC, Ondo WG, Allen FH, Jr., Scott BL, Goetz CG, Small GW, Mastaglia F, Stajich JM, Zhang F, Booze MW, Winn MP, Middleton LT, Haines JL, Pericak-Vance MA, Vance JM (2003) Parkin mutations and susceptibility alleles in late-onset Parkinson's disease. Ann Neurol 53 (5):624-629. doi:10.1002/ana.10524

53. Illarioshkin SN, Periquet M, Rawal N, Lucking CB, Zagorovskaya TB, Slominsky PA, Miloserdova OV, Markova ED, Limborska SA, Ivanova-Smolenskaya IA, Brice A (2003) Mutation analysis of the parkin gene in Russian families with autosomal recessive juvenile parkinsonism. Mov Disord 18 (8):914-919. doi:10.1002/mds.10467

54. Pigullo S, De Luca A, Barone P, Marchese R, Bellone E, Colosimo A, Scaglione C, Martinelli P, Di Maria E, Pizzuti A, Abbruzzese G, Dallapiccola B, Ajmar F, Mandich P (2004) Mutational analysis of parkin gene by denaturing high-performance liquid chromatography (DHPLC) in essential tremor. Parkinsonism & related disorders 10 (6):357-362. doi:10.1016/j.parkreldis.2004.04.012

55. Scherfler C, Khan NL, Pavese N, Eunson L, Graham E, Lees AJ, Quinn NP, Wood NW, Brooks DJ, Piccini PP (2004) Striatal and cortical pre- and postsynaptic dopaminergic dysfunction in sporadic parkin-linked parkinsonism. Brain 127 (Pt 6):1332-1342. doi:10.1093/brain/awh150

56. Bertoli-Avella AM, Giroud-Benitez JL, Akyol A, Barbosa E, Schaap O, van der Linde HC, Martignoni E, Lopiano L, Lamberti P, Fincati E, Antonini A, Stocchi F, Montagna P, Squitieri F, Marini P, Abbruzzese G, Fabbrini G, Marconi R, Dalla Libera A, Trianni G, Guidi M, De Gaetano A, Boff Maegawa

G, De Leo A, Gallai V, de Rosa G, Vanacore N, Meco G, van Duijn CM, Oostra BA, Heutink P, Bonifati V, Italian Parkinson Genetics N (2005) Novel parkin mutations detected in patients with early-onset Parkinson's disease. Mov Disord 20 (4):424-431. doi:10.1002/mds.20343

57. Bardien S, Keyser R, Yako Y, Lombard D, Carr J (2009) Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease. Parkinsonism & related disorders 15 (2):116-121. doi:10.1016/j.parkreldis.2008.04.005

58. Pankratz N, Kissell DK, Pauciulo MW, Halter CA, Rudolph A, Pfeiffer RF, Marder KS, Foroud T, Nichols WC, Parkinson Study Group PI (2009) Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations. Neurology 73 (4):279-286. doi:10.1212/WNL.0b013e3181af7a33

59. La Cognata V, Iemmolo R, D'Agata V, Scuderi S, Drago F, Zappia M, Cavallaro S (2014) Increasing the Coding Potential of Genomes Through Alternative Splicing: The Case of PARK2 Gene. Curr Genomics 15 (3):203-216. doi:10.2174/1389202915666140426003342

60. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 392 (6676):605-608

61. Matsumine H, Saito M, Shimoda-Matsubayashi S, Tanaka H, Ishikawa A, Nakagawa-Hattori Y, Yokochi M, Kobayashi T, Igarashi S, Takano H, Sanpei K, Koike R, Mori H, Kondo T, Mizutani Y, Schaffer AA, Yamamura Y, Nakamura S, Kuzuhara S, Tsuji S, Mizuno Y (1997) Localization of a gene for an autosomal recessive form of juvenile Parkinsonism to chromosome 6q25.2-27. American journal of human genetics 60 (3):588-596

62. Xu L, Lin DC, Yin D, Koeffler HP (2014) An emerging role of PARK2 in cancer. J Mol Med (Berl) 92 (1):31-42. doi:10.1007/s00109-013-1107-0

63. Ikeuchi K, Marusawa H, Fujiwara M, Matsumoto Y, Endo Y, Watanabe T, Iwai A, Sakai Y, Takahashi R, Chiba T (2009) Attenuation of proteolysis-mediated cyclin E regulation by alternatively spliced Parkin in human colorectal cancers. Int J Cancer 125 (9):2029-2035. doi:10.1002/ijc.24565

64. Kitada T, Asakawa S, Minoshima S, Mizuno Y, Shimizu N (2000) Molecular cloning, gene expression, and identification of a splicing variant of the mouse parkin gene. Mamm Genome 11 (6):417-421. doi:MG99-682 [pii]

65. Humbert J, Beyer K, Carrato C, Mate JL, Ferrer I, Ariza A (2007) Parkin and synphilin-1 isoform expression changes in Lewy body diseases. Neurobiol Dis 26 (3):681-687. doi:S0969-9961(07)00063-0 [pii]

10.1016/j.nbd.2007.03.007

66. Tan EK, Shen H, Tan JM, Lim KL, Fook-Chong S, Hu WP, Paterson MC, Chandran VR, Yew K, Tan C, Yuen Y, Pavanni R, Wong MC, Puvan K, Zhao Y (2005) Differential expression of splice variant and wild-type parkin in sporadic Parkinson's disease. Neurogenetics 6 (4):179-184. doi:10.1007/s10048-005-0001-5

67. Dagata V, Cavallaro S (2004) Parkin transcript variants in rat and human brain. Neurochem Res 29 (9):1715-1724

68. Sunada Y, Saito F, Matsumura K, Shimizu T (1998) Differential expression of the parkin gene in the human brain and peripheral leukocytes. Neurosci Lett 254 (3):180-182. doi:S0304394098006971 [pii]

69. Solano SM, Miller DW, Augood SJ, Young AB, Penney JB, Jr. (2000) Expression of alpha-synuclein, parkin, and ubiquitin carboxy-terminal hydrolase L1 mRNA in human brain: genes associated with familial Parkinson's disease. Ann Neurol 47 (2):201-210

70. Kasap M, Akpinar G, Sazci A, Idrisoglu HA, Vahaboglu H (2009) Evidence for the presence of full-length PARK2 mRNA and Parkin protein in human blood. Neurosci Lett 460 (3):196-200. doi:10.1016/j.neulet.2009.05.079

71. Pawlyk AC, Giasson BI, Sampathu DM, Perez FA, Lim KL, Dawson VL, Dawson TM, Palmiter RD, Trojanowski JQ, Lee VM (2003) Novel monoclonal antibodies demonstrate biochemical variation of

brain parkin with age. J Biol Chem 278 (48):48120-48128. doi:10.1074/jbc.M306889200

72. Horowitz JM, Myers J, Stachowiak MK, Torres G (1999) Identification and distribution of Parkin in rat brain. Neuroreport 10 (16):3393-3397

73. Stichel CC, Augustin M, Kuhn K, Zhu XR, Engels P, Ullmer C, Lubbert H (2000) Parkin expression in the adult mouse brain. Eur J Neurosci 12 (12):4181-4194

74. Gu WJ, Abbas N, Lagunes MZ, Parent A, Pradier L, Bohme GA, Agid Y, Hirsch EC, Raisman-Vozari R, Brice A (2000) Cloning of rat parkin cDNA and distribution of parkin in rat brain. J Neurochem 74 (4):1773-1776

75. Huynh DP, Dy M, Nguyen D, Kiehl TR, Pulst SM (2001) Differential expression and tissue distribution of parkin isoforms during mouse development. Brain Res Dev Brain Res 130 (2):173-181

76. D'Agata V, Grimaldi M, Pascale A, Cavallaro S (2000) Regional and cellular expression of the parkin gene in the rat cerebral cortex. Eur J Neurosci 12 (10):3583-3588

77. Pogson JH, Ivatt RM, Whitworth AJ (2011) Molecular mechanisms of PINK1-related neurodegeneration. Curr Neurol Neurosci Rep 11 (3):283-290. doi:10.1007/s11910-011-0187-x

78. Nuytemans K, Theuns J, Cruts M, Van Broeckhoven C (2010) Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update. Hum Mutat 31 (7):763-780. doi:10.1002/humu.21277

79. Marongiu R, Brancati F, Antonini A, Ialongo T, Ceccarini C, Scarciolla O, Capalbo A, Benti R, Pezzoli G, Dallapiccola B, Goldwurm S, Valente EM (2007) Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum. Hum Mutat 28 (1):98. doi:10.1002/humu.9472

80. Samaranch L, Lorenzo-Betancor O, Arbelo JM, Ferrer I, Lorenzo E, Irigoyen J, Pastor MA, Marrero C, Isla C, Herrera-Henriquez J, Pastor P (2010) PINK1-linked parkinsonism is associated with Lewy body pathology. Brain 133 (Pt 4):1128-1142. doi:10.1093/brain/awq051

81. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S, Ali Z, Del Turco D, Bentivoglio AR, Healy DG, Albanese A, Nussbaum R, Gonzalez-Maldonado R, Deller T, Salvi S, Cortelli P, Gilks WP, Latchman DS, Harvey RJ, Dallapiccola B, Auburger G, Wood NW (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304 (5674):1158-1160. doi:10.1126/science.1096284

82. Silvestri L, Caputo V, Bellacchio E, Atorino L, Dallapiccola B, Valente EM, Casari G (2005) Mitochondrial import and enzymatic activity of PINK1 mutants associated to recessive parkinsonism. Human molecular genetics 14 (22):3477-3492. doi:10.1093/hmg/ddi377

83. Beilina A, Van Der Brug M, Ahmad R, Kesavapany S, Miller DW, Petsko GA, Cookson MR (2005) Mutations in PTEN-induced putative kinase 1 associated with recessive parkinsonism have differential effects on protein stability. Proceedings of the National Academy of Sciences of the United States of America 102 (16):5703-5708. doi:10.1073/pnas.0500617102

84. d'Amora M, Angelini C, Marcoli M, Cervetto C, Kitada T, Vallarino M (2011) Expression of PINK1 in the brain, eye and ear of mouse during embryonic development. J Chem Neuroanat 41 (2):73-85. doi:10.1016/j.jchemneu.2010.11.004

85. Lin W, Kang UJ (2008) Characterization of PINK1 processing, stability, and subcellular localization. J Neurochem 106 (1):464-474. doi:10.1111/j.1471-4159.2008.05398.x

86. Deas E, Plun-Favreau H, Gandhi S, Desmond H, Kjaer S, Loh SH, Renton AE, Harvey RJ, Whitworth AJ, Martins LM, Abramov AY, Wood NW (2011) PINK1 cleavage at position A103 by the mitochondrial protease PARL. Human molecular genetics 20 (5):867-879. doi:10.1093/hmg/ddq526

87. Lockhart PJ, Lincoln S, Hulihan M, Kachergus J, Wilkes K, Bisceglio G, Mash DC, Farrer MJ (2004) DJ-1 mutations are a rare cause of recessively inherited early onset parkinsonism mediated by loss of protein function. J Med Genet 41 (3):e22

88. Moore DJ, West AB, Dawson VL, Dawson TM (2005) Molecular pathophysiology of Parkinson's

disease. Annu Rev Neurosci 28:57-87. doi:10.1146/annurev.neuro.28.061604.135718

89. Bonifati V, Rizzu P, Squitieri F, Krieger E, Vanacore N, van Swieten JC, Brice A, van Duijn CM, Oostra B, Meco G, Heutink P (2003) DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism. Neurol Sci 24 (3):159-160. doi:10.1007/s10072-003-0108-0

90. van Duijn CM, Dekker MC, Bonifati V, Galjaard RJ, Houwing-Duistermaat JJ, Snijders PJ, Testers L, Breedveld GJ, Horstink M, Sandkuijl LA, van Swieten JC, Oostra BA, Heutink P (2001) Park7, a novel locus for autosomal recessive early-onset parkinsonism, on chromosome 1p36. American journal of human genetics 69 (3):629-634. doi:10.1086/322996

91. Corti O, Lesage S, Brice A (2011) What genetics tells us about the causes and mechanisms of Parkinson's disease. Physiol Rev 91 (4):1161-1218. doi:10.1152/physrev.00022.2010

92. Tarantino P, Civitelli D, Annesi F, De Marco EV, Rocca FE, Pugliese P, Nicoletti G, Carrideo S, Provenzano G, Annesi G, Quattrone A (2009) Compound heterozygosity in DJ-1 gene non-coding portion related to parkinsonism. Parkinsonism & related disorders 15 (4):324-326. doi:10.1016/j.parkreldis.2008.07.001

93. Hedrich K, Djarmati A, Schafer N, Hering R, Wellenbrock C, Weiss PH, Hilker R, Vieregge P, Ozelius LJ, Heutink P, Bonifati V, Schwinger E, Lang AE, Noth J, Bressman SB, Pramstaller PP, Riess O, Klein C (2004) DJ-1 (PARK7) mutations are less frequent than Parkin (PARK2) mutations in early-onset Parkinson disease. Neurology 62 (3):389-394

94. Lev N, Roncevic D, Ickowicz D, Melamed E, Offen D (2006) Role of DJ-1 in Parkinson's disease. J Mol Neurosci 29 (3):215-225

95. Ariga H, Takahashi-Niki K, Kato I, Maita H, Niki T, Iguchi-Ariga SM (2013) Neuroprotective function of DJ-1 in Parkinson's disease. Oxid Med Cell Longev 2013:683920. doi:10.1155/2013/683920

96. Xu J, Zhong N, Wang H, Elias JE, Kim CY, Woldman I, Pifl C, Gygi SP, Geula C, Yankner BA (2005) The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. Human molecular genetics 14 (9):1231-1241. doi:10.1093/hmg/ddi134

97. Bandopadhyay R, Kingsbury AE, Cookson MR, Reid AR, Evans IM, Hope AD, Pittman AM, Lashley T, Canet-Aviles R, Miller DW, McLendon C, Strand C, Leonard AJ, Abou-Sleiman PM, Healy DG, Ariga H, Wood NW, de Silva R, Revesz T, Hardy JA, Lees AJ (2004) The expression of DJ-1 (PARK7) in normal human CNS and idiopathic Parkinson's disease. Brain 127 (Pt 2):420-430. doi:10.1093/brain/awh054

98. Besong Agbo D, Klafki H, Poschmann G, Seyfarth K, Genius J, Janssen C, Stuhler K, Wurst W, Meyer HE, Klingenspor M, Wiltfang J (2013) Development of a capillary isoelectric focusing immunoassay to measure DJ-1 isoforms in biological samples. Anal Biochem 443 (2):197-204. doi:10.1016/j.ab.2013.09.013

99. Kumaran R, Kingsbury A, Coulter I, Lashley T, Williams D, de Silva R, Mann D, Revesz T, Lees A, Bandopadhyay R (2007) DJ-1 (PARK7) is associated with 3R and 4R tau neuronal and glial inclusions in neurodegenerative disorders. Neurobiol Dis 28 (1):122-132. doi:10.1016/j.nbd.2007.07.012

100. Kumaran R, Vandrovcova J, Luk C, Sharma S, Renton A, Wood NW, Hardy JA, Lees AJ, Bandopadhyay R (2009) Differential DJ-1 gene expression in Parkinson's disease. Neurobiol Dis 36 (2):393-400. doi:10.1016/j.nbd.2009.08.011

101. Lin X, Cook TJ, Zabetian CP, Leverenz JB, Peskind ER, Hu SC, Cain KC, Pan C, Edgar JS, Goodlett DR, Racette BA, Checkoway H, Montine TJ, Shi M, Zhang J (2012) DJ-1 isoforms in whole blood as potential biomarkers of Parkinson disease. Sci Rep 2:954. doi:10.1038/srep00954

102. Zhong N, Kim CY, Rizzu P, Geula C, Porter DR, Pothos EN, Squitieri F, Heutink P, Xu J (2006) DJ-1 transcriptionally up-regulates the human tyrosine hydroxylase by inhibiting the sumoylation of pyrimidine tract-binding protein-associated splicing factor. J Biol Chem 281 (30):20940-20948. doi:10.1074/jbc.M601935200

103. Vilarino-Guell C, Soto AI, Lincoln SJ, Ben Yahmed S, Kefi M, Heckman MG, Hulihan MM, Chai H, Diehl NN, Amouri R, Rajput A, Mash DC,

Dickson DW, Middleton LT, Gibson RA, Hentati F, Farrer MJ (2009) ATP13A2 variability in Parkinson disease. Hum Mutat 30 (3):406-410. doi:10.1002/humu.20877

104. Murphy KE, Cottle L, Gysbers AM, Cooper AA, Halliday GM (2013) ATP13A2 (PARK9) protein levels are reduced in brain tissue of cases with Lewy bodies. Acta Neuropathol Commun 1 (1):11. doi:10.1186/2051-5960-1-11

105. Ugolino J, Fang S, Kubisch C, Monteiro MJ (2011) Mutant Atp13a2 proteins involved in parkinsonism are degraded by ER-associated degradation and sensitize cells to ER-stress induced cell death. Human molecular genetics 20 (18):3565-3577. doi:10.1093/hmg/ddr274

106. Lu CS, Lai SC, Wu RM, Weng YH, Huang CL, Chen RS, Chang HC, Wu-Chou YH, Yeh TH (2012) PLA2G6 mutations in PARK14-linked young-onset parkinsonism and sporadic Parkinson's disease. Am J Med Genet B Neuropsychiatr Genet 159B (2):183-191. doi:10.1002/ajmg.b.32012

107. Di Fonzo A, Dekker MC, Montagna P, Baruzzi A, Yonova EH, Correia Guedes L, Szczerbinska A, Zhao T, Dubbel-Hulsman LO, Wouters CH, de Graaff E, Oyen WJ, Simons EJ, Breedveld GJ, Oostra BA, Horstink MW, Bonifati V (2009) FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. Neurology 72 (3):240-245. doi:10.1212/01.wnl.0000338144.10967.2b

108. Deng H, Liang H, Jankovic J (2013) F-box only protein 7 gene in parkinsonian-pyramidal disease. JAMA Neurol 70 (1):20-24. doi:10.1001/jamaneurol.2013.572

109. Gomez-Garre P, Jesus S, Carrillo F, Caceres-Redondo MT, Huertas-Fernandez I, Bernal-Bernal I, Bonilla-Toribio M, Vargas-Gonzalez L, Carballo M, Mir P (2014) Systematic mutational analysis of FBXO7 in a Parkinson's disease population from southern Spain. Neurobiology of aging 35 (3):727 e725-727. doi:10.1016/j.neurobiolaging.2013.09.011

110. Ilyin GP, Rialland M, Pigeon C, Guguen-Guillouzo C (2000) cDNA cloning and expression analysis of new members of the mammalian F-box protein family. Genomics 67 (1):40-47. doi:10.1006/geno.2000.6211

111. Zhao T, De Graaff E, Breedveld GJ, Loda A, Severijnen LA, Wouters CH, Verheijen FW, Dekker MC, Montagna P, Willemsen R, Oostra BA, Bonifati V (2011) Loss of nuclear activity of the FBXO7 protein in patients with parkinsonian-pyramidal syndrome (PARK15). PLoS One 6 (2):e16983. doi:10.1371/journal.pone.0016983

112. Nolte D, Niemann S, Muller U (2003) Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. Proceedings of the National Academy of Sciences of the United States of America 100 (18):10347-10352. doi:10.1073/pnas.1831949100

113. Herzfeld T, Nolte D, Grznarova M, Hofmann A, Schultze JL, Muller U (2013) X-linked dystonia parkinsonism syndrome (XDP, lubag): disease-specific sequence change DSC3 in TAF1/DYT3 affects genes in vesicular transport and dopamine metabolism. Human molecular genetics 22 (5):941-951. doi:10.1093/hmg/dds499

114. Korvatska O, Strand NS, Berndt JD, Strovas T, Chen DH, Leverenz JB, Kiianitsa K, Mata IF, Karakoc E, Greenup JL, Bonkowski E, Chuang J, Moon RT, Eichler EE, Nickerson DA, Zabetian CP, Kraemer BC, Bird TD, Raskind WH (2013) Altered splicing of ATP6AP2 causes X-linked parkinsonism with spasticity (XPDS). Human molecular genetics 22 (16):3259-3268. doi:10.1093/hmg/ddt180

115. Myhre R, Klungland H, Farrer MJ, Aasly JO (2008) Genetic association study of synphilin-1 in idiopathic Parkinson's disease. BMC Med Genet 9:19. doi:10.1186/1471-2350-9-19

116. Keyser RJ, Oppon E, Carr JA, Bardien S (2011) Identification of Parkinson's disease candidate genes using CAESAR and screening of MAPT and SNCAIP in South African Parkinson's disease patients. Journal of neural transmission 118 (6):889-897. doi:10.1007/s00702-011-0591-z

117. Marx FP, Holzmann C, Strauss KM, Li L, Eberhardt O, Gerhardt E, Cookson MR, Hernandez D, Farrer MJ, Kachergus J, Engelender S, Ross CA, Berger K, Schols L, Schulz JB, Riess O, Kruger R (2003) Identification and functional

characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. Human molecular genetics 12 (11):1223-1231

118. Eyal A, Szargel R, Avraham E, Liani E, Haskin J, Rott R, Engelender S (2006) Synphilin-1A: an aggregation-prone isoform of synphilin-1 that causes neuronal death and is present in aggregates from alpha-synucleinopathy patients. Proceedings of the National Academy of Sciences of the United States of America 103 (15):5917-5922. doi:10.1073/pnas.0509707103

119. Eyal A, Engelender S (2006) Synphilin isoforms and the search for a cellular model of lewy body formation in Parkinson's disease. Cell Cycle 5 (18):2082-2086

120. Szargel R, Rott R, Engelender S (2008) Synphilin-1 isoforms in Parkinson's disease: regulation by phosphorylation and ubiquitylation. Cell Mol Life Sci 65 (1):80-88. doi:10.1007/s00018-007-7343-0

121. Lei P, Ayton S, Finkelstein DI, Adlard PA, Masters CL, Bush AI (2010) Tau protein: relevance to Parkinson's disease. Int J Biochem Cell Biol 42 (11):1775-1778. doi:10.1016/j.biocel.2010.07.016

122. Spatola M, Wider C (2014) Genetics of Parkinson's disease: the yield. Parkinsonism & related disorders 20 Suppl 1:S35-38. doi:10.1016/S1353-8020(13)70011-7

123. Svobodova E, Mrazova L, Luksan O, Elstein D, Zimran A, Stolnaya L, Minks J, Eberova J, Dvorakova L, Jirsa M, Hrebicek M (2011) Glucocerebrosidase gene has an alternative upstream promoter, which has features and expression characteristic of housekeeping genes. Blood Cells Mol Dis 46 (3):239-245. doi:10.1016/j.bcmd.2010.12.011

124. Jakubauskiene E, Janaviciute V, Peciuliene I, Soderkvist P, Kanopka A (2012) G/A polymorphism in intronic sequence affects the processing of MAO-B gene in patients with Parkinson disease. FEBS Lett 586 (20):3698-3704. doi:10.1016/j.febslet.2012.08.028

125. Thomas T (2000) Monoamine oxidase-B inhibitors in the treatment of Alzheimer's disease. Neurobiology of aging 21 (2):343-348

126. Stern G (1998) Neuroprotection by selegiline and other MAO inhibitors. J Neural Transm Suppl 52:99-107

127. Ho SL, Kapadi AL, Ramsden DB, Williams AC (1995) An allelic association study of monoamine oxidase B in Parkinson's disease. Ann Neurol 37 (3):403-405. doi:10.1002/ana.410370318

128. Kurth JH, Kurth MC, Poduslo SE, Schwankhaus JD (1993) Association of a monoamine oxidase B allele with Parkinson's disease. Ann Neurol 33 (4):368-372. doi:10.1002/ana.410330406

129. Sobell JL, Lind TJ, Hebrink DD, Heston LL, Sommer SS (1997) Screening the monoamine oxidase B gene in 100 male patients with schizophrenia: a cluster of polymorphisms in African-Americans but lack of functionally significant sequence changes. Am J Med Genet 74 (1):44-49

130. Shehadeh LA, Yu K, Wang L, Guevara A, Singer C, Vance J, Papapetropoulos S (2010) SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. PLoS One 5 (2):e9104. doi:10.1371/journal.pone.0009104

131. Stamper C, Siegel A, Liang WS, Pearson JV, Stephan DA, Shill H, Connor D, Caviness JN, Sabbagh M, Beach TG, Adler CH, Dunckley T (2008) Neuronal gene expression correlates of Parkinson's disease with dementia. Mov Disord 23 (11):1588-1595. doi:10.1002/mds.22184

132. Soreq L, Bergman H, Israel Z, Soreq H (2012) Exon arrays reveal alternative splicing aberrations in Parkinson's disease leukocytes. Neurodegener Dis 10 (1-4):203-206. doi:10.1159/000332598

133. Soreq L, Bergman H, Israel Z, Soreq H (2013) Deep brain stimulation modulates nonsense-mediated RNA decay in Parkinson's patients leukocytes. BMC Genomics 14:478. doi:10.1186/1471-2164-14-478

134. Potashkin JA, Santiago JA, Ravina BM, Watts A, Leontovich AA (2012) Biosignatures for Parkinson's disease and atypical parkinsonian disorders patients. PLoS One 7 (8):e43595. doi:10.1371/journal.pone.0043595

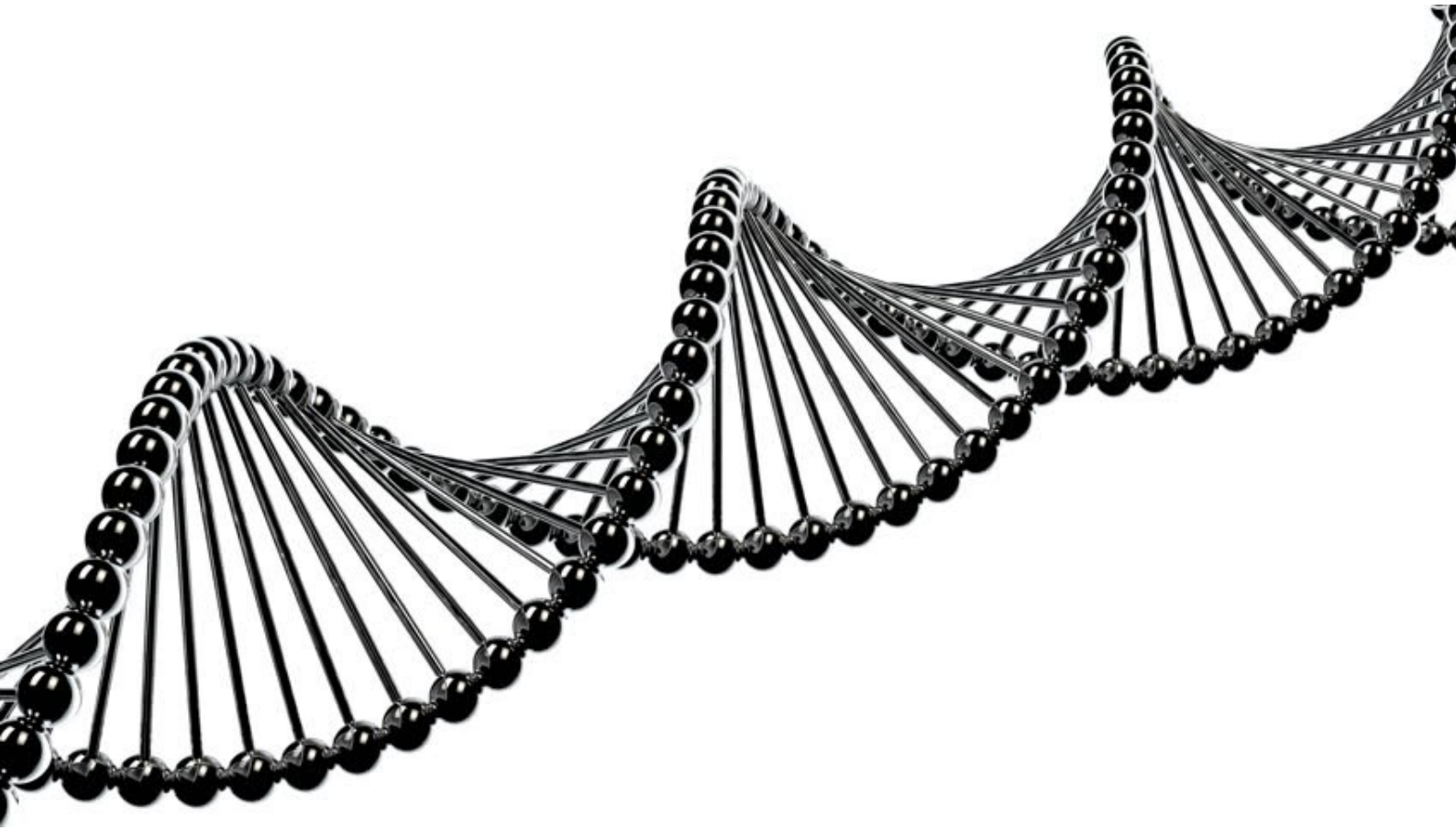
135. Santiago JA, Scherzer CR, Harvard Biomarker S, Potashkin JA (2013) Specific splice variants are associated with Parkinson's disease. Mov Disord 28 (12):1724-1727. doi:10.1002/mds.25635

136. Soreq L, Guffanti A, Salomonis N, Simchovitz A, Israel Z, Bergman H, Soreq H (2014) Long non-coding RNA and alternative splicing modulations in Parkinson's leukocytes identified by RNA sequencing. PLoS Comput Biol 10 (3):e1003517. doi:10.1371/journal.pcbi.1003517

137. Soreq L, Salomonis N, Bronstein M, Greenberg DS, Israel Z, Bergman H, Soreq H (2013) Small RNA sequencing-microarray analyses in Parkinson leukocytes reveal deep brain stimulation-induced splicing changes that classify brain region transcriptomes. Front Mol Neurosci 6:10. doi:10.3389/fnmol.2013.00010

CHAPTER 5

Increasing the coding potential of genomes through alternative splicing: the case of PARK2 gene



La Cognata V. et al., Curr Genomics. 2014 Jun;15(3):203-16. doi: 10.2174/1389202915666140426003342.

Increasing the coding potential of genomes through alternative splicing: the case of *PARK2* gene

Valentina La Cognata[1,2], Rosario Iemmolo[1], Velia D'Agata[2], Soraya Scuderi[2], Filippo Drago[3], Mario Zappia[4] and Sebastiano Cavallaro[1]

[1]Institute of Neurological Sciences, Italian National Research Council, Catania, Italy; [2]Department of Bio-Medical Sciences, Section of Anatomy and Histology, University of Catania, Italy; [3]Department of Clinical and Molecular Biomedicine, Section of Pharmacology and Biochemistry, University of Catania, Italy; [4]Department "G.F. Ingrassia", Section of Neuroscience, University of Catania, Italy

Corresponding author. Sebastiano Cavallaro, M.D., Ph.D., Functional Genomics Center, Institute of Neurological Sciences, Italian National Research Council, Via Paolo Gaifami 18, 95125, Catania, Italy. Phone: +39-095-7436634; Fax: + 39-095-7122426; E-mail: sebastiano.cavallaro@cnr.it.

ABSTRACT

The completion of the Human Genome Project aroused renewed interest in alternative splicing, an efficient and widespread mechanism that generates multiple protein isoforms from individual genes. Although our knowledge about alternative splicing is growing exponentially, its real impact in cellular life is still to be clarified. Connecting all splicing features (genes, splice transcripts, isoforms and relative functions) may be useful to resolve this tangle. Herein, we will start from the case of a single gene, Parkinson protein 2, E3 ubiquitin protein ligase (*PARK2*). This gene is one of the largest genes in the human genome. It is involved in the pathogenesis of autosomal recessive juvenile Parkinsonism and it has been recently linked to cancer, leprosy, autism, type 2 diabetes mellitus and Alzheimer disease. *PARK2* primary transcript undergoes an extensive alternative splicing, which enhances transcriptomic diversification and protein diversity in tissues and cells. This review will provide an update of all human *PARK2* alternative splice transcripts and isoforms presently known, and correlate them to those in rat and mouse, two common animal models for

studying human disease genes. Alternative splicing relies upon a complex process that could be easily altered by both *cis* and *trans-acting* mutations. Although the contribute of *PARK2* splicing in human disease remains to be fully explored, some evidences show disruption of this versatile form of genetic regulation may have pathological consequences.

KEYWORDS

Alternative splicing; mRNA; *PARK2;* Protein isoforms; Splice variants; Splice expression patterns.

INTRODUCTION

With the completion of the Human Genome Project, it came as a surprise to discover that human genome contains only a fraction of genes than originally predicted. It was clear that the small number of human genes could not account for the complexity of the proteome. Therefore, the biological paradigm “one gene to one protein” was groundless. Among several proposed mechanisms, alternative splicing is considered the major driving force for transcriptome and proteome diversity. Because of its ability in increasing the coding potential of a genome, alternative splicing represents a cheap and powerful tool that allows cells to expand their proteome, producing multiple protein products from a single gene. Although our knowledge about alternative splicing is growing exponentially, its real impact in cellular life is still under debate. Connecting each genes to its splice transcripts, corresponding isoforms and relative functions may be useful to resolve this tangle and decipher how splicing acts in physiological and pathological conditions. Indeed, this is not an easy task. In this review, we will represent the splicing features of a single gene, Parkinson protein 2, E3 ubiquitin protein ligase gene (PARK2), one of the largest genes in human genome [1].

Mutations in PARK2 gene are responsible for the development of a form of autosomal recessive juvenile Parkinsonism (AR-JP) characterized by all the classical symptoms of Parkinson disease (PD), such as bradykinesia, rigidity, and tremor [2]. In addition to AR-JP, PARK2 has been recently linked to cancer [3, 4], leprosy [5], autism [6], type 2 diabetes mellitus [7] and Alzheimer disease [8].

PARK2 gene spans more than 1.38 Mb of genomic DNA in the long arm of chromosome 6 (6q25.2-q27) [1, 9]. To date, homologous PARK2 genes have been characterized in twelve different organisms, including rat [10], mouse [11, 12], fruit fly [13], zebrafish [14] and worm [15].

The first isolated human PARK2 transcript was of 2,960 bases with an open reading frame (1,395 bases) encoding a protein of 465 amino acids [1]. Based on this transcript, the genomic organization and exon/intron boundary sequences of PARK2 consisted of 12 exons [1]. In the last fifteen years, these 12 exons have been the focus of hundreds of different screenings. The Parkinson Disease Mutation Database (http://www.molgen.vib-ua.be/PDmutDB) [16] currently lists 214 PARK2 mutations: exon rearrangements (deletions, duplications) or, more often, point mutations.

While many studies concentrated on the genetic variations present in the 12 originally established PARK2 exons and in their exon/intron boundaries, there is now enough evidence that additional exonic sequences exist in human and other species, and that they can be alternatively spliced to produce different variants [11, 17-22]. These transcripts show different patterns of expression and encode proteins with different functions [11, 17-22].

This review will provide an update of all human PARK2 alternative splice variants presently known and correlate them to those in rat and mouse, two common animal models for studying human disease genes. Before describing PARK2 splice variants, the next paragraph will briefly introduce the process of alternative splicing.

ALTERNATIVE SPLICING, BASIC CONCEPTS

A crucial regulatory stage in the pathway of gene expression is splicing of precursor mRNA (pre-mRNA). In this process, introns are removed and exons are joined to form a mature mRNA, which is then polyadenilated and exported from nucleus to cytoplasm, where it can be translated into protein. The basis of splicing is the recognition of introns and exons by a splicing machinery. A large macromolecular complex, the spliceosome, recognizes exons and removes introns while the pre-mRNA is synthesized by RNA polymerase II in the nucleus (Figure 1). The spliceosome is composed of five snRNAs (small nuclear RNAs - U1, U2, U4, U5 and U6) that assemble

with proteins to form small nuclear ribonucleoproteins (snRNPs) [23]. The four conserved sequences that enable recognition of RNA by the spliceosome are: the 5' (GU) and the 3' (AG) splice sites, the branch point located upstream of the 3' splice site and the polypyrimidine tract (PPT) located between the 3' splice site and the branch site (Figure 1). The spliceosome recognizes these elements and assembles in a tidy manner on the nascent pre-mRNA. Early in spliceosome assembly (E complex), U1 forms a base-pairing interaction with the 5'-splice site, whereas U2 base-pairs with the branch-point. Then, a tri-snRNP complex containing U4, U5 and U6 associates with the forming spliceosome (B complex) and U4 is removed from the complex. This allows U6 to replace U1 at the 5' splice site and leads to a U6–U2 interaction that brings the 5'-splice site and the branch point close together, allowing for a transesterification step. By forming non-canonical interactions, U5 brings the two exons into close proximity and allows for the second step of splicing, joining the two exons [24].

Splicing of exons does not always proceed in the same manner and different combinations of exons can be joined by a process known as alternative splicing. Alternative splicing of coding exons may generate protein isoforms with different biological properties, protein-protein interactions, subcellular localization, signaling pathway or catalytic ability [25]. Alternative splicing in non-coding sequences, instead, can affect mRNA translation efficiency, stability or localization [26]. By changing the reading frame or adding premature stop codons, some splicing events lead to truncated proteins or "Nonsense-Mediated mRNA Degradation" (NMD mechanism) [27]. This allows splicing to act as an on–off switch in gene expression. Alternative splicing is regulated in time and space, allowing a particular mRNA to be expressed in a specific cell and physiological condition [28]. This process, therefore, represents an extremely economical mean of increasing protein diversity, which can finely tune genomic information to meet the unique needs of each cell. The process of alternative splicing increases the diversity of the mRNA expressed from the genome and explains the discrepancy between the estimated 24,000 protein-coding genes in the human genome and the 100,000 different proteins that are postulated to be synthesized [29].

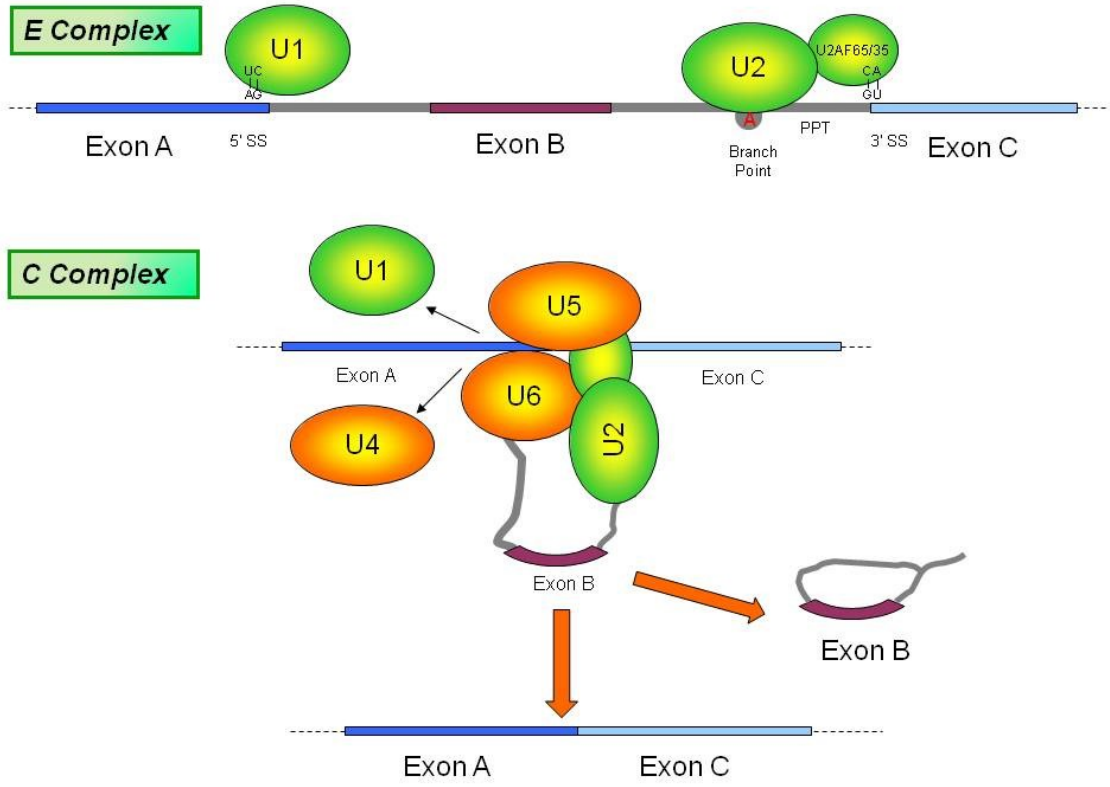


Figure 1. The alternative splicing mechanism. The spliceosome machinery (U1, U2, U4, U5 and U6) assembles on the nascent pre-mRNA. The conserved sequences that enable recognition of the mRNA by the spliceosome are: the 5' splice site (GU), the 3' splice site (AG), the branch point and the polypyrimidine tract (PPT). In E complex, U1 forms a base-pairing interaction with the 5'-splice site, whereas U2 base-pairs with the branch-point. Then, a tri-snRNP complex containing U4, U5 and U6 associates with the forming spliceosome, removing U1 and U4 (C complex). These steps allow the two transesterification reactions and join the exons.

ALTERNATIVE SPLICING OF *PARK2*

In the following paragraphs we will first describe all *PARK2* alternative splice variants presently known in human, followed by those in rat and mouse, two common animal models for studying human disease genes.

HUMAN PARK2 ALTERNATIVE SPLICE VARIANTS

PARK2 is one of the largest genes in the human genome, spanning more than 1.38 Mb of genomic DNA in the long arm of chromosome 6 (6q25.2-q27).

To date, GenBank (Unigene cluster Hs.132954) currently lists 26 human PARK2 transcripts corresponding to 21 different alternative splice variants. Each of these alternative splice variants is shown in Table 1. The alignment of genomic and transcript sequences (Figure 2) indicates these alternative splice variants are composed by 17 exons, whose exact length and specific coordinates on PARK2 gene

are indicated in Table 2. With the exception of few exons (9, 11, and 14), which are each exclusively expressed in a single splice variant, the others are not associated to specific transcripts (Figure 2). The joining of different exons does not seem to follow any specific order. Indeed, no exonic cluster (i.e. exonic sequences lying close in gene and spliced always, in or out, together) is evident.

The cDNA clone (2,960 bp) submitted by Kitada et al. (Accession number AB009973.1) [1] represents the longest transcript sequence present in GenBank, although the same repository contains three records of PARK2 Reference Sequences with a length between 3 and 4 Kbp. These RefSeq sequences have been generated by NCBI staff by assembling transcript and genomic sequences, and their existence remains uncertain (for further details see Table 1). The shortest human PARK2 transcript variant is 454 bp. This means that PARK2 undergoes a complicated pattern of splicing assembling that greatly reduces primary transcript length up to 3000 times.

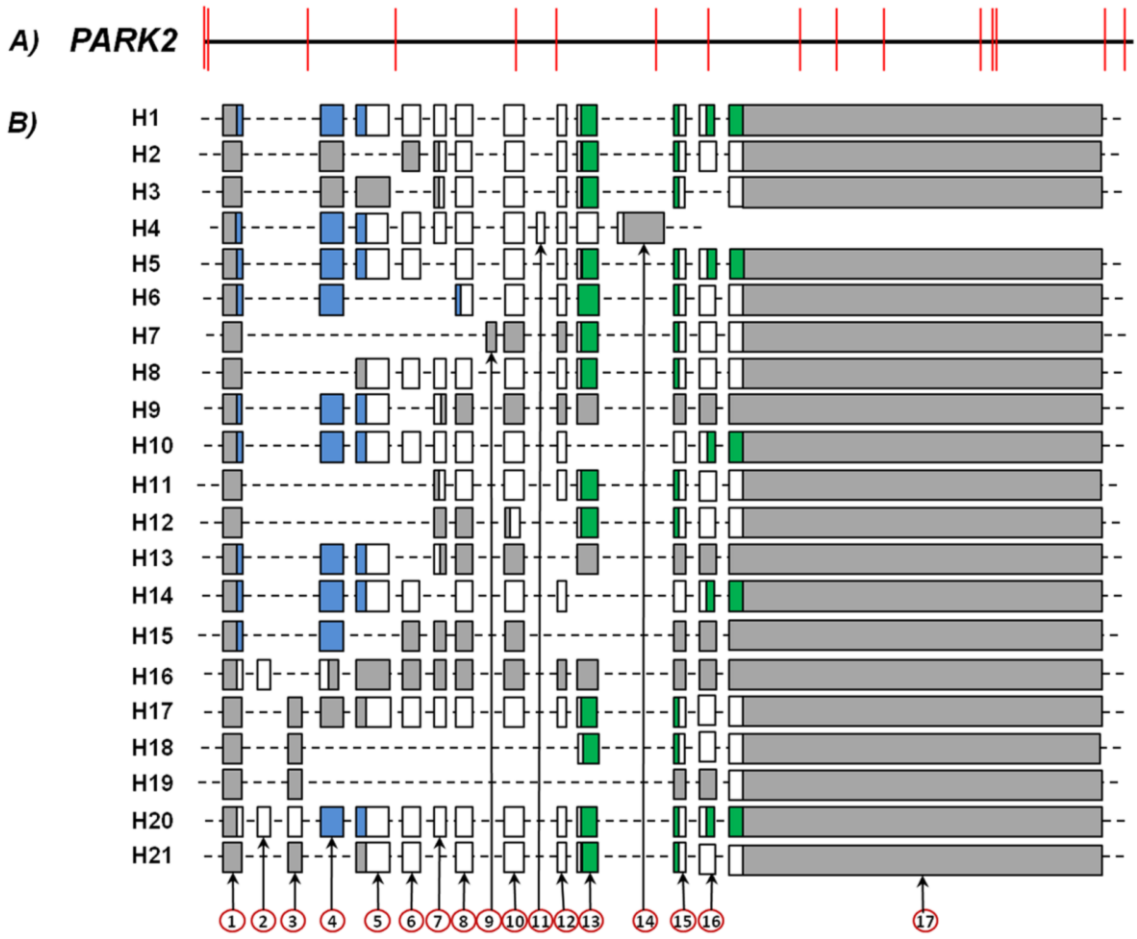


Figure 2. Human *PARK2* gene and exonic structure of alternative splice variants. **A**: Exonic and intronic organization of human *PARK2* gene. Exons are represented as red bars. The size of introns (black line) is proportional to their length. **B**: Exon organization map of the 21 human *PARK2* splice variants currently

known. Exons are represented by shaded boxes (gray for non coding sequence, white for coding sequence, blue for UBQ domain and green for IBR domains) with a size proportional to their length. The first (1) and last (17) exons are represented entirely, although their sequence is partial in some variants (H1-H5, H7-H21).

GI *PARK2* mRNAs	Code Identifier	Corresponding homologs
3063387; 121308969; 125630744; 158258616; 169790968*	**H1**	**R1 - M1**
20385797	**H2**	**R2 - M7**
20385801	**H3**	**R6**
34191069	**H4**	
169790970*; 284468410	**H5**	**R8**
169790972*	**H6**	
194378189	**H7**	
284468407	**H8**	**R14 - M5**
284468408	**H9**	
284468412	**H10**	
284516981	**H11**	
284516982	**H12**	
284516983	**H13**	
284516985	**H14**	
284516987	**H15**	
284516989	**H16**	
284516991	**H17**	
284516993	**H18**	
469609974	**H19**	
469609976	**H20**	
520845529	**H21**	

Gene identifiers corresponding to human (H) *PARK2* splice variants currently known are reported. For convenience, a new code identifier based on submission date has been assigned to each variant. In some cases, different submissions (with different Gene Identifiers) exist for the same splice variant. Three records (marked with an asterisk) are Reviewed Reference Sequences (RefSeq) that have been curated by NCBI staff, by assembling transcript and genomic sequences derived from DB023187.1, AK292590.1 and AL32982.12. These RefSeq records include a subset of the publications that are avaible for *PARK2* gene. However, we have no certainty of these full lengths, because they are not supported by direct cloning and submitted

sequences. Homologous transcripts in rat (R) and mouse (M) are reported in the table; for their Gene Identifiers the reader is referred to Tables 3 and 4.

Homo sapiens			*Rattus norvegicus*			*Mus musculus*		
Exon	**Length**	**Exon coordinates**	**Exon**	**Length**	**Exon coordinates**	**Exon**	**Length**	**Exon coordinates**
1	141	5.001-5.141	**1**	77	178.383-178.459	**1**	106	183.449-183.554
2	97	61.313-61.409						
3	98	161.506-161.603						
4	164	289.330-289.493	**2**	164	403.508-403.671	**2**	164	410.194-410.357
			3	72	427.602-427.673			
			4	154	438.685-438.838			
			5	156	503.452-503.607			
5	241	470.038-470.278	**6**	241	564.217-564.457	**3**	241	580.527-580.767
6	122	531.551-531.672	**7**	122	627.620-627.741	**4**	122	646.655-646.776
			8	237	659.191-659.427			
			9	90	673.136-673.225			
7	84	678.629-678.712	**10**	84	754.421-754.504	**5**	84	777.615-777.698
8	116	759.386-759.501	**11**	116	819.899-820.014	**6**	116	835.744-835.859
						7	587	899.713-900.299
			12	35	950.478-950.512			
9	70	895.786-895.855						
10	137	946.895-947.031	**13**	137	953.668-953.804	**8**	137	978.375-978.511
11	57	1016.672-1016.728						
12	62	1163.387-1163.448	**14**	62	1159.331-1159.392	**9**	62	1181.667-1181.728
						10	154	1188.275-1188.428
13	150	1183.800-1183.949	**15**	150	1174.875-1175.024	**11**	150	1197.792-1197.941
14	326	1187.366-1187.691						

						12	895	1209.731-1210.625
15	84	1345.926-1345.009	**16**	84	1323.441-1323.524	**13**	84	1347.110-1347.193
16	118	1372.598-1372.715	**17**	118	1356.481-1356.598	**14**	118	1393.691-1393.808
			18	8	1362.100-1362.107			
			19	172	1362.156-1362.327			
17	2654	1382.592-1385.245	**20**	209	1367.384-1367.592	**15**	1821	1404.606-1406.426

Names, coordinates on human (NG_008289.1), rat (NC_005100.3, selected region from base 49505976 to 51051947) and mouse (NC_000083.6, selected region from base 10656936 to 12246807) genes and length (bp) of *PARK2* exons are reported. For convenience, in this work, exons have been renamed consecutively and in ascending order. Homologous exons among different species are reported on the same row.

RAT Park2 ALTERNATIVE SPLICE VARIANTS

Rat Park2 gene is located in the long arm of chromosome 1 and spans more than 1.18 Mb. The cloned rat Park2 transcripts range between 1670 and 534 bp, and thus the splicing assembling strongly reduces the length of the primary transcript more than 2200 times.

GenBank (Unigene cluster Rn.207194) currently lists 27 rat Park2 transcripts (Table 3) that correspond to 20 unique alternative splice variants (Figure 3). These alternative splice variants are composed by 20 exons, which are reported in Table 2. As in human, some exons are uniquely present in a single splice variant (exons 3, 4, 5, 8, 9, 12, 18, and 19). Based on the exonic composition of transcript variants, no exonic cluster or preferential rearrangement is evident.

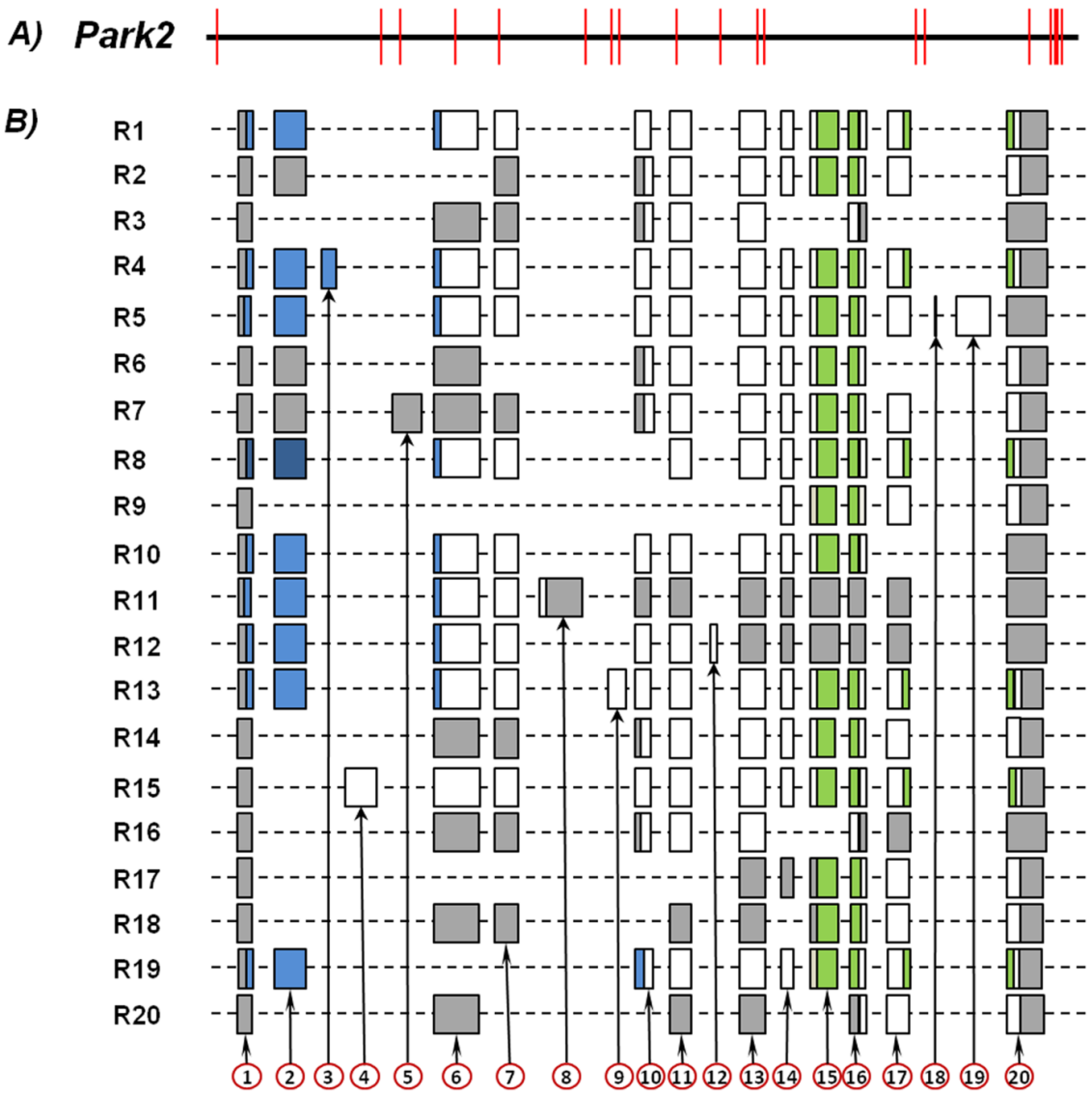


Figure 3. Rat *Park2* gene and exonic structure of alternative splice variants. **A**: Exons and introns organization of rat *Park2* gene. **B**: Exon organization map of the 20 rat *Park2* splice variants currently known. For details, see Fig. **2**. Exon 1 and exon 20 sequences are partial in some variants (R1-R13).

GI *Park2* mRNAs	Code Identifier	Corresponding homologs
7001383; 7229096; 7717034; 11464986*; 11527823	**R1**	**H1 - M1**
18478865	**R2**	**H2 - M7**
18478869	**R3**	
20385787	**R4**	
20385789	**R5**	
20385791	**R6**	**H3**
20385793; 284810436	**R7**	
20385795; 284066979	**R8**	**H5**
20385803	**R9**	
284066981	**R10**	
284468403	**R11**	
284468405	**R12**	
284810438	**R13**	
520845525; 520845527	**R14**	**H8 - M5**
520845531	**R15**	
520845533	**R16**	
520845535	**R17**	
520845537	**R18**	
520845539	**R19**	
520845541	**R20**	

Gene identifiers corresponding to rat (R) *Park2* splice variants currently known are reported. For convenience, a new code identifier based on submission date has been assigned to each variant. In some cases, different submissions (with different Gene Identifiers) exist for the same splice variant. One record (marked with an asterisk) is a Provisional Reference Sequence (RefSeq) identical to 7229096, which has not yet been subjected to NCBI final review. Homologous transcripts in human (H) and mouse (M) are reported in the table; for their Gene Identifiers the reader is referred to Tables 1 and 4.

MOUSE Park2 ALTERNATIVE SPLICE VARIANTS

Park2 splicing has been less investigated in mouse. To date, only 12 transcripts have been cloned and are collected in Unigene cluster Mm.311110 (Table 4). These transcripts arise from splicing of 15 different exons (Table 2) and correspond to 9 unique alternative splice variants (Figure 4). Among these, a splice variant (M3) is

generated by the use of an alternative 5' donor site inside exon 7 (Figure 4). Exons 10 and 12 are exclusively expressed in a single splice variant.

Mouse *Park2* gene is located in chromosome 17 where it spans 1.22 Mb. Cloned transcripts range from 3226 to 793 bp and, therefore, alternative splicing process reduces primary transcript length by about 1500 times.

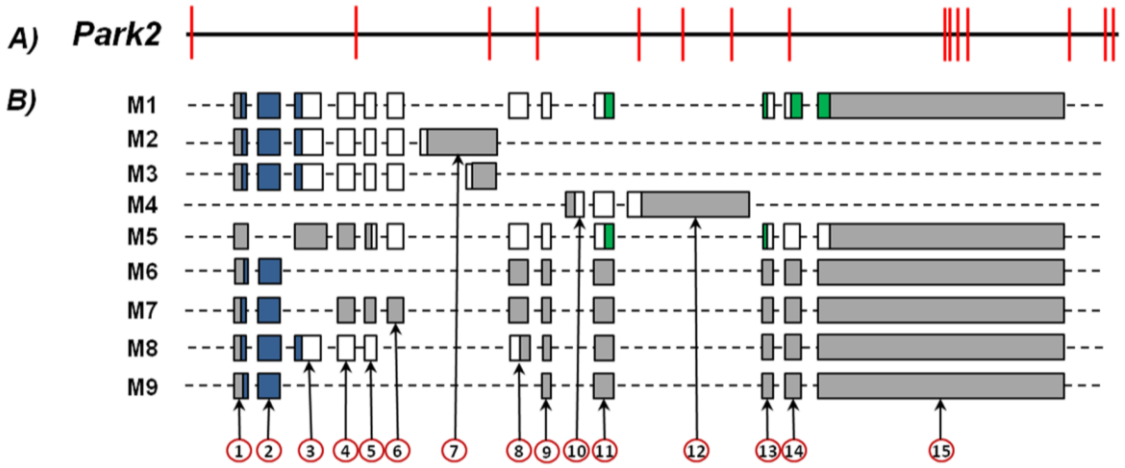


Figure 4. Mouse *Park2* gene and exonic structure of splice variants. A: Exons and introns organization of mouse *Park2* gene. **B**: Exon organization map of the 9 mouse *Park2* splice variants currently known. For details, see Fig. **2**. Exon 1 and exon 15 sequences are partial in some variants (M5-M9).

GI *Park2* mRNAs	Code Identifier	Corresponding homologs
5456929; 86577675; 10179808; 118131140*	**M1**	**H1 - R1**
10179810	**M2**	
10179812	**M3**	
74227131	**M4**	
220961631	**M5**	**H8 - R14**
220961633	**M6**	
220961635	**M7**	**H2 - R2**
220961637	**M8**	
284829878	**M9**	

Gene identifiers corresponding to mouse (M) *Park2* splice variants currently known are reported. For convenience, a new code identifier based on submission date has been assigned to each variant. In some cases, different submissions (with different Gene Identifiers) exist for the same splice variant. One records (marked with an asterisk) is a Provisional Reference Sequence (RefSeq), derived from AC105305.8, AC091254.77, AC091484.8, AC091777.26, AC093450.20, AC122259.2, CT009575.8 and AC163687.5, that has

not yet been subjected to NCBI final review. Although this record has been generated by genomic sequence alignments, it perfectly matches to 10179808. Homologous transcripts in human (H) and rat (R) are reported in the table; for their Gene Identifiers the reader is referred to Tables 1 and 3.

SPECIES-SPECIFIC ALTERNATIVE SPLICING OF *PARK2* IN HUMAN, RAT AND MOUSE

Alternative splicing is thought to be the major source of phenotypic diversity in higher eukaryotes, especially in mammals. It contributes to enhance transcriptomic diversification, and thus plays an important role in speciation and in the dynamic evolution of genome structure [30].

Inter-species comparison of *PARK2* genes is useful to identify the role of alternative cassette exons during evolution. In addition, investigating the analogies and the divergences between species may be fundamental in creating a valid animal model for Parkinson's disease. To this regard, knocking out *Park2* function in mice has been accomplished by deletion of exon 2, 3, and 8 in *Park2* gene, but no loss of nigrostriatal dopaminergic neurons has been reported [31-34].

The genomic structures of human *PARK2* gene and its homologs in rat and mouse are showed in Figure 5. Most of the exons have been conserved during evolution (e.g rat exon 2, human exon 4 and mouse exon 2; rat exon 14, human exon 12 and mouse exon 9), and their sequences have a high degree of homology (up to 95%). Differently, other exons are specie-specific (e.g. human exons 2, 3, 9, 11 and 14).

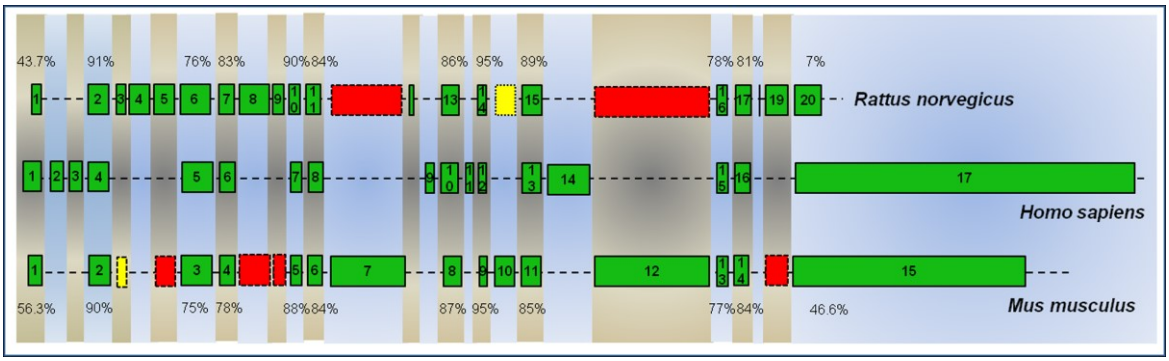


Figure 5. **Exonic structures of human, rat and mouse *PARK2* genes.**
Homologous sequences are shown in the same column. Green boxes represent the known exons of *PARK2* genes. Yellow boxes represent homologous sequences that could be potentially expressed, since they are provided with splice sites (AG/GT). Red boxes represent homologous sequences without splice sites and, therefore, not expressible. Numbers in the top of the figure indicate percentages of homology between human and rat exons, while those at the bottom denote percentages of homology between human and mouse exons.

Species-specific cassette exons may originate by two different mechanisms: i) exonization of common ancestral intronic sequences, and ii) exon shuffling, in which a new exon is inserted into an existing gene [30]. Rat and mouse specific *Park2* exons may have originated by the first mechanism (exonization). Gene-comparison, in fact, reveals that some rat-specific *Park2* exons (e.g., rat exon 3, 5, 8, 9 and 19) have a high degree of identity level with intronic sequences of mouse and *vice versa* (e.g. mouse exon 7, 10 and 12). During evolution, therefore, these sequences may have gained splice sites (AG/GT) and been expressed in one specie but not in the other one. Conversely, human-specific *PARK2* exons might have originated by exon shuffling, since their sequences are unique in human *PARK2* gene and do not match to any of the corresponding *Park2* intronic regions of rat and mouse (Figure 5).

Analyzing the exonic structures of *PARK2* alternative splice variants of the three species (Figures 2, 3 and 4), we can separate exons with a low (e.g. human exons 2, 3, 9, 11 and 14) or high inclusion level (e.g. human exons 1, 4-8, 10, 12, 13, 15-17) in splice variants. Exons with a high inclusion level in splice variants coincide with conserved exons, while those having a low inclusion level coincide with species-specific exons (Figure 5).

ALTERNATIVE SPLICING OF *PARK2* PRODUCES DIVERSITY

The combinatorial arrangement of PARK2 exons generates a large number of alternatively spliced mRNAs that may diverge for untranslated regions, half-life, regulation or degradation modality (e.g. the NMD mechanism seen above). Moreover, their different coding regions may lead the expression of distinct PARK2 protein isoforms, which are different in amino acid sequence, post-translational modifications and functional domain composition (Figure 6).

The original (canonical) PARK2 protein (Accession number BAA25751.1) [1] comprises an N-terminal ubiquitin-like (UBQ) domain and two C-terminal in-between ring fingers (IBR) domains, encoded by specific PARK2 exons (Figures 2, 3 and 4). The UBQ domain targets specific protein substrates for degradation by the proteasome, whereas IBR domains occur between pairs of ring fingers and play a role in protein quality control. PARK2 isoforms, encoded by the alternative splice transcripts currently known, structurally diverge from the canonic one for the presence or absence of the UBQ domain and for one or both IBR domains (Figure 6). Moreover,

when UBQ domain is present, it often differs in length from the canonic one. Interestingly, some isoforms miss some of these domains (Figure 6).

Alternative splicing also affects intrinsically disordered protein regions (e.g. regions lacking of stable tertiary structure), thus playing a critical role in remodeling protein-protein interactions [35]. Alternative splicing events on intrinsically disordered protein regions could regulate interactions of PARK2 isoforms with specific cellular targets. In addition, PARK2 isoforms generated by different alternative splice transcripts could interact with each other mutually regulating their functions, as it has been reported for RBCK1, a protein with IBR and E3 ubiquitin ligase domains, whose migration in the nucleus is inhibited by interaction with RBCK2, an isoform lacking of IBR domain [36].

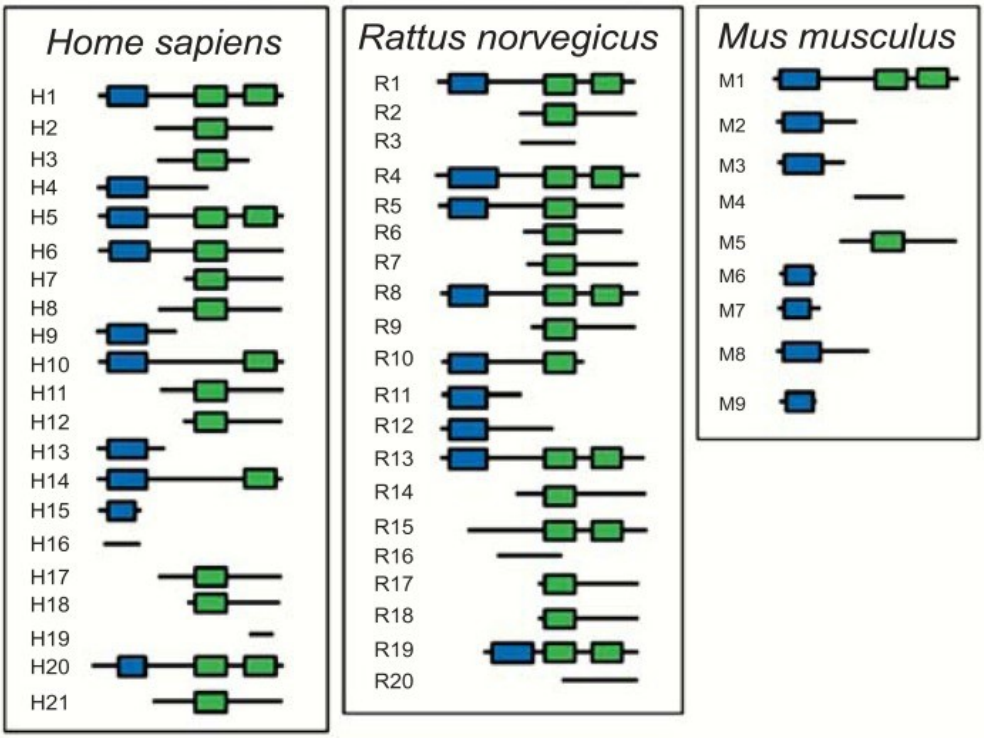


Figure 6. Predicted molecular architecture of *PARK2* isoforms.*PARK2* isoforms contain one or more of the following domains: an N-terminal Ubiquitin-like domain (UBQ, in blue) and one or two In Between Ring finger domains (IBR, in green). The code identifier of each isoform, reported on their left, corresponds to that of the encoding splice variants listed in (Tables **1**, **3** and **4**).

In addition to molecular architectures, alternative splicing may also influence stability, localization and catalytic efficiency of PARK2 isoforms. Although scientific

evidences concerning this are still few, preliminary studies reported the identification of a PARK2 isoform, missing exons 5-8, with a defective degradation activity of Cyclin E and control of cellular cycle [19]. Another study detected a splice variant of pdr-1 (a *Caenorhabditis elegans* homologous of PARK2) with an in frame deletion, characterized by altered solubility and intracellular localization [15].

Besides the well-known involvement in proteasome-dependent degradation of target proteins [37, 38], PARK2 has been implicated in apoptosis regulation [39], mitochondrial homeostasis, mitophagy [40, 41] and mitochondrial DNA stability [42]. In addition, UBQ proteins such as PARK2, are implicated in endocytosis, cellular trafficking, signal transduction, transcriptional regulation and DNA repair, in an ubiquitin degradation independent manner [43]. It may not be excluded that PARK2 alternative splice isoforms (included those missing functional domains) act each one in a different cellular context, operating in a still, not yet characterized manner.

ALTERNATIVE SPLICING OF *PARK2* IS TISSUE AND CELL SPECIFIC

Alternative splicing events are finely regulated in time and space, and thus contribute to cell specialization and tissue definition. Although we are not yet able to define the tissue and cell specific spectrum of expression of PARK2 and its homologs, the few evidences reported to date undoubtedly demonstrate that a regional and cellular differential expression of transcripts exist. One of these evidences [18], shown in Figure 7, for example, clearly shows the different expression patterns of Park2 splice variants in different rat brain areas and isolated cell types (cerebellar granule cells, cortical neurons and type I astrocytes).

Although a little number of studies has compared mRNA and protein expression, the few evidences present in the literature show that expression pattern of Park2 transcripts is mirrored also at the protein level. As shown in Figure 8 and previously reported [44], for example, Park2 isoforms are clearly differentially expressed in two cell types, rat cortical neurons and type I astrocytes. Both western blot and immunofluorescence assays demonstrate not only a quantitative difference in Park2 expression levels between these cell types, but also a different expression pattern of splice isoforms.

When the extensive PARK2 alternative splicing was still unknown, correlation between mRNA and protein expression patterns was not an easy task. The presence of unexpected immunoreactive bands on western blot, for example, was very often explained as the result of post-translational modifications [45]. Interestingly, the datasheet of many commercially available PARK2 antibodies (see below for further details) show the presence of multiple immunoreactive bands without providing sufficient explanations. Unlike the past, splicing of PARK2 can now be investigated at the protein level in more details by the use of different antibodies. To date more than 160 PARK2 antibodies are commercially available. They are generally raised from rabbit or mouse and commercialized by various companies. Table 5 list 35 commercially available PARK2 antibodies whose immunogens used are known. Some of them have been raised against the same immunogen, and thus recognize common epitopes. These 35 antibodies may allow recognizing 15 different PARK2 epitopes. Although no epitope is probably isoform specific, the combinatorial use of antibodies targeting different protein regions, together with the use of different techniques such as two dimensional gel assays, may provide a precious aid to decode the exact spectrum of PARK2 isoforms expressed in tissues and cells.

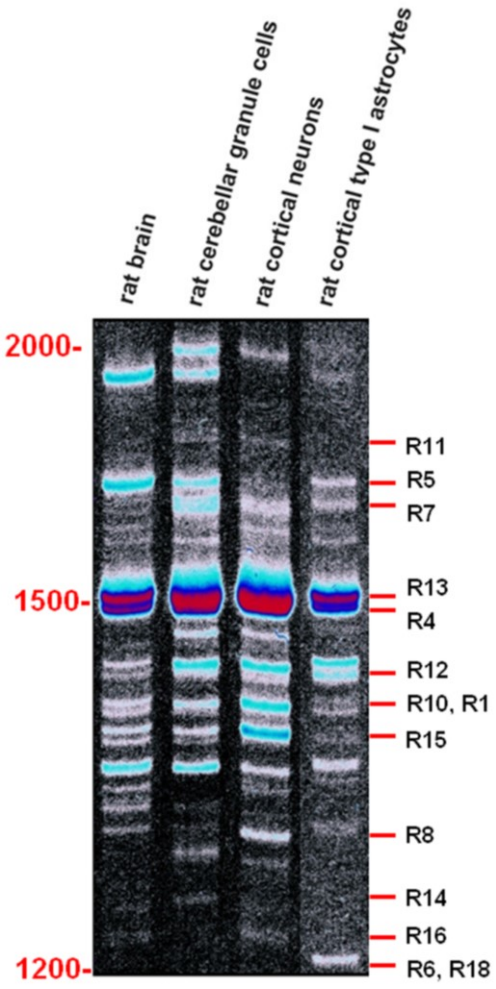


Figure 7. **Differential expression of *Park2* transcript variants in rat brain and isolated cells**

Single-stranded cDNAs from adult rat brain, rat cortical neurons, rat cerebellar granule cells and rat cortical type I astrocytes mRNAs was PCR amplified with primers flanking the start and the stop codon of *Park2*. The resulting splicing patterns clearly show a regional and cellular differential expression among different rat neuronal cells, producing spatial and functional diversification. Marker length (bp) is shown on the left. Known rat splice variants with a length between 1200 and 2000 bp are shown on the right (for codes, see Table 3). All the other products have not yet been cloned and remain unknown.

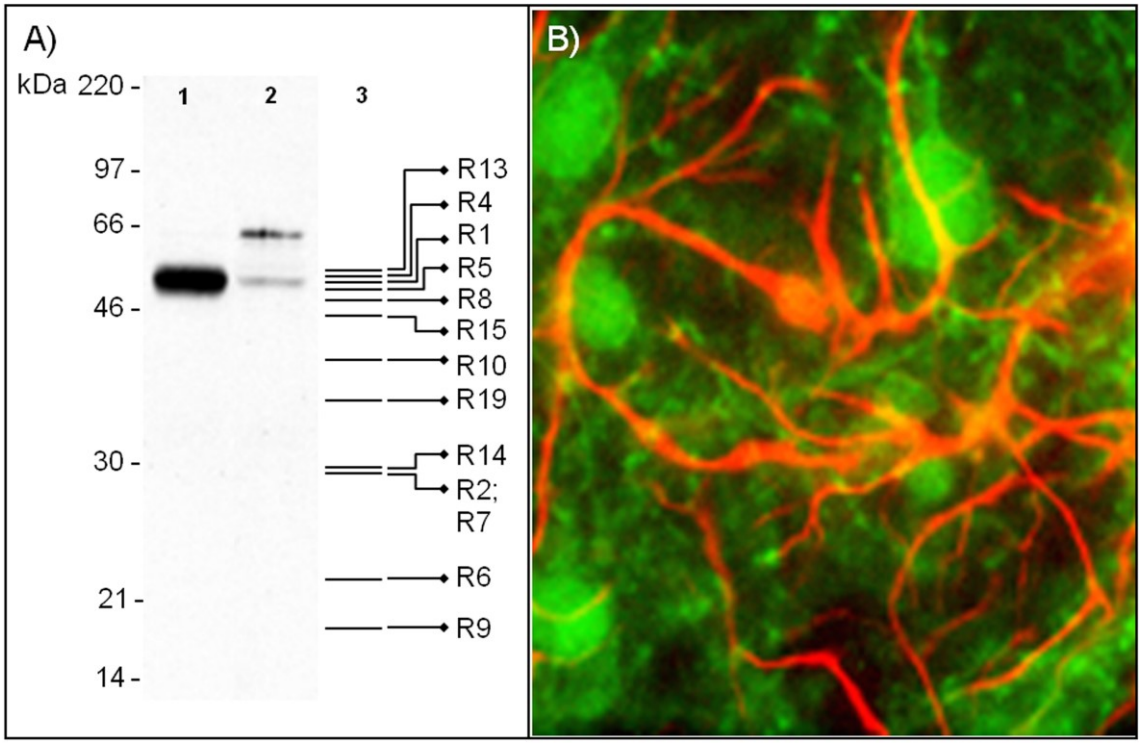


Figure 8. **Differential expression of *Park2* isoforms in rat cortical neurons and type I astrocytes**

A: Western Blot analysis of *Park2* isoforms in rat neurons (lane 1) and type I astrocytes (lane 2). The assay was performed using a rabbit anti-*Park2* polyclonal antibody (AB5112, Millipore) as previously described [44]. Expected molecular weights of Park2 isoforms potentially recognized by this antibody are drawn on lane 3. The very intense immunoreactive band of about 51 kDa in cortical neurons overlaps with the expected molecular weight of several isoforms (R1, R2, R4-R10, R13-R15, and R19). Instead, type I astrocytes express a faint band of 51 kDa which may correspond to R1 isoform, and a 65 kDa band that corresponds to a still uncharacterized variant. It should be noted that the antibody used in this western blot analysis recognizes only one epitope, which is not present in all isoforms. Other isoforms, therefore, may be expressed in these cell types and not be immunoreactive to the antibody used. **B**: *Park2* and GFAP immunoreactivity in mixed cortical cultures. Primary cultures of cortical neurons and type I astrocytes were double labeled with antibodies against *Park2* protein (green signal) and GFAP (red signal), as previously described [44], and observed by fluorescence microscopy.

Product Number	Provider	Immunogen (aa)	Recognized Isoforms
H00005071-B01P	Abnova	1-387	H4
H00005071-D01P	Abnova		
H00005071-D01	Abnova		
OASA06385	Aviva Systems biology	83-97	H1, H4, H5, H8, H9, H10, H13, H14, H17, H20, H21
AHP495	AbD Serotec		
MD-19-0144	Raybiotech		
DS-PB-01562	Raybiotech		
PAB14022	Abnova		
MCA3315Z	AbD Serotec	288-388	H4
H00005071-M01	Abnova		
PAB1105	Abnova	62-80	H1, H4, H5, H9, H10, H13, H14, H20
70R-PR059	Fitzgerald		
PAB0714	Abnova	305-323	H1-H6, H8, H11, H17, H20, H21, R1, R2, R4-R10, R13-R15, R19
AB5112	Millipore		
R-113-100	Novus biologicals		
P5748	Sigma	298-313	H1-H6, H8, H11, H17, H20, H21, R1, R2, R4-R10, R13-R15, R19, M1, M5
GTX25667	GeneTex		
ABIN122870	Antibodies online		
PA1-751	Thermo Scientific		
R-114-100	Novus biologicals	295-311	H1-H6, H8, H10, H11, H14, H17, H20, H21, R1, R2, R4-R10, R13-R15, R19, M1, M5
AB5978	Millipore		
MAB5512	Millipore	399-465	H1, H2, H5-H8, H10-H12, H14, H17-H21
05-882	Millipore		
sc-32282	Santa Cruz		
sc-30130	Santa Cruz	61-360	H1-H3, H6, H9, H11-H13
sc-133167	Santa Cruz		
sc-136989	Santa Cruz		
EB07439	Everest Biotech	394-409	H2, H6, H7, H11, H12, H18
GTX89242	Gene Tex		
NB100-53798	Novus biologicals		
GTX113239	GeneTex	28-258	H1
10R-3061	Fitzgerald	390-406	H1, H2, H5-H8, H10-H12, H14, H17, H18, H20, H21

A01250-40	GenScript	300-350	H1-H6, H8, H11, H17, H20, H21
NB600-1540	Novus biologicals	399-412	H1, H2, H5-H7, H10-H12, H18, H20
ARP43038_P050	Aviva Systems biology	311-360	H2, H3, H6, H7, H11, H12, H18, M1, M5

Table 5. List of commercially available antibodies against human, rat and mouse PARK2. Some PARK2 antibodies have been raised against the same immunogen. Amino acids positions refer to PARK2 isoform NP_004553.2. Human (H), rat (R), and mouse (M) recognized isoforms are indicated in the right column.

ALTERNATIVE SPLICING OF *PARK2* AND PATHOLOGY

Alternative splicing process is a key element in *PARK2* gene expression and could be easily disrupted through multiple errors. An aberrant alternative splicing may arise from changes of regulatory sequences required for correct pre-mRNA processing, such as splice sites, branch point, polypyrimidine tract, exonic splicing enhancers/silencers and intronic enhancers/silencers, which are called *cis-acting* mutations. To this regard, a number of *PARK2 cis-acting* mutations identified in patients with Parkinson's disease have been collected in Parkinson Disease Mutations database. To date, point mutations localized in splice sites of *PARK2* introns 1, 6, 7, 10, 12, 13 and 16 have been investigated [46-52], while c*is-acting* mutations in splice sites of exons 2-5, 8, 9, 11, 14, 15, 17 and in the other splicing regulatory regions have not yet been explored. Moreover, deregulation of alternative splicing may be the result of changes in the components of spliceosome machinery required for splicing regulation (*trans-acting* mutations). Further studies are needed to discover their possible pathogenic influence on *PARK2* alternative splicing, and to elucidate the exact splice pattern/phenotype correlations. Disruption of *PARK2* alternative splicing by both *cis-* and *trans-acting* mutations, in fact, could result in a functionally harmful *PARK2* expression pattern, creating aberrant events with pathological consequences that may provide an explanation for the broad spectrum of phenotypic abnormalities observed in patients with *PARK2* mutations [53-55].

Both *cis-* and *trans-acting* mutations have been already associated to human diseases. For example, *cis-acting* mutations have been found in genes involved in Alzheimer Disease (*PSEN1*, *MAPT*, *GRN*) [56], while *trans-acting* mutations have been detected in Retinitis pigmentosa (U4/U5/U6 protein complex) [57], in Spinal Muscle

Atrophy [58], and in Distrophia Myotonica [59]. In addition to *PARK2*, five other genes (*SNCAIP, LRRK2, SNCA, SRRM2, MAPT*) associated with Parkinsonism are subjected to extensive splice processing [60]. Therefore, the hypothesis that aberrant alternative splicing regulation may play a crucial role in human diseases strongly seems to take hold.

CONCLUSIONS

PARK2 gene provides a fascinating example of the use of alternative splicing to create different variants within single cell types. Disruption of this versatile form of genetic regulation may alter the fine-tuning of the encoded proteins to suit specific cellular needs. Investigating the full spectrum of *PARK2* alternative spliced mRNAs, studying the complete pattern of expression of alternative splicing transcripts and isoforms in tissue and cell-types, and determining each of their functions will require additional studies. Furthermore, investigating *cis*- and *trans-acting* mutations may provide novel insights into the pathogenesis of human diseases associated to *PARK2* and their cure.

REFERENCES

[1] 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

[2] Takahashi, H., Ohama, E., Suzuki, S., Horikawa, Y., Ishikawa, A., Morita, T., Tsuji, S. & Ikuta, F. Familial juvenile parkinsonism: clinical and pathologic study in a family, *Neurology,* 1994, *44,* 437-441

[3] Cesari, R., Martin, E.S., Calin, G.A., Pentimalli, F., Bichi, R., McAdams, H., Trapasso, F., Drusco, A., Shimizu, M., Masciullo, V., D'Andrilli, G., Scambia, G., Picchio, M.C., Alder, H., Godwin, A.K. & Croce, C.M. Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27, *Proceedings of the National Academy of Sciences of the United States of America,* 2003, *100,* 5956-5961

[4] Veeriah, S., Taylor, B.S., Meng, S., Fang, F., Yilmaz, E., Vivanco, I., Janakiraman, M., Schultz, N., Hanrahan, A.J., Pao, W., Ladanyi, M., Sander, C., Heguy, A., Holland, E.C., Paty, P.B., Mischel, P.S., Liau, L., Cloughesy, T.F., Mellinghoff, I.K., Solit, D.B. & Chan, T.A. Somatic mutations of the Parkinson's disease-associated gene PARK2 in glioblastoma and other human malignancies, *Nat Genet,* 2010, *42,* 77-82

[5] Mira, M.T., Alcais, A., Nguyen, V.T., Moraes, M.O., Di Flumeri, C., Vu, H.T., Mai, C.P., Nguyen, T.H., Nguyen, N.B., Pham, X.K., Sarno, E.N., Alter, A., Montpetit, A., Moraes, M.E., Moraes, J.R., Dore, C., Gallant, C.J., Lepage, P., Verner, A., Van De Vosse, E., Hudson, T.J., Abel, L. & Schurr, E.

Susceptibility to leprosy is associated with PARK2 and PACRG, *Nature,* 2004, *427,* 636-640

[6] Glessner, J.T., Wang, K., Cai, G., Korvatska, O., Kim, C.E., Wood, S., Zhang, H., Estes, A., Brune, C.W., Bradfield, J.P., Imielinski, M., Frackelton, E.C., Reichert, J., Crawford, E.L., Munson, J., Sleiman, P.M., Chiavacci, R., Annaiah, K., Thomas, K., Hou, C., Glaberson, W., Flory, J., Otieno, F., Garris, M., Soorya, L., Klei, L., Piven, J., Meyer, K.J., Anagnostou, E., Sakurai, T., Game, R.M., Rudd, D.S., Zurawiecki, D., McDougle, C.J., Davis, L.K., Miller, J., Posey, D.J., Michaels, S., Kolevzon, A., Silverman, J.M., Bernier, R., Levy, S.E., Schultz, R.T., Dawson, G., Owley, T., McMahon, W.M., Wassink, T.H., Sweeney, J.A., Nurnberger, J.I., Coon, H., Sutcliffe, J.S., Minshew, N.J., Grant, S.F., Bucan, M., Cook, E.H., Buxbaum, J.D., Devlin, B., Schellenberg, G.D. & Hakonarson, H. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes, *Nature,* 2009, *459,* 569-573

[7] Wongseree, W., Assawamakin, A., Piroonratana, T., Sinsomros, S., Limwongse, C. & Chaiyaratana, N. Detecting purely epistatic multi-locus interactions by an omnibus permutation test on ensembles of two-locus analyses, *BMC Bioinformatics,* 2009, *10,* 294

[8] Burns, M.P., Zhang, L., Rebeck, G.W., Querfurth, H.W. & Moussa, C.E. Parkin promotes intracellular Abeta1-42 clearance, *Human molecular genetics,* 2009, *18,* 3206-3216

[9] Matsumine, H., Saito, M., Shimoda-Matsubayashi, S., Tanaka, H., Ishikawa, A., Nakagawa-Hattori, Y., Yokochi, M., Kobayashi, T., Igarashi, S., Takano, H., Sanpei, K., Koike, R., Mori, H., Kondo, T., Mizutani, Y., Schaffer, A.A., Yamamura, Y., Nakamura, S., Kuzuhara, S., Tsuji, S. & Mizuno, Y. Localization of a gene for an autosomal recessive form of juvenile Parkinsonism to chromosome 6q25.2-27, *Am J Hum Genet,* 1997, *60,* 588-596

[10] D'Agata, V., Zhao, W. & Cavallaro, S. Cloning and distribution of the rat parkin mRNA, *Brain Res Mol Brain Res,* 2000, *75,* 345-349

[11] Kitada, T., Asakawa, S., Minoshima, S., Mizuno, Y. & Shimizu, N. Molecular cloning, gene expression, and identification of a splicing variant of the mouse parkin gene, *Mamm Genome,* 2000, *11,* 417-421

[12] Tomac, A.C. & Hoffer, B.J. Assignment of the mouse Park2 (PARKIN), the homologue to a new human Parkinson candidate gene, to the telomeric region of mouse 17A3.2-3.3, by in situ hybridization, *Cytogenet Cell Genet,* 2001, *95,* 120-121

[13] Bae, Y.J., Park, K.S. & Kang, S.J. Genomic organization and expression of parkin in Drosophila melanogaster, *Exp Mol Med,* 2003, *35,* 393-402

[14] Fett, M.E., Pilsl, A., Paquet, D., van Bebber, F., Haass, C., Tatzelt, J., Schmid, B. & Winklhofer, K.F. Parkin is protective against proteotoxic stress in a transgenic zebrafish model, *PloS one,* 2010, *5,* e11783

[15] Springer, W., Hoppe, T., Schmidt, E. & Baumeister, R. A Caenorhabditis elegans Parkin mutant with altered solubility couples alpha-synuclein aggregation to proteotoxic stress, *Human molecular genetics,* 2005, *14,* 3407-3423

[16] Cruts, M., Theuns, J. & Van Broeckhoven, C. Locus-specific mutation databases for neurodegenerative brain diseases, *Human mutation,* 2012, *33,* 1340-1344

[17] Sunada, Y., Saito, F., Matsumura, K. & Shimizu, T. Differential expression of the parkin gene in the human brain and peripheral leukocytes, *Neurosci Lett,* 1998, *254,* 180-182

[18] Dagata, V. & Cavallaro, S. Parkin transcript variants in rat and human brain, *Neurochem Res,* 2004, *29,* 1715-1724

[19] Ikeuchi, K., Marusawa, H., Fujiwara, M., Matsumoto, Y., Endo, Y., Watanabe, T., Iwai, A., Sakai, Y., Takahashi, R. & Chiba, T. Attenuation of proteolysis-mediated cyclin E regulation by alternatively spliced Parkin in human colorectal cancers, *International journal of cancer. Journal international du cancer,* 2009, *125,* 2029-2035

[20] Beyer, K., Domingo-Sabat, M., Humbert, J., Carrato, C., Ferrer, I. & Ariza, A. Differential expression of alpha-synuclein, parkin, and synphilin-1 isoforms in Lewy body disease, *Neurogenetics,* 2008, *9,* 163-172

[21] Humbert, J., Beyer, K., Carrato, C., Mate, J.L., Ferrer, I. & Ariza, A. Parkin and synphilin-1 isoform expression changes in Lewy body diseases, *Neurobiol Dis,* 2007, *26,* 681-687

[22] Tan, E.K., Shen, H., Tan, J.M., Lim, K.L., Fook-Chong, S., Hu, W.P., Paterson, M.C., Chandran, V.R., Yew, K., Tan, C., Yuen, Y., Pavanni, R., Wong, M.C., Puvan, K. & Zhao, Y. Differential expression of splice variant and wild-type parkin in sporadic

Parkinson's disease, *Neurogenetics,* 2005, *6,* 179-184

[23] Behzadnia, N., Golas, M.M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C.L., Urlaub, H., Stark, H. & Luhrmann, R. Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes, *The EMBO journal,* 2007, *26,* 1737-1748

[24] Tazi, J., Bakkour, N. & Stamm, S. Alternative splicing and disease, *Biochimica et biophysica acta,* 2009, *1792,* 14-26

[25] Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A. & Soreq, H. Function of alternative splicing, *Gene,* 2005, *344,* 1-20

[26] Hughes, T.A. Regulation of gene expression by alternative untranslated regions, *Trends Genet,* 2006, *22,* 119-122

[27] Lejeune, F. & Maquat, L.E. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells, *Curr Opin Cell Biol,* 2005, *17,* 309-315

[28] Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M. & Stamm, S. Function of alternative splicing, *Gene,* 2012,

[29] Modrek, B. & Lee, C. A genomic view of alternative splicing, *Nat Genet,* 2002, *30,* 13-19

[30] Lev-Maor, G., Goren, A., Sela, N., Kim, E., Keren, H., Doron-Faigenboim, A., Leibman-Barak, S., Pupko, T. & Ast, G. The "alternative" choice of constitutive exons throughout evolution, *PLoS genetics,* 2007, *3,* e203

[31] 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., Gajendiran, M., Roth, B.L., Chesselet, M.F., Maidment, N.T., Levine, M.S. & Shen, J. Parkin-deficient mice exhibit nigrostriatal deficits but not loss of dopaminergic neurons, *The Journal of biological chemistry,* 2003, *278,* 43628-43635

[32] Perez, F.A. & Palmiter, R.D. Parkin-deficient mice are not a robust model of parkinsonism, *Proceedings of the National Academy of Sciences of the United States of America,* 2005, *102,* 2174-2179

[33] Itier, J.M., Ibanez, P., Mena, M.A., Abbas, N., Cohen-Salmon, C., Bohme, G.A., Laville, M., Pratt, J., Corti, O., Pradier, L., Ret, G., Joubert, C., Periquet, M., Araujo, F., Negroni, J., Casarejos, M.J., Canals, S., Solano, R., Serrano, A., Gallego, E., Sanchez, M., Denefle, P., Benavides, J., Tremp, G., Rooney, T.A., Brice, A. & Garcia de Yebenes, J. Parkin gene inactivation alters behaviour and dopamine neurotransmission in the mouse, *Human molecular genetics,* 2003, *12,* 2277-2291

[34] Von Coelln, R., Thomas, B., Savitt, J.M., Lim, K.L., Sasaki, M., Hess, E.J., Dawson, V.L. & Dawson, T.M. Loss of locus coeruleus neurons and reduced startle in parkin null mice, *Proceedings of the National Academy of Sciences of the United States of America,* 2004, *101,* 10744-10749

[35] Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C.M., Wilson, M.D., Kim, P.M., Odom, D.T., Frey, B.J. & Blencowe, B.J. The evolutionary landscape of alternative splicing in vertebrate species, *Science,* 2012, *338,* 1587-1593

[36] Tatematsu, K., Yoshimoto, N., Okajima, T., Tanizawa, K. & Kuroda, S. Identification of ubiquitin ligase activity of RBCK1 and its inhibition by splice variant RBCK2 and protein kinase Cbeta, *The Journal of biological chemistry,* 2008, *283,* 11575-11585

[37] Jiang, H., Ren, Y., Zhao, J. & Feng, J. Parkin protects human dopaminergic neuroblastoma cells against dopamine-induced apoptosis, *Human molecular genetics,* 2004, *13,* 1745-1754

[38] Huynh, D.P., Scoles, D.R., Nguyen, D. & Pulst, S.M. The autosomal recessive juvenile Parkinson disease gene product, parkin, interacts with and ubiquitinates synaptotagmin XI, *Human molecular genetics,* 2003, *12,* 2587-2597

[39] da Costa, C.A., Sunyach, C., Giaime, E., West, A., Corti, O., Brice, A., Safe, S., Abou-Sleiman, P.M., Wood, N.W., Takahashi, H., Goldberg, M.S., Shen, J. & Checler, F. Transcriptional repression of p53 by parkin and impairment by mutations associated with autosomal recessive juvenile Parkinson's disease, *Nature cell biology,* 2009, *11,* 1370-1375

[40] Yoshii, S.R., Kishi, C., Ishihara, N. & Mizushima, N. Parkin mediates proteasome-dependent protein degradation and rupture of the outer mitochondrial membrane, *The Journal of biological chemistry,* 2011, *286,* 19630-19640

[41] Chen, Y. & Dorn, G.W., 2nd. PINK1-phosphorylated mitofusin 2 is a Parkin receptor for

culling damaged mitochondria, *Science,* 2013, *340,* 471-475

[42] Rothfuss, O., Fischer, H., Hasegawa, T., Maisel, M., Leitner, P., Miesel, F., Sharma, M., Bornemann, A., Berg, D., Gasser, T. & Patenge, N. Parkin protects mitochondrial genome integrity and supports mitochondrial DNA repair, *Human molecular genetics,* 2009, *18,* 3832-3850

[43] Winklhofer, K.F., Tatzelt, J. & Haass, C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases, *The EMBO journal,* 2008, *27,* 336-349

[44] D'Agata, V., Grimaldi, M., Pascale, A. & Cavallaro, S. Regional and cellular expression of the parkin gene in the rat cerebral cortex, *The European journal of neuroscience,* 2000, *12,* 3583-3588

[45] Horowitz, J.M., Myers, J., Stachowiak, M.K. & Torres, G. Identification and distribution of Parkin in rat brain, *Neuroreport,* 1999, *10,* 3393-3397

[46] Illarioshkin, S.N., Periquet, M., Rawal, N., Lucking, C.B., Zagorovskaya, T.B., Slominsky, P.A., Miloserdova, O.V., Markova, E.D., Limborska, S.A., Ivanova-Smolenskaya, I.A. & Brice, A. Mutation analysis of the parkin gene in Russian families with autosomal recessive juvenile parkinsonism, *Movement disorders : official journal of the Movement Disorder Society,* 2003, *18,* 914-919

[47] Pigullo, S., De Luca, A., Barone, P., Marchese, R., Bellone, E., Colosimo, A., Scaglione, C., Martinelli, P., Di Maria, E., Pizzuti, A., Abbruzzese, G., Dallapiccola, B., Ajmar, F. & Mandich, P. Mutational analysis of parkin gene by denaturing high-performance liquid chromatography (DHPLC) in essential tremor, *Parkinsonism & related disorders,* 2004, *10,* 357-362

[48] Scherfler, C., Khan, N.L., Pavese, N., Eunson, L., Graham, E., Lees, A.J., Quinn, N.P., Wood, N.W., Brooks, D.J. & Piccini, P.P. Striatal and cortical pre- and postsynaptic dopaminergic dysfunction in sporadic parkin-linked parkinsonism, *Brain : a journal of neurology,* 2004, *127,* 1332-1342

[49] Bertoli-Avella, A.M., Giroud-Benitez, J.L., Akyol, A., Barbosa, E., Schaap, O., van der Linde, H.C., Martignoni, E., Lopiano, L., Lamberti, P., Fincati, E., Antonini, A., Stocchi, F., Montagna, P., Squitieri, F., Marini, P., Abbruzzese, G., Fabbrini, G., Marconi, R., Dalla Libera, A., Trianni, G., Guidi, M., De Gaetano, A., Boff Maegawa, G., De Leo, A., Gallai, V., de Rosa, G., Vanacore, N., Meco, G., van Duijn, C.M., Oostra, B.A., Heutink, P., Bonifati, V. & Italian Parkinson Genetics, N. Novel parkin mutations detected in patients with early-onset Parkinson's disease, *Movement disorders : official journal of the Movement Disorder Society,* 2005, *20,* 424-431

[50] Bardien, S., Keyser, R., Yako, Y., Lombard, D. & Carr, J. Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease, *Parkinsonism & related disorders,* 2009, *15,* 116-121

[51] Nuytemans, K., Meeus, B., Crosiers, D., Brouwers, N., Goossens, D., Engelborghs, S., Pals, P., Pickut, B., Van den Broeck, M., Corsmit, E., Cras, P., De Deyn, P.P., Del-Favero, J., Van Broeckhoven, C. & Theuns, J. Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population, *Human mutation,* 2009, *30,* 1054-1061

[52] Pankratz, N., Kissell, D.K., Pauciulo, M.W., Halter, C.A., Rudolph, A., Pfeiffer, R.F., Marder, K.S., Foroud, T., Nichols, W.C. & Parkinson Study Group, P.I. Parkin dosage mutations have greater pathogenicity in familial PD than simple sequence mutations, *Neurology,* 2009, *73,* 279-286

[53] Abbas, N., Lucking, C.B., Ricard, S., Durr, A., Bonifati, V., De Michele, G., Bouley, S., Vaughan, J.R., Gasser, T., Marconi, R., Broussolle, E., Brefel-Courbon, C., Harhangi, B.S., Oostra, B.A., Fabrizio, E., Bohme, G.A., Pradier, L., Wood, N.W., Filla, A., Meco, G., Denefle, P., Agid, Y. & Brice, A. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease, *Human molecular genetics,* 1999, *8,* 567-574

[54] Hattori, N., Kitada, T., Matsumine, H., Asakawa, S., Yamamura, Y., Yoshino, H., Kobayashi, T., Yokochi, M., Wang, M., Yoritaka, A., Kondo, T., Kuzuhara, S., Nakamura, S., Shimizu, N. & Mizuno, Y. Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals, *Annals of neurology,* 1998, *44,* 935-941

[55] Lucking, C.B., Durr, A., Bonifati, V., Vaughan, J., De Michele, G., Gasser, T., Harhangi, B.S., Meco, G., Denefle, P., Wood, N.W., Agid, Y., Brice, A., French Parkinson's Disease Genetics Study, G. &

European Consortium on Genetic Susceptibility in Parkinson's, D. Association between early-onset Parkinson's disease and mutations in the parkin gene, *The New England journal of medicine,* 2000, *342,* 1560-1567

[56] Mills, J.D. & Janitz, M. Alternative splicing of mRNA in the molecular pathology of neurodegenerative diseases, *Neurobiology of aging,* 2012, *33,* 1012 e1011-1024

[57] Mordes, D., Luo, X., Kar, A., Kuo, D., Xu, L., Fushimi, K., Yu, G., Sternberg, P., Jr. & Wu, J.Y. Pre-mRNA splicing and retinitis pigmentosa, *Molecular vision,* 2006, *12,* 1259-1271

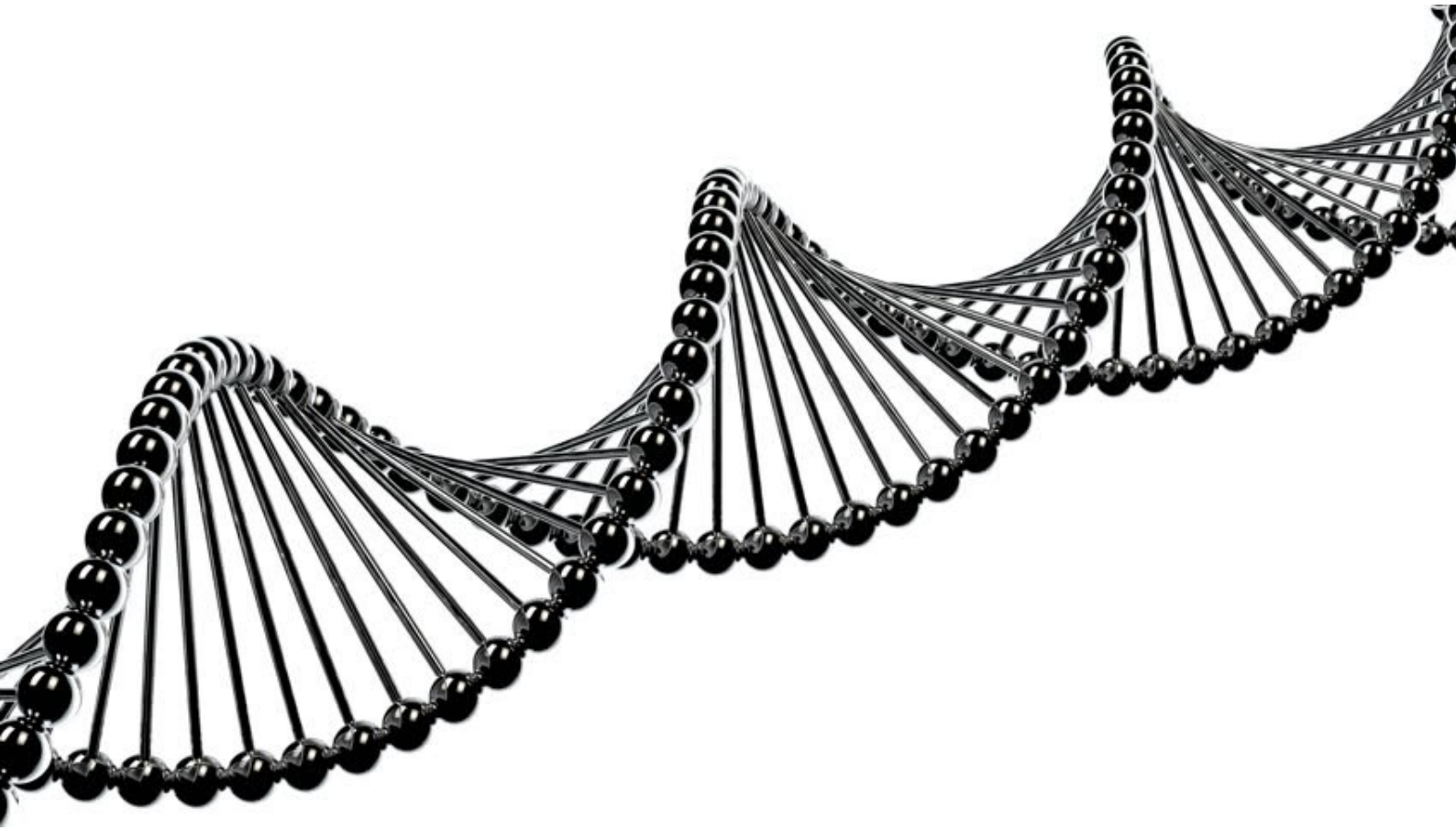
[58] Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M. & et al. Identification and characterization of a spinal muscular atrophy-determining gene, *Cell,* 1995, *80,* 155-165

[59] Faustino, N.A. & Cooper, T.A. Pre-mRNA splicing and human disease, *Genes & development,* 2003, *17,* 419-437

[60] Fu, R.H., Liu, S.P., Huang, S.J., Chen, H.J., Chen, P.R., Lin, Y.H., Ho, Y.C., Chang, W.L., Tsai, C.H., Shyu, W.C. & Lin, S.Z. Aberrant alternative splicing events in Parkinson's disease, *Cell transplantation,* 2013, *22,* 653-661

CHAPTER 6

Alternative splicing generates different parkin protein isoforms: evidences in human, rat and mouse brain



Scuderi S. et al., Biomed Res Int. 2014;2014:690796. doi: 10.1155/2014/690796

Alternative splicing generates different parkin protein isoforms: evidences in human, rat and mouse brain

Soraya Scuderi[1], Valentina La Cognata[1,2], Filippo Drago[3], Sebastiano Cavallaro[2], Velia D'Agata[1*]

[1]Department of Bio-Medical Sciences, Section of Anatomy and Histology, University of Catania; [2]Functional Genomics Center, Institute of Neurological Sciences, Italian National Research Council, Via Paolo Gaifami, 18, Catania 95125, Italy; [3]Department of Clinical and Molecular Biomedicine, Section of Pharmacology and Biochemistry, University of Catania.

Corresponding Author: Velia D'Agata, M.D. Ph.D., Department of Bio-Medical Sciences, Via S. Sofia, 87, 95123 Catania, Tel: +39-095-3782147, Fax: +39-095-3782046, e-mail: vdagata@unict.it

ABSTRACT

Parkinson protein 2, E3 ubiquitin protein ligase (PARK2) gene mutations are the most frequent causes of autosomal recessive early onset Parkinson's disease and juvenile Parkinson disease. Parkin deficiency has also been linked to other human pathologies, for example, sporadic Parkinson disease, Alzheimer disease, autism, and cancer. PARK2 primary transcript undergoes an extensive alternative splicing, which enhances transcriptomic diversification. To date several PARK2 splice variants have been identified; however, the expression and distribution of parkin isoforms have not been deeply investigated yet. Here, the currently known PARK2 gene transcripts and relative predicted encoded proteins in human, rat, and mouse are reviewed. By analyzing the literature, we highlight the existing data showing the presence of multiple parkin isoforms in the brain. Their expression emerges from conflicting results regarding the electrophoretic mobility of the protein, but it is also assumed from discrepant observations on the cellular and tissue distribution of parkin. Although the characterization of each predicted isoforms is complex, since they often diverge only for few amino acids, analysis of their expression patterns in the brain might account for the different pathogenetic effects linked to PARK2 gene mutations.

INTRODUCTION

Homozygous or compound heterozygous mutations of Parkinson protein 2, E3 ubiquitin protein ligase (PARK2) gene are cause (50% of cases) of autosomal recessive forms of PD, usually without atypical clinical features. PARK2 mutations also explain ~15% of the sporadic cases with onset before 45 [1, 2] and act as susceptibility alleles for late-onset forms of Parkinson disease (2% of cases) [3]. Along with Parkinsonism forms, PARK2 gene has been linked to other human pathologies, such as Alzheimer disease [4], autism [5], multiple sclerosis [6], cancer [7, 8], leprosy [9], type 2 diabetes mellitus [10], and myositis [11].

PARK2 gene is located in the long arm of chromosome 6 (6q25.2-q27) and spans more than 1.38 Mb [12, 13]. From the cloning of the first human cDNA [12, 13], PARK2 genomic organization was thought to include only 12 exons encoding one transcript. Many evidences now demonstrate the existence of additional exonic sequences, which can be alternatively included or skipped in mature mRNAs. To date, dozens of PARK2 splice transcripts have been described [14] and have been demonstrated to be differentially expressed in tissue and cells [15–21]. These multiple PARK2 splice variants potentially encode for a wide range of distinct protein isoforms with different structures and molecular architectures. However, the characterization and the distribution of these isoforms have not been deeply detailed yet. While studying PARK2 splice variants mRNAs is relatively simple, differentiating protein isoforms is more complex, since they often diverge only for few amino acids. The complexity of this task could explain the small number of scientific papers on this topic. However, solving this riddle is fundamental to comprehend the precise role of PARK2 in human diseases. The tissue and cell specific expression pattern of PARK2 isoforms, in fact, might account for the different pathogenetic effects linked to this gene.

In this review, we briefly describe the structure of PARK2 gene, its currently known transcript products, and the predicted encoded protein isoforms expressed in human, rat and mouse; the latter are two commonly used animal models for studying human diseases. Then, we illustrate the expression of these isoforms by recapitulating the major literature evidences already available, which have previously unknowingly demonstrated their existence. We focus on the expression and cellular distribution of parkin isoforms in the brain. Finally, we collect in a panel the different parkin

antibodies, commercially available, which could be useful for the characterization of the isoforms expression and distribution.

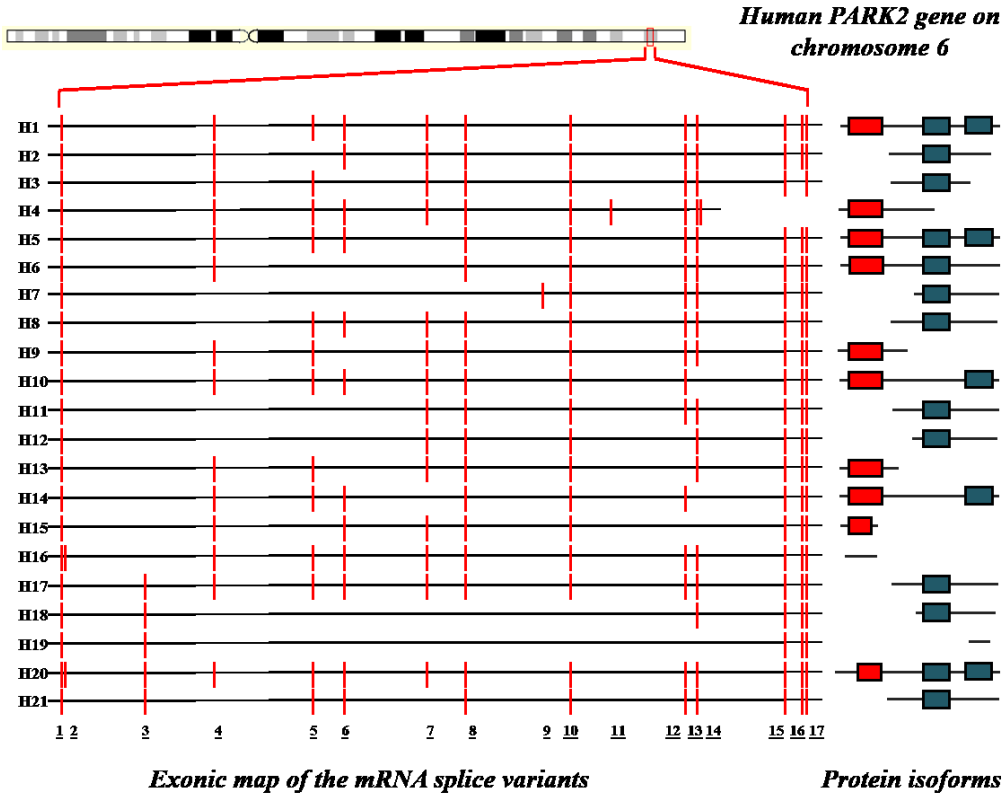


Figure 1. Chromosomal localization, exonic structure of alternative splice variants and corresponding predicted protein isoforms of human *PARK2*.A) Cytogenetic location of human *PARK2* gene (6q26). **B)** Exon organization map of the 21 human *PARK2* splice variants currently known. Exons are represented as red bars. The size of introns (black line) is proportional to their length. The codes on left refer to Gene Identifiers reported in Table 1. **C)** Predicted molecular architecture of *PARK2* isoforms. Red boxes represent UBQ domain and blue boxes represent IBR domains.

New code identifier	GI	Protein accession number	aa sequence	Predicted MW	pI
H20	469609976	AGH62057.1	530 aa	58,127	6,41
H1	3063387 121308969 158258616 169790968 125630744	BAA25751.1 BAF43729.1 BAF85279.1 NP_004553.2 ABN46990.1	465 aa	51,65	6,71
H5	284468410 169790970	ADB90270.1 NP_054642.2	437 aa	48,713	7,12
H10	284468412	ADB90271.1	415 aa	46,412	6,91
H14	284516985	ADB91979.1	387 aa	43,485	7,43
H4	34191069	AAH22014.1	387 aa	42,407	8,15
H8	284468407	*	386 aa	42,52	6,65
H17	284516991	*	386 aa	42,52	6,65
H21	520845529	AGP25366.1	358 aa	39,592	7,08
H6	169790972	NP_054643.2	316 aa	35,63	6,45
H11	284516981	*	274 aa	30,615	6,3
H2	20385797	AAM21457.1	270 aa	30,155	6,05
H3	20385801	AAM21459.1	203 aa	22,192	5,68
H12	284516982	*	172 aa	19,201	6,09
H9	284468408	ADB90269.1	143 aa	15,521	5,54
H13	284516983	ADB91978.1	143 aa	15,521	5,54
H7	194378189	BAG57845.1	139 aa	15,407	6,41
H18	284516993	*	139 aa	15,393	6,41
H15	284516987	ADB91980.1	95 aa	10,531	8,74
H19	469609974	AGH62056.1	61 aa	6,832	10,09
H16	284516989	ADB91981.1	51 aa	5,348	7,79

Table 1: Homo sapiens parkin isoforms. H1 represents the canonical sequence cloned by Kitada et al., 1998 [12]. * The protein accession number is not present in database.

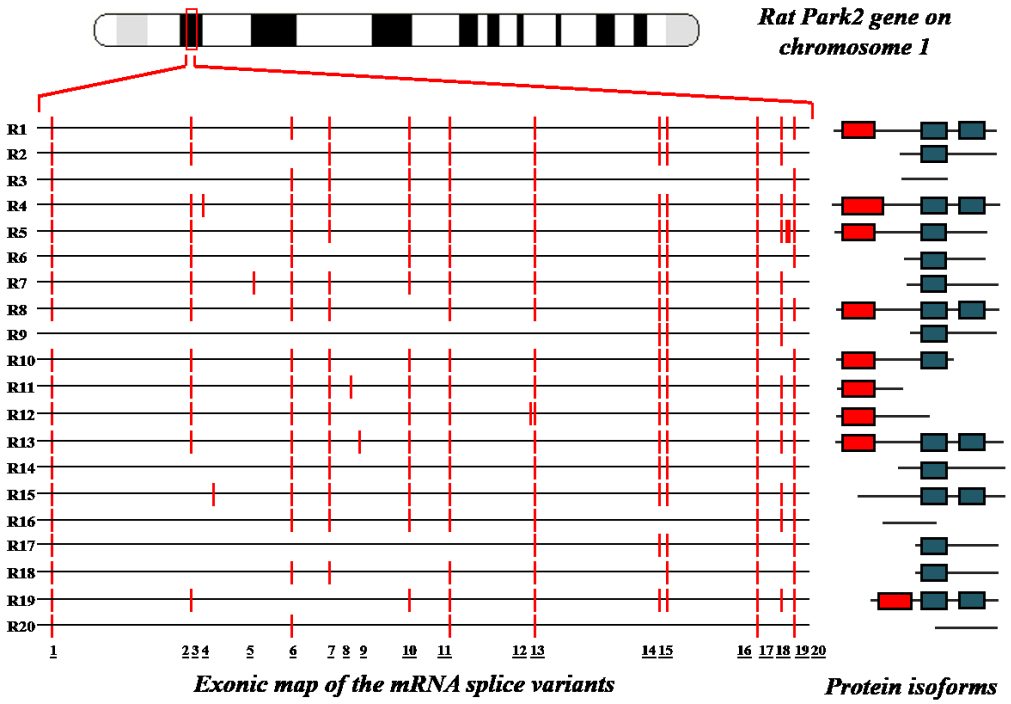


Figure 2. Chromosomal localization, exonic structure of alternative splice variants and corresponding predicted protein isoforms of rat *Park2*. Cytogenetic location of rat *Park2* gene (1q11). **B)** Exon organization map of the 20 rat *Park2* splice variants currently known. Exons are represented as red bars. The size of introns (black line) is proportional to their length. The codes on left refer to Gene Identifiers reported in Table 2. **C)** Predicted molecular architecture of *Park2* isoforms. Red boxes represent UBQ domain and blue boxes represent IBR domains.

New code identifier	GI	Protein accession number	aa sequence	Predicted MW	pI
R13	284810438	ADB96019.1	494 aa	54,829	6,46
R4	20385787	AAM21452.1	489 aa	54,417	6,46
R1	7229096 7717034 11464986 11527823 7001383	BAA92431.1 AAF68666.1 NP_064478.1 AAG37013.1 AAF34874.1	465 aa	51,678	6,59
R5	20385789	AAM21453.1	446 aa	49,367	6,59
R8	20385795 284066979	AAM21456.1 ADB77772.1	437 aa	48,734	6,74
R15	520845531	AGP25367.1	421 aa	46,854	6,59
R10	284066981	ADB77773.1	394 aa	43,297	6,06
R19	520845539	AGP25371.1	344 aa	38,558	6,13
R2	18478865	AAL73348.1	274 aa	30,641	6,2
R7	20385793 284810436	AAM21455.1 ADB96018.1	274 aa	30,641	6,2
R14	520845525 520845527	AGP25364.1 AGP25365.1	274 aa	30,669	6,2
R12	284468405	ADB90268.1	256 aa	28,006	6,44
R6	20385791	AAM21454.1	203 aa	22,288	5,42
R11	284468403	ADB90267.1	193 aa	21,253	8,54
R9	20385803	AAM21460.1	177 aa	19,84	5,97
R17	520845535	AGP25369.1	139 aa	15,404	6,29
R18	520845537	AGP25370.1	139 aa	15,404	6,29
R3	18478869	AAL73349.1	111 aa	12,329	6,92
R16	520845533	AGP25368.1	111 aa	12,329	6,92
R20	520845541	AGP25372.1	86 aa	9,929	7,5

Table 2: Rattus Norvegicus parkin isoforms.

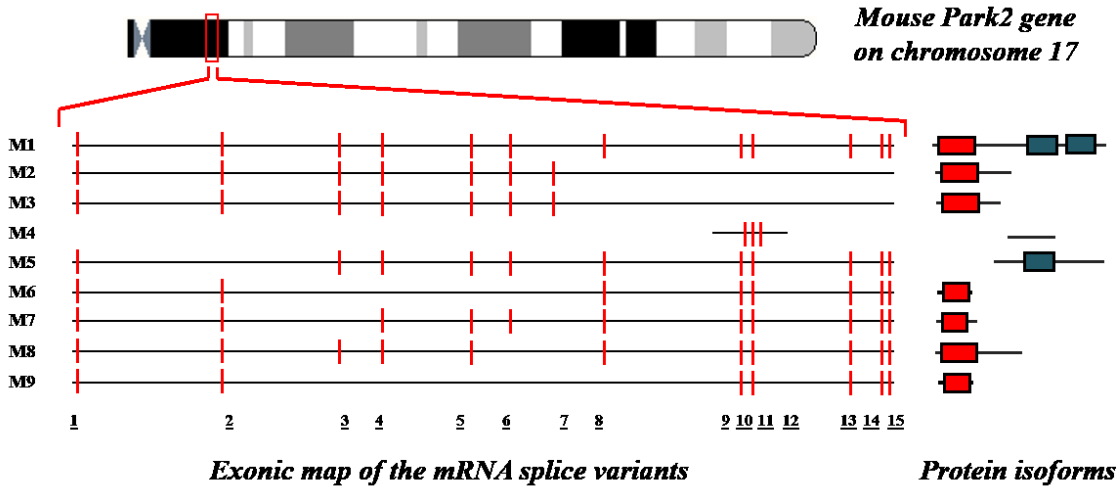


Figure 3. **Chromosomal localization, exonic structure of alternative splice variants and corresponding predicted protein isoforms of mouse *Park2*. A)** Cytogenetic location of mouse *Park2* gene (A3.2-A3.3). **B)** Exon organization map of the 9 mouse *Park2* splice variants currently known. Exons are represented as red bars. The size of introns (black line) is proportional to their length. The codes on left refer to Gene Identifiers reported in Table 3. **C)** Predicted molecular architecture of *Park2* isoforms. Red boxes represent UBQ domain and blue boxes represent IBR domains.

New code identifier	GI	Protein accession number	aa sequence	Predicted MW	pI
M1	10179808 118131140 5456929 86577675	AAG13890.1 NP_057903.1 BAA82404.1 AAI13205.1	464 aa	51,617	6,9
M5	220961631	*	274 aa	30,631	6,54
M2	10179810	AAG13891.1	262 aa	28,7	7,57
M3	10179812	AAG13892.1	255 aa	28,154	8,49
M8	220961637	ACL93283.1	214 aa	23,388	6,51
M7	220961635	ACL93282.1	106 aa	11,482	9,3
M4	74227131	*	75 aa	8,053	8,85
M6	220961633	ACL93281.1	65 aa	7,181	5,62
M9	284829878	ADB99567.1	63 aa	6,967	6,53

Table 3: Mus musculus parkin isoforms. The protein accession number is not present in database.

PARK2 ALTERNATIVE SPLICE TRANSCRIPTS PRODUCE ISOFORMS WITH DIFFERENT STRUCTURES AND FUNCTIONS

To date, 26 human different cDNAs, corresponding to 21 unique PARK2 alternative splice variants, have been described and are summarized in Figure 1 and Table 1. These mature transcripts are derived from the combination of 17 different exonic regions. Similarly, 20 PARK2 transcripts (20 exons) have been characterized in rat (Figure 2 and Table 2) and 9 (15 exons) in mouse (Figure 3 and Table 3). All of them have been carefully described in our previous paper [14]. For each of these variants, the encoded protein isoform, the corresponding molecular weight, and isoelectric point have been predicted and reported in Tables 1, 2, and 3. H8/H17, H9/H13, and H7/H18 isoforms show the same molecular weight and isoelectric point (Table 1), since they have the same amino acid composition; similarly, R2/R7/R14, R17/R18, and R3/R16 show the same primary structure, as shown in Table 2. Although equal, these proteins are encoded by different splice variants which probably produce the same protein with different efficiency.

In addition to primary structures, molecular architectures and domains composition have also been evaluated (Figures 1, 2, and 3 panels (b) and (c)). As previously described, the original (canonical) PARK2 protein (Accession number BAA25751.1) [12] comprises an N-terminal ubiquitin-like (UBQ) domain and two C-terminal in-between ring fingers (IBR) domains. The UBQ domain targets specific protein substrates for proteasome degradation, whereas IBR domains occur between pairs of ring fingers and play a role in protein quality control. PARK2 encoded isoforms structurally diverge from the canonic one for the presence or absence of the UBQ domain and for one of or both IBR domains. Moreover, when the UBQ domain is present, it often differs in length from that of the canonical sequence. Interestingly, some isoforms miss all of these domains.

The different molecular architectures and domain composition of isoforms might roughly alter also their functions. Parkin protein acts as an E3 ubiquitin ligase and is responsible of substrates recognition for proteasome-mediated degradation. PARK2 tags various types of proteins, including cytosolic (Synphilin-1, Pael-R, CDCrel-1 and 2a, α-synuclein, p22, and Synaptotagmin XI) [25–29], nuclear (Cyclin E) [15], and mitochondrial ones (MFN1 and MFN2, VDAC, TOM70, TOM40 and TOM20, BAK, MIRO1 and MIRO2, and FIS1) [30–34]. The number of targets is so high that parkin

protein results involved in numerous molecular pathways (proteasome-degradation, mitochondrial homeostasis, mitophagy, mitochondrial DNA stability, and regulation of cellular cycle). To date it is unknown if all these functions are mediated by a single protein or by different isoforms. However, considering that parkin mRNAs have a different expression and distribution in tissues and cells [14], which should be also mirrored at the protein level, it is reasonable to hypotisize that these distinct isoforms might perfom specific functions and could be differentially expressed in each cellular phenotype. Each PARK2 splice variants may acts in different manner to suit cell specific needs. This hypothesis is supported by previous evidences showing different and even opposite functions of other splice variants, such as BCl2L12 pattern expression related to cellular phenotype [35]. Finally, based on the extensive alternative splicing process of PARK2 gene, we cannot rule out that additional splice variants with different functions (beyond those listed) may exist.

EVIDENCES OF MULTIPLE PARKIN ISOFORMS IN BRAIN

A remarkable number of papers have demonstrated the existence, in human and other species, of different mRNA parkin variants [15–21]. However, few of them have investigated parkin isoforms existence, and some have done it without the awareness of PARK2 complex splicing [23, 36, 37]. In fact, although many mRNA parkin splice variants have been cloned, the corresponding proteins have been only deduced through the analysis of the longest open reading frame and uploaded on protein databases as predicted sequences. To date many questions are still unanswered: Are all mRNA parkin splice variants translated? Does a different expression pattern of parkin proteins, in tissue and cells, exist? Does each protein isoform have a specific function? In the following paragraphs we try to answer these questions by summarizing the knowledge accumulated over the last three decades on parkin expression and distribution in human, rat, and mouse brain. Existing data are reinterpreted by considering the complexity level of PARK2 gene splicing described above.

Many conflicting data emerges in the literature regarding the number and relative electrophoretic mobility of parkin proteins. While the majority of papers reported only a band of ~52 kDa corresponding to the canonical parkin isoform, also known as full length parkin, additional bands (from ~22 kDa to ~100 kDa) both in

rodent [23, 28, 36–41] and human brain regions were also detected [22–25, 39, 42–45].

Parkin was observed both in rat central and peripheral nervous system. Two major bands of ~50 and ~44 kDa were recognized in cell extracts from rat Substantia Nigra (SN) and cerebellum by western blot analysis. In adrenal glands there were visualized several immunoreactive bands of 50, 69–66, and 89 kDa [36]. Additional bands were also observed in primary cultures of cortical type I astrocytes [37].

Similar result was observed in mouse brain homogenate: a major band of 50 kDa and fainter bands of ~40 and 85/118 kDa were identified on immunoblot. In all these papers, lower and higher molecular weight bands were described as posttranslational modification or proteolytic cleavage of 52 kDa canonical protein or heterodimers resulting from the interaction of parkin with other proteins [42]. However, we speculate that they might correspond to multiple parkin isoforms with different molecular weight.

In knocked-out mice for parkin exon 2, several unexpected bands were also observed on immunoblot. This was interpreted as antibody cross-reactivity with nonauthentic parkin protein [46]. However, as shown in Figure 3, these bands might represent isoforms encoded by splice variants not containing the deleted exon (i.e., M5 and M4).

Parkin expression was also demonstrated in human brains of normal and sporadic Parkinson disease (PD) subjects, but it was absent in any regions of AR-JP brain [22, 23]. A major band of 52 kDa and a second fainter band of ~41 kDa were observed on immunoblot from human frontal cortex of PD patients and control subjects [22]. Parkin expression was also observed in Lewy bodies (LBs), characteristic neuronal inclusions in PD brain. However, in this regard we highlight widely varying results. Initially, the parkin protein expression was reported in neurons of the SN, locus coeruleus, putamen, and frontal lobe cortex of sporadic PD and control individuals but no parkin-immunoreactivity (IR) was found in SN LBs of PD patients [22, 23]. Later on, parkin-IR was described in nigral LBs of four related human disorders, sporadic PD, α-synuclein-linked PD, LB positive parkin-linked PD, and dementia with LBs (DBL) [24]. These discrepant results might be due to the antibodies used. In fact, as shown in Table 4, aligning the epitope sequence recognized by the antibody to each isoform sequence, we discovered that every antibody identifies a pool of different isoforms.

In accord with this hypothesis, we also explain discordant results observed by Schlossmacher et al. (2002) regarding the cellular distribution of the protein. In fact, they described strongly labeled cores of classical intracellular LBs in pigmented neurons of the SN in PD and DLB patients by using HP2A antibody, whereas HP1A and HP7A antibodies intensively labeled cytoplasmic parkin, in a granular pattern, of cell bodies and proximal neurites of dopaminergic neurons in both diseased and normal brains [24]. These results might represent a different cellular expression profile of parkin isoforms in healthy and diseased human brains.

Name	Target	Recognized Parkin isoforms
M73 (Shimura et al., 1999) [22]	124–137	H1, H4, H5, H8, H9, H10, H13, H14, H17, H20, H21
M74 (Shimura et al., 1999) [22]	293–306	H1, H2, H3, H4, H5, H6, H8, H10, H11, H14, H17, H20, H21
ParkA (Huynh et al., 2000) [23]	96–109	H1, H2, H3, H4, H5, H6, H8, H9, H10, H11, H13, H14, H17, H20, H21
ParkB (Huynh et al., 2000) [23]	440–415	H1, H2, H5, H6, H7, H8, H10, H11, H12, H14, H17, H18, H20, H21
HP6A (Schlossmacher et al., 2002) [24]	6–15	H1, H4, H5, H6, H9, H10, H13, H14, H16, H20
HP7A (Schlossmacher et al., 2002) [24]	51–62	H1, H4, H5, H6, H9, H10, H13, H14, H15, H20
HP1A (Schlossmacher et al., 2002) [24]	84–98	H1, H2, H3, H4, H5, H6, H8, H9, H10, H11, H13, H14, H17, H20, H21
HP2A (Schlossmacher et al., 2002) [24]	342–353	H1, H2, H3, H4, H5, H6, H7, H8, H11, H12, H17, H18, H20, H21
HP5A (Schlossmacher et al., 2002) [24]	453–465	H1, H2, H5, H6, H7, H8, H10, H11, H12, H14, H17, H18, H20, H21

Table 4: Parkin isoforms recognized by antibodies used in some studies.

This hypothesis is supported by another study demonstrating a different expression profile of parkin mRNA splice variants in frontal cortex of patients with common dementia with LB, pure form of dementia with LB, and Alzheimer disease

suggesting the direct involvement of isoform-expression deregulation in the development of such neurodegenerative disorders [17]. To date there exists only one paper that has dealt with parkin amino acid sequencing [47]. Trying to ensure that the signal observed on human serum by western blot analysis belongs to parkin protein, they cut off the area on the blot between 50 and 55 kDa in two separate pieces and performed a MALDI-TOF analysis on each. Peptides peaks analysis revealed the presence of six other proteins with similar sequence to canonical one. However, authors did not even speculate that they could represent additional parkin isoforms.

Further evidences on the existence of multiple isoforms come from the conflicting data on their tissue and cellular distribution. Parkin protein is particularly abundant in the mammalian brain and retina [22, 23, 36, 48, 49]. In human, parkin immunoreactivity (IR) has been observed in SN, locus coeruleus, putamen, and frontal lobe cortex [22, 23]. Similarly, it has been strongly measured in rat hippocampus, amygdaloid nucleus, endopiriform nucleus, cerebral cortex, colliculus, and SN (pars compacta and pars reticulata) [37, 50].

Analog parkin distribution was reported in mouse. Most immunoreactive cells were found in the hindbrain. In the cerebellum only the cells within the cerebellar nuclei were positive, while the structures located in the mesencephalon presented moderate to strong immunopositivity. In the ventral part of the mesencephalon the red nucleus showed large strongly stained cells. In the SN moderate parkin immunoreactivity was confined to the pars reticulate. In the dorsal mesencephalon, immunopositive cells were found in the intermediate and deep gray layer of the superior colliculus and in all parts of the inferior colliculus [12, 36, 41, 51].

Although in most brain regions good correlations between parkin-IR and mRNA were observed, incongruent data emerged from some paper in rat SNc (substantia nigra pars compacta), hippocampus, and cerebellar Purkinje cells distribution, where mRNA was detected but no parkin-IR was revealed [23, 36].

Furthermore, in an early study, parkin was described in cytoplasm, in granular structure, and in neuronal processes but was absent in the nucleus [22]. Subsequently other studies reported also its nuclear localization [23, 37, 48, 52–54]. Finally, some papers have also observed a small mitochondrial pool of the protein [55, 56]. All these evidences have suggested that protein could localize to specific subcellular structure under some circumstances. However, it is also reasonably hypothesized that a specific pattern of subcellular distribution of parkin isoforms is related to each cellular

phenotype, since in all these papers, protein immunolocalization was performed by using antibodies recognizing different epitopes. Some discrepancies are also observed in the expression of parkin in the SNc of patients affected by other forms of parkinsonism [23].

Brain isoforms might have different species-specific biochemical characteristics, when comparing murine versus human parkin. In fact, it has been shown that mouse protein is easily extracted from brain by high salt buffer, instead human parkin is only extracted with harsher buffers, especially in elderly. This suggested that human parkin becomes modified or interacts with other molecules with age, and this alters its biochemical properties [42]. However, we cannot rule out that this may correlate to a specific expression pattern of isoforms with different biochemical properties in the brains of rodents and humans relative to age.

All of these observations were also supported by contradictory results emerging from clinical studies. Initially, recessive mutations in the parkin gene were related to sporadic early onset parkinsonism [2]; however, the mode of transmission was subsequently rejected by other genetic studies with not only homozygous or compound heterozygous mutations, but also single heterozygous mutations, affecting only one allele of the gene [2, 57–61]. It has been suggested that haploinsufficiency is a risk factor for disease, but certain mutations are dominant, conferring dominant-negative or toxic gain of functions of parkin protein [61]. However, in light of the evidence outlined above, it is possible that some single heterozygous mutation might affect gene expression by inducing loss of function of some isoforms and gain of function of other.

THE DIVERSIFIED PANEL OF ANTIBODIES COMMERCIALLY AVAILABLE AGAINST *PARK2*

To date more than 160 PARK2 antibodies are commercially available. They are obtained from different species (generally rabbit or mouse) and commercialized by various companies. Table 5 lists 32 commercially available PARK2 antibodies whose immunogens used are specified by providers in datasheet. Some of them recognize a common epitope, therefore, have been included in the same group. Tables 6, 7, and 8 report, respectively, human, rat, and mouse parkin isoforms recognized by these antibodies. When the amino acid sequence recognized by the antibody perfectly match

with the sequence of the protein, it is very likely to get a signal by western blot or immunohistochemistry analysis (this is indicated in the table by “Yes”). Instead, if the antibody recognizes at least 8 consecutive amino acids on the protein, it is likely to visualize a signal both by western blot or immunohistochemistry analysis (this is indicated in the table by “May be”). Finally, if the antibody recognizes less than 8 consecutive amino acids, it could rule out the possibility to visualize a signal on immunoblot or immunohistochemistry analysis (this is indicated in the table by “No”). The use of these 32 antibodies may allow the identification of at least 15 different PARK2 epitopes (Table 5). Although no epitope is isoform specific, the combinatorial use of antibodies targeting different protein regions may provide a precious aid to decode the exact spectrum of PARK2 isoforms expressed in tissues and cells. An example of combinatorial use of antibodies has been reported in Figure 4. On rat brain homogenate, these five antibodies raised against different parkin epitopes, revealed the canonical ~50 kDa band, but additional putative bands of higher and lower molecular weight were visualized. This experimental data reinforce the existence of more than one parkin isoform and confirm that the investigation of parkin expression profile should not be restricted to the use of a single antibody. The latter approach, in fact, could not reveal the entire spectrum of parkin variants.

Antibody group #	Generic name		Target domain
	Trade name	Companies	
#1	H00005071-B01P	Abnova	1 aa–387 aa
	H00005071-D01P	Abnova	
	H00005071-D01	Abnova	
#2	OASA06385	Aviva System biology	83 aa–97 aa
	AHP495	AbD Serotec	
	MD-19-0144	Raybiotech, Inc.	
	DS-PB-01562	Raybiotech, Inc.	
	PAB14022	Abnova	
#3	MCA3315Z	AbD Serotec	288 aa–388 aa

	H00005071-M01	Abnova	
#4	PAB1105	Abnova	62 aa–80 aa
	70R-PR059	Fitzgerald	
#5	PAB0714	Abnova	305 aa–323 aa
	AB5112	Millipore Chemicon	
	R-113-100	Novus biologicals	
#6	P5748	Sigma	298 aa–313 aa
	GTX25667 Parkin antibody CR20121213_GTX25667	GeneTex International Corporation	
	ABIN122870	Antibodies on-line	
	PA1-751	Thermo Fisher Scientific, Inc.	
#7	R-114-100	Novus biologicals	295 aa–311 aa
	Anti-Parkin, aa295-311 h Parkin; C-terminal	Millipore Chemicon	
#8	MAB5512	Millipore Chemicon	399 aa–465 aa
	Anti-Parkin antibody, clone PRK8/05882	Millipore Upstate	
	Parkin (PRK8): sc-32282	Santa Cruz	
#9	Parkin (H-300): sc-30130	Santa Cruz	61 aa–360 aa
	Parkin (D-1): sc-133167	Santa Cruz	
	Parkin (H-8): sc-136989	Santa Cruz	
#10	EB07439	Everest Biotech	394 aa–409 aa
	GTX89242 PARK2 antibody, internal CR20121213_GTX89242	GeneTex International Corporation	
	NB100-53798	Novus biologicals	

#11	GTX113239 Parkin antibody [N1C1] CR20121213_GTX113239	GeneTex International Corporation	28 aa–258 aa
#12	10R-3061	Fitzgerald	390 aa–406 aa
#13	A01250-40	GenScript	300 aa–350 aa
#14	NB600-1540	Novus biologicals	399 aa–412 aa
#15	ARP43038_P050	Aviva System biology	311 aa–360 aa

Table 5: List of antibodies targeting PARK2 isoforms.. Antibodies against canonical PARK2 isoform (NP_004553.2) were grouped if they recognize the same epitope. To each group was assigned a new identification code (#).

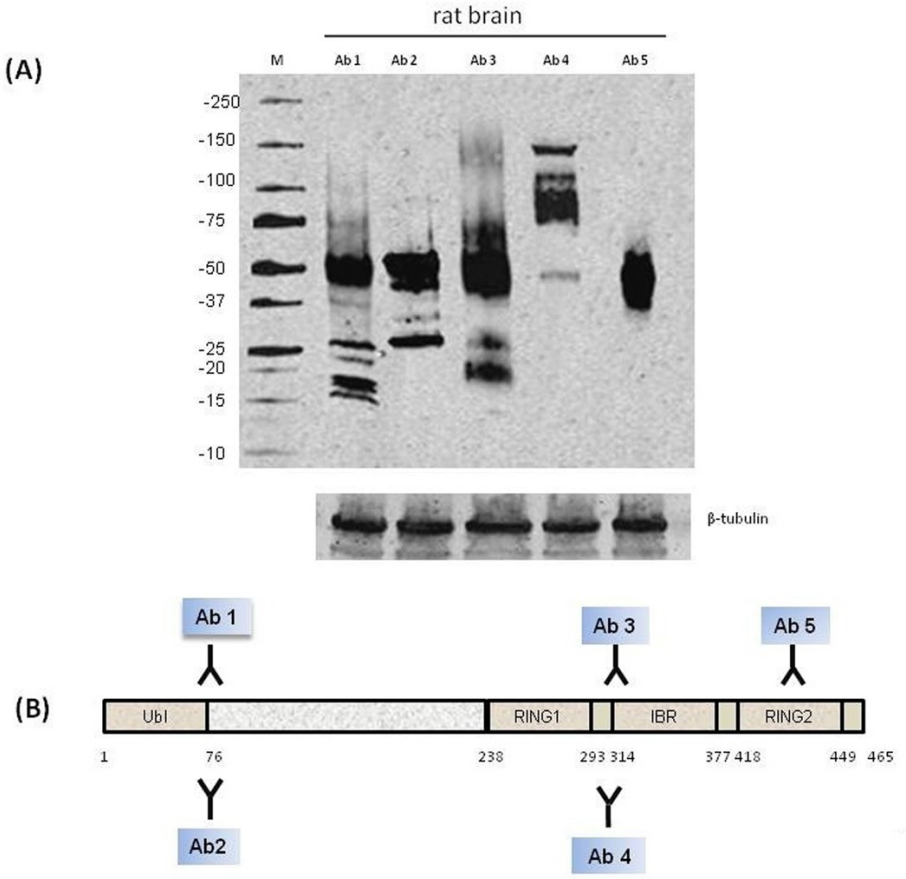


Figure 4. **Differential detection of parkin isoforms in rat brain using five anti-parkin antibodies.**

Representative immunoblot of parkin isoforms in rat brain visualized by using five different antibodies. Ab1, Ab2, Ab3, Ab4, Ab5 correspond to group #3,#4,#5,#8,#9 of table 5. Immunoblot for β-tubulin was used as loading control. **B)** Canonical parkin sequence domains recognized by the five antibodies.

New code identifier	Ab #1	Ab #2	Ab #3	Ab #4	Ab #5	Ab #6	Ab #7	Ab #8	Ab #9	Ab #10	Ab #11	Ab #12	Ab #13	Ab #14	Ab #15
H20	May be (360 aa)	Yes	May be (64 aa)	Yes	Yes	Yes	Yes	Yes	May be (299 aa)	May be (17 aa)	May be (230 aa)	Yes	Yes	Yes	May be (47 aa)
H1	May be (360 aa)	Yes	May be (64 aa)	Yes	Yes	Yes	Yes	Yes	Yes	May be (17 aa)	Yes	Yes	Yes	Yes	May be (47 aa)
H5	May be (333 aa)	Yes	May be (64 aa)	Yes	Yes	Yes	Yes	Yes	May be (271 aa)	May be (17 aa)	May be (202 aa)	Yes	Yes	Yes	May be (47 aa)
H10	May be (311 aa)	Yes	May be (22 aa)	Yes	No	May be (14 aa)	Yes	Yes	May be (250 aa)	May be (17 aa)	May be (230 aa)	Yes	No	Yes	No
H14	May be (283 aa)	Yes	May be (22 aa)	Yes	No	May be (14 aa)	Yes	Yes	May be (222 aa)	May be (17 aa)	yes (partial match 202 aa/231 aa)	Yes	May be (12 aa)	May be (15 aa)	No
H4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	May be (299 aa)	No	May be (230 aa)	No	Yes	No	May be (47 aa)
H8	May be (274 aa)	Yes	May be (64 aa)	No	Yes	Yes	Yes	Yes	May be (281 aa)	May be (17 aa)	May be (178 aa)	Yes	Yes	May be (15 aa)	May be (47 aa)
H17	May be (274 aa)	Yes	May be (64 aa)	No	Yes	Yes	Yes	Yes	May be (280 aa)	May be (17 aa)	May be (178 aa)	Yes	Yes	May be (15 aa)	May be (47 aa)
H21	May be (254 aa)	Yes	May be (64 aa)	No	Yes	Yes	Yes	Yes	May be (252 aa)	May be (17 aa)	May be (150 aa)	Yes	Yes	May be (15 aa)	May be (47 aa)

H6	May be (148 aa)	No	May be (64 aa)	No	Yes	Yes	Yes	Yes	Yes	Yes	May be (52 aa)	Yes	Yes	Yes	Yes
H11	May be (162 aa)	No	May be (64 aa)	No	Yes	Yes	Yes	Yes	Yes	Yes	May be (66 aa)	Yes	Yes	Yes	Yes
H2	May be (161 aa)	No	May be (64 aa)	No	Yes	Yes	Yes	Yes	Yes	Yes	May be (67 aa)	Yes	Yes	Yes	Yes
H3	May be (161 aa)	No	May be (64 aa)	No	Yes	Yes	Yes	No	Yes	No	May be (67 aa)	No	Yes	No	Yes
H12	May be (42 aa)	No	May be (42 aa)	No	May be (12 aa)	No	No	Yes	Yes	Yes	No	Yes	May be (39 aa)	Yes	Yes
H9	May be (137 aa)	Yes	No	Yes	No	No	No	No	Yes	No	May be (110 aa)	No	No	No	No
H13	May be (137 aa)	Yes	No	Yes	No	No	No	No	Yes	No	May be (110 aa)	No	No	No	No
H7	May be (27 aa)	No	May be (27 aa)	No	No	No	No	Yes	May be (30 aa)	Yes	No	Yes	May be (24 aa)	Yes	Yes
H18	May be (27 aa)	No	May be (27 aa)	No	No	No	No	Yes	May be (30 aa)	Yes	No	Yes	May be (24 aa)	Yes	Yes
H15	May be (65 aa)	No	No	No	No	No	No	No	No	No	May be (38 aa)	No	No	No	No
H19	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No	No

H16	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No

Table 6: Homo sapiens. Yes = perfect match between predicted protein sequence and antibody epitope. May be = partial match between predicted protein sequence and antibody epitope; in parenthesis number of amino acid matching/total number of amino acid recognized by antibody epitope.No = matching between predicted protein sequence and antibody epitope is less than 8 consecutive amino acids

New code identifier	Ab #1	Ab #2	Ab #3	Ab #4	Ab #5	Ab #6	Ab #7	Ab #8	Ab #9	Ab #10	Ab #11	Ab #12	Ab #13	Ab #14	Ab #15
R13	May be (306 aa)	May be (5 aa)	May be (69 aa)	May be (14 aa)	Yes	Yes	Yes	May be (66 aa)	May be (248 aa)	May be (14 aa)	May be (180 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (48 aa)
R4	May be (307 aa)	May be (5 aa)	May be (69 aa)	May be (14 aa)	Yes	Yes	Yes	May be (66 aa)	May be (247 aa)	May be (14 aa)	May be (179 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (49 aa)
R1	May be (307 aa)	May be (5 aa)	May be (70 aa)	May be (14 aa)	Yes	Yes	Yes	May be (66 aa)	May be (249 aa)	May be (14 aa)	May be (180 aa)	May be (15 aa)	May be (49 aa)	May be (13 aa)	May be (49 aa)
R5	May be (305 aa)	May be (5 aa)	May be (69 aa)	May be (14 aa)	Yes	Yes	Yes	May be (31 aa)	May be (247 aa)	May be (14 aa)	May be (179 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (48 aa)

R8	May be (279 aa)	May be (5 aa)	May be (69 aa)	May be (14 aa)	Yes	Yes	Yes	May be (66 aa)	May be (221 aa)	May be (14 aa)	May be (153 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (48 aa)
R15	May be (254 aa)	May be (5 aa)	May be (73 aa)	May be (14 aa)	Yes	Yes	Yes	May be (66 aa)	May be (248 aa)	May be (14 aa)	May be (153 aa)	May be (15 aa)	May be (49 aa)	May be (13 aa)	May be (49 aa)
R10	May be (173 aa)	May be (5 aa)	May be (69 aa)	May be (14 aa)	Yes	Yes	Yes	May be (9 aa)	May be (248 aa)	No	May be (180 aa)	No	May be (48 aa)	No	May be (48 aa)
R19	May be (162 aa)	No	May be (70 aa)	No	Yes	Yes	Yes	May be (68 aa)	May be (173 aa)	May be (14 aa)	May be (74 aa)	May be (15 aa)	May be (49 aa)	May be (13 aa)	May be (49 aa)
R2	May be (147 aa)	No	May be (72 aa)	No	Yes	Yes	Yes	May be (68 aa)	May be (156 aa)	May be (14 aa)	May be (55 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (48 aa)
R7	May be (147 aa)	No	May be (72 aa)	No	Yes	Yes	Yes	May be (68 aa)	May be (153 aa)	May be (14 aa)	May be (55 aa)	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (49 aa)
R14	May be (149 aa)	No	May be (73 aa)	No	Yes	Yes	Yes	May be (68 aa)	May be (155 aa)	May be (14 aa)	May be (56 aa)	May be (15 aa)	May be (49 aa)	May be (13 aa)	May be (49 aa)
R12	May be (196 aa)	May be (5 aa)	No	May be (14 aa)	No	No	No	May be (9 aa)	May be (138 aa)	No	May be (168 aa)	No	No	No	No

R6	May be (147 aa)	No	May be (69 aa)	No	Yes	Yes	Yes	No	May be (153 aa)	No	May be (55 aa)	No	May be (48 aa)	No	May be (48 aa)
R11	May be (139 aa)	May be (5 aa)	No	May be (14 aa)	No	No	No	No	May be (82 aa)	No	May be (112 aa)	No	No	No	No
R9	May be (60 aa)	No	May be (68 aa)	No	Yes	Yes	Yes	May be (68 aa)	May be (67 aa)	May be (14 aa)	No	May be (15 aa)	May be (48 aa)	May be (13 aa)	May be (48 aa)
R17	May be (25 aa)	No	May be (33 aa)	No	No	No	No	May be (68 aa)	May be (32 aa)	May be (14 aa)	No	May be (15 aa)	May be (22 aa)	May be (13 aa)	May be (35 aa)
R18	May be (25 aa)	No	May be (33 aa)	No	No	No	No	May be (68 aa)	May be (32 aa)	May be (14 aa)	No	May be (15 aa)	May be (22 aa)	May be (13 aa)	May be (35 aa)
R3	May be (87 aa)	No	No	No	No	No	No	May be (8 aa)	May be (86 aa)	No	May be (55 aa)	No	No	No	No
R16	May be (87 aa)	No	No	No	No	No	No	May be (8 aa)	May be (86 aa)	No	May be (55 aa)	No	No	No	No
R20	No	No	No	No	No	No	No	May be (66 aa)	No	May be (14 aa)	No	May be (15 aa)	No	May be (13 aa)	No

Table 7: Rattus Norvegicus. Please, refer for legend to Table 6.

New code identifier	Ab #1	Ab #2	Ab #3	Ab #4	Ab #5	Ab #6	Ab #7	Ab #8	Ab #9	Ab #10	Ab #11	Ab #12	Ab #13	Ab #14	Ab #15
M1	May be (294 aa)	No	May be (61 aa)	May be (13 aa)	May be (18 aa)	Yes	Yes	May be (70 aa)	May be (244 aa)	No	May be (176 aa)	May be (15 aa)	May be (48 aa)	May be (14 aa)	Yes
M5	May be (147 aa)	No	May be (62 aa)	No	May be (18 aa)	Yes	Yes	May be (70 aa)	May be (153 aa)	No	May be (55 aa)	May be (15 aa)	May be (48 aa)	May be (14 aa)	Yes
M2	May be (191 aa)	No	No	May be (13 aa)	No	No	No	No	May be (134 aa)	No	May be (164 aa)	No	No	No	No
M3	May be (192 aa)	No	No	May be (13 aa)	No	No	No	No	May be (135 aa)	No	May be (165 aa)	No	No	No	No
M8	May be (161 aa)	No	No	May be (13 aa)	No	No	No	No	May be (106 aa)	No	May be (136 aa)	No	No	No	No
M7	May be (53 aa)	No	No	No	No	No	No	No	No	No	May be (27 aa)	No	No	No	No
M4	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
M6	May be (53 aa)	No	No	No	No	No	No	No	No	No	May be (27 aa)	No	No	No	No
M9	May be (53 aa)	No	No	No	No	No	No	No	No	No	May be (27 aa)	No	No	No	No

Table 8: Mus musculus. Please, refer for legend to Table 6.

CONCLUSION

Alternative splicing is a complex molecular mechanism that increases the functional diversity without the need for gene duplication. Alternative splicing performs a crucial regulatory role by altering the localization, function, and expression level of gene products, often in response to the activities of key signaling pathways [62]. PARK2 gene, as the vast majority of multiexon genes in humans, undergoes alternative splicing [14, 63, 64]. The importance of alternative splicing in the regulation of diverse biological processes is highlighted by the growing list of human diseases associated with known or suspected splicing defects, including PD [65].

Mutations that affect PARK2 splicing could modify the levels of correctly spliced transcripts, alter their localization, and lead to a loss of function of some of them and/or gain of function of others in time- and cell-specific manner. Even if few, some evidences supporting this hypothesis have been already described. Preliminary studies reported PARK2 isoforms with defective degradation activity of cyclin E and control of cellular cycle [15] or characterized by altered solubility and intracellular localization [66]. No evidence of gain of function has been reported, but it is plausible, because a functional screen of the PARK2 splice variants has not been done yet. The huge number of molecular targets attributed to full-size parkin protein could be shared by the others parkin isoforms which could have additional biological activities that until now are uncosidered. In light of this consideration, alteration of the natural splicing of PARK2 and deregulation in the expression of parkin isoforms might lead to the selective degeneration of dopaminergic neurons in SN of ARJP. However this is a hypothesis, since the functional screen of the PARK2 splice variants is not available and this field is still unexplored.

All these could, at least in part, justifying the conflicting and heterogeneous data of studies revised in this work, which preceded the knowledge of PARK2 alternative splicing and expression of multiple isoforms for this gene. Understanding PARK2 alternative splicing could open up new scenarios for the resolution of some Parkinsonian syndrome.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. V. Bonifati, “Autosomal recessive parkinsonism,” Parkinsonism and Related Disorders, vol. 18, supplement 1, pp. S4–S6, 2012.
2. C. B. Lücking, A. Dürr, V. Bonifati et al., “Association between early-onset Parkinson's disease and mutations in the parkin gene,” The New England Journal of Medicine, vol. 342, no. 21, pp. 1560–1567, 2000.
3. S. A. Oliveira, W. K. Scott, E. R. Martin et al., “Parkin mutations and susceptibility alleles in late-onset Parkinson's disease,” Annals of Neurology, vol. 53, no. 5, pp. 624–629, 2003.
4. M. P. Burns, L. Zhang, G. W. Rebeck, H. W. Querfurth, and C. E.-. Moussa, “Parkin promotes intracellular Aβ1-42 clearance,” Human Molecular Genetics, vol. 18, no. 17, pp. 3206–3216, 2009.
5. J. T. Glessner, K. Wang, G. Cai et al., “Autism genome-wide copy number variation reveals ubiquitin and neuronal genes,” Nature, vol. 459, no. 7246, pp. 569–573, 2009.
6. M. E. Witte, J. G. J. M. Bol, W. H. Gerritsen et al., “Parkinson's disease-associated parkin colocalizes with Alzheimer's disease and multiple sclerosis brain lesions,” Neurobiology of Disease, vol. 36, no. 3, pp. 445–452, 2009.
7. R. Cesari, E. S. Martin, G. A. Calin et al., “Parkin, a gene implicated in autosomal recessive juvenile parkinsonism, is a candidate tumor suppressor gene on chromosome 6q25-q27,” Proceedings of the National Academy of Sciences of the United States of America, vol. 100, no. 10, pp. 5956–5961, 2003.
8. S. Veeriah, B. S. Taylor, S. Meng et al., “Somatic mutations of the Parkinson's disease-associated gene PARK2 in glioblastoma and other human malignancies,” Nature Genetics, vol. 42, no. 1, pp. 77–82, 2010.
9. M. T. Mira, A. Alcaïs, H. van Thuc et al., “Susceptibility to leprosy is associated with PARK2 and PACRG,” Nature, vol. 427, no. 6975, pp. 636–640, 2004.
10. W. Wongseree, A. Assawamakin, T. Piroonratana, S. Sinsomros, C. Limwongse, and N. Chaiyaratana, “Detecting purely epistatic multi-locus interactions by an omnibus permutation test on ensembles of two-locus analyses,” BMC Bioinformatics, vol. 10, article 294, 2009.
11. K. M. Rosen, V. Veereshwarayya, C. E. Moussa et al., “Parkin protects against mitochondrial toxins and β-amyloid accumulation in skeletal muscle cells,” The Journal of Biological Chemistry, vol. 281, no. 18, pp. 12809–12816, 2006.
12. T. Kitada, S. Asakawa, N. Hattori et al., “Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism,” Nature, vol. 392, no. 6676, pp. 605–608, 1998.
13. H. Matsumine, M. Saito, S. Shimoda-Matsubayashi et al., “Localization of a gene for an autosomal recessive form of juvenile parkinsonism to chromosome 6q25.2-27,” American Journal of Human Genetics, vol. 60, no. 3, pp. 588–596, 1997.
14. V. la Cognata, R. Iemmolo, V. D'Agata et al., “Increasing the coding potential of genomes through alternative splicing: the case of PARK2 gene,” Current Genomics: Bentham Science, vol. 15, no. 3, pp. 203–216, 2014.
15. K. Ikeuchi, H. Marusawa, M. Fujiwara et al., “Attenuation of proteolysis-mediated cyclin E regulation by alternatively spliced Parkin in human colorectal cancers,” International Journal of Cancer, vol. 125, no. 9, pp. 2029–2035, 2009.
16. T. Kitada, S. Asakawa, S. Minoshima, Y. Mizuno, and N. Shimizu, “Molecular cloning, gene expression, and identification of a splicing variant of the mouse parkin gene,” Mammalian Genome, vol. 11, no. 6, pp. 417–421, 2000.
17. K. Beyer, M. Domingo-Sàbat, J. Humbert, C. Carrato, I. Ferrer, and A. Ariza, “Differential expression of alpha-synuclein, parkin, and synphilin-1 isoforms in Lewy body disease,” Neurogenetics, vol. 9, no. 3, pp. 163–172, 2008.

18. J. Humbert, K. Beyer, C. Carrato, J. L. Mate, I. Ferrer, and A. Ariza, “Parkin and synphilin-1 isoform expression changes in Lewy body diseases,” Neurobiology of Disease, vol. 26, no. 3, pp. 681–687, 2007.

19. E. K. Tan, H. Shen, J. M. M. Tan et al., “Differential expression of splice variant and wild-type parkin in sporadic Parkinson's disease,” Neurogenetics, vol. 6, no. 4, pp. 179–184, 2005.

20. V. D'agata and S. Cavallaro, “Parkin transcript variants in rat and human brain,” Neurochemical Research, vol. 29, no. 9, pp. 1715–1724, 2004.

21. Y. Sunada, F. Saito, K. Matsumura, and T. Shimizu, “Differential expression of the parkin gene in the human brain and peripheral leukocytes,” Neuroscience Letters, vol. 254, no. 3, pp. 180–182, 1998.

22. H. Shimura, N. Hattori, S. Kubo, et al., “Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients,” Annals of Neurology, vol. 45, no. 5, pp. 668–672, 1999.

23. D. P. Huynh, D. R. Scoles, T. H. Ho, M. R. del Bigio, and S. M. Pulst, “Parkin is associated with actin filaments in neuronal and nonneural cells,” Annals of Neurology, vol. 48, no. 5, pp. 737–744, 2000.

24. M. G. Schlossmacher, M. P. Frosch, W. P. Gai et al., “Parkin localizes to the Lewy bodies of Parkinson disease and dementia with Lewy bodies,” The American Journal of Pathology, vol. 160, no. 5, pp. 1655–1667, 2002.

25. Y. Imai, M. Soda, and R. Takahashi, “Parkin suppresses unfolded protein stress-induced cell death through its E3 ubiquitin-protein ligase activity,” The Journal of Biological Chemistry, vol. 275, no. 46, pp. 35661–35664, 2000.

26. H. Shimura, N. Hattori, S. Kubo et al., “Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase,” Nature Genetics, vol. 25, no. 3, pp. 302–305, 2000.

27. J. F. Staropoli, C. McDermott, C. Martinat, B. Schulman, E. Demireva, and A. Abeliovich, “Parkin is a component of an SCF-like ubiquitin ligase complex and protects postmitotic neurons from kainate excitotoxicity,” Neuron, vol. 37, no. 5, pp. 735–749, 2003.

28. K. K. K. Chung, Y. Zhang, K. L. Lim et al., “Parkin ubiquitinates the α-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease,” Nature Medicine, vol. 7, no. 10, pp. 1144–1150, 2001.

29. Y. Zhang, J. Gao, K. K. K. Chung, H. Huang, V. L. Dawson, and T. M. Dawson, “Parkin functions as an E2-dependent ubiquitin-protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1,” Proceedings of the National Academy of Sciences of the United States of America, vol. 97, no. 24, pp. 13354–13359, 2000.

30. N. C. Chan, A. M. Salazar, A. H. Pham et al., “Broad activation of the ubiquitin-proteasome system by Parkin is critical for mitophagy,” Human Molecular Genetics, vol. 20, no. 9, pp. 1726–1737, 2011.

31. S. R. Yoshii, C. Kishi, N. Ishihara, and N. Mizushima, “Parkin mediates proteasome-dependent protein degradation and rupture of the outer mitochondrial membrane,” The Journal of Biological Chemistry, vol. 286, no. 22, pp. 19630–19640, 2011.

32. M. R. Cookson, “Parkinsonism due to mutations in PINK1, parkin, and DJ-1 and oxidative stress and mitochondrial pathways.,” Cold Spring Harbor Perspectives in Medicine, vol. 2, no. 9, Article ID a009415, 2012.

33. S. M. Jin and R. J. Youle, “PINK1-and Parkin-mediated mitophagy at a glance,” Journal of Cell Science, vol. 125, part 4, pp. 795–799, 2012.

34. D. Narendra, A. Tanaka, D. Suen, and R. J. Youle, “Parkin is recruited selectively to impaired mitochondria and promotes their autophagy,” The Journal of Cell Biology, vol. 183, no. 5, pp. 795–803, 2008.

35. C. K. Kontos and A. Scorilas, “Molecular cloning of novel alternatively spliced variants of BCL2L12, a new member of the BCL2 gene family, and their expression analysis in cancer cells,” Gene, vol. 505, no. 1, pp. 153–166, 2012.

36. J. M. Horowitz, J. Myers, M. K. Stachowiak, and G. Torres, “Identification and distribution of Parkin in rat brain,” NeuroReport, vol. 10, no. 16, pp. 3393–3397, 1999.

37. V. D'Agata, M. Grimaldi, A. Pascale, and S. Cavallaro, “Regional and cellular expression of the parkin gene in the rat cerebral cortex,”

European Journal of Neuroscience, vol. 12, no. 10, pp. 3583–3588, 2000.

38. W. Gu, N. Abbas, M. Z. Lagunes et al., “Cloning of rat parkin cDNA and distribution of parkin in rat brain,” Journal of Neurochemistry, vol. 74, no. 4, pp. 1773–1776, 2000.

39. J. M. Horowitz, V. A. Vernace, J. Myers et al., “Immunodetection of Parkin protein in vertebrate and invertebrate brains: a comparative study using specific antibodies,” Journal of Chemical Neuroanatomy, vol. 21, no. 1, pp. 75–93, 2001.

40. S. I. Kubo, T. Kitami, S. Noda et al., “Parkin is associated with cellular vesicles,” Journal of Neurochemistry, vol. 78, no. 1, pp. 42–54, 2001.

41. C. C. Stichel, M. Augustin, K. Kühn et al., “Parkin expression in the adult mouse brain,” European Journal of Neuroscience, vol. 12, no. 12, pp. 4181–4194, 2000.

42. C. Pawlyk, B. I. Giasson, D. M. Sampathu et al., “Novel monoclonal antibodies demonstrate biochemical variation of brain parkin with age,” The Journal of Biological Chemistry, vol. 278, no. 48, pp. 48120–48128, 2003.

43. P. Choi, N. Golts, H. Snyder et al., “Co-association of parkin and α-synuclein,” NeuroReport, vol. 12, no. 13, pp. 2839–2843, 2001.

44. Y. Imai, M. Soda, H. Inoue, N. Hattori, Y. Mizuno, and R. Takahashi, “An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin,” Cell, vol. 105, no. 7, pp. 891–902, 2001.

45. M. Zarate-Lagunes, W. Gu, V. Blanchard et al., “Parkin immunoreactivity in the brain of human and non-human primates: An immunohistochemical analysis in normal conditions and in parkinsonian syndromes,” The Journal of Comparative Neurology, vol. 432, no. 2, pp. 184–196, 2001.

46. B. A. van der Reijden, C. A. J. Erpelinck-Verschueren, B. Löwenberg, and J. H. Jansen, “TRIADs: a new class of proteins with a novel cysteine-rich signature,” Protein Science, vol. 8, no. 7, pp. 1557–1561, 1999.

47. M. Kasap, G. Akpinar, A. Sazci, H. A. Idrisoglu, and H. Vahaboğlu, “Evidence for the presence of full-length PARK2 mRNA and Parkin protein in human blood,” Neuroscience Letters, vol. 460, no. 3, pp. 196–200, 2009.

48. V. D'Agata, W. Zhao, A. Pascale, O. Zohar, G. Scapagnini, and S. Cavallaro, “Distribution of parkin in the adult rat brain,” Progress in Neuro-Psychopharmacology and Biological Psychiatry, vol. 26, no. 3, pp. 519–527, 2002.

49. J. Esteve-Rudd, L. Campello, M. Herrero, N. Cuenca, and J. Martín-Nieto, “Expression in the mammalian retina of parkin and UCH-L1, two components of the ubiquitin-proteasome system,” Brain Research, vol. 1352, pp. 70–82, 2010.

50. V. D'Agata, W. Zhao, and S. Cavallaro, “Cloning and distribution of the rat parkin mRNA,” Molecular Brain Research, vol. 75, no. 2, pp. 345–349, 2000.

51. S. M. Solano, D. W. Miller, S. J. Augood, A. B. Young, and J. B. Penney Jr., “Expression of alpha-synuclein, parkin, and ubiquitin carboxy-terminal hydrolase L1 mRNA in human brain: genes associated with familial Parkinson's disease,” Annals of Neurology, vol. 47, no. 2, pp. 201–210, 2000.

52. M. R. Cookson, P. J. Lockhart, C. McLendon, C. O'Farrell, M. Schlossmacher, and M. J. Farrer, “RING finger 1 mutations in Parkin produce altered localization of the protein,” Human Molecular Genetics, vol. 12, no. 22, pp. 2957–2965, 2003.

53. C. Hampe, H. Ardila-Osorio, M. Fournier, A. Brice, and O. Corti, “Biochemical analysis of Parkinson's disease-causing variants of Parkin, an E3 ubiquitin—protein ligase with monoubiquitylation capacity,” Human Molecular Genetics, vol. 15, no. 13, pp. 2059–2075, 2006.

54. S. R. Sriram, X. Li, H. S. Ko et al., “Familial-associated mutations differentially disrupt the solubility, localization, binding and ubiquitination properties of parkin,” Human Molecular Genetics, vol. 14, no. 17, pp. 2571–2586, 2005.

55. F. Darios, O. Corti, C. B. Lücking et al., “Parkin prevents mitochondrial swelling and cytochrome c release in mitochondria-dependent cell death,” Human Molecular Genetics, vol. 12, no. 5, pp. 517–526, 2003.

56. Y. Kuroda, T. Mitsui, M. Kunishige et al., “Parkin enhances mitochondrial biogenesis in proliferating cells,” Human Molecular Genetics, vol. 15, no. 6, pp. 883–895, 2006.

57. M. Farrer, P. Chan, R. Chen et al., “Lewy bodies and parkinsonism in families with parkin mutations,” Annals of Neurology, vol. 50, no. 3, pp. 293–300, 2001.

58. R. Hilker, C. Klein, M. Ghaemi et al., “Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene,” Annals of Neurology, vol. 49, no. 3, pp. 367–376, 2001.

59. C. Klein, P. P. Pramstaller, B. Kis, et al., “Parkin deletions in a family with adult-onset, tremor-dominant parkinsonism: expanding the phenotype,” Annals of Neurology, vol. 48, no. 1, pp. 65–71, 2000.

60. M. Maruyama, T. Ikeuchi, M. Saito, et al., “Novel mutations, pseudo-dominant inheritance, and possible familial affects in patients with autosomal recessive Juvenile Parkinsonism,” Annals of Neurology, vol. 48, no. 2, pp. 245–250, 2000.

61. West, M. Periquet, S. Lincoln et al., “Complex relationship between Parkin mutations and Parkinson disease,” American Journal of Medical Genetics—Neuropsychiatric Genetics, vol. 114, no. 5, pp. 584–591, 2002.

62. Shin and J. L. Manley, “Cell signalling and the control of pre-mRNA splicing,” Nature Reviews Molecular Cell Biology, vol. 5, no. 9, pp. 727–738, 2004.

63. Q. Pan, O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe, “Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing,” Nature Genetics, vol. 40, no. 12, pp. 1413–1415, 2008.

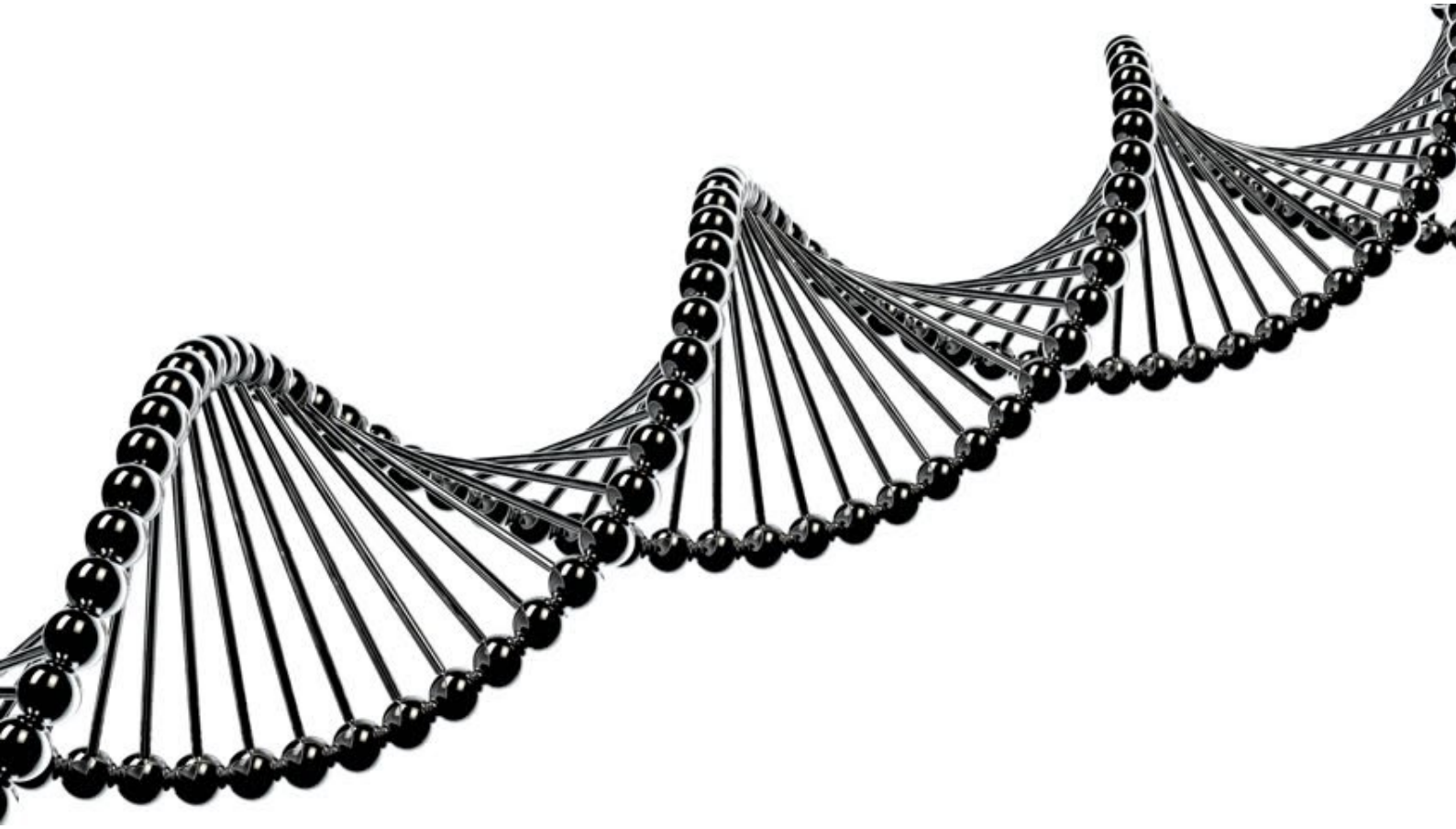
64. E. T. Wang, R. Sandberg, S. Luo et al., “Alternative isoform regulation in human tissue transcriptomes,” Nature, vol. 456, no. 7221, pp. 470–476, 2008.

65. G. S. Wang and T. A. Cooper, “Splicing in disease: disruption of the splicing code and the decoding machinery,” Nature Reviews Genetics, vol. 8, no. 10, pp. 749–761, 2007.

66. W. Springer, T. Hoppe, E. Schmidt, and R. Baumeister, “A Caenorhabditis elegans Parkin mutant with altered solubility couples α-synuclein aggregation to proteotoxic stress,” Human Molecular Genetics, vol. 14, no. 22, pp. 3407–3423, 2005.

CHAPTER 7

General discussion & Concluding remarks



GENERAL DISCUSSION

During the last years, the advent of innovative and extremely powerful methods for analyzing nucleic acids (DNA and RNA molecules) has produced a major boost in our understanding of the role of genes and genetic variability in biology and pathology, and is offering novel research and clinical opportunities for all fields of medicine (Olgiati, Quadri et al. 2016). While incredible successes have been achieved for mono and oligogenic diseases, for multigenic neurological disorders our genetic knowledge is still quite limited. A perfect example is represented by the complex genetic architecture of PD, which is characterized by a strongly familial inheritance firstly identified through pedigree-based linkage mapping analysis and more recently by genome-wide association studies (GWAS) (Ferreira and Massano 2016).

After many years of scientific research, what we currently know is that a large number of genetics factors is implicated in the etiology of PD, including highly penetrant Mendelian mutations, moderately-rare factors with moderate effect size, and variants with low effect size. The identification of such genetic factors has been very important, and it has allowed the scientific community to formulate novel hypotheses about the disease mechanisms. The small effect-size factors are very common, but have limited clinical relevance and, for these reasons, are poorly investigated. On the other side, high-risk alleles are often good diagnostic and prognostic markers, but they explain only small minorities of patients. All together, these variants explain only a small proportion of the heritable component of PD. Our comprehension about the genetic architecture of PD is still incompletely understood and a lot of work remains to do in order to identify additional sources of missing heritability or to assign a precise causal mechanism to the growing number of discovered loci.

With these premises in mind, we decided to investigate the genetics of PD by a different point of view, i.e. by evaluating the role of structural variations in the exonic regions of PD-linked genes at both genomic and transcriptomic level. To achieve this goal, we focused on two different but strictly connected cellular phenomena, copy number variations and alternative splicing.

Both genetic copy number changes and alternative splicing contribute to increase the complexity of the genome. CNVs represent a first order mechanism

because they act directly on the genome sequence, while alternative splicing is a "second order" mechanism since it operates directly on transcription control. It has been proposed that these two mechanisms had a synergistic effect on genomes evolution. One proposed theory discusses that in the beginning of the genome evolution, the genome organization was still relatively simple, and the first order mechanisms (such as duplication) might be preferred to effectively increase genome complexity. As genome complexity increased, the transcription/translation regulation network was getting too complex to allow a further increase of genome complexity by the first order mechanisms. Thus, the second-order regulation mechanisms (i.e., the alternative splicing) were starting to play a more important role (Jin, Kryukov et al. 2008). Thus, while CNVs provide the ingredients for the alternative splicing mechanisms, alternative splicing is able to utilize the complexity created through CNVs, and further multiplies it by producing a variety of products from the same source sequence.

Going beyond the purely evolutionary aspect of CNVs and alternative splicing, both of them doubtless contribute to enhance exonic variation and functional diversity by increasing the diversification of gene products and are able to regulate PD-related genes.

In Chapters 2 and 3, we emphasized the role of copy number variants in the pathogenesis of PD and showed the importance of rare individual variants and gene-dosage screening assays. Although it is clear that CNVs constitute a prevalent source of genomic variations and substantially participate to each individual's genomic make-up and phenotypic outcome, most of the GWAS have currently focused their attention on common variants (minor allele frequency, MAF>~1% in the general population) reaching a significant statistical association. We showed that the small-effect disregarded CNVs act on commonly deregulated biological processes relevant for PD pathogenesis and could be direct or indirect target of Mendelian genes. Additionally, GWAS are often based on commercially available genome-wide stations that are not able to characterize minimal copy number changes at exonic level. As a result, GWAS result in large genomic regions associated with the disease (not single genes nor exonic regions) and it is difficult to pinpoint the exact genomic trait responsible for the association. For this reason, we developed a customized method (the *NeuroArray* platform) able to detect both single and multi-exon CNVs in a large panel of PD-related

genes. This exon-centric aCGH platform could be useful for new GWAS in case-control population, or in investigating potentially pathogenic CNVs in familial pedigree.

In Chapters 4, 5 and 6, we focused on the alternative splicing regulation of PD-linked genes, with a special focus on *PARK2* gene. Given the impact of splicing in the regulation of PD-related genes, it is not so unlikely that the splicing alterations may contribute to PD onset or modulate disease severity. Moreover, recent discoveries in amyotrophic lateral sclerosis and frontotemporal lobar degeneration have demonstrated that impaired RNA processing can be directly involved in the onset of these neurodegenerative disorders (Kalbfuss, Mabon et al. 2001, Kar, Kuo et al. 2005, van Blitterswijk and Landers 2010, Polymenidou, Lagier-Tourenne et al. 2012, Arnold, Ling et al. 2013, Highley, Kirby et al. 2014). Lastly, in Alzheimer's disease, deposition of a long non-coding RNA involved in splicing has also been reported (Pastori and Wahlestedt 2012, Love, Hayden et al. 2015). All these data suggest that RNA anomalies, especially splicing defects, are probably more frequent than originally thought and need to be evaluated in PD. Studying the global dysregulation of the splicing mechanisms or specific alterations in individual genes could aid to better understand the selective degeneration of neurons of the *substantia nigra*, and provide interesting clues to identify new biomarkers and new therapeutic targets. It is therefore evident the necessity to take care of splicing phenomena, and for this task, the high-resolution NGS technologies such as the RNA sequencing (RNA-Seq) are available to delineate such perturbations.

In a near future, the application of NGS technologies will allow the study of rare variants or the alternative splicing regulation of PD genes. It is important to remember that the identification of disease-causing variants should also rely on the important biological information that can be gathered by studying their effect in *in vitro* and *in vivo* models. Most importantly, these studies are fundamental to understand the pathogenetic mechanism of the disease and to translate the genetic findings into novel therapeutic strategies.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The genetic findings of the past years helped us to understand better the pathogenetic mechanisms of neurological disorders, including PD. However, in order to fully understand the contribution of genetic architecture on PD, further work remains ahead. A complete knowledge of the genetic changes and splicing regulation of this pathology could be the cornerstone for mechanistic studies suggesting additional molecular pathways. A clear understanding of the molecular events underlying PD may provide hints for the development of novel and personalized therapies. In the coming years, the era of genomics technologies will fuel a new wave of genomic discovery, holding great promises for the PD field.

GENERAL LIST OF REFERENCES

Dauer, W. and S. Przedborski (2003). "Parkinson's disease: mechanisms and models." Neuron **39**(6): 889-909.

Kalinderi, K., S. Bostantjopoulou and L. Fidani (2016). "The genetic background of Parkinson's disease: current progress and future prospects." Acta Neurol Scand.

Jin, L., K. Kryukov, J. C. Clemente, T. Komiyama, Y. Suzuki, T. Imanishi, K. Ikeo and T. Gojobori (2008). "The evolutionary relationship between gene duplication and alternative splicing." Gene 427(1-2): 19-31.

Olgiati, S., M. Quadri and V. Bonifati (2016). "Genetics of movement disorders in the next-generation sequencing era." Mov Disord 31(4): 458-470.

Ferreira, M. and J. Massano (2016). "An updated review of Parkinson's disease genetics and clinicopathological correlations." Acta Neurol Scand.

Arnold, E. S., S. C. Ling, S. C. Huelga, C. Lagier-Tourenne, M. Polymenidou, D. Ditsworth, H. B. Kordasiewicz, M. McAlonis-Downes, O. Platoshyn, P. A. Parone, S. Da Cruz, K. M. Clutario, D. Swing, L. Tessarollo, M. Marsala, C. E. Shaw, G. W. Yeo and D. W. Cleveland (2013). "ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43." Proc Natl Acad Sci U S A 110(8): E736-745.

Highley, J. R., J. Kirby, J. A. Jansweijer, P. S. Webb, C. A. Hewamadduma, P. R. Heath, A. Higginbottom, R. Raman, L. Ferraiuolo, J. Cooper-Knock, C. J. McDermott, S. B. Wharton, P. J. Shaw and P. G. Ince (2014). "Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones." Neuropathol Appl Neurobiol 40(6): 670-685.

Kalbfuss, B., S. A. Mabon and T. Misteli (2001). "Correction of alternative splicing of tau in frontotemporal dementia and parkinsonism linked to chromosome 17." J Biol Chem 276(46): 42986-42993.

Kar, A., D. Kuo, R. He, J. Zhou and J. Y. Wu (2005). "Tau alternative splicing and frontotemporal dementia." Alzheimer Dis Assoc Disord 19 Suppl 1: S29-36.

Polymenidou, M., C. Lagier-Tourenne, K. R. Hutt, C. F. Bennett, D. W. Cleveland and G. W. Yeo (2012). "Misregulated RNA processing in amyotrophic lateral sclerosis." Brain Res 1462: 3-15.

van Blitterswijk, M. and J. E. Landers (2010). "RNA processing pathways in amyotrophic lateral sclerosis." Neurogenetics 11(3): 275-290.

Love, J. E., E. J. Hayden and T. T. Rohn (2015). "Alternative Splicing in Alzheimer's Disease." J Parkinsons Dis Alzheimers Dis 2(2).

Pastori, C. and C. Wahlestedt (2012). "Involvement of long noncoding RNAs in diseases affecting the central nervous system." RNA Biol 9(6): 860-870.

LIST OF PUBLICATIONS AND SCIENTIFIC CONTRIBUTIONS

Articles:

1. **La Cognata V.**, Maugeri G., D'Amico AG, Saccone S., Federico R., Cavallaro S., D'Agata V., "Differential expression of *PARK2* splice isoforms in a model, in vitro, of dopaminergic neurons exposed to various toxic insults mimicking Parkinson's disease", *under review*

2. Maugeri G., D'Amico AG., Rasà DM., **La Cognata V.**, Saccone S., Federico R., Cavallaro S., D'Agata V., "Caffeine prevents blood retinal barrier damage in a model, in vitro, of diabetic macular edema", J Cell Biochem. 2017 Jan 20

3. Morello G., Guarnaccia M., Spampinato AG., **La Cognata V**., D'Agata V., Cavallaro S., "Copy number variations in amyotrophic lateral sclerosis: piecing the mosaic tiles together through a systems biology approac". Mol Neurobiol (2017). doi:10.1007/s12035-017-0393-x

4. **La Cognata, V**., G. Morello, V. D'Agata and S. Cavallaro (2017). "Copy number variability in Parkinson's disease: assembling the puzzle through a systems biology approach." Hum Genet. 2017 Jan;136(1):13-37

5. **La Cognata V.**, Morello G., Gentile G., D'Agata V., Criscuolo C., Cavalcanti F., Cavallaro S., "A customized high-resolution array-comparative genomic hybridization to explore copy number variations in Parkinson's disease", Neurogenetics, 2016 Oct;17(4):233-244

6. **La Cognata V**, Cavallaro S., PARK7 (Parkinsonism associated deglycase), Card for the Atlas of Genetics and Cytogenetics in Oncology and Hematology, *in press*. On line version : http://AtlasGeneticsOncology.org/Genes/PARK7ID41639ch1p36.html

7. Maugeri G., **La Cognata V**., Scuderi S., D'Amico AG., Reitano R., Saccone S., Federico C., Cavallaro S. e D'Agata V., " Expression Profile of Human PARK2 Splicing Isoforms in Peripheral Blood Cells and Isolated Lymphomonocytes", ARC Journal of Neuroscience, Volume 1, Issue 3, Pagg. 1-9, Giugno 2016

8. **La Cognata V**., D'Agata V., Cavalcanti F. e Cavallaro S., "Parkin alternative splicing: not only Parkinsonism". ARC Journal of Neuroscience, Volume 1, Issue 1, Pagg. 9-18, Gennaio 2016 (https://www.arcjournals.org/journal-of-neuroscience/volume-1-issue-1)

9. **La Cognata V**., Cavallaro S., PARK2 (RBR E3 ubiquitin protein ligase), Card for the Atlas of Genetics and Cytogenetics in Oncology and Hematology, *in press*. Online version: http://AtlasGeneticsOncology.org/Genes/PARK2ID46408ch6q26.html

10. **La Cognata, V**.; D'Agata, V.; Cavalcanti, F.; Cavallaro, S., Splicing: is there an alternative contribution to Parkinson's disease? Neurogenetics. 2015 Oct;16(4):245-63. doi: 10.1007/s10048-015-0449-x. Epub 2015 May 16.

11. Scuderi, S.; **La Cognata, V**.; Drago, F.; Cavallaro, S.; D'Agata, V., Alternative splicing generates different parkin protein isoforms: evidences in human, rat, and mouse brain. BioMed research international 2014, 690796.

12. **La Cognata, V**.; Iemmolo, R.; D'Agata, V.; Scuderi, S.; Drago, F.; Zappia, M.; Cavallaro, S., Increasing the Coding Potential of Genomes through Alternative Splicing: The Case of PARK2 Gene. Current genomics 2014, 15 (3), 203-16.

13. Ruggeri M., Salpietro V., **La Cognata V**., Morello G., Gentile G., Concolino D., Polizzi A., Cavallaro S., Scienze "omiche" e biologia dei sistemi complessi: applicazioni in neurologia pediatrica, Neurologia Pediatrica, 2014

Chapters book:

1. **La Cognata V**., D'Agata V.,Cavalcanti F., Cavallaro S., "Genetics of Parkinson's disease: the role of Copy Number Variations", Chapter book for the "Parkinson's Disease" book, InTech publications, DOI: 10.5772/62881, August 2016

2. D'Agata V., **La Cognata V**., Cavallaro S. "Alternative splicing and neurological disorders: focus on Parkinson's disease". Opening chapter of the "Alternative Splicing and Disease" book, NovaScience Publishers, Chapter 1, pp. 1-38, November 2015

3. D'agata V., **La Cognata V**., Cavallaro S., "Genes and Memory", LA MEMORIA E LE MEMORIE, ATTI DELLE "GIORNATE INTERDISCIPLINARI DI BIOETICA", Aracne, Novembre 2014

Posters:

1. G. Gentile, **V. La Cognata**, G. Morello, M. Muglia, A. Magariello, A. Patitucci, M. Ruggieri, S. Cavallaro, "Una piattaforma array CGH custom per l'identificazione di CNV in pazienti affetti da Sindromi Neurocutanee ed altri disordini neurologici: validazione della regione d'interesse per la Neurofibromatosi di tipo 1", XIX Congresso Nazionale SIGU, Torino 23 - 26 Novembre 2016

2. **La Cognata V**., Morello G., Gentile G., D'Agata V., Cavallaro S., "A customized Comparative Genomic Hybridization array for the analysis of Copy Number Variations and exon-dosage anomalies in neurological disorders.", Society for Neuroscience 2016, San Diego (California), 12-16 November 2016

3. **La Cognata V**., D'Agata V., Cavalcanti F., Criscuolo C., Cavallaro S. "Una piattaforma array CGH personalizzata per analizzare Variazioni del Numero di Copie in pazienti con Morbo di Parkinson", 56° SNO (Società di Neurologia Ospedaliere), Maggio 2016, Catania

4. **La Cognata V**., Morello G., Gentile G., D'Agata V., Drago F., Cavallaro S., *NeuroArray*: a customized Comparative Genomic Hybridization array for the analysis of Copy Number Variations in neurological disorders, Summer School of Neuroscience 2015, Catania, Italy

5. Maugeri G., **La Cognata V.**, Scuderi S., Cavallaro S., D'Agata V., Expression of Parkin Isoforms in Human Lymphomonocytes, 69° Congresso Nazionale S.I.A.I. - Società Italiana di Anatomia e Istologia, 2015, Ferrara, Italy

Novel *PARK2* alternative spliced variants identified in human and rat, and deposited in GenBank:

1. Rattus norvegicus clone PP22 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774177.1)
2. Rattus norvegicus clone PP20 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774176.1)
3. Rattus norvegicus clone PP17 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774175.1)
4. Rattus norvegicus clone PP15 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774174.1)
5. Rattus norvegicus clone PP14 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774173.1)
6. Rattus norvegicus clone PP10 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774172.1)
7. Rattus norvegicus clone PP13 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774170.1)
8. Rattus norvegicus clone PP9 E3 ubiquitin-protein ligase parkin isoform (Park2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774169.1)
9. Homo sapiens clone PP24 PARK2 splice variant (PARK2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC357595.1)
10. Homo sapiens clone PP32 E3 ubiquitin-protein ligase parkin isoform (PARK2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC774171.1)
11. Homo sapiens clone PP47/PP50 PARK2 splice variant (PARK2) mRNA, complete cds, alternatively spliced (http://www.ncbi.nlm.nih.gov/nuccore/KC357594.1)