Genetics and epigenetics in Parkinson´s disease.

Luis Navarro Sánchez.

Thesis supervisor: Jordi Pérez Tur.
Unitat de Genètica Molecular,
Institut de Biomedicina de València, CSIC.

Doctoral studies: Biotechnology.
Facultat de Ciències Biològiques,
Universitat de València.

Tras preguntas y más preguntas y miedo y nervios y ansiedad y más ansiedad y bloqueos y periodos depresivos y tristeza y mucha, pero mucha, mierda… acabas. Tarde y reventado, pero acabas. Y lo único que quieres es pasar página y que esto no te pase *más nunca*.

¿Y?

Pues que te surgen más preguntas y más miedo y más nervios…. Y te das cuenta de que vives en un puto bucle científico que se está comiendo tu vida.

¿Y?

No lo sé, estoy en ello. Terapia y medicación, supongo. Aunque la autodestrucción y los *realities* son muy apetecibles.

En primer lugar, las formalidades.

Gracias Jordi por dejarme que trabajara en la UGM. A estas alturas, creo que los dos nos habremos planteado muchas veces si fue una buena decisión, pero ya da igual.

A continuación, el momento pasteleo, aunque si no sé dar un abrazo, no penséis que esto me va a salir fluido. Y menos aún, estando sobrio. Y que no se me enfade nadie si no he puesto su nombre o si no he escrito un parrafazo lacrimógeno y azucarado exaltando la belleza de nuestra relación. Si dejo que leas esto, significa que también eres importante para mí.

Lo mejor de todo este tiempo sois, de calle, vosotros y vosotras:

Bea, Carolina, Eva, Fer, LauraA, LauraG, María, Marta, Pili, Silvia, Vicente/Andreu y VicenteH. Mis amig@s, con l@s que he vivido, he sentido y he crecido…. y me he *desafasao*, sí, que seguro que estáis pensando en el dragón y/o en otros momentazos. Lo mejor de la UGM era su gente.

Ali, Isa, Lou, Elena, Manoli, Sole, Nuria, l@s narrativ@s... y muchas más. Personas maravillosas que forman parte de mi vida, lo cual no es fácil.

Y para acabar los agradecimientos:

Les doy las gracias a mis padres y a mi hermana, porque son los tres pilares fundamentales de mi vida. Son mi refugio (literalmente en estos últimos meses).

Esta tesis me la dedico a mí. Sí, a mí. Por lo malo, por los errores, por crecer, por saber perder, por aguantar, por terminar y por intentar aprender durante este largo camino tanto a nivel personal como profesional todo lo que he podido.

Hasta aquí llega la parte suculenta, donde enseño y escondo, como siempre.

Ahora llega la ciencia, a la que, aunque tenga mis dudas, creo que tengo algo que aportar, y de la que quiero vivir trabajando con dignidad.

P.D. Negaré tener cualquier relación con este texto; tengo una fachada que mantener.

Este trabajo pude llevarlo a cabo gracias a la beca JAEPre del CSIC que disfruté del 2008 al 2011. De no ser por ella, todo este proceso, ya difícil de por sí, lo hubiera sido muchísimo más.

ABBREVIATIONS.

- 5mC: 5-methylcytosine.
- 6-OHDA: 6-hydroxidopamine.
- AD: Alzheimer´s disease.
- ALS: amyotrophic lateral sclerosis.
- αSyn: α-synuclein.
- CBD: corticobasal degeneration.
- CI: confidence interval.
- C.I.: cognitive impairment.
- CMA: chaperone-mediated autophagy.
- df: degrees of freedom.
- DLB: dementia with Lewy bodies.
- ER: endoplasmic reticulum.
- FTD: frontotemporal dementia.
- FTLD: frontotemporal lobar degeneration.
- GD: Gaucher disease.
- GWAS: genome-wide association study.
- Hcy: homocysteine.
- HD: Huntington´s disease.
- HWE: Hardy-Weinberg equilibrium.
- LB: Lewy bodies.
- LD: linkage disequilibrium.
- L-dopa: levodopa.
- M.C.I.: mild cognitive impairment.
- MPTP: 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine.
- N.C.: normal cognition.

- NE: nuclear envelope.
- OR: odds ratio.
- PD: Parkinson´s disease.
- PSP: progressive supranuclear palsy.
- ROS: reactive oxygen species.
- SN: *substantia nigra*.
- SNpc: *substantia nigra pars compacta*.
- SNpr: *substantia nigra pars reticulata*.
- TSS: transcription start site.

INDEX.

"…the unhappy sufferer has considered
it has an evil, from the domination of
which he had no prospect of escape."

(James Parkinson, 1817)

In 1817, the British surgeon James Parkinson published "An essay on the shaking palsy" [1]. He described an unknown malady ("*it has not yet obtained a place in the classification of nosologists*") that he observed in 6 people. The features of the disease (*"tremor at rest, bradykinesia, tendency to fall…"*) were so recognizable that he was sure all patients were suffering from shaking palsy, even when they weren´t anatomically examined and were perhaps at different stages of the disease.

James Parkinson thought that, due to the parts of the body involved in the disease, there was a disordered state in some part of the medulla. He was wrong. However, almost two-hundred years after the publication of his book, most part of the reflections he pointed out are still useful:

"So slight and nearly imperceptible are the first inroads of this malady, and so extremely slow is its progress, that it rarely happens, that the patient can form any recollection of the precise period of its commencement. The first symptoms perceived are, a slight sense of weakness, with a proneness to trembling in some particular part; sometimes in the head, but most commonly in one of the hands and arms."

There are previous descriptions in literature of a disease similar to shaking palsy. Nevertheless, after the studies that the French neurologist Jean-Martin Charcot carried out between 1868 and 1881, the name shaking palsy (or paralysis agitans) changed to Parkinson´s disease (PD) on behalf of James Parkinson.

I. INTRODUCTION.

I.1. Definition of PD.

Parkinson´s disease is a progressive neurodegenerative disorder

- anatomopathologically defined by nigral degeneration and presence of numerous Lewy bodies[1] in surviving neurons [2],
- clinically defined by postural instability, tremor at rest, bradykinesia and rigidity.

PD can also be called **primary parkinsonism**[2]. There are other parkinsonisms:

- Secondary parkinsonism is due to environmental factors (drugs, toxins, head trauma, brain tumors…).
- In parkinsonism-plus syndromes (multiple system atrophy, progressive supranuclear palsy, diffuse Lewy body disease[3]…) there is parkinsonism plus other motoric neurologic features.
- In heredodegenerative disorders, parkinsonism is only one feature of a hereditary degenerative disorder: Alzheimer´s disease, Huntington´s disease, frontotemporal dementia… [3, 4]

[1] **Lewy bodies** are abnormal filamentous intracytoplasmic protein inclusions with α-synuclein as their major component. Other components are ubiquitin, phosphorylated neurofilaments, molecular chaperons…
These deposits are the predominant lesions in PD, dementia with Lewy bodies (DLB) and multiple system atrophy (MSA) although they can also be present in other diseases like Alzheimer´s disease, prion diseases… as secondary lesions.

[2] The term **parkinsonism** refers to a clinical syndrome comprising combinations of motor problems: postural instability, tremor at rest, bradykinesia, rigidity, flexed posture and the freezing phenomenon.

[3] Diffuse Lewy body disease is also called dementia with Lewy bodies (DLB).

I.2. Anatomical perspective.

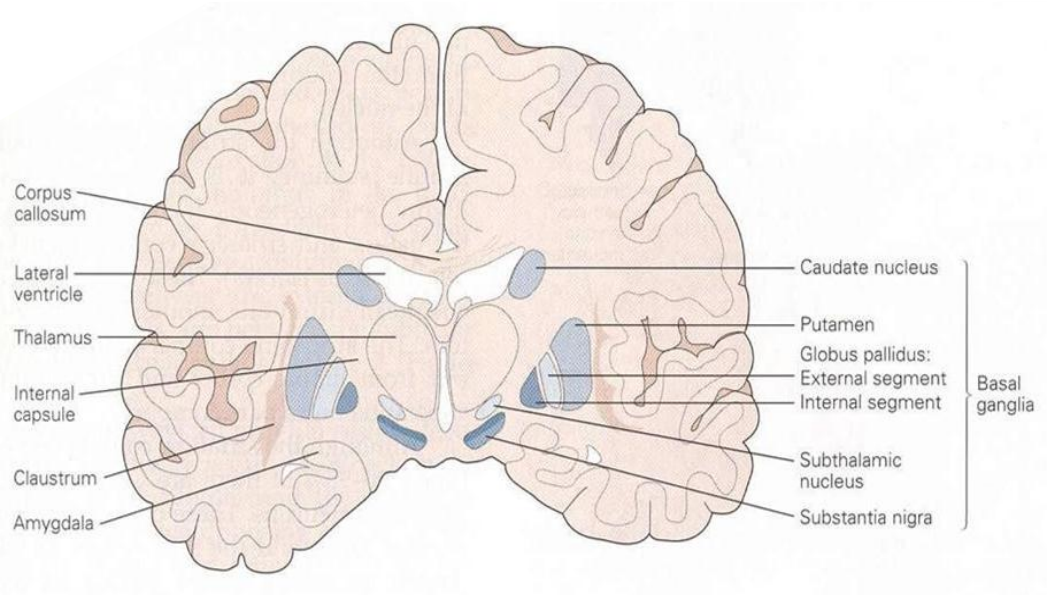


Fig. 1. Frontal section of a human brain.
Basal ganglia are formed by five different interconnected nuclei: *substantia nigra*, subthalamic nucleus, *globus pallidus, putamen* and caudate nucleus. Caudate nucleus and *putamen* form the **striatum (neostriatum)**.

Modified from [5]

The ***substantia nigra*** **(SN)** is a large midbrain structure which can be divided into two different parts: the *substantia nigra pars compacta* (SNpc) and the *substantia nigra pars reticulata* (SNpr).

SNpc is a densely populated area formed by neurons that contain dopamine and neuromelanin (this pigment gives the characteristic dark color to the region); whereas SNpr is a cell-sparse portion, located ventrally to the SNpc and formed by GABAergic neurons.

SN, as a component of basal ganglia, takes part in the circuits that link the five nuclei shown in Figure 1 to the cortex and thalamus to control voluntary movement execution by influencing its planning and initiation. SNpc and SNpr have very different connections and functions in that process. It is called the **direct and indirect pathway model**.

How does it work in normal conditions (healthy people)?

The striatum, the main *input* nucleus of the circuit, transmits the flow of information received from the cortex to the *output* nuclei (SNpr and internal *globus pallidus*, GPi). The transmission is mediated by two different groups of striatal neurons that express either the D_1 dopamine receptor (direct pathway) or the D_2 dopamine receptor (indirect pathway). In the direct pathway, GABAergic striatal neurons project to SNpr and GPi, whereas in the indirect pathway, the route is more complicated and D_2-expressing neurons project to the external *globus pallidus* (GPe), which sends GABAergic projections to the subthalamic nucleus that, on its turn, sends its glutamatergic efferents to GPi/SNpr. The loop is closed via the GABAergic projections that connect the *output* nuclei to the ventral lateral and ventral anterior nuclei of the motor thalamus which finally sends its glutamatergic efferents to the cortex [6] (Figure 2).

The striatum receives dopamine from the SNpc. This dopamine has opposite effects in D_1 and D_2-expressing neurons:

- striatal D_1 receptors are activated by dopamine and, consequently, the GABAergic transmission to the *output* nuclei via the direct pathway is increased: GPi/SNpr are less active via the direct pathway.

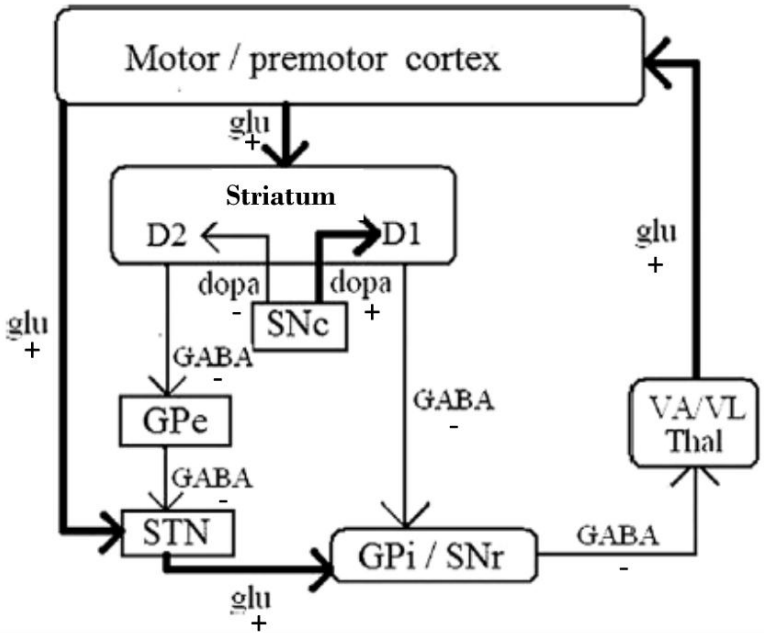


Fig. 2. Direct and indirect pathway model in healthy people.
Excitatory projections are depicted as thick black arrows and + symbols, and inhibitory projections as thin black arrows and - symbols. GPe: external *globus pallidus*; GPi: internal *globus pallidus*; SNc: *substantia nigra pars compacta*; SNr: *substantia nigra pars reticulata*; STN: subthalamic nucleus; VA/VL Thal: ventral anterior and ventrolateral nuclei of the thalamus; D1: dopaminergic receptor type 1; D2: dopaminergic receptor type 2; dopa: dopamine; glu: glutamate; GABA: gamma-aminobutyric acid.

Modified from [7]

- striatal D_2 receptors are inhibited by dopamine and, therefore, the GABAergic transmission to GPe decreases. Being less inhibited, GPe transmits more GABA to the subthalamic nucleus which is inhibited and decreases its glutamatergic transport to

the *output* nuclei: GPi/SNpr are less active via the indirect pathway.

Both mechanisms have the same consequence, the *output* nuclei is less active and, as a consequence, there is disinhibition of the thalamus and activation of the cortex. That means that movement is initiated.

What happens in PD patients?

However, in Parkinson´s disease there is loss of dopamine-containing neurons in the *substantia nigra pars compacta*. Due to that fact, the *output* nuclei become hyperactive and such hyperactivity is sustained by the enhanced glutamatergic inputs that GPi/SNpr receive from the subthalamic nucleus. Therefore, the ventral lateral and ventral anterior nuclei of the motor thalamus are inhibited and, thus, motor cortex and prefrontal cortex are less active (Figure 3). As a result, PD patients lose their ability to control voluntary movements.

Although the profound changes in the circuit basal ganglia-thalamus-cortex explain the difficulty to initiate movements, the cause for the tremor and rigidity is less clear.

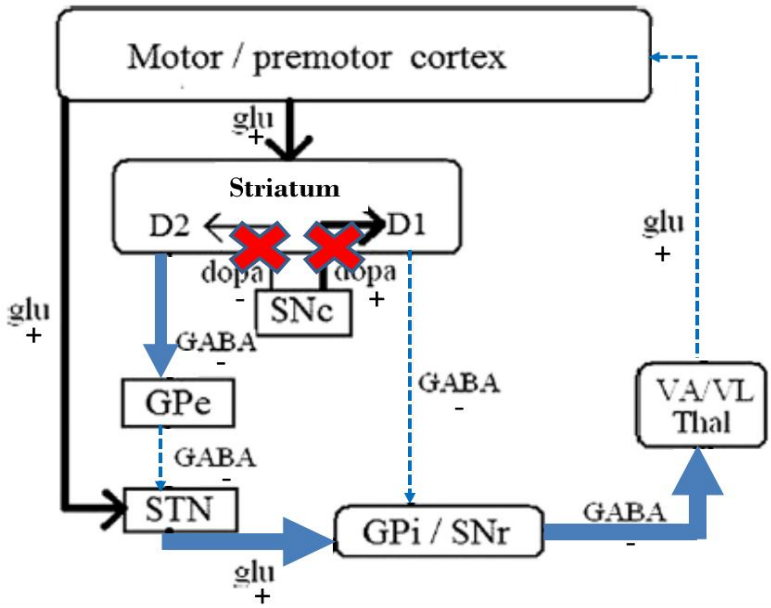


Fig. 3. Direct and indirect pathway model in a PD patient.
Now, excitatory projections are depicted by + symbol, and inhibitory projections by - symbol. The width of the arrow does not mean either excitatory or inhibitory. Increased transmissions are illustrated by thick blue arrows whereas decreased transmissions are illustrated by dashed blue arrows. GPe: external *globus pallidus*; GPi: internal *globus pallidus*; SNc: *substantia nigra pars compacta*; SNr: *substantia nigra pars reticulata*; STN: subthalamic nucleus; VA/VL Thal: ventral anterior and ventrolateral nuclei of the thalamus; D1: dopaminergic receptor type 1; D2: dopaminergic receptor type 2; dopa: dopamine; glu: glutamate; GABA: gamma-aminobutyric acid.

Modified from [7]

SNpc can be subdivided into different areas, although the definition and terminology of these subregions vary considerably. Based on [8], [9] and [10], [11] split the *substantia nigra pars compacta* in four main regions (Figure 4):

- A small *pars medialis* (medioventral group).
- A small *pars lateralis* (nigrosome 3).
- A large dorsal tier (nigrosomes 4 and 5 and the matrix).
- A large ventral tier (nigrosomes 1 and 2).

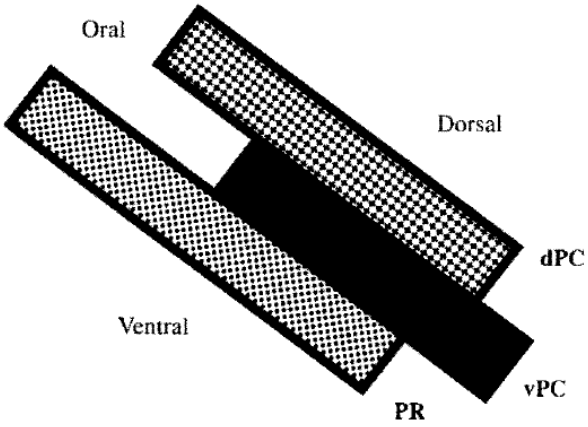


Fig. 4. *Substantia nigra* tiers. Sagittal representation of the *substantia nigra*. PR = *pars reticulata* (dots). vPC ventral tier *pars compacta* (black). dPC dorsal tier *pars compacta* (crosses).

Taken from [8]

The terms nigrosome and matrix are defined in [9] and [10] using calbindin D_{28K} immunostaining. SNpc shows two nigral compartments: the matrix, which is a wide calbindin-rich region and contains the 60% of dopamine-containing neurons, and the nigrosomes, which are small invaginated calbindin-poor pockets embedded in the matrix, that contain the remaining 40%.

In PD patients nigral death follows a strict order, beginning in nigrosome 1 and spreading to nigrosomes 2, 4, 3, 5 and finally to the matrix. Depletion begins in the main pocket (nigrosome 1) and then goes to other nigrosomes and the matrix along rostral, medial and dorsal axes of progression.

For this reason, the ventral tier (70-90% of neuronal loss) is considerably more affected than the dorsal tier (25-70% of neuronal loss) (Figure 5) [11]. This neuronal pattern of death in Parkinson´s disease is completely opposite to the observed in aging, where the dorsal tier is more affected than the ventral tier.

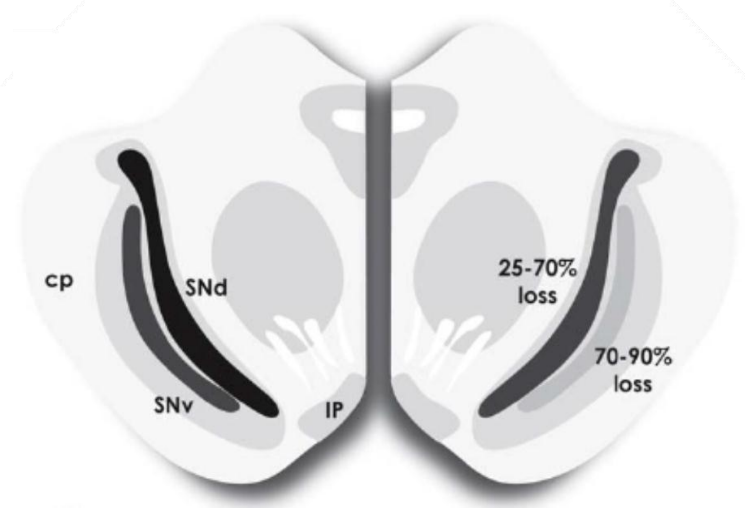


Fig. 5. Schematic representation of the variability within the SN in the susceptibility to degeneration.
Schematic representation of the position of the dorsal (SNd) and ventral (SNv) tiers of the human SN (left) and the variable degree of neurodegeneration in these tiers in patients with Parkinson's disease (right) (cp = cerebral peduncle, IP = interpeduncular nucleus, red nucleus, exiting 3rd nerve fibres and periaqueductal grey are highlighted).

Taken from [11].

Despite this great loss, initially, there is an increased efficiency of residual dopamine-containing neurons which results in an increased release of dopamine at the striatal level that mask PD symptoms in an early phase of the disease. There are no clinical signs until the percentage of nigral death is superior to 50% and the loss of striatal dopamine reaches the 80% [6, 8].

The loss of dopamine-containing neurons in SNpc is spatially driven and increases with the duration in PD. Nevertheless, there are other factors than nigral location that control neuronal survival because, after decades of degeneration, there still are neurons alive in the *substantia nigra pars compacta* and other basal ganglia and brain areas are affected too.

I.3. Clinical perspective.

The four main clinical features in Parkinson´s disease are motor symptoms:

- **Postural instability;** the loss of balance makes the individual feel unsteady.

- **Tremor,** defined as involuntary movements that are rhythmic or fairly rhythmic, whether of large amplitude or barely visible. This tremor is only present at rest, i.e. it is reduced or eliminated by movement. It is distal and in one or more planes and it is present in hands, legs, tongue, lips and lightly closed eyes.

- **Rigidity** appears throughout the full range of movement in limbs, trunk and neck.

- **Bradykinesia** (slowness of movement). It encompasses difficulties with planning, initiating and executing movement and with performing sequential and simultaneous tasks. As a consequence, affected people show:
  - Hypomimia (reduced degree of facial expression) and reduced blinking.
  - Dysarthria (poorly articulated speech, with quiet and monotonous voices).
  - Dysphagia (difficulty in swallowing).
  - Shuffling and slow gait (reduced arm swing while walking).

- Problems to do some fine tasks like buttoning, using utensils or writing. In this respect, micrographia, handwriting that decreases in size from normal to minute, is common.

Tremor is the most embarrassing trait for PD patients although rigidity and bradykinesia are more disabling. Any of these features may occur in isolation or in any combination and their onset can be unilateral or bilateral (symmetric or asymmetric) [12]. It is noteworthy that neuronal loss occurs contralaterally to the affected body side.

There are other clinical signs that can also be present in PD patients. These signs can be divided in:

- Secondary motor symptoms, as for example:
  - Sialorrhoea (excessive secretion of saliva).
  - Festination (involuntary tendency to increase speed of gait).
  - Freezing (sudden and transient inability to move; it is a common cause of fall).
  - Dystonia.

- Non-motor symptoms, which can be classified into different categories:
  - Sleep disorders: vivid dreams where PD patients kick, grab, swear or punch while sleeping, daytime drowsiness, sleep fragmentation, REM behavior disorder…
  - Sensory abnormalities: pain, anosmia (loss of sense of smell), paresthesias (sensation of tingling, burning, pricking or

numbness of a person´s skin with no apparent physical effect)…

- Cognitive/neurobehavioral abnormalities: cognitive impairment, bradyphrenia (slowness of thinking), word finding difficulties (tip-of-the-tongue phenomenon), depression, apathy, anhedonia (loss of sensation of pleasure), dementia, obsessive-compulsive disorders (like punding, an intense fascination with repetitive handling and examining of mechanical objects)…
- Autonomic dysfunctions: constipation, urinary and sexual dysfunction, abnormal sweating, seborrhea, weight loss… [13]

Non-motor symptoms are very common. In fact, up to 60% of patients suffer from more than one, and 25% have four or more. They can be due to the progression of the disease itself that affects other parts of the brain or to its treatment [14].

If the onset of the disease is unilateral, the initial body side commonly remains more affected than the later-involved side. In addition, it has been shown that there are also differences in the cognitive and perceptual functions of PD patients due to this differential beginning. The hemispheres of the brain are specialized in different tasks, being the right hemisphere, that controls the left body side, related to visuospatial function, and the left hemisphere, that controls the right body side, related to verbally-based actions. Therefore, PD patients will show different impairment on visual memory or poorer verbal memory performance depending on where the disease started [15].

I.4. Diagnosis.

There is no conclusive test for diagnosis in Parkinson´s disease. The only "definite" way to demonstrate that a person is suffering from Parkinson´s disease is a *post-mortem* analysis of the brain to determine nigral degeneration and Lewy body presence.

Despite this, the presence of a combination of some typical motor symptoms is the common way to identify PD. Although it will not be definite, its accuracy will improve with time and repeated assessments.

The diagnosis follows some steps:

- Step **1**: Diagnosis of parkinsonian syndrome.

  There is bradykinesia and at least one of the following: muscular rigidity, tremor at rest or postural instability not caused by primary visual, vestibular, cerebellar or proprioceptive dysfunction.

- Step **2**: Exclusion criteria for Parkinson´s disease.

  The presence of any of the following implies that the disease is not PD: history of repeated strokes with stepwise progression of parkinsonian features, history of repeated head injury, history of definite encephalitis, sustained remission, strictly unilateral features after 3 years, supranuclear gaze palsy, cerebellar signs, early severe autonomic involvement, early severe dementia, presence of a cerebral tumour, MPTP exposure or negative response to large doses of L-dopa (if malabsorption excluded).

- Step **3**: Supportive prospective positive criteria of Parkinson´s disease.

> When the patient has three or more of the following, it is more probable that it is suffering PD: tremor at rest, unilateral onset, progressive disorder, hyposmia, visual hallucination, L-dopa response for 5 years or more, persistent asymmetry affecting the side onset most, clinical course of 10 years or more [16].

Based upon the analysis of brain tissue from patients at different stages of the disease, [17] proposed the most suitable scheme at the present time to explain the onset and the evolution of symptoms in PD patients. Nevertheless, it might be taken into account that at least 15% of PD patients do not follow it. In this model, known as **Braak stages,** neuronal damage follows a predetermined sequence (Table 1, Figure 6).

Table 1. Stages in the evolution of PD-related pathology.

Stage **1**	medulla oblongata	Lesions in the dorsal IX/X motor nucleus and/or intermediate reticular zone and, frequently, in the anterior olfactory nucleus.
Stage **2**	medulla oblongata and pontine tegmentum	Pathology of stage 1 plus lesions in caudal raphe nuclei, gigantocellular reticular nucleus, and coeruleus–subcoeruleus complex.
Stage **3**	midbrain	Pathology of stage 2 plus midbrain lesions, in particular in the *pars compacta* of the *substantia nigra*.
Stage **4**	basal prosencephalon and mesocortex	Pathology of stage 3 plus prosencephalic lesions. Cortical involvement is confined to the temporal mesocortex (transentorhinal region) and allocortex (CA2-plexus).
Stage **5**	neocortex	Pathology of stage 4 plus lesions in high order sensory association areas of the neocortex and prefrontal neocortex.
Stage **6**	neocortex	Pathology of stage 5 plus lesions in first order sensory association areas of the neocortex and premotor areas, occasionally mild changes in primary sensory areas and the primary motor field.

Taken from [17]

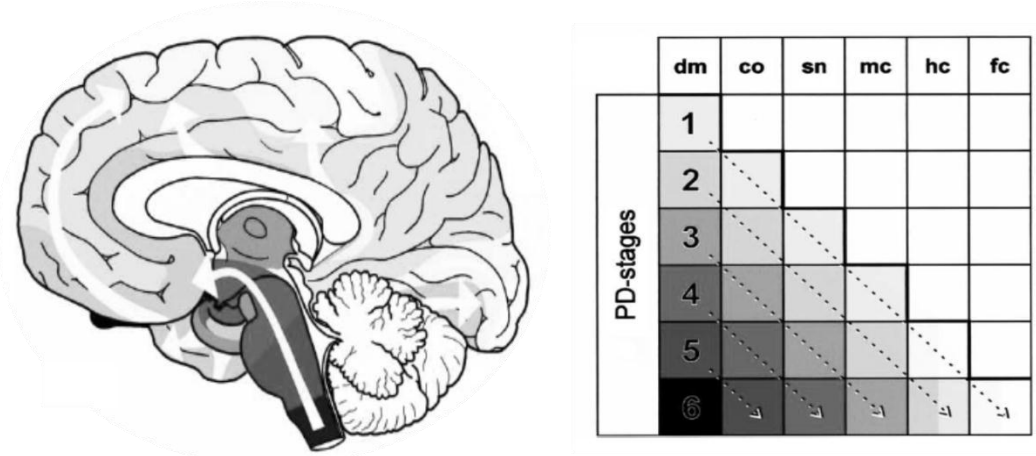


Fig. 6. Progression of PD.
Both representations show how Parkinson´s disease evolves according to Braak stages.
dm, dorsal motor nucleus of the glossopharyngeal and vagal nerves; co, coeruleus–subcoeruleus complex; sn, *substantia nigra*; mc, anteromedial temporal mesocortex; hc, high order sensory association areas and prefrontal fields; fc, first order sensory association areas, premotor areas, as well as primary sensory and motor fields.

Modified from [17]

The current clinical diagnosis identifies the malady once it has evolved to, at least, stage 3, when there is damage in the *substantia nigra* and the motor symptoms begin. It would be very useful to identify the disease in its presymptomatic phase, stages 1 and 2, which is characterized by hyposmia, depression and sleep disorders. The identification of this silent phase is one of the most challenging aspects of the malady because if it could be recognized, there would be the possibility to administer a neuroprotective therapy to PD patients to delay the beginning of motor symptoms. Although there have been some attempts to find biomarkers[4] [18], those assessed do not add relevant information to our knowledge of PD initiation and progression to date. It has also been postulated that some genes change its expression in blood in the

[4] Biomarker or biological marker: a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.

early phase or Parkinson´s disease [19]. However, there is still controversy about the number and identity of those genes.

Once the clinical diagnosis has been done, there are some rating scales to evaluate the motor impairment and disability in PD patients: The Hoehn and Yahr scale, The Unified Parkinson´s Disease Rating Scale (UPDRS), The Short Parkinson Evaluation Scale (SPES) and the SCales for Outcomes in PArkinson´s disease (SCOPA).

Most of these scales have not been fully evaluated for validity and reliability. However, the Hoehn and Yahr scale is one of the most widely used when rating the evolution of the disease.

I.5. Treatment.

Parkinson´s disease is incurable and pharmacological and non-pharmacological treatments only improve quality of life and functional capacity temporarily. Treatment is symptomatic, not neuroprotective, and does not modify the progression of the disease. For that reason, nowadays, it is necessary to find new and more effective treatments without disabling adverse effects.

- **I.5.a. Pharmacological therapy**: it is the most common option. These agents only improve motor symptoms. Adverse effects may be due to presynaptic and postsynaptic adaptations in neurotransmitter and receptor interactions or in neurotransmitter release or in signaling cascades [20]. It is noteworthy that dopamine cannot be used as a therapeutic agent because it does not cross the blood-brain barrier (Figures 7 and 8).

  i. Precursor of dopamine: levodopa (L-dopa). More than 40 years after its discovery, it still is the best option for treating PD. After 5 years of treatment more than 50% of patients develop adverse effects: acute side effects (nausea, tachycardia) and chronic side effects (obsessive-compulsive disorders -punding, gambling, compulsive shopping, hypersexuality, compulsive eating and compulsive medication use- and motor complications -dyskinesias and motor fluctuations: wearing off, on-off[5]-) [21].

[5] Dyskinesia: involuntary, purposeless, irregular but sometimes repetitive movements.

ii. Dopamine agonists: bromocriptine, pergolide, pramipexole, ropinirole, apomorphine, cabergoline. They directly stimulate dopamine receptors by decreasing presynaptic dopamine synthesis. They may be used alone to delay the need for levodopa or may be used with levodopa to increase their effectiveness. Complications: hallucinations, confusion, drowsiness, psychosis.

iii. Peripheral decarboxylase inhibitors: carbidopa, benserazide. Mainly used with levodopa to increase the percentage of L-dopa that enters to the brain (from 5% without them to 25% when using these inhibitors).

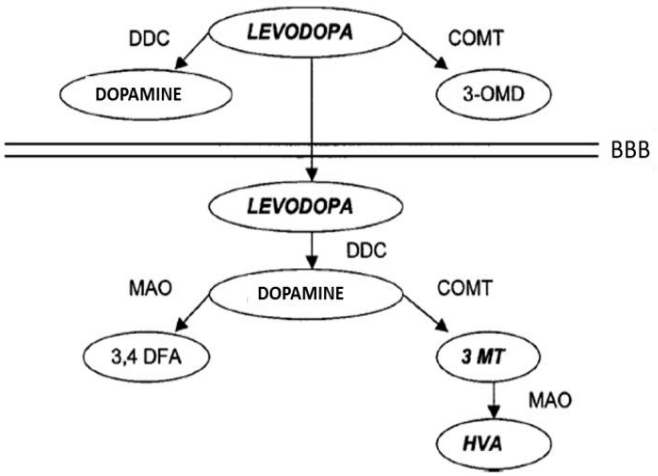


Fig. 7. Metabolism of L-dopa.
DDC: L-dopa decarboxylase; COMT: catechol-O-methyltransferase; 3-OMD: 3-O-methyldopa; MAO: monoamine oxidase; 3,4 DFA: 3,4-dihydrophenylacetic acid; 3 MT: 3-methoxytramine; HVA: homovanillic acid; BBB: blood-brain barrier.

Modified from [22]

iv. Dopamine releasers: amantadine. Side effects: restlessness, depression, confusion and hallucinations.

Wearing off: the benefits of a levodopa dose fade off gradually and do not last until the next dose.

On-off: sudden, sometimes unpredictable, changes in PD patient´s symptoms.

v. Catechol-O-methyltransferase (COMT) inhibitors: tolcapone, entacapone. Mainly used in combination with L-dopa because they increase its half-life. Complications: sleep disturbances, insomnia, confusion and dyskinesia.

vi. MAOB inhibitors: selegiline, rasagiline. Prevent *in vivo* metabolism of dopamine, i.e. they prolong the action of dopamine at the synapse. They are mainly used in combination with L-dopa to enhance its antiparkinsonian effect thus allowing a reduction in the dose of levodopa.

vii. Anticholinergics: trihexyphenidyl, benztropine. They are especially effective against tremor. Side effects: confusion, agitation, hallucinations and drowsiness [23].

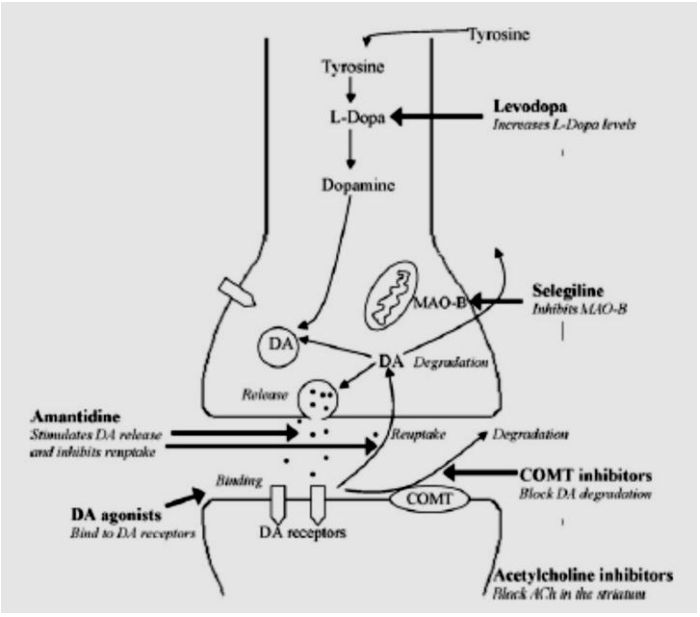


Fig. 8. Sites of action of pharmacological therapies. Taken from [23]
Effect of pharmacological therapies on dopamine (DA) synthesis, release, reuptake and degradation.

There is some controversy between neurologists about which pharmacological agents should be prescribed and when and how much should be administered. The main question in the current treatment approaches is finding the appropriate time to include L-dopa in the therapy because, although levodopa is the most effective compound to treat Parkinson´s disease, there are some doubts about its toxicity and adverse effects. For that reason, the common option at the present time, especially if patient is younger than 65 years old, is to start the treatment with dopamine agonists till they cannot control motor symptoms. At that point, L-dopa is included in the therapy.

There have been some studies in the last years to clarify the relation between L-dopa, toxicity and dyskinesia:

- In *in vitro* studies, high concentrations of levodopa cause degeneration of cultured dopaminergic neurons; however, there are neither *in vivo* animal models nor PD patients studies that show harm in neurons after L-dopa treatment [24, 25].
    - Neuronal death in *in vitro* studies can be explained by the absence of glial cells and trophic factors (always present in *in vivo* systems) that could protect cells against harmful effects of reactive oxygen species, ROS, originated from the auto-oxidation of levodopa [24].
    - The ELLDOPA study [26], one of the most important trials carried out in the last years to explain the effect of levodopa in the progression of PD disease, concluded that there was a

significant positive effect in the evolution of patients who received levodopa but neuroimaging techniques showed opposite results. The authors of the study proposed that levodopa did not increase neuronal death, but it interfered in marker binding.

- The cause of motor complications can be related to the abnormal pulsatile stimulation that the dopamine receptors receive during treatment. In the normal brain, the striatal dopamine levels and the activation of the dopamine receptors on striatal neurons remain approximately constant along the day, but this situation changes as the disease progresses because there is lack of striatal dopamine and the quantity of dopamine depends on external levodopa. As a consequence, dopamine receptors are exposed to alternating high and low dopamine concentrations (depending on L-dopa dose) and it is more and more complicated to buffer uptake and release of dopamine to maintain its concentration stable [25]. It is also thought that dyskinesia is an integral part of the antiparkinsonian response originated by L-dopa treatment and that this pulsatile stimulation could originate the beginning of the motor complications and the reduction of the positive effects of levodopa (Figure 9) [27]. A continuous infusion of L-dopa is the solution, but this is impractical.

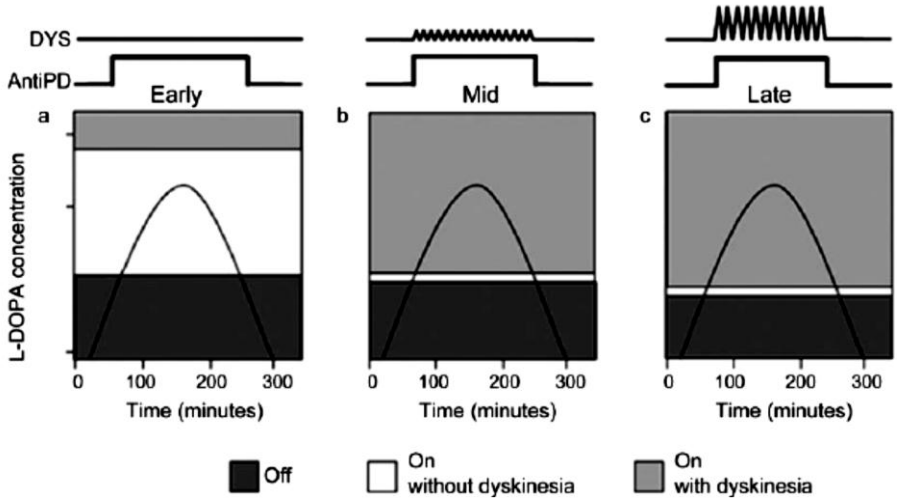


Fig. 9. Diagram illustrating how dyskinesia emerges during long-term levodopa therapy.

Panels illustrate early (a), mid (b), and late (c) stages of long-term levodopa therapy. The lines above the panels illustrate when the antiparkinsonian response (AntiPD) and dyskinesia (DYS) occur during the dose cycle.

The magnitude of the dyskinesia response progressively increases from a subtle effect in early treatment to a more disabling form. The threshold for dyskinesia and the antiparkinsonian response are identical when dyskinesia first appears and both progressively decrease during long-duration levodopa therapy.

Taken from [27]

- **I.5.b. Non-pharmacological therapy**: it includes approaches that are less frequently used.

  i. Cell therapy: in the late 80s, human fetal mesencephalic tissue was used for grafting in PD patients to recover the striatum. Although these initial trials had promising results, some patients showed dyskinesia and long term studies about the benefits should have been done. Lately, cell replacement has almost been blocked due to ethical problems related to the use of

human fetuses, political problems such as changes in legislation about stem cells and funding [28], and technical problems as human fetuses are a limited cell source (it is necessary to use at least 6 for a single patient), there is poor rate of transplanted cell survival -5 to 10%- and there is risk of uncontrolled cellular development -teratomas-, so dopaminergic cells must be completely differentiated *in vitro* before the transplant [29]. To avoid some of these problems, current research is focused on finding new sources of dopaminergic neurons: embryonic stem cells, adult neural stem cells, iPS[6] cells in other tissues (blood, bone marrow, skin or umbilical cord blood), foreign cells which produce dopamine or GDNF… [30].

ii. Surgery: surgical procedures are only considered for people with intolerable adverse side effects from medication and without mental disorders. It means that only about 5 to 10% of PD patients are good candidates. Long term studies are necessary to establish the longevity of the benefits, that can range from almost imperceptible to nearly complete. In any case, after surgery it is necessary to adjust the medications. There are two types of procedures: lesioning (irreversible destruction of hyperactive neurons; -otomy) or stimulation (reversible stunning of hyperactive neurons; DBS or deep brain stimulation). The parts of the brain that are subject of surgery are the thalamus (ventral intermediate nucleus), the internal *globus pallidus* and

[6] iPS: Induced pluripotent stem.

the subthalamic nucleus. There can be side effects: infections, brain hemorrhage, seizures and even death (Table 2, Figure 10) [31].

Table 2. Surgery selection by primary problem.

Primary problem	Type of surgery
Severe tremor: unilateral	Thalamotomy Thalamic stimulation
Sever tremor: bilateral	Bilateral thalamic stimulation Thalamotomy and thalamic stimulation Bilateral subthalamic stimulation
Severe bradykinesia, rigidity: unilateral	Pallidotomy
Levodopa-induced dyskinesia	Pallidotomy
Prominent motor fluctuations	Pallidotomy
Severe bradykinesia, rigidity: bilateral	Bilateral subthalamic stimulation Bilateral pallidal stimulation Billateral pallidotomy
Falling, postural instability, unintelligible speech	None

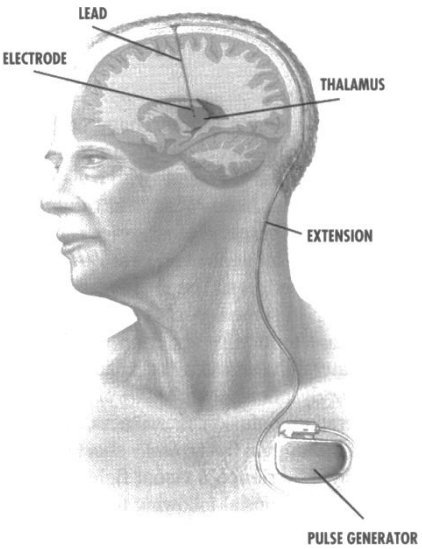


Fig. 10. Example of DBS: Thalamic deep brain stimulaton.

Modified from [31]

iii. Therapy with trophic factors[7] [32]: GDNF is one of the most powerful dopaminergic neurotrophic factors. It protects dopaminergic cells against toxins: in animal PD models it is neuroprotective and neuroregenerative against MPTP in mice and 6-OHDA in rats. The main problem in this approach is how to maintain and to transport GDNF to the appropriate cell type

[7] Trophic factors are important proteins for the survival and function of specific cellular subpopulatons.

[33]. Nevertheless, its therapeutic effect in humans is controversial because some recent studies have not observed significant improvement in PD patients after GDNF intra-putamenal treatment [34].

iv. Other options in development: PD vaccine (to stimulate immune system against the abnormal form of αSyn), gene therapy, controlled doses of nicotine, caffeine, coenzyme Q_{10},…

I.6. Incidence and prevalence.

The term incidence refers to the number of new cases of patients diagnosed on a given period of time, usually each year, whereas the term prevalence refers to the estimated population of individuals who are managing a disease at any given time [35].

Incidence rates are theoretically not affected by differences in survival of patients and therefore better measures of the risk of disease than prevalence estimates [36].

In Parkinson´s disease, both parameters increase with age and are influenced by ethnicity (Asian population is less affected by PD) [35, 37]. Nevertheless, the effect of gender is controversial: although it seems that men are more affected than women, results are not always significant. The protective effect of estrogens is still on debate [36, 38].

The values of incidence in PD differ in various orders depending on the type of study: in cohort studies [38] (Figure 11), the evolution of the population is followed for a period of time during which all participants are interviewed, screened and diagnosed by specialists, but in studies based on medical records [37] (Figure 12) only patients which have visited the hospital, have a medical record and have been diagnosed are quantified. Consequently, cohort studies report higher incidence values.

Prevalence of Parkinson´s disease is ≈ 1% at the age of 65 and increases to 4-5% by the age of 85: PD is the second most common neurodegenerative disorder (Figure 13).

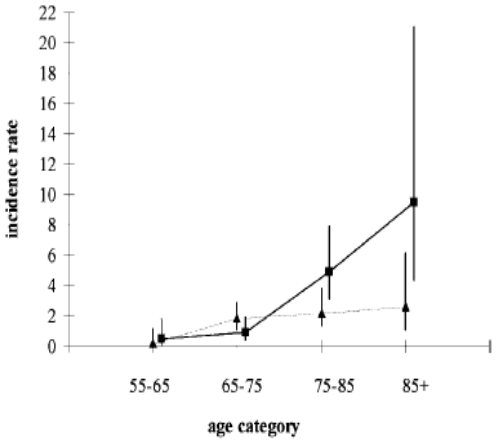


Fig. 11. Incidence rates per 1000 person-years of PD by age and gender: men (□) and women (Δ). $CI_{95\%}$ represented by vertical lines.

Taken from [38]

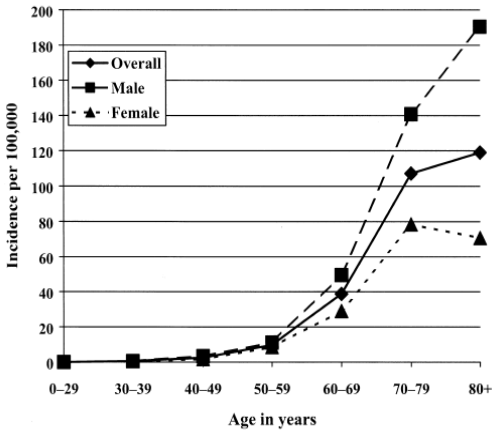


Fig. 12. Incidence rates per 100000 person-years of PD by age and gender.

Taken from [37]

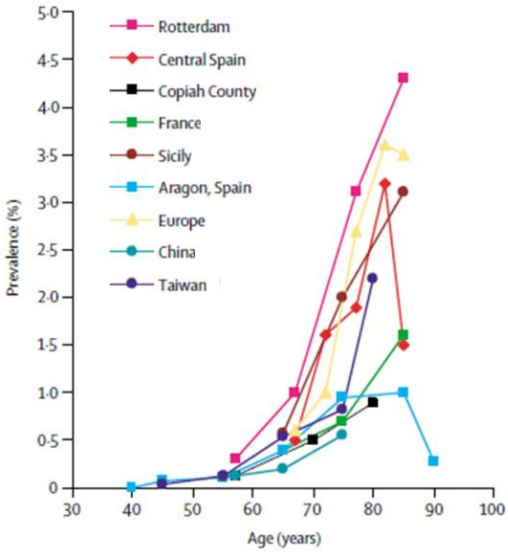


Fig. 13. Population-based prevalence studies of Parkinson´s disease.

Modified from [36]

I.7. Etiology.

At the present time, the etiology of Parkinson´s disease remains unknown.

The factor or combination of factors that trigger the neurodegenerative process need to be elucidated yet. Environment and/or genetics have been postulated as those factors. However, their role has been changing for the last 30 years.

The specific pathological pathway is also unknown: oxidative stress[8], mitochondrial dysfunction, protein aggregation, inflammation, altered Ca^{2+} homeostasis… are some of the processes that seem to be altered in PD although their specific role (cause or consequence) is not clear.

I.7.a. Environmental factors.

In 1983, four addict people developed severe parkinsonism just one week after injecting intravenously the first dose of a new "synthetic heroine" in northern California. This drug contained **MPTP**[9] plus variable amounts of MPPP[10]. These people showed the typical PD symptoms (flexed posture, immobility, reduced blinking, drooling, cogwheel rigidity in upper limbs…) and good response to L-dopa plus carbidopa with subsequent development of severe adverse effects [39].

[8] Oxidative stress can be defined as an imbalance between the production of ROS and the antioxidant capacity of the cell.
[9] MPTP is commercially available as a chemical intermediate.
[10] MPPP: 1-methyl-4-phenyl-4-propion-oxy-piperidine.

MPTP freely crosses the blood-brain barrier and, once in the brain, it is oxidized by MAOB to MPP^+[11] which is the active toxin (Figure 14) that selectively affects dopamine-containing neurons in the SN. MPP^+ enters to the cell via the dopamine transporter and accumulates in the mitochondria where it binds to complex I (NADH-dehydrogenase), blocking mitochondrial respiration: this energy crisis leads to a massive cell death [6]. MPP^+ also translocates into synaptic vesicles and stimulates the release of dopamine in the cytoplasm where it readily undergoes autoxidation, originating a burst of oxidative stress due to ROS [23].

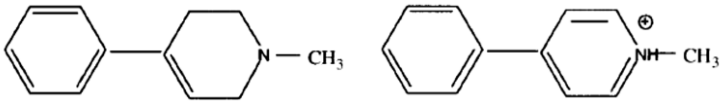


Fig. 14. MPTP (left) and MPP^+ (right) structures. Taken from [4]

This chance fact, where an exogenous substance induced a phenotype extremely close to that observed in Parkinson´s disease patients, potentiated the research on environmental factors.

Nowadays, there is no knowledge about any other compound with the same properties than MPTP. Neither the cellular specificity nor the similar phenotype and response to treatment have been found for any other substance.

[11] MPP^+: 1-methyl-4-phenylpyridinium.

Pesticides have been widely studied because their chemical structure is similar to MPTP and because it is known that acute exposures to them originate neurological dysfunction. However, the consequences of chronic and moderate exposures are less clear: changes in mood, sleeping, movement, cognition and other aspects have been investigated and the overall conclusion is that the earliest or most general response to pesticide neurotoxicity is a general malaise lacking in specificity and related to mild cognitive dysfunction [40]. With respect to PD, it is considered that pesticides increase risk of developing the disease, especially the herbicide paraquat and the insecticide rotenone[12] [41]. Due to the important methodological differences between studies with respect to the size of the groups, the doses considered and the type of exposition, there are discrepancies about the size of this risk.

Plenty of products (metals, medicines, food…) and habits (physical exercise, jobs, stress…), even other diseases (diabetes, anemia, depression…), have also been studied but the most consistent and significant results are for [42]:

- Uric acid is a natural antioxidant that may reduce oxidative stress and, consequently, reduce the risk of developing PD and slow its progression. There are strong evidences that serum urate is a protective factor [43] [44].
- Coffee and tobacco can also be considered neuroprotective factors because caffeine is an antagonist of A2A receptors and

[12] Rotenone and paraquat, like MPTP, are inhibitors of complex I in mitochondria.
Rotenone, MPTP and 6-OHDA are used to reproduce PD phenotype in animal models and cellular cultures.

nicotine inhibits the formation of αSyn deposits, improves PD symptoms and stimulates dopamine-containing neurons.

At present, although plenty of studies and effort have been done trying to find some environmental factor which could cause Parkinson´s disease, there is no definite candidate, just some potential neuroprotective or toxic compounds. That is why the contribution of environment to the development of the disease is still controversial [45].

I.7.b. Genetic factors.

To address the heritability of the disease, epidemiological, case-control and twin studies have been conducted. These studies support that there is a genetic contribution to the development of PD, although the strength of familial aggregation remains uncertain because there is clear familial clustering only in a minority of cases.

Monozygotic twins (MZ) have the same genome and, therefore, both will have the same phenotype in pure genetic diseases. Calculating the genetic contribution of a disease should be possible by comparing the concordance

rate[13] in MZ versus the concordance rate in other siblings, usually dizygotic twins (DZ).

For Parkinson´s disease, which is a multifactorial malady, results have been controversial: [46] showed that incidence rates of PD were not significantly different in MZ and DZ pairs, thus ruling out the genetic role in Parkinson´s disease. However, when considering age at onset, there was a significant result for those pairs with onset before 50 years that pointed out to a genetic effect. This study was based on clinical diagnosis. It was cross-sectional, without follow-up and only considered white North-American men. Nevertheless, [47] analyzed the dopaminergic function of twin pairs that were clinically discordant for PD and observed that the concordance for striatal dysfunction was significantly higher in MZ pairs than in DZ pairs. Moreover, after a follow-up, this concordance was even higher pointing out to a role for inheritance in sporadic PD. Recently, another twin study [48], cross-sectional but also longitudinal, with Swedish men and women, found a modest but significant genetic effect for same-sex couples.

The disparity in results observed could be due to the fact that in PD there is a long preclinical phase and that may distort the concordance rate observed in cross-sectional studies.

[13] The concordance rate is a quantitative statistical expression for the concordance of a given genetic trait, especially in pairs of twins in genetic studies. It is the proportion of pairs where both individuals share a certain characteristic.

In a meta-analysis of different studies about familial aggregation in PD, it was shown that there was an increased relative risk for first-degree relatives, ranging from 2.7 for child-parent pairs to 4.9 for sibling pairs [49]. Although it was complicated to compare the studies due to the different methodological methods used (different familial relationships, diagnostic procedures and follow-up), the tendency showed that heritability is present in PD.

Nowadays, ≈10% of Parkinson´s disease is familial [50], i.e. monogenic forms of the disease. It is due to mutations in 5 genes: *SNCA*, *PRKN*, *PINK1*, *DJ-1* and *LRRK2*. Maybe, the traces of heritability found in the previously referred articles reflect the contribution of these or other genes to the malady.

The remaining 90% of PD patients suffer from idiopathic or sporadic forms of PD[14]. Are those PD cases exposed to unknown environmental factors? Are there other genetic factors which have not been discovered yet? Nowadays, the majority opinion is that Parkinson´s disease will develop in those people because they have a specific genetic background which enhances their susceptibility to some environmental factors. When the external influences are present, the combination triggers the development of PD.

[14] Juvenile Parkinson´s disease (age at onset < 20 years) and early onset PD (20-50 years) are more common in familial forms of PD whereas idiopathic Parkinson´s disease usually develops lately (late onset > 50 years).

I.8. Familial PD.

So far, five genes are responsible for the familial forms of PD:

- **I.8.a. *SNCA* (*PARK1*, *PARK4*): 4q22.1**

The *SNCA* gene has 6 exons and codifies for a protein, α-synuclein or αSyn, with 140 amino acids.

- Protein structure: the N-terminal domain includes six 11-amino acid imperfect repeats with a highly conservative motif (KTKEGV); the middle region is hydrophobic and contains the non-amyloid component (NAC)[15] which is supposed to give the protein its proneness to aggregate; and the C-terminal domain is acidic and includes residues that are post-translationally modified like 129S, whose phosphorylation increases propensity to fibrillize, and 125Y, whose phosphorylation prevents fibrillation (Figure 15) [51].

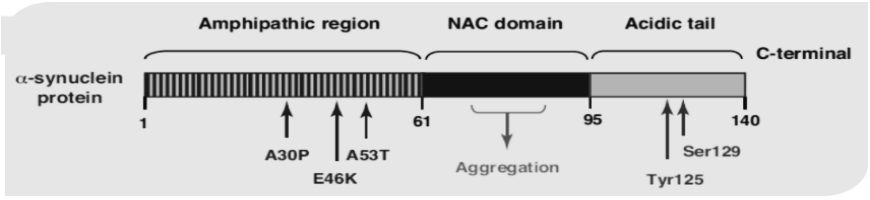


Fig. 15. SNCA structure.
On the left, the three pathogenic missense mutations that have been described.

Modified from [51]

The mutants are more prone to form aggregates, especially p.A53T which forms aggregates more readily than p.A30P [52]. Nevertheless,

[15] The reason why this name for the central domain is because αSyn was first found in brain amyloid plaques, where it is the second major component, in AD patients.

the protein, even when mutated, is natively unfolded, i.e. it has no secondary structure under physiological conditions. Despite this, and due to a reason that still remains unknown, αSyn might change its unfolded structure and organize into more complex formations: *in vitro* it forms fibrils of highly organized secondary structure under low pH conditions or high temperature. In addition, environmental factors, metals and pesticides (like paraquat), can induce aggregation [53]. The proposed mechanism (Figure 16) supposes that different physiological factors, such as oxidative stress and exposition to pesticides, or non-physiological factors, such as low pH and high temperature, increase the proportion of a partially folded intermediate and, depending on how the conditions evolve, the intermediate unfolds or forms highly organized structures [54].

Taken from [54]

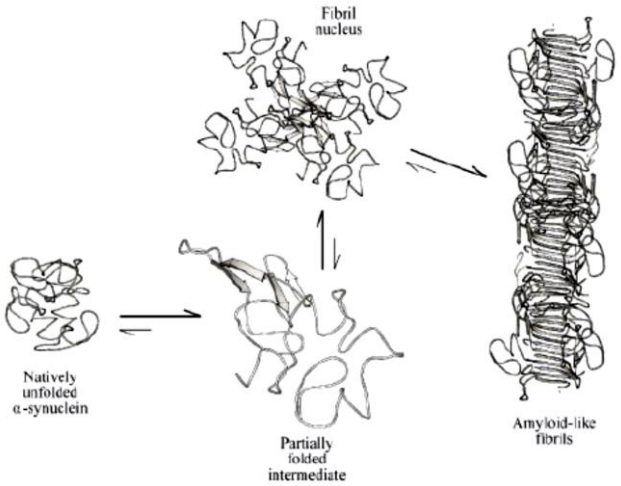


Fig. 16. Model for α-synuclein fibrillation.
Oligomerization of the intermediate leads either to fibrils via a critical nucleus or to soluble oligomers, resulting in amorphous aggregates. Additional conformational changes occur between the aggregation-competent intermediate and the fibrils.

Moreover, *in vitro*, αSyn can fold into other different complex structures (Figure 17).

Taken from [53]

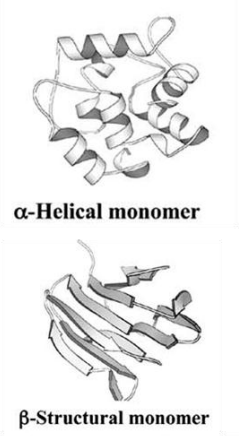


Fig. 17.More αSyn structures.

- Tissue expression: α-synuclein is expressed almost ubiquitously in human body, although its highest expression is in brain where it is present in presynaptic nerve terminals [55].

  Furthermore, αSyn is the main component of Lewy bodies [56]: in PD, LB are present in surviving neurons in SN but also in other affected brain regions.

- Cellular localization: αSyn is mainly cytosolic, although sometimes is bound to the membrane of synaptic vesicles.

  αSyn seems to have a prion-like behavior. Monomeric and aggregated α-synuclein are secreted and endocyted by neighboring cells. Once in the receptor cell, aggregates are transported through the endosomal pathway and finally degraded by lysosomes. However, if lysosomal capacity is compromised (aging, mutations,…) internalized αSyn accumulate in the receptor cell forming Lewy-like inclusions. The mechanisms why donor and receptor cells transfer αSyn are unknown [57, 58]. That explains why long-term transplanted

PD patients developed LB in their fetal grafted mesencephalic tissue [59, 60].

- Pathogenic gene variants: changes in αSyn sequence or quantity cause PD with autosomal dominant inheritance; however, the only current explanation to that fact is that α-synuclein has higher tendency to aggregate if it is mutated or if it is more abundant (duplications, triplications). Mutations in this gene are extremely rare and it seems that they do not change either protein structure or cellular localization or any other characteristic.
  - p.A53T (missense mutation): a highly penetrant mutation (85%) described by [61] in an Italian-American kindred (the Contursi family); other Greek families have also been described with this mutation. All carriers of the mutation but one had typical PD symptoms with early age at onset (30 to 50 years) and rapid progression.
  - p.A30P (missense mutation): described by [62] in a German family. Carriers presented typical PD symptoms.
  - p.E46K (missense mutation): described by [63] in a Basque kindred. Patients showed typical PD features although cognitive decline (especially dementia and hallucinations) and early onset were more frequent. This characteristics plus the widespread presence of LB in the brain resembled dementia with Lewy bodies, a heterogeneous disorder that clinically and pathologically overlaps with PD, making DLB more probable than PD in those cases.

- Gene duplication and triplication: a linkage analysis conducted in a large kindred with typical PD features, rapid progression and early onset -although some members suffered from dementia and others postural tremor instead of parkinsonism- revealed that the responsible for the phenotype was the triplication of a region that included *SNCA* and 16 other genes [64, 65]. αSyn levels were increased in these PD patients due to the increase in gene dosage: protein and mRNA levels doubled normal values in blood and brain as a consequence of the double number of *SNCA* alleles, which were all expressed. Moreover, the increased expression was correlated with a higher proneness to form aggregates in brain [66]. In a large Swedish kindred (Lister family complex) changes in gene dosage were also identified: in one branch, the Swedish-American branch, the carriers of triplicated *SNCA* showed similar clinical features to these previously described, whereas at the Swedish branch, carriers of a duplication in a smaller region including *SNCA* and one more gene [67] developed parkinsonism but with a wider range of features when comparing with previous reported families which showed typical PD with late onset [68, 69]. Gene dosage seems to be correlated with disease progression, age at onset and symptom severity.

- Function: unknown. It has been suggested to act in multiple processes including synaptic vesicle biogenesis, brain lipid metabolism, cytoskeleton stability and lysosomal function. α-synuclein aggregates are degraded by autophagy whereas monomers are degraded through

the proteasome [70]. α-synuclein also participates in dopamine homeostasis in the presynaptic terminal (Figure 18).

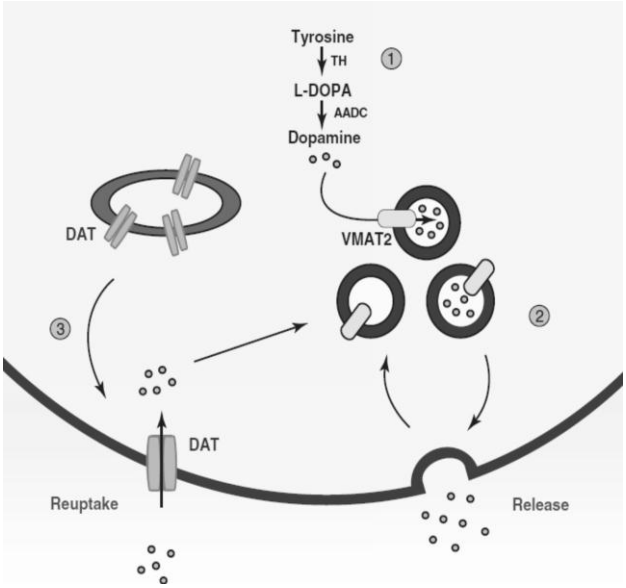


Fig. 18. αSyn and dopamine homeostasis.
1.α-synuclein regulates dopamine synthesis by controlling the activity of tyrosine hydroxylase (TH) and L-dopa decarboxylase (DDC, also known as AADC: amino acid decarboxylase).
2.α-synuclein is involved in regulating synaptic vesicle function and dopamine release into the synaptic cleft. Vesicular monoamine transporter 2 (VMAT2) sequesters dopamine into vesicles.
3.α-synuclein is necessary for the trafficking of dopamine transporter (DAT) to the cell surface, which is necessary to reuptake dopamine and finish dopamine signaling.

Modified from [51]

- Polymorphisms: two are the most studied, although none of them is located inside the gene sequence. Both are considered possible susceptibility PD factors with influence in αSyn expression.
  - **Rep1** (D4S3481): is a polymorphic microsatellite repeat located ≈10kb upstream of the translation start site. Its sequence repeat is mixed, $(TC)_{10-11}(TT)(TC)_{8-11}(TA)_{7-9}(CA)_{10-13}$ [71], and different dinucleotide composition can originate the same microsatellite

length [72]. Although this could be considered as an additional complexity factor, [73] concluded that, *in vitro*, the overall sequence repeat length, and not the number of any specific dinucleotide, changes SNCA expression: taking allele 267 expression as the reference, allele 271 increased 1.5 times the expression, allele 273 2.5 times and allele 269 3 times in SH-SY5Y cells (from human neuroblastoma) but not in 293T cells (from human kidney) [74]. However, [75] found a trend, not a significant result, that pointed out that allele 269 was related to lower mRNA levels in frontal cortex, temporal cortex and SN from control cases. Therefore, the real effect of Rep1 alleles in SNCA expression remains unclear.

[76] were the first who described the microsatellite and its alleles and, lately, [77] concluded that allele 273 was more prevalent in PD cases than in controls in a German population. Other articles, [71] in an Italian population or [78] in a Singaporean population, did not obtain the same conclusion. The works by [79] in a Greek population, [80] a meta-analysis and [81, 82] in a population from USA determined that allele 271 was the real risk factor instead of 273; moreover, [83] in Australians with European ancestries, [80] a meta-analysis and [81] in a population from USA[16] observed that allele 267 could be neuroprotective. However, it is difficult to

[16] Depending on the article used as a reference, allele name may change: [76] described alleles 0 (267), 1 (269), 2 (271), 3 (273)… Nevertheless, alleles in [77] are 10 bp shorter than those. These are the most common nomenclatures, although some articles ([80] and [82] for example) mix both: their allele 263 is actually 271 in [76].

explain how allele 271 can be a risk factor for PD when it is related to less αSyn expression than allele 269, the most common. The frequency of the different alleles changes depending on ethnicity, being allele 273 more frequent in Asian population than in Caucasian, although allele 269 is always the most frequent [78]. Maybe, the actual risk factor is in linkage disequilibrium to Rep1 but not itself.

- **rs356219** (A>G): is located 9kb downstream of the gene. Some articles, [84] in a Norwegian population, [85] in a Caucasian population from the USA (this same group also found that the G allele is associated with higher levels of plasmatic αSyn in cases), [86] in an Italian population and [87] in Han Chinese, concluded that the G variant is a PD risk factor, although there are opposite results, as in [88] in Swedish. In addition, in frontal cortex, C allele is related to higher mRNA quantities of SNCA112, an splicing variant without exon 5 which is considered to enhance αSyn aggregation [89].
- Although there is no correlation between Rep1 alleles and rs356219 alleles or genotypes, their interaction with other supposed PD risk factors such as the H1/H2 haplotype in *MAPT,* has been studied: the majority of results conclude that there is no joint effect neither for Rep1-H1 [90, 91] nor for rs356219-H1 [92-94]. All these studies concluded that, separately, the G allele and the H1 haplotype are PD risk factors but the risk does not increase with the presence of both, except for [95] who found that the

interaction between rs356219 and H1, i.e. the presence of the G allele plus the H1 haplotype, doubles the risk of developing PD.

- Animal models: although any rodent model presents all the key features that are present in PD cases originated by *SNCA* [96], they closely resemble the situation:
  - The knockout mice are viable and fertile and have normal brain structure (presynaptic terminals and dopamine neurons are normal). Therefore, α-synuclein is not essential for neuronal development. However, there is a reduction in striatal dopamine [97].
  - Overexpression of wild type αSyn with heterologous promoter in mice recapitulates many features of PD: accumulation of insoluble α-synuclein aggregates, loss of dopaminergic terminals and motor abnormalities (fine motor skills are altered) [98].
  - Transgenic mice expressing mutated αSyn show the common PD features, although p.A53T causes more toxic effects than p.A30P [99].

I.8.b. *PRKN* (*PARK2*): 6q26

PRKN gene has 12 exons and encodes for a protein, parkin, with 465 amino acids. It is the second biggest gene in the human genome.

- Protein structure: parkin has a N-terminal ubiquitin-like domain (Ubl), a linker region and a C-terminal RING (Really Interesting New Gene)

box that can be divided into three domains: RING1, RING2 and IBR (In-Between-RING) (Figure 19) [100].

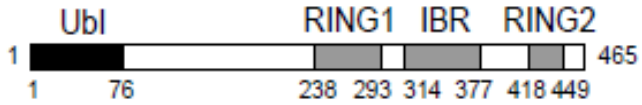


Fig.19. Parkin structure.

Taken from [100]

- Tissue expression: ubiquitous.
- Cellular localization: PRKN is mainly cytosolic but it is also present in nucleus and even in mitochondria because parkin can be recruited to the mitochondrial membrane when it loses membrane potential.
- Function: parkin is an E3-ubiquitin ligase enzyme so it works in the **unfolded protein response**.

  Misfolded or short-lived proteins are targeted for degradation in the 26S proteasome by covalent attachment of ubiquitin. Three enzymes participate successively in this process:

  - E1 (ubiquitin-activating enzyme) activates the ubiquitin in an ATP-dependent manner.
  - E2 (ubiquitin-conjugating enzyme) binds ubiquitin.
  - E3 (ubiquitin-protein ligase) transfers ubiquitin from E2 to the target protein.

  Parkin interacts with E2 proteins (UbcH7, UbcH8) via its RING box and with its target proteins via its Ubl domain (Figure 20) [100].

  It is known that PRKN labels some proteins for degradation (K-48 linked poliubiquitination). However, to confirm which are its real

targets *in vivo* is controversial (*in vitro*, tubulins, CDCrel-1, LRRK2 aggregates [101] and more proteins are its targets [102]):

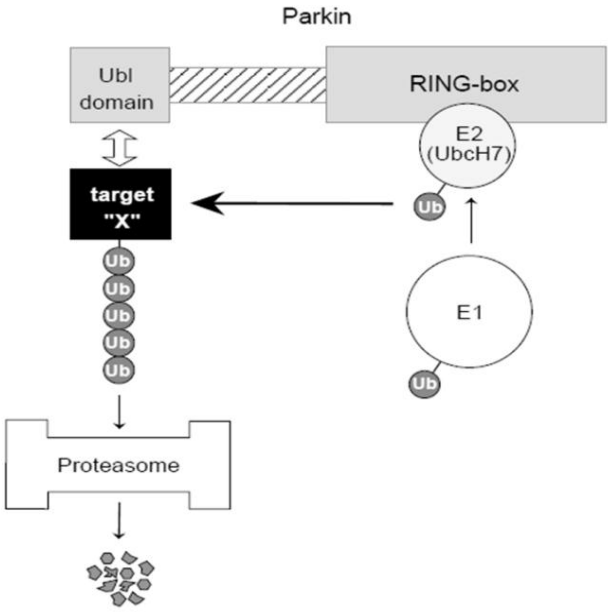


Fig.20. Model of the parkin-directed ubiquitination pathway.

Modified from [100]

- αSp22, is a O-glycosylated human αSyn form, only present in human brains (22KDa), that accumulates in LB in parkin-deficient cases. Only this form, and not the common αSyn (16KDa), interacts with, and is ubiquitinated by, parkin [103].
- in human brain, parkin, CHIP (E3 enzyme which, at least *in vitro*, could be another parkin substrate), and Hsp70 (molecular chaperone) work together to ubiquitinate and degrade Pael-R (parkin-associated endothelin receptor-like receptor; this protein has unknown function and encodes for a multipass endoplasmic

reticulum transmembrane protein) when there is some folding problem [104].

- synphilin-1 is an α-synuclein-interacting protein without known function that interacts with parkin via the RING2 domain, not by the Ubl domain, and is ubiquitinated in rat brain [105].

In addition, Parkin catalyzes monoubiquitination and K-63 linked polyubiquitination, which are proteasome-independent mechanisms to regulate endocytosis, aggresome[17] formation and NF-κB signaling[18], in conjunction with the heterodimeric E2 enzyme UbcH13/Uev1a. This process is mediated by PINK1 phosphorylation of parkin which increases its affinity to UbcH13/Uev1a [106]. Mutations in parkin or PINK1 block this process and that can be related to PD pathogenesis.

– Pathogenic gene variants: parkin was first time associated with PD in Japanese population by [107]. They found that changes in this gene cause autosomal recessive Parkinson´s disease with slow progression, frequent dystonia, sleep benefit, good response to L-dopa, early onset (although there is a wide range of ages at onset, varying from early childhood to late adulthood) and without Lewy bodies presence in brain [108, 109]. This later aspect is on debate: not too many brains have been analyzed and, even when the majority of *PRKN*-related PD

[17] When there is an overload of damaged proteins in the cytosol and proteasome cannot degrade it, an aggresome (a proteinaceus inclusion body) is formed. This protective mechanism gives the cell more time to eliminate abnormal proteins.

[18] This is a ubiquitously expressed transcription factor which mediates transcription of a number of pro-survival genes.

patients have no LB but neurofibrillary tangles[19] [110], some of them do [111]; in addition, parkin has been identified in LB [112].

Changes in gene sequence affect PRKN function and, therefore, misfolded target proteins accumulate, but the molecular mechanisms that connect this protein aggregation and neuronal death remain unclear.

The most frequent gene variants in PD cases are deletion of exon 4, deletion of exon 3 and deletion of exons 3 and 4 [113], although not only exonic deletions have been described: rearrangements, duplications, missense and nonsense mutations can be present too [114]. Homozygous and compound heterozygous for mutations in *PRKN* are affected by PD, but some carriers of just one mutation in parkin have also developed the malady and positron emission tomography studies suggest that parkin heterozygous, although usually asymptomatic, may exhibit nigrostriatal dysfunction [115]. The relation between haploinsufficiency or dominant-negative effect and pathogenicity has an unknown significance in those people, especially as [116, 117] have analyzed *PRKN* mutations in PD cases and control populations and have found that there is no evidence for association of heterozygous parkin sequence variants, mutations or polymorphisms, with risk or protection against AR-JP[20].

- Animal models: in mice, exon 3 knockout animals have normal brain morphology without reduction of nigrostriatal dopamine neurons.

[19] Tau-positive protein inclusions.
[20] Autosomal recessive juvenil parkinsonim.

However, the absence of protein originates subtle motor, behavioral and cognitive deficits, increased dopamine level, reduced synaptic excitability and inhibition of glutamate neurotransmission [118]. In *Drosophila*, null mutants show reduced life span, locomotor defects, muscle fiber degeneration, sterility and mitochondrial pathology but no neuronal degeneration [119].

Any of the models show a clear parkinsonism phenotype; that fact together with the wide range of phenotypes observed in PD cases seem to imply that parkin acts together with other factors to induce the disease.

- **I.8.c. *PINK1* (*PARK6*): 1p36.12**

PINK1 gene has 8 exons and encodes for a protein (PTEN-induced putative kinase 1) with 581 amino acids.

- Protein structure: two characteristics define the structure of the protein
  - the first 34 amino acids, i.e. the N-terminal end, target the protein to the mitochondria.
  - there is a highly conserved protein kinase domain (amino acids 156 to 509) that shows a high degree of homology to the serine/threonine kinases of the Ca^{2+}/calmodulin family [120].
- Tissue expression: ubiquitous.
- Cellular localization: in mitochondria, where its localization is unclear: some studies show its kinase domain facing the cytosol whereas some

others place it on the inter-membrane space. PINK1 is also found in cytosol, where a fraction of it is exported after processing in the mitochondria [121].

- Pathogenic gene variants: [120] described for the first time that PINK1 causes recessively inherited forms of Parkinson´s disease. They showed it in Italian and Spanish families that presented typical PD with slow progression and good response to L-dopa but early onset.

  In PD patients, frameshifts, missense, nonsense, intronic and synonymous mutations have been described (Figure 21). However, deletions and multiplications are rare [122]. Individuals homozygous and compound heterozygous for mutations in PINK1 develop Parkinson´s disease, and some heterozygous too (familial or sporadic) although with delayed onset. Positron emission tomography studies support that fact: they have concluded that there is decreased dopaminergic function in asymptomatic heterozygous carriers of PINK1 suggesting that some mutations might predispose to PD [123]. Nevertheless, no study has found differences in the overall frequency of PINK1 mutations between PD cases and control population: maybe a specific mutation, possibly not described yet, is more prevalent in controls or in cases and can shed light to this confusing situation.

  Haploinsufficiency[21], dominant-negative effect[22] or dominant gain-of-function mutations could not be enough to explain the different phenotypes observed in heterozygous. Those processes could be just

[21] There is a lack of active protein, the 50% is missing.
[22] Mutated proteins block wild type proteins function.

PD risk factors and maybe it is a combination of PINK1 mutations plus environmental factors or mutations in PINK1 plus mutations that have not been yet described in other genes only in brain (somatic mosaicism) the real cause [124].

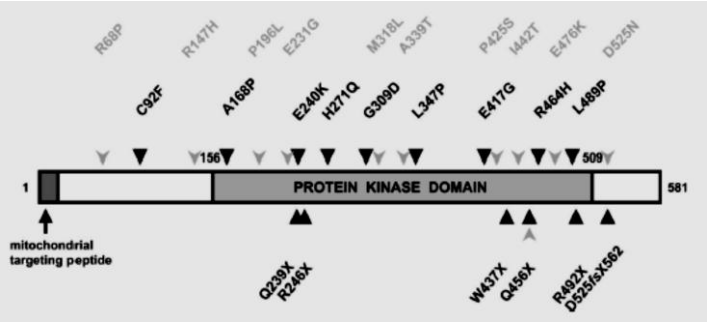


Fig. 21. PINK1 mutations reported in patients with parkinsonism. Missense and truncating mutations are depicted above and below the protein bar. Mutations found in homozygous or compound heterozygous state are in black. Mutations found in heterozygous state are in gray.

Taken from [124]

- Function: it is a kinase. Further studies to establish the natural substrates of the protein are required although it is known that parkin and HtrA2 are some of them [125, 126]. Mutated proteins lose their activity (changes affect kinase activity or substrate recognition) but not their localization [121]: *in vitro*, p.G309D and p.L347P mutants showed reduced kinase activity (the effect was drastic in p.L347P whereas more modest in p.G309D). In both cases mutations affected neither mRNA levels nor cellular localization. Nevertheless, those mutated proteins can be less stable and more degraded.

PINK1 also has an important role in mitochondrial protection: parkin interacts with PINK1 and both functionally cooperate to identify and label damaged mitochondria for selective degradation via autophagy (mitophagy) [127, 128]. The model (Figure 22) has some steps:

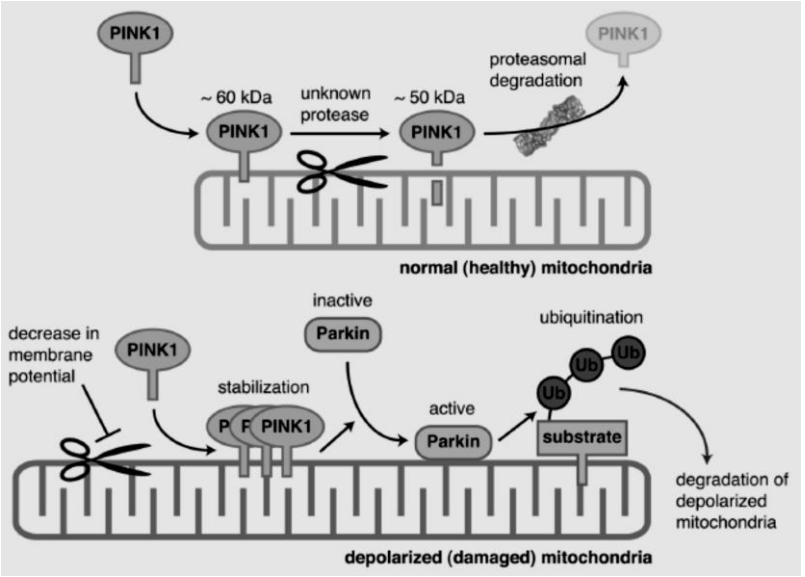


Fig. 22. Quality control model for depolarized mitochondria.

Modified from [127]

- In normal conditions, the mature PINK1 is constantly cleaved by an unknown protease to its intermediate form that is degraded. PINK1 cleavage is voltage-sensitive.
- However, if the mitochondria are depolarized, there is a decrease in membrane potential, mature PINK1 is stabilized at the outer membrane and not cleaved. PINK1 recruits parkin from the cytoplasm in a membrane potential-dependent manner and activates it. Parkin ubiquitinates an unknown mitochondrial

substrate(s) and, as a consequence, damaged mitochondria are degraded via mitophagy.

- Mutations in both genes block the mechanism thus originating recessive familial PD.

– Animal models: in mice, knockout *PINK1*$^{-/-}$ had no nigral degeneration but impaired mitochondrial respiration in striatum in early life (3-4 months) and in cerebral cortex in late life (24 months). There was a selective increase in larger mitochondria although mitochondrial structure was not affected [129]. In *Drosophila*, *PINK1* is 60% similar to the human gene including the mitochondrial targeting motif and the kinase domain. PINK1 mutants exhibited dopaminergic neuronal degeneration accompanied by locomotion defects but parkin rescued the phenotype demonstrating that parkin acts downstream of PINK1. There was mitochondrial dysfunction in these neurons and also, their size was increased. Mitochondrial integrity was damaged not only in dopaminergic neurons, also in sperm and flight muscle [130]. PINK1 could be considered a critical factor required in dopaminergic neurons for maintaining mitochondrial integrity as well as neuronal function. [131] showed that, in *Drosophila*, PINK1 and parkin work in the same pathway regulating mitochondrial morphology (fission). The mechanism is unknown but fission regulation could provide the cell the way to segregate small damaged mitochondrial units that will be eliminated through autophagy thus eliminating oxidative stress (which is one of the prime suspects in PD).

Results in *Drosophila* resemble more those obtained in humans whereas in knockout mice the only similarities are that mitochondrial size and respiration are altered: the presence of nigral degeneration and the mechanism to regulate mitochondrial fission, which can be homolog to the quality control model for depolarized mitochondria described in humans, represent better the pathological phenotype observed in PD cases.

➢ I.8.d. *DJ-1* (*PARK7*): 1p36.23

DJ-1 gene has 8 exons and encodes for a protein with 189 amino acids.

– Protein structure: DJ-1 forms homodimers (Figure 23) [132].

It belongs to the ThiJ/PfpI superfamily. In this group, proteins with very different biochemical and cellular functions are included: intracellular proteases, amidotransferases, kinases…

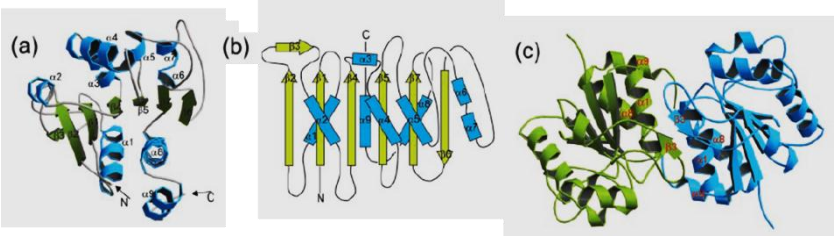


Fig. 23. DJ-1 structure.
a. Ribbon diagram of the DJ-1 monomer. The secondary structure elements are blue (α-helices) and green (β-strands). Positions of the N and C termini are indicated with arrows.
b. Topology diagrams of DJ-1 (color-coded as in a). The α-helices are represented by blue rectangles, and the β-strands are represented by green arrows.
c. Ribbon diagram of the DJ-1 dimer. Monomer A is blue and monomer B is green.
Modified from [133]

- Tissue expression: ubiquitous.
- Cellular localization: the major part of the protein is cytosolic and nuclear although there is a pool in mitochondria (mitochondrial matrix and inter-membrane space) [134].
- Pathogenic gene variants: mutations in DJ-1 are uncommon. Nevertheless, [135] described families from Italy and the Netherlands who developed autosomal recessive PD with slow progression, good response to L-dopa (typical PD features) but early onset due to homozygous mutations in this gene. Phenotypically, DJ-1 patients are indistinguishable from parkin and PINK1-linked PD cases.

  Homozygous and compound heterozygous for DJ-1 mutations have been described, but there are not enough data to draw conclusions about the role of heterozygous mutations in this gene. Exonic deletions, frameshifts, nonsense and missense mutations have been reported.

  Mutations alter DJ-1 cellular distribution and cause loss of function and decreased protein stability. For example, the p.L166P change abrogates dimerization and, consequently, increases protein turnover through the ubiquitin-proteasome system (this mutant is present mainly in mitochondria and almost absent in cytosol and nucleus).

  Mutations in parkin account for the majority of changes in recessive Parkinson´s disease, whereas mutations in PINK1, and even more in DJ-1, are not so common. Moreover, it is unknown whether the absence of protein aggregates in brain is a common feature in recessive PD or not, as it is the case for parkin-caused PD, because few

anatomical studies have been conducted in brains of PD patients who carry mutations in *DJ-1* or *PINK1* genes to determine the presence of LB.

- Function: DJ-1 has been involved in tumorigenesis, as a putative oncogene, and fertilization, but it also acts as a protective protein against oxidative stress. *In vitro*, DJ-1 increases its expression to protect against oxidative stress-induced cell death caused by agents like H_2O_2, MPP^+ or 6-OHDA. If mutated, this scavenger activity is reduced and cells are more susceptible to death [136]. This protective mechanism may involve Nrf2, an antioxidant transcriptional master regulator that triggers the protective cascade against oxidative conditions, because, *in vitro*, DJ-1 stabilizes it by preventing its association with KEAP1 (its inhibitor) and subsequent ubiquitination and proteasome-mediated degradation. It has not been determined how DJ-1 avoids the association of Nrf2 with Keap1 because physical interaction between DJ-1 and Nrf2 has not been reported [137]. Another explanation of this protective role of DJ-1 involves Cu/Zn-Superoxide Dismutase 1 (SOD1) because, *in vitro* and *in vivo*, DJ-1 participates in a pathway leading to the induction of SOD1 expression: under oxidative stress conditions, DJ-1 interacts with Erk1/2 (MAPK kinases) and translocates them to the nucleus to phosphorylate the transcriptional factor Elk1 which, on its turn, triggers SOD1 expression to counteract oxidative stimulation [138].

  The redox sensitivity of DJ-1 also relates the protein to mitochondrial maintenance (because it is recruited to the mitochondria when there

are oxidative conditions inside the cell [139]) and to clearance of αSyn aggregates (because DJ-1 can act as a redox-dependent chaperone that eliminates αSyn aggregates [140]).

In addition, DJ-1 interacts *in vitro* with other proteins related to familial PD:

- DJ-1 interacts with and stabilizes PINK1: one family was described whose members developed Parkinson´s disease due to digenic mutations, i.e. affected people carried one mutation in DJ-1 and other in PINK1. Heterozygous individuals for any of the mutations did not develop the disease. The relation between genotype and phenotype assumed that there was a functional interaction between both genes, interaction that was confirmed [141].
- Parkin, PINK1 and DJ-1 form the PPD complex in the cytosol: this is an E3 complex that promotes ubiquitination and degradation of aberrantly expressed and heat-shock induced parkin substrates. DJ-1 stabilizes PINK1, although the presence of DJ-1 at the complex is not essential, and PINK1 potentiates parkin activity. This complex has also been described *in vivo* in human brain lysates [142].
- Finally DJ-1 also interacts *in vitro* with two nuclear proteins: p54nrb and pyrimidine tract-binding protein-associated splicing factor (PSF) which are multifunctional regulators of transcription. PSF induces apoptosis by reducing the expression of neuroprotective proteins or anti-apoptotic genes. DJ-1 works with p54nrb to inhibit the repressing activity of PSF. However, mutated DJ-1 proteins are not

in the nucleus and, therefore, cannot protect the cell [143]. DJ-1 is a transcriptional co-activator that regulates transcription without directly binding any promoter.

- Animal models: all knockout mice (DJ-$1^{-/-}$) analyzed were viable and fertile and showed normal brain morphology. Their phenotypes supported conclusions obtained in previous studies even when any of the rodent models developed all typical PD features [119]:
  - Some mice showed age-dependent and task-dependent motoric behavioral deficits, increased dopamine reuptake rates and elevated tissue dopamine content.
  - Others had reduced sensitivity of nigral neurons to dopamine, reduced dopamine overflow due to increased reuptake and decreased spontaneous locomotor activity (generalized hypokinesia). However, the number of dopamine-containing neurons in the SN at 3 and 12 months of age were normal. Nevertheless aged mice were not analyzed and cannot be discarded the fact that they could had neuronal degeneration. Taken together, these findings indicated that the nigrostriatal pathway was intact but dysfunctional [144].
  - The influence of oxidative stress in DJ-$1^{-/-}$ mice was also studied and the *in vivo* results demonstrated previous *in vitro* conclusions. DJ-1 is a protective protein against oxidative stress, its loss confers increased susceptibility to oxidative stress-induced neuronal death. Loss of DJ-1 exacerbated oxidative stress-induced neuronal death, by rotenone and MPTP, in primary cortical and dopaminergic

neurons. Neuronal death was only increased as a consequence of oxidative insults; non-oxidative insults had no influence on cell survival and restoration of DJ-1 mitigated the phenotypes. In this study, dopamine levels in the striatum were normal [145].

➢ I.8.e. *LRRK2* (*PARK8*): 12q12

LRRK2 (leucine-rich repeat kinase 2) gene has 51 exons and encodes for a protein, dardarin, with 2527 amino acids.

- Protein structure: the protein contains several functional domains: ARM (armadillo), ANK (ankyrin repeat), LRR (leucine-rich repeat), ROC (Ras of complex proteins; with GTPase activity), COR (carboxy terminal of ROC), kinase and WD40 (rich in tryptophan and aspartate repeats). Four of these domains are related to protein-protein interactions: ARM, ANK, LRR and WD40 [146] (Figure 24).

Taken from [147]

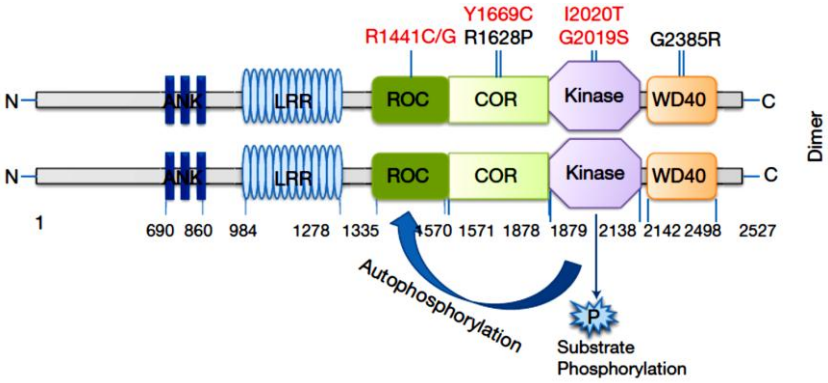


Fig. 24. Domains and mutations in LRRK2.
The LRRK2 protein is shown diagrammatically with amino acid numbers below each domain. Above the dimer are positions of known pathogenic mutations (in red) and some possible risk variants (in black). The dimer is present in a head-to-head orientation although it is not known.

- Tissue expression: dardarin is expressed in brain regions of direct relevance to the pathogenesis of PD like cerebral cortex, SNpc or striatum [148] and also in heart, kidney, lung, liver and peripheral leukocytes.
- Cellular localization: *in vitro* studies suggest that LRRK2 is a cytosolic protein [101]. However, LRRK2 is also associated with a variety of membrane and vesicular structures (i.e. outer mitochondrial membrane) and the microtubule network [149, 150].
- Pathogenic gene variants: [151] described that there was a region associated with PD (12p11.2-q13.1) at the Sagamihara family from Japan. Affected cases presented typical idiopathic symptoms, good response to L-dopa, autosomal dominant inheritance and late onset. Nevertheless, there was incomplete penetrance: the "pathogenic" haplotype was shared by PD cases and healthy individuals within the family. Two years later, [152] and [153] found that it was due to mutations in *LRRK2* gene.

  To date, more than 40 missense or nonsense mutations have been described. Only point mutations, no deletions nor duplications, have been found causing PD. Six of the described mutations are recognized as disease-causing variants and segregate with disease in large families: p.N1437H, p.R1441C, p.R1441G, p.Y1699C, p.G2019S, p.I2020T. Mutations in *LRRK2* gene are the most common cause of Parkinson´s disease. They are present in up to 5-13% of familial cases and in 1-2% of sporadic PD patients.

p.N1437H mutation is located on the ROC domain. It has been found in a large Norwegian family segregating with autosomal dominant parkinsonism. However, more research is required to confirm that this family is suffering from PD [154].

Mutations p.R1441C and p.R1441G are located on the ROC domain and have been described in familial and sporadic cases. Both decrease GTPase activity [155, 156] due to the change they produce in the folding properties of the domain [157], but the localization, turnover and protein steady-state levels remain unaltered [158]. p.R1441C is the second most recurrent PD mutation and is found in different ethnic races with high penetrance. Nevertheless, p.R1441G is most common in the Basque country and is rare outside Northern Spain [159]; its penetrance is high, increases with age and is independent of sex [160]. This position can be considered a hotspot because another mutation has been described (p.R1441H) in different populations (Greek, Italian, Taiwanese…) but it has never been found to co-segregate with PD in a large family. For this reason it is not considered a pathogenic variant [161].

p.Y1699C mutation is located on the COR domain. It was found in the initial reports about LRRK2: [152, 153].

p.G2019S mutation is the most frequent cause of familial and sporadic PD. Its frequency is heterogeneous around the world with the highest values for Ashkenazi Jews and in North African Arab countries [162] (Figure 25).

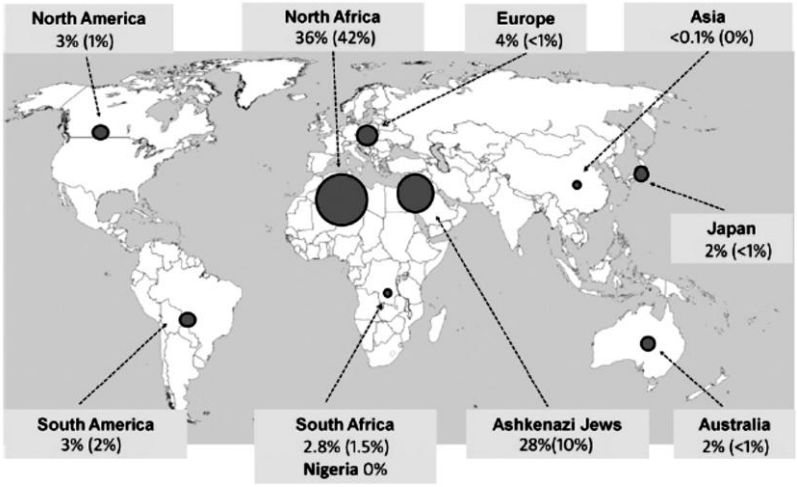


Fig. 25. Worldwide distribution of p.G2019S mutation.
The frequencies are for familial cases and, in brackets, for sporadic cases.

Taken from [161]

Its penetrance is incomplete and smaller than for p.R1441C and p.R1441G: it is estimated to be ≈75% at the age of 79 years, age-dependent and sex-independent [163]. The mutation is located on the kinase domain and it increases the kinase activity by two to three folds [158]. This is the only effect it has; it does not modify the protein steady-state levels, its localization or its turnover.

p.I2020T mutation is located on the kinase domain. Its influence on kinase activity remains controversial: some articles conclude that the activity is increased by the mutation [149] whereas others conclude the opposite, that it is decreased by the mutation [164, 165].

In general, PD cases that carry these mutations are clinically indistinguishable from idiopathic PD patients.

Furthermore, two polymorphic variants have been described as risk factors, only in Asian population, both increasing PD risk by approximately two-fold:

p.G2385R is located on the WD40 domain. This mutation is absent in Caucasian subjects but it is considered a common risk factor for PD in Chinese population [166]. [167] observed that mutated protein does not change its cellular localization but it is considered to increase cellular susceptibility to oxidative stress.

p.R1628P is located on the COR domain and it also increases the risk of Parkinson´s disease among Chinese [168].

- Function: LRRK2 is a kinase [158] and also a GTPase [169]. GTPase activity controls kinase activity (LRRK2 needs to bind GTP to act as a kinase) but GTPase activity is independent of kinase activity [155], i.e. kinase activity does not modulate GTP-binding affinity [165]. In conclusion, LRRK2 is a GTP/GDP-regulated protein kinase [170]. The various mutations present in the protein are scattered throughout the protein with some evidence of clustering into these two enzymatic domains.

  The function of dardarin is unknown to date: it is suggested to have a role as GTPase, as kinase involved in cellular signaling… amongst others.

  Dardarin can be involved in microtubules formation or stability because it is known that wild type and mutated p.R1441C LRRK2 interact *in vitro* with α/β-tubulin heterodimers through the ROC domain, an interaction that is GTP- or GDP-independent. Although

these experiments were done *in vitro,* endogenous LRRK2 and α/β-tubulin were found to co-localize in primary hippocampal neurons *in vivo* [171]. In addition, in bovine brains, as is also the case in mice brain, LRRK2 preferentially phosphorylates β-tubulin and this phosphorylation is three times higher due to p.G2019S mutation. The increase in phosphorylation results in decreased microtubule dynamics and cumulative dysfunction/damage over time. Furthermore, LRRK2 enhances the polymerization of tubulin (isolated from bovine brain) in the presence of microtubule-associated proteins. Maintaining microtubule dynamics within a physiological range is essential for neuronal function and survival. This stability is necessary to avoid problems in axonal transport and synapse formation but mutations in LRRK2 could originate changes in the microtubules dynamics and cause neuronal dysfunction that could eventually lead to PD [172].

LRRK2 also regulates mitochondrial dynamics through its direct interaction with DLP1, a fission protein. LRRK2 expression *in vitro* caused mitochondrial fragmentation, mitochondrial dysfunction and neuronal toxicity; LRRK2 mutants increased DLP1 expression and, therefore, increased fragmentation. LRRK2 kinase activity plays a critical role in this process [173].

There is some controversy about if dardarin dimerizes or not. Some studies found that LRRK2 migrated in polyacrylamide gels at the double of its expected size (600KDa) and, consequently, they

concluded that LRRK2 was a dimer. It is supposed to be a homodimer although that is controversial too because there is the possibility that the dimer is formed by a truncated LRRK2 plus one complete protein. It is known that the WD40 domain is necessary for the formation of the dimer, and also for the kinase activity [174], and, that the interaction between both parts needs the LRR domain and the N-terminal region too. Moreover, the dimer undergoes intra-molecular phosphorylation in its ROC domain that potentiates the kinase activity [175-177].

However, recent studies have characterized "the dimer" (called p600) and have concluded that LRRK2 seems to be predominantly monomeric within cells. The monomer possesses the kinase activity as well as the GTP binding activity and these activities do not change even when it dimerizes [178]. The homodimer could be just a minor subespecie in the cell [179] and the unexpected migration of the protein could be a characteristic of high molecular weight proteins as it has been reported in other proteins [180].

To explain the relation between dimer and monomers some models have been proposed (Figures 26 and 27). Nevertheless, the effect of mutations on dimer formation remains unclear.

- Animal models: knockout mice (*LRRK2*$^{-/-}$) are viable and fertile, show normal development and brain morphology with a normal dopaminergic system. Knockout mice are not more sensitive to MPTP than wild type mice and have the same life span [181]. That is, they are indistinguishable from wild type mice. The only difference is on

their kidneys because aged mice have proteinaceous aggregates in their kidneys composed of αSyn and ubiquitinated proteins. Moreover, they show a smaller size and weight and a granular aspect [182]. These results imply that LRRK2 is not important for survival or development of dopaminergic neurons.

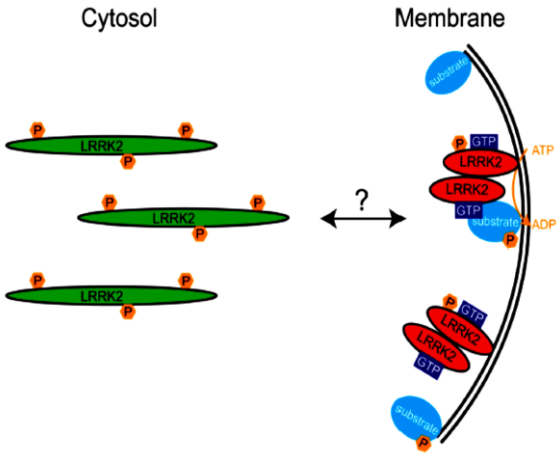


Fig. 26. Proposed model for LRRK2 dimer assembly.
LRRK2 exists mostly as a monomer in the cytosol and can translocate to membrane where it dimerizes, becomes more active and subsequently phosphorylates its substrates. Membrane-associated LRRK2 possesses greater kinase activity, an increased propensity to bind GTP, and is relatively dephosphorylated, compared to cytosolic LRRK2.

Taken from [179]

However, mice expressing human wild type or mutant LRRK2 showed reduced extracellular dopamine levels in the striatum [183], and although over-expression of human LRRK2 protein did not cause degeneration, it promoted αSyn aggregation in p.A53T mice. This effect was stronger when expressing the p.G2019S mutant [184]. Over-expression of human p.R1441G LRRK2 in mice caused age-dependent and progressive motor-activity deficits and they were

responsive to L-dopa, features that resembled the human phenotype for PD [185]. They showed diminished dopamine release and there was more axonal fragmentation although the number of neurons in SNpc was normal and their anatomical organization too. These results stress how difficult is to obtain a good animal model for LRRK2.

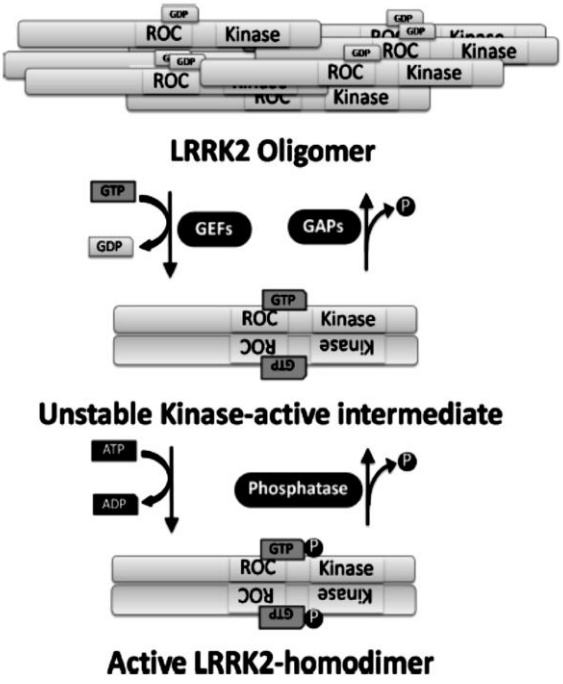


Fig. 27. Hypothetical model of kinase activation.
A major fraction of LRRK2 protein in cells may reside in large oligomers with low or no kinase activity. LRRK2 oligomers are dissociated through conformational changes induced by GTP binding within the ROC domain, which may lead to the formation of a dimer structure initially stabilized by a ROC-ROC interaction. LRRK2 autophosphorylation may lead to the stabilization of the kinase-active dimer, which can be destabilized by competing phosphatase activity, GTPase hydrolytic activity, or stochastic interactions with LRRK2 oligomers.

Modified from [186]

I.9. Other *PARK loci*.

There are more *PARK loci*, but their contribution to the monogenic forms of Parkinson´s disease is not as clear as it is for the 5 previous genes:

- **I.9.a. *PARK3*: 2p13**

  [187] described a possible new *PARK locus* with autosomal dominant inheritance pattern. They found that different families with typical idiopathic PD symptoms (parkinsonism, response to L-dopa and LB presence) shared a haplotype in chromosome 2. However, it had reduced penetrance, i.e. unaffected people carried the "pathogenic" haplotype too, and the familial phenotype was wider than PD (some of the affected people also had dementia, cortical plaques and neurofibrillary tangles in their brains). This *locus* was lately associated to age at onset in Parkinson´s disease in the studies that [188, 189] conducted in families affected by PD. Their *PARK3* regions partially overlapped with the original one and included the marker D2S1394 which showed the highest association for age at onset in PD.

  In an attempt to refine the *locus* and find the concrete gene, [190] sequenced the coding region of the 14 genes located on *PARK3* but could not find any pathogenic marker segregating with the disease. Other later studies have repeated the genotype for some of those 14 genes without success.

- **I.9.b. UCHL1 (*PARK5*): 4p14**

  UCHL1 (ubiquitin carboxyl-terminal esterase L1 or ubiquitin C-terminal hydrolase) gene has 9 exons and encodes for a protein with 223 amino acids.

  UCHL1 is highly specific for neurons. It removes small amides and esters at the C-terminal extreme of ubiquitin and also hydrolyzes polymeric ubiquitin chains into monomers: it is involved in the ubiquitin proteasome system acting as a deubiquitinating enzyme [191]. It is also a dimerization-dependent ubiquityl ligase [192].

  [193] found that members of one German family who were suffering from PD (tremor at rest, rigidity, bradykinesia, postural instability and good response to L-dopa) carried the mutation p.I93M, although penetrance was incomplete. This mutation implies partial loss of the hydrolase activity of the protein and, as a consequence, it could cause altered cleavage and turnover of the substrates, which remain unknown, leading to aggregation of proteins and cellular death; however, this decreased activity is not considered enough to cause PD and recent studies have concluded that, actually, its pathogenicity is due to a gain of function rather than to a loss of function: the mutation changes the structure of the protein and thus allows UCHL1 to be more prone to aggregation and to interact aberrantly with tubulin and with components of the chaperone-mediated autophagy inhibiting the process [194, 195]. The p.I93M mutation also increases the K-63 ubiquitination of αSyn, stabilizing it; both facts make that αSyn aggregates can accumulate inside the cell causing neuronal death.

UCHL1 is considered a good candidate to explain autosomal dominant Parkinson´s disease in the family. Its functions, its presence in Lewy bodies and its high abundance in human brain support the hypothesis. However, just two siblings of the family were alive at that moment thus making the subsequent research impossible. In addition, other studies have tried to confirm the association between UCHL1 and PD but the p.I93M mutation, nor any different mutation at the same gene, have not been reported again either for familial or for sporadic PD cases: this mutation is either a rare cause of PD or it has no influence on the disease and its presence in the two siblings was coincidental [196, 197].

Other polymorphism in UCHL1, p.S18Y, was discovered due to all the analyses carried out to determine the influence of UCHL1 in PD [197]. This variant does not change the structure of the protein or its hydrolase activity but it reduces the dimerization-dependent ubiquityl ligase activity of UCHL1: p.S18Y in UCHL1 reduces the K-63 ubiquitination in αSyn and, therefore, αSyn aggregates are not stabilized [192]. It is on debate whether or not the S>Y change is a protective factor against PD. The Y allele is considered a protective factor in some studies: [196] on German PD cases, familial and sporadic, and unrelated controls, [198] a meta-analysis of 11 previous studies with Asian and Caucasian populations analyzed together, or [199] the biggest meta-analysis to date with Asian and Caucasian populations analyzed separately (although results were significant only when using a recessive model for Asian population and a dominant model for Caucasian population; only a trend was found when using

other models). On the other hand, for other groups, the Y allele has no influence on PD risk: [200] in a French population, although they find association between Y and age at onset, [201] a case-control study plus meta-analysis of 8 previous studies (only Caucasian populations), [202] in Chinese sporadic PD cases and unrelated controls, or [203] in Japanese sporadic PD cases and unrelated controls. In the latter, there was a trend of association for those carrying the Y variant and a reduction of risk for early-onset disease. The opposite conclusions obtained could be explained by differences between studies in sample size and ethnicity as, for example, the Y variant is more common in Asians than in Caucasians. UCHL1 remains as a controversial *PARK locus*.

- **I.9.c. ATP13A2 *(PARK9)*: 1p36**

  After analyzing one Jordanian consanguineous family, [204] linked markers D1S436 and D1S2843 in region 1p36 with Kufor-Rakeb syndrome. Kufor-Rakeb syndrome is an autosomal recessive disorder characterized by juvenil onset, L-dopa responsive parkinsonism (rigidity, bradykinesia but not tremor at rest), rapid progression, dementia and pyramidal signs [205]. The causative gene was found in a non-consanguineous Chilean family who was also suffering this rare disease: *ATP13A2* gene has 29 exons and encodes for an ATPase with 1180 amino acids and 10 transmembrane domains. This ubiquitously expressed protein localizes in the lysosomal membrane. However, truncated proteins are retained in the endoplasmic reticulum (ER) and

are degraded by the proteasome [206]. Its function and substrates remain unknown although as a member of the P-type ATPase superfamily is supposed to use ATP to maintain an ion gradient and use the membrane potential to produce ATP. Deletions, duplications, insertions, splicing site mutations and nonsense mutations, that originate frameshifts and deleterious proteins, have been described in homozygosis or compound heterozygosis in Kufor-Rakeb cases.

Even when the phenotype of Kufor-Rakeb disease partially overlaps with Parkinson´s disease, some studies have been conducted to analyze the possible relation between ATP13A2 mutations and juvenile or early onset PD. The influence seems to be minimal although there are discordant opinions:

- No ATP13A2 mutation segregates with familial PD cases [207].
- Missense mutations in ATP13A2 have been described in heterozygosity in idiopathic early onset PD cases but also in healthy controls [208].
- However, another study conducted in familial and idiopathic juvenile and early onset PD cases found one homozygous carrier for a missense mutation in ATP13A2 plus some heterozygous for other missense mutations. Their unaffected parents were heterozygous for these mutations. The authors considered the homozygous mutation as a proof of the role of ATP13A2 in PD pathogenesis [209].

- **I.9.d. *PARK10*: 1p32**

The *PARK10 locus* was described by [210] after studying 51 Icelandic families (117 patients) with late onset Parkinson´s disease. Their genome-wide linkage analysis reported the highest value for marker D1S231 in 1p32 region. After that, some studies have been conducted to find the candidate gene related to late onset PD: [211] used iterative association mapping in singleton and multiplex families[23] and concluded that some SNPs in *HIVEP3* gene were related to risk for late onset PD whereas the *EIF2B3* gene, and, to a lesser extent, *USP24* were associated with age at onset. [212] obtained significant results for *CDCP2* gene in their whole genome association study (using 2 tiers, case-control and sibling pairs). [213] analyzed *HIVEP3* and *CDCP2* genes in multiplex and singleton families and confirmed that *HIVEP3*, that encodes for a protein that regulates transcription of viral genes and genes involved in immunity and inflammation which are processes affected in PD, was associated with PD. Nevertheless, their significant results were obtained for SNPs that were different from those associated with PD by [211] and did not found the association with *CDCP2*, confirming thus previous studies conducted with unrelated cases and controls. [214] did a 3 stage study with Norwegian, Irish and North-American unrelated cases and controls and concluded that *USP24* gene was a PD risk factor.

[23] The definition of singleton and multiplex families is based on the total number of parent-child triads and discordant sibpairs (DSP) in a family that can contribute to association tests. Singleton families contain only either a triad or a DSP. Multiplex families consist of at least two pairs of either triads or DSPs.

This *locus* could connect familial forms of the disease and late onset, which is characteristic of the idiopathic forms. However, this association still remains uncertain maybe due to the use of different methodologies and the diverse characteristics of the populations such as size and familial aggregation.

- **I.9.e. GIGYF2 (*PARK11*): 2q36-37 and *PARK12*: Xq21-25.**

  [215] conducted a genome-wide linkage study in multiplex families that did not carry parkin mutations. Under two different models, stringent diagnosis of PD or broader criteria for inclusion, regions in chromosome 2 (2q36-37) and chromosome X (Xq21-25) were linked to PD. The study was later extended to include more multiplex families and the results were unchanged. However, in this case, the region in chromosome 2 was related to PD only when considering PD patients diagnosed with a restrictive criteria whereas the region in chromosome X was only related to PD when using a broader criteria for inclusion of PD patients [216]. In both studies, the major part of the association for the region in X chromosome was due to brother-brother pairs, with sister-sister pairs and mixed sex sibships having lower linkage values, thus introducing the possibility that sex has some influence on PD susceptibility.

  These two regions were named as *PARK11* (2q36-37) and *PARK12* (Xq21-25). There has been no update on *PARK12* and the gene responsible for

the *PARK11*-associated PD remains unclear: [217] tried to validate the previous results in Caucasian European multiplex families ([215, 216] used Caucasian and Hispanic North-American families). They analyzed the linkage in the putative causative region limited by markers D2S126 and D2S125 but they could not replicate the values.

There are tens of genes in the 2q36-37 region, but the highest linkage score was obtained for a microsatellite marker located in the *GIGYF2* gene. Therefore, the later studies have focused on it. The protein has an unknown function. It interacts with Grb10 which is a growth factor receptor-binding protein with potential role in insulin and insulin-like growth factor signaling. The sequencing of its 27 coding exons has shown that there are plenty of mutations and polymorphisms in the gene. However, the frequency of these variants has been reported as different or equal between familial cases and healthy controls depending on the articles, thus originating opposite conclusions: [218] concluded that GIGYF2 could be considered responsible for autosomal dominant familial forms of PD with incomplete penetrance in Italian and French populations, whereas [219] in a Spanish population, [220] in Portuguese and USA populations, including sporadic PD cases too, and [221] in a Japanese population, including sporadic PD cases too, found no evidence to consider GIGYF2 responsible for familial PD. Moreover, [222] found that the frequency for common variants and haplotypes was not different between Australian sporadic PD cases and controls. At the present time, this *locus* could be considered, at best, a rare PD cause.

- **I.9.f. HtrA2 (*PARK13*): 2p13.1**

The mouse mutant *mnd2* (motor neuron degeneration 2) was spontaneously generated in 1990. Its phenotype includes altered gait, muscle wasting, repetitive movements, akinesia, degeneration of striatal neurons followed by widespread neuronal death in the late stages, and death by the day 40 of age. The p.S276C mutation in the *Omi* gene is responsible for this phenotype.

Omi, also known as HtrA2, is a nuclear-encoded mitochondrial serine protease that localizes to the mitochondrial intermembrane space. The p.S276C change is located on the protease domain of the protein and causes the loss of access to the active site pocket and the consequent decrease in the protease activity.

It has been postulated that loss of the protease activity of HtrA2 increases sensitivity to stress-induced cell death and is probable responsible for the massive loss of striatal neurons in the mice: HtrA2 could be a sensor of unfolding stress in mitochondria and when it is mutated there would be accumulation of misfolded proteins leading to mitochondrial dysfunction and permeability. In this case HtrA2 would leave the mitochondria and would have an apoptotic role [223].

Due to the neurodegeneration and parkinsonism observed in this rodent model, some studies have been conducted to analyze the relation between HtrA2 and PD in humans. [224] found that HtrA2 colocalized at Lewy bodies in brains of PD patients. They sequenced the coding regions of the gene plus the adjacent intronic sequences in German cases and controls and found that 4 sporadic PD cases with

typical PD features carried the p.G399S mutation in heterozygosis, whereas no control had it. This mutation decreases the protease activity of the protein because it is localized in its PDZ domain, which regulates this function. In addition, the p.A141S polymorphism was reported as a PD risk factor, only found in heterozygous state. The S variant also decreased the protease activity. Both variants are supposed to trigger the same pathological mechanism than the p.S276C mutation, although p.A141S could do it with a smaller effect. None of these three changes modifies the expression, stability or localization of the protein [224].

These conclusions could not been replicated in a study with North-American PD cases and controls where the p.G399S mutation was found in PD cases but also in healthy people and with the same frequency, there was no association between the p. A141S variant and PD, and, in addition, there was no association for any of the other polymorphisms they discovered by the sequencing of the complete gene [225]. There was no association even when stratifying the population in early or late onset. Recently, the analysis of the 5 most informative SNPs spanning the complete gene in a large number of cases and controls from populations collected worldwide showed again no association for this gene and PD [226].

It is known that PINK1 phosphorylates HtrA2 *in vitro*, thus increasing its protease activity. In human brains, when PINK1 is mutated the levels of phosphorylation of HtrA2 are decreased [126]. Considering this interaction and modification and the hypothetical pathological

mechanism that involves HtrA2 in *mnd2* neurodegeneration, it has been postulated that PINK1 and HtrA2 could work together in a mitochondrial pathway against stress. However, at least in *Drosophila*, HtrA2 is not essential in the pathway that involves parkin and PINK1 to control mitochondria integrity [227].

In conclusion, HtrA2 seems that does not have an essential role in PD pathogenesis.

- **I.9.g. PLA2G6 (*PARK14*): 22q13.1**

  The function of the protein is essential for membrane homeostasis: it is a calcium-independent phospholipase A2 that catalyzes the hydrolysis of glycerophospholipids.

  Mutations in *PLA2G6* were associated with homogeneous clinical presentations: NBIA (neurodegeneration associated with brain iron accumulation) INAD (infantile neuroaxonal dystrophy) and Karak syndrome. However, recently it has been associated with more heterogeneous phenotypes: in 5 families from India and Pakistan with young adult onset levodopa-responsive dystonia-parkinsonism cases, pyramidal signs and cognitive/psychiatric features but no iron accumulation, a genome-wide homozygosity study was conducted. Three of the families showed the highest values for homozygosity at chromosome 22 and a deeper analysis in the region found that three members of two different consanguineous families were homozygous for *PLA2G6* mutations in the coding region [228].

Other study analyzed the presence of mutations in *PLA2G6* in patients that have early onset parkinsonism plus other features like dementia, dystonia, psychosis, i.e. a broad phenotype. Two compound heterozygous were described after sequencing the 17 coding exons plus the exon-intron boundaries in 29 affected Japanese people. Although few patients with a very severe and atypical phenotype were included, the high frequency of the mutations observed raised up the possibility that this gene was related to those complex phenotypes [229].

Nevertheless, the situation is more complicated. At present time, there are confusing results about the relation between mutations in *PLA2G6* and autosomal recessive PD because the mutations have also been described in patients with typical forms of the disease (without severe and broad phenotypes). In addition there is controversy about the influence of heterozygosis and specific mutations in conserved coding regions. Furthermore, although all the studies have analyzed early onset PD patients, there are different results for sporadic and familial forms of the disease.

All these facts are reflected in the latest studies conducted in Asian population:

- One early onset Chinese PD patient (without atypical features) was homozygous for a mutation in the gene. However, neither his homozygous sister nor any of the heterozygous members of the family presented any PD feature [230].
- An heterozygous carrier for a mutation in PLA2G6 was reported after sequencing the 17 coding exons plus the exon-intron boundaries of

the gene in a Chinese population. This woman was suffering early onset sporadic PD without atypical features like dementia or dystonia [231].

- A case-control study conducted in a Japanese population could not find any association between any of the three mutations analyzed and the risk of developing sporadic Parkinson´s disease [232].

I.9.h. FBXO7 (*PARK15*): 22q12.3

Parkinsonian-pyramidal syndrome is a rare disorder that exhibits both parkinsonian and pyramidal-associated symptoms. Symptoms start in young adulthood, progress relatively slow and may culminate in severe movement incapacity. The response to levodopa of the parkinsonism is variable.

A genome-wide linkage study was conducted in an Iranian family with parkinsonian-pyramidal syndrome. The highest values were observed at chromosome 22 in a wide region that included 34 genes. Considering that the inheritance of the disease was autosomal recessive, the linked region was redefined and only 4 genes were included in the deeper analysis. After the sequencing of their coding regions, some variants were discovered in homozygosis. However, only one of them, p.R378G, located on *FBXO7* gene, segregated with the disease [233].

This gene encodes for a member of the F-box family of proteins. It includes an F-box motif, a 40 amino acids motif that interacts with Skp1 and acts as the scaffold in the SCF ubiquitin ligase complex which plays

a role in the ubiquitin-mediated proteasomal degradation, an N-terminal ubiquitin-like domain and a C-terminal proline rich region, both necessary to its target specificity. FBOX7 do not only works in SCF-mediated functions, it also interacts with proteins like HURP (mitotic protein), cIAP1 (inhibitor of apoptosis protein 1) and PI31 (proteasome inhibitor protein). FBOX7 also enhances the interaction of CDK6 with its targets. Nevertheless, its pathological mechanism and the hypothetical proteins involved remain unknown.

FBXO7 was confirmed as a pathological gene in other study with two European families whose members had early onset, pyramidal tract signs and progressive parkinsonism. Three novel mutations that segregated with the disease were reported in it: p.R498X in homozygosity in an Italian family and the compound heterozygous p.T22M with c.1144+1G/T in a Dutch family [234].

FBXO7 has two isoforms (Figure 28).

It was observed *in vitro* that loss of isoform 1 is pathogenic because Dutch patients developed the disease even when they were expressing the isoform 2. The Italian and Iranian cases showed before, did not express any of the isoforms due to their homozygous mutation, i.e. the mutations decrease the stability of the protein. One possible explanation for that is that FBXO7 has lost its proper cellular localization: isoform 2 lacks the N-terminal part of the protein which is necessary for its nuclear localization [235].

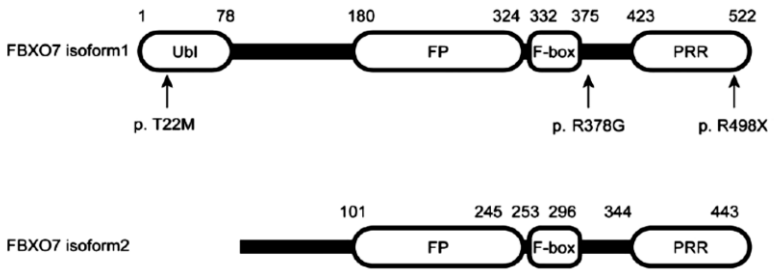


Fig. 28. Schematic representation of the FBXO7 protein isoforms.
Here, it is represented the domain organization of both proteins. The previously described mutations are indicated. Ubl: ubiquitin-like domain; FP: FBXO7/PI31 domain; F-box: F-box motif; PRR: proline rich region.

Taken from [235]

Thousands of studies have been conducted to find more *PARK loci* or susceptibility genes. There have been regions associated in almost all the chromosomes and some genes have been postulated as risk factors, mainly those related to dopaminergic transmission (*D2R*, *D3R*, *TH*, *MAOA*, *MAOB*…) or xenobiotic metabolism (*CYP2D6*, cytochrome P4501A1…). However, it has been complicated to replicate the results except for *SNCA*, *MAPT* and *GBA*. Even the mitochondrial DNA (mtDNA) has been widely analyzed: complex I is impaired in PD patients and it is known that there is a close relation between some pathological genes and mitochondrial function or integrity. Moreover, some PD patients have mutations in their mtDNA.

Two strategies have been employed:

- linkage disequilibrium studies: there are families were the disease is inherited due to the presence of a specific mutation. Genome-wide analyses and gene candidate approaches have been conducted to find regions where recombination in meiosis is lower than the expected. It is achieved by genotyping microsatellites which define the pathological region. However, due to the effect of penetrance, the reduced number of members in current families and the low percentage of cases caused by familial forms of Parkinson´s disease, nowadays this approach is less used.
- association studies with PD cases and healthy controls to compare the frequency of specific variants looking for significant differences. Lots of studies with different sample sizes and populations from all over the world with familial and/or sporadic PD have been conducted. Nowadays this is the best technique to find candidates in multifactorial diseases, specially the genome-wide association studies (GWAS): in this approach, the entire genome is scanned using densely distributed genetic markers thus being a powerful approach to identify common genetic variants of weak effect that underlie the risk of a common disease like PD [51].

The latest *PARK loci* have been found by GWAS:

- **I.9.i. *PARK16*: 1q32**

  [236] concluded that there were 4 regions associated with PD in a Japanese population: two of them included already known PD *loci*

(*SNCA* and *LRRK2*) but there were two new associated regions: *PARK16* (1q32) and other including the gene *BST1* (4q15).

[237] replicated the association for *PARK16* in a European population (and also for *SNCA* and *LRRK2*). Nevertheless, their fourth region included *MAPT* and no *BST1,* highlighting the possibility that these two genes could be related to ethnic-specific susceptibility.

PARK16 includes 5 genes: *SLC45A3*, *NUCKS1*, *RAB7L1, SLC41A1* and *PM20D1*. Some of them are functionally interesting candidate genes for PD etiology: *SLC41A1* is a magnesium (Mg^{2+}) transporter, *RAB7L1* is a small GTP-binding protein that plays an important role in regulation of exo-and endocytotic pathways and *NUCKS1* is a nuclear DNA-binding protein that contains several consensus phosphorylation sites for casein kinase II and cyclin-dependent kinases.

Some case-control studies, mainly using sporadic PD cases, have been conducted to confirm the association and to refine the *locus*. The most studied polymorphisms have been those described in the GWAS that defined the *locus*: rs823128 and rs947211 are located on the putative promoter of *RAB7L1* and in the intergenic region between *SLC41A1* and *RAB7L1* respectively. Significant results have been obtained in Taiwanese [238], in Chilean [239], in Ashkenazi Jews [240] and in Han Chinese [241], but not in a Spanish population [242].

In conclusion, more studies are required to clarify the concrete identity of *PARK16* which could be influenced by ethnicity.

- **I.9.j. *PARK17* and *PARK18***

PARK17 and *PARK18 loci* are still on debate: *BST1*, *GAK*, *HLA*, *VPS35* and *EIF4G1* are the candidates.

The first GWAS in familial PD had no significant results, although the highest values were found for regions in 4q22 (*SNCA*), 17q21 (*MAPT*) and 4p (*GAK/DGKQ*) (rs1564282 and rs11248051 are located on *GAK* whereas rs11248060 is in *DGKQ*) [243].

Other GWAS done in European population concluded that *SNCA*, *MAPT*, *BST1* (rs4698412) and a region in chromosome 12 (12q24) were associated with PD [244].

[245] concluded that there was a significant result for the SNP rs3129882 in intron 1 of *HLA-DRA* gene in their case-control study conducted in a Caucasian population from USA.

And another GWAS in Dutch cases and controls found an initial association for *SNCA* and *BST1* (rs12502586) and, after *post-hoc* analysis, also for *MAPT*, *GAK/DGKQ* (rs1564282 and rs2242235 and rs4690296 which are located far away from *GAK* and *DGKQ*, in gene *PCGF3*) and *HLA* (rs4248166)[24] [246].

It is noteworthy that although the same regions have been reported as possible PD *loci* in different GWAS, the highest values are not always obtained for the same markers and those can be in different genes

[24] There were 36 SNPs with significant results, with the lowest value for rs4248166, spanning a region of 440.1 kb that included *HLA-DRA*.

separated for tens of kilobases. To confirm the results is necessary to do case-control studies:

- *BST1*: [241] could not find any significant results for any of the 4 SNPs described by [236] in a case-control study in Han Chinese (rs11931532, rs12645693, rs4698412, rs4538475).
- *GAK*: it has been observed that there is association for rs1564282 and PD in case-control studies in a Caucasian population from USA [247] and in Han Chinese [248]. However, [242] could not replicate the results for rs11248051 in a Spanish population.
- *HLA*: [242] could not replicate the results for rs3129882 in a Spanish population.

More GWAS in sporadic PD have found that markers in regions containing *SNCA* (4q22) and/or *MAPT* (17q21) *loci* are related to PD susceptibility: [249] in a Caucasian population and [250] in a British population. Both genes are considered associated with PD risk without any doubt although it remains unclear which is/are the precise susceptibility variant/s.

[250] also found weak but consistent association for 4p15 (*BST1*) (rs4698412) and 4p16 (*GAK*) (rs1564282).

At present, *BST1* remains as a controversial *PARK locus* candidate because although some GWAS have found significant results in markers located on this gene, mainly for rs4698412, further replication of those associations in case-control studies has not been obtained.

BST1 (4p15.32) encodes for a protein that catalyzes the formation of cyclic ADP-ribose (cADPR). cADPR mobilizes calcium from ryanodine-sensitive intracellular Ca^{2+} stores in the endoplasmic reticulum.

GAK (4p16.3) encodes for a kinase that regulates the cell cycle. The results observed in rs1564282 support the idea that this (and not DGKQ that encodes for a diacylglycerol kinase) could be *PARK17 locus*.
However, *VPS35* (16q11.2) has also been postulated as *PARK17*: VPS35 is involved in the recycling of membrane proteins between endosomes and the trans-Golgi network, and it is evolutionarily highly conserved. An exome sequencing analysis was conducted in a Swiss family with autosomal dominant late onset PD. The p.N620N mutation in *VSP35* gene was described in heterozygosis in all the affected members of the family. It was also observed in three more families and in one patient with sporadic PD, but not in any of the controls analyzed [251]. The same results were obtained after the exome sequencing analysis of an Austrian family with the same phenotype than the Swiss family: tremor predominant dopa-responsive PD, similar to idiopathic Parkinson´s disease. The p.N620N mutation was present in all the affected family members and it was cosegregating with the disease in an autosomal dominant mode. It had high but incomplete age-dependent penetrance because unaffected carriers were reported too [252]. In addition, p.N620N was detected in 3 patients with autosomal dominant PD and also in 1 patient with sporadic PD but not in controls in a Japanese population [253].

The presence of the amino acid asparagine in position 620 is highly conserved among species. For all these reasons, the p.N620N change has been suggested as a pathogenic mutation.

HLA-DRA (6p21.3) encodes for the major histocompatibility complex class II and it has been proposed as *PARK18 locus*. This could represent the link between inflammation and Parkinson´s disease which has been widely studied.

Nevertheless, *EIF4G1* (3q27.1) has also been named as *PARK18*: the genome-wide linkage analysis of a French family with autosomal dominant late onset Parkinson´s disease (phenotypically similar to idiopathic PD) revealed a significant linkage for the region 3q26-q28. The subsequent deeper analysis found that the p.R1205H mutation in *EIF4G1* gene segregated with the disease in all the affected family members. The mutation was absent in unrelated control subjects but it was identified in heterozygosis in seven families from USA, Canada, Ireland, Italy and Tunisia.

EIF4G1 is the core scaffold of the translation initiation complex (eIF4F) that regulates the translation initiation of mRNAs encoding mitochondrial, cell survival and growth genes in response to different stresses. The p.R1205H mutation impairs complex formation, consistent with a dominant-negative loss of function, and, therefore, it impairs the ability of cells to rapidly and dynamically respond to stress, presumably through changes in the translation of existing mRNAs essential to cell survival [254].

I.10. Genetic susceptibility factors.

○ I.10.a. *MAPT*: 17q21.31

Microtubule-associated protein tau (MAPT) is a microtubule binding protein that is particularly abundant in axons. It interacts with tubulin and promotes its assembly into microtubules. Moreover, tau stabilizes the structure of microtubules and is also involved in axonal transport along this network through its interaction with the motor proteins dynein and kinesin.

There are six isoforms in adult human brain which can be differentiated by the number of N-terminal repeats (N repeats, each one with 29 amino acids; there can be 0, 1 or 2) and C-terminal repeats (R repeats or microtubule-binding repeats, each one with 31 or 32 amino acids; there can be 3 or 4). Alternative splicing in exons 2 and 3 determines the number of N repeats whereas alternative splicing in exon 10 , which encodes for the second R repeat, determines the number of R repeats (4R bind more efficiently to microtubules than 3R) (Figure 29) [255]. In a normal brain, the levels of 4R and 3R forms are similar.

All these functions are lost in tauopathies[25] where the levels of phosphorylation of tau are higher than the normal and the ratio 4R/3R is altered. This originates deficits in microtubule stability and in microtubule-dependent trafficking and subsequent neuronal death. In addition, tau

[25] Alzheimer´s disease (AD), progressive supranuclear palsy (PSP) or corticobasal degeneration (CBD) are some of these neurodegenerative disorders.

assembles into filaments through its N and R repeats. After cell death, the aggregates remain in the extracellular space. Those neurofibrillary tangles are the characteristic hallmark of tauopathies.

Nevertheless, these deposits are abnormal in Parkinson´s disease which is a synucleinopathy (it is characterized by αSyn deposits).

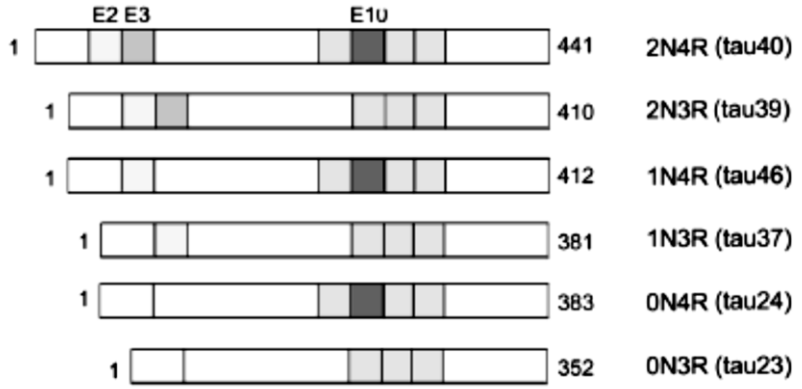


Fig. 29. Tau isoforms.
Gray boxes represent the N repeats (coded by exons 2 and 3) or the R repeats (coded by exons 9, 10, 11 and 12). The second microtubule-binding repeat is highlighted in dark gray.

Taken from [255]

The relation between *MAPT* and Parkinson´s disease is nowadays genetic:

There is a ≈2Mb region in 17q21.31, centered in *MAPT*, with high linkage disequilibrium that includes other genes like *CRHR1*, *IMP5* and *Saitohin* [256]. Inside this region there is an inverted fragment of 900kb that defines two different haplotypes called H1 and H2. There has been no recombination between both haplotypes since they diverged over 3 million years ago.

H1 haplotype (direct orientation) is distributed worldwide and shows a normal pattern of recombination and genetic variability, whereas H2 haplotype (inverted orientation) is almost exclusive of Caucasian population [257] and shows an extreme homogeneity, i.e. there are some H1 sub-haplotypes but no H2 sub-haplotypes. These two facts support the idea that the H2 haplotype suffered positive selection: its frequency increased rapidly starting just from few founder chromosomes [258].

The frequency of H1 haplotype has been analyzed in different neurodegenerative diseases: it is higher in PSP and CBD patients and also in PD cases.

The major part of studies have used the 238bp intronic deletion present in *MAPT* to differentiate H1 and H2 [259], usually with the same conclusion: homozygous H1H1 carriers have increased risk to develop PD [260-263]. And that has been obtained for familial and sporadic PD cases, independently of sex, age at onset and even ethnicity, although some opposite results have been reported [264] (small sample size is the common given reason for those divergences).

Some sub-haplotypes containing H1 plus other polymorphisms have been proposed to increase the probability to develop PD. However, the identity of those polymorphisms remains unclear. There is no validation for these results because there are different conclusions between studies (each study proposes a different SNP: rs242562, rs2435207, rs3785883…) and, moreover, GWAS have also pointed to other variants along *MAPT* gene or even further. Maybe, the high linkage disequilibrium (LD) along the chromosomal region

makes difficult to differentiate between the polymorphisms that are contributing to the disease from those that are merely in LD with the "functional" polymorphism.

In conclusion, there is currently insufficient evidence to refine the disease association to a specific region within *MAPT* or neighboring genes. *MAPT* is the supposed candidate but it is possible that genes near it are also involved in PD. Furthermore, it is difficult to explain why the frequent H1 haplotype is related to different pathologies (PD, PSP, CBD).

There is also a proteic connection between tau and PD which includes α-synuclein and can be related to the formation of aggregates:

αSyn and tau are abundant neuronal proteins. Normally both adopt an unfolded conformation but they can polymerize: αSyn selfpolymerizes but tau requires cofactors and αSyn can act as such. [265] observed that *in vitro* tau and αSyn synergistically promote and propagate each other's polymerization into fibrils. αSyn initiates the process by the formation of amyloidogenic "seeds" which may be degraded after the initiation of tau polymerization. This induces a conformational change in tau and triggers the process. Tau and αSyn preferentially form homopolymers, although they can also associate in the same filament.

Moreover, *in vitro*, αSyn stimulates the phosphorylation of tau by GSK3β, which is one of the over 20 kinases that have been found to phosphorylate it. Those three proteins form a heterotrimeric complex but heat

shock protein 70 (Hsp70) suppresses αSyn-induced phosphorylation of tau by GSK3β through its direct binding to αSyn [266].

o I.10.b. *GBA*: 1q22

Gaucher disease (GD) is the most prevalent lysosomal storage disorder. It is an autosomal recessive maladie caused by homozygous or compound heterozygous mutations in the glucocerebrosidase (*GBA*) gene. Glucocerebrosidase (GCase) catalyzes the cleavage of the glycolipid glucocerebroside in glucose and ceramide. However, the mutated protein losses this ability thus leading to the accumulation of glucocerebroside mainly in macrophages: they increase their size and acquire a diagnostically characteristic appearance (they are called "Gaucher cells"). These macrophages accumulate in the spleen and liver, which causes organ enlargement and inflammation.

Gaucher disease is clinically divided in 3 types:

- Type 1 or nonneuronopathic is the most common. This is the mildest form of the disease with no primary central nervous system involvement. There is enough residual enzymatic activity to prevent glucocerebroside accumulation in other cells rather than macrophages. It is panethnic (although Ashkenazi Jews are the most affected population).
- Type 2 or acute neuronopathic is the rarer and most severe type. It is associated with fatal progressive neurological manifestations that cause death usually in the first year of life. There is complete deficiency in GCase activity and glucocerebroside accumulates in neurons and other cell types.

- Type 3 or juvenile subacute neuronopathic has a slower progressive course with neurologic manifestations. People develop the disease in their adolescence and die when they are in their 30s.

Some GD patients develop parkinsonism. They present LB with mutated GCase in their brains. Their phenotype is wide, ranging from idiopathic PD symptoms with good response to L-dopa to early onset PD with dementia. In addition, relatives of those patients (that usually are carriers of *GBA* mutations in heterozygosis) develop PD with higher frequency than general population. Furthermore, there is a higher frequency of *GBA* mutations in Parkinson´s disease patients and [267] has also reported that there is a significant deficiency of GCase activity in *substantia nigra* and cerebellum in PD cases without *GBA* mutations.

Around 300 mutations have been described throughout the gene: missense mutations are the most frequent, although nonsense, intronic, splicing mutations, frameshifts, insertions, deletions and even recombinations have been described (there is a pseudogen only 16kb downstream the gene, just 2kb shorter and with 96% of homology[26]).

Ashkenazi Jews are the most affected by Gaucher disease. The 70% of the GD cases carry the p.N370S mutation. However in other populations, were GD is less prevalent, p.L444P is more frequent than p.N370S. The frequency of mutations is supposed to be different between ethnic groups.

[26] It can influence the genotype and cause misleading results: for that reason genotyping should be done by sequencing all the exons with primers specific for the gene.

It is noteworthy that there is no correlation between clinical phenotype and genotype: vast phenotypic variations among patients with the same genotypes have been reported, even in sibling pairs and twins. Moreover, there are genotypic differences between patients with the same phenotype. There is also no correlation between phenotype and residual enzymatic activity.

Most of the studies that have analyzed the frequency of *GBA* mutations in PD patients and in controls have concluded that *GBA* mutations can be considered as a PD risk factor (even for familial and early onset Parkinson´s disease): in Ashkenazi Jews [268], in Caucasians from Canada [269], in people from different ethnicities from USA [270], in Portuguese [271], in Italian [272], in Brazilian [273], in Chinese [274, 275] and in Korean [276]. Nevertheless, some studies (in Tunisian [277] and in Norwegian [278] for example) did not found significant association for *GBA* mutations and Parkinson´s disease. The large, collaborative, international multicenter study, with thousands of PD patients and controls, conducted by [279] shed light to this controversy. They observed that there is an increased probability to develop PD for carriers of *GBA* mutations, and that is not exclusive for a specific ethnicity or a specific mutation. Moreover they confirmed a previous conclusion obtained by [280]: the age at onset of Parkinson´s disease was significantly lower among patients with *GBA* mutations as compared with those without mutations.

Nowadays, *GBA* is considered a PD risk factor, although there are important differences in methodology and in ethnicity between studies: for example, some studies sequenced the complete gene whereas others only analyzed concrete mutations, and the number of samples analyzed in some

cases was not enough to obtain true statistically significant results. Those divergences preclude the proper quantification of the risk and the frequency for specific mutations or in overall.

Some models have been proposed to explain the possible connection between *GBA* mutations and PD:

1. There could be a gain-of-novel-function in mutated GCase: the unstable or misfolded protein would contribute to enhance the aggregation of αSyn by direct or indirect interaction. That would explain the presence of mutant GCase in LB although other explanation could be that αSyn oligomers trap the misfolded GCase.
2. Glucocerebrosidase is a membrane-associated lysosomal protein, but when mutated, it could alter the lysosomal or autophagic pathway thus leading to αSyn accumulation and subsequent aggregation, or leading to dysfunctional mitophagy and, consequently, damaged mitochondria would accumulate thus causing cellular death.
3. Mutant GCase, rather than the wild type protein, is a substrate for parkin-mediated ERAD [281][27]. As a consequence, GCase would block the interactions between parkin and its other substrates, leading to ER stress and cell death.

[27] Endoplasmic reticulum-associated degradation: misfolded proteins are detected by the ER quality control machinery and after several attempts to refold them by the ER chaperones, the misfolded proteins are retrotranslocated from the ER to the cytosol, ubiquitinated and eliminated by the ubiquitin–proteasome system.

4. The accumulation of glucocerebroside originates vesicles that are suitable for αSyn aggregation because α-synuclein changes its structure in a lipid environment and tends to aggregate on the surface of lipid vesicles.
5. In a "Gaucher cell", large amounts of exogenously ingested or endogenous αSyn could aggregate and acquire the prion form. Therefore, it has been proposed that a second hit might occur in the *GBA* gene to produce a disease-causing somatic mutation in a subset of macrophages. In those patients that have the mutation, cellular death or enhanced transfer via exosomes[28] could cause that these "Gaucher cells" unload the prion form of αSyn into the extracellular environment near neurons (Figure 30) [282].

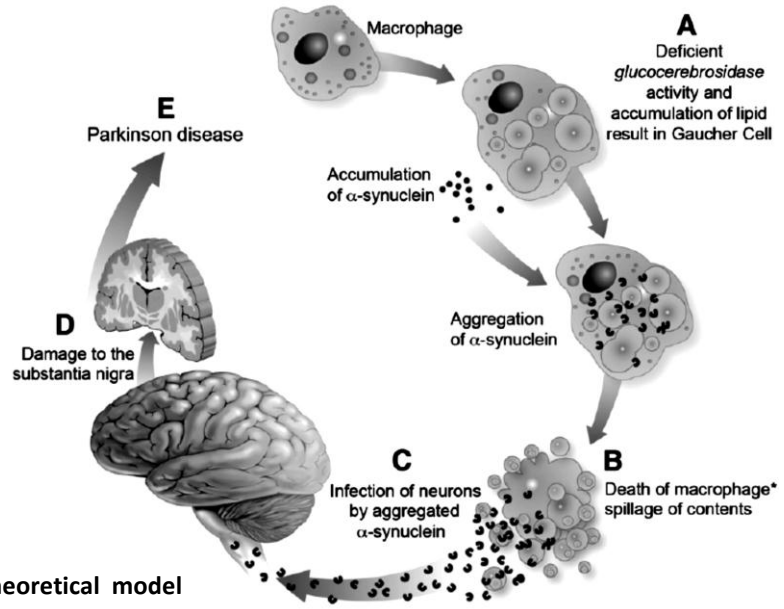


Fig. 30. A theoretical model for αSyn as a prion in *GBA*-associated parkinsonism.
The asterisk denotes the less probable process.

Modified from [283]

[28] Exosomes are small intraluminal membranous vesicles (50–100 nm) that are released into the extracellular environment.

Any of the models is definitive. Any can explain why only some GD patients develop parkinsonism and why only some PD patients are carriers of mutations in *GBA* gene. It is supposed that *GBA* just contributes to, but not initiates, the development of *SNCA* pathology: the mutations just exacerbate and accelerate the process.

o I.10.c. *NR4A2*: 2q24.1

NR4A2, also known as Nurr1, is essential for the development and survival of dopaminergic neurons. It is important not only during development but also in adulthood.

Nurr1 does not work alone; it works with other transcriptional factors and neurotrophic growth factors in the different development and maturation stages of mesencephalic dopaminergic neurons.

Nurr1 belongs to the nuclear receptor family of transcription factors, concretely to the subgroup of nuclear orphan receptors, because unlike the others, it functions independently of ligands: thus, the regulation of the activity of Nurr1 might be mediated by the control of its expression or by post translational modifications of the protein.

Nurr1 acts as a monomer or as a dimer and, for example, activates the transcription of tyrosine hydroxylase and enhances the expression of dopamine transporter. It is highly expressed in the *substantia nigra* but also in other parts of the brain and the body.

In addition, it is related to some *PARK* genes:

- Nurr1 transcriptionally regulates the expression of αSyn: *in vitro*, in SH-SY5Y cells, there is an increased expression of α-synuclein when the expression of Nurr1 is blocked. However, the expression of αSyn is just slightly decreased when Nurr1 is overexpressed [284]. Moreover, the expression of Nurr1 is compromised in cells that contain αSyn inclusions like dopamine-containing cells present in SN in brains from PD cases [285].
- DJ-1 activates Nurr1 via the ERK1/2 pathway to upregulate the expression of tyrosine hydroxylase [286].

Nurr1 has also been associated with neuroprotection against stress (induced by CREB) [287] and with promotion of cell survival (it regulates the expression of Bax via an interaction-dependent repression of p53 [288]; and, once the NMDA receptor stimulation of neurons is triggered, CREB activates Nurr1 that increases the expression of BDNF to prevent apoptosis [289]).

Some genetic studies have been conducted:

[290] analyzed the exon sequence of the gene in German PD patients (familial and sporadic cases) and controls, and described two new mutations in heterozygosis in 10 familial cases. These patients had clinical features similar to those seen in idiopathic PD. These changes were in the noncoding exon 1 (c.-291delT and c.-245T>G) and caused decreased expression of the gene, that is, lower levels of mRNA *in vitro*, in cell lines and in lymphocytes of affected individuals. They concluded that both mutations could be considered the cause of familial Parkinson´s disease with an autosomal dominant mode of

inheritance, because both mutations were described in heterozygosis. However, [291] and [292] tried to confirm this conclusions in 50 and 44 familial PD cases with European origin by sequencing the exon 1 and neither of them found any mutation, even new mutations. [293] also sequenced the exon 1 in 108 PD cases with familial history and an apparent autosomal dominant PD, mainly French, and again did not find any mutation.

The 7048G7049 variant in intron 6 (c.1361+16insG) was previously reported in a study about mental and neurologic disease. [294] analyzed it by sequencing and restriction enzyme analysis in familial and sporadic PD cases and in controls from USA. They found that this mutation in homozygosis could be considered a risk factor for PD (familial and sporadic cases). The carriers did not differ in their clinical features from those of typical PD. Nevertheless, [295] obtained a different conclusion: in white controls and PD cases found that heterozygous carriers of the mutation had an increased risk to develop PD. And [296], also in a population from USA, found no significant differences for this intronic variant neither in homozygosis nor in heterozygosis between healthy controls and familial Parkinson´s disease patients. In addition, they could not find any of the two previously described mutations in exon 1 in this group. [297] obtained the same results than [296] for intron 6 and exon 1 in a group composed of controls and sporadic PD cases mainly from Germany, although they described two new mutations in exon 1: c.-253C>T and c.-223C>T.

At the present time, genetic alterations at the NR4A2 *locus* are neither a major cause of familial PD in Europe nor a significant PD risk factor.

I.11. Dementia.

Dementia is a syndrome characterized by progressive deterioration of intellectual functions (cognition and behavior).

Some risk factors have been described [298]:

- Age: there are early cases of dementia but the majority of people develop it after the age of 65; the prevalence for people older than 65 years is around 6%.
- Sex: there are controversial results; the majority of studies have not found any relation but some of them have concluded that women are more affected, particularly at older ages [299].
- High education level and physical activity, as it implicates reduced vascular risk and obesity and enhanced fitness, are considered protective factors, whereas controversial conclusions have been reported for body mass index, alcohol consumption and smoking.

Dementia is one of the non-motor symptoms that a PD patient can develop. Its risk increases with longer evolution periods: 20 years after the diagnosis, almost 80% of PD patients will show dementia [300].

Clinicopathological studies have shown that dementia is more common in PD patients whose main feature is akinesia or rigidity than in those with tremor dominant phenotypes [301]. Early indicators associated with cognitive decline include increasing apathy, impaired attention and concentration, forgetfulness and excessive daytime sleepiness. However, the diagnosis of

Parkinson´s disease with dementia (PDD) is challenging because the presence of dementia and parkinsonism describes also other disease: dementia with Lewy bodies (DLB).

PDD and DLB could be considered part of a continuous spectrum of syndromes mainly characterized by parkinsonism and dementia, but they are considered two different clinicopathological syndromes.

It is considered that a person is affected by DLB if there are 24 or less months of difference between the beginning of motor symptoms and the beginning of cognitive decline. If dementia occurs later in the disease it is considered PDD. This definition is arbitrary because it is known that the onset of dementia is earlier in DLB than in PDD and that the progression is faster and shorter, but this earlier onset is not always before the 24^{th} month after the beginning of motor symptoms. Other features that differentiate PDD and DLB are that the frequency of visual hallucinations is greater in DLB than in PDD and that PDD patients possess LB deposits in their brains but DLB cases possess LB deposits in a more widespread pattern plus cortical Aβ plaques, in a situation that is reminiscent of AD patients that present Aβ aggregates but also neurofibrillary tangles [302].

Not all PD patients will develop dementia. Nevertheless, mild cognitive impairment (M.C.I.) is present since the earliest stages. M.C.I. can be defined as a cognitive decline from previous performance baseline, that is considered abnormal for the patient´s age, but with retention of normal daily functioning

[301]. This cognitive impairment can stabilize, evolve to dementia or recover to normal cognition. The fact that people showing M.C.I. recover their normal cognition argues against the supposed consecutive steps in cognitive decline: normal cognition → M.C.I. → dementia.

The most widely test used to determine the cognitive state of PD patients is the Mini-Mental State Examination (MMSE) [303]: this test lasts ≈10 minutes and measures multiple cognitive domains like orientation, registration, attention and calculation, recall, language and visuospatial function. However, the Montreal Cognitive Assessment (MoCA) test [304] is more sensitive to identify mild cognitive impairment [305, 306], also in PD [307], and for that reason it is nowadays more recommended than MMSE especially in patients with cognitive complaints and functional impairment whose MMSE score is ≥26[29]. Nevertheless, MMSE is adequate to measure progression once patients develop dementia.

MoCA is very similar to MMSE in duration (≈10 minutes), maximum score (30 points) and cognitive domains measured (orientation, recall, visuospatial function, attention, language but also executive functions).

Dementia with Lewy bodies (DLB), Parkinson´s disease (PD) and other maladies like frontotemporal dementia (FTD), Huntington´s disease (HD), Creutzfeldt-Jakob disease (CJD)… only represent ≈30% of dementia cases.

[29] It is considered that there is normal cognition function if MMSE score is ≥26. Sometimes the cutoff is 24 instead of 26.

Dementia is mainly caused by Alzheimer´s disease (AD), ≈50% of cases, and vascular dementia (VaD)[30], ≈20% of cases [308].

It is essential to select a well-characterized affected population and do longitudinal studies (better than cross-sectional ones) to avoid inaccurate results due to the heterogeneous etiology of dementia.

I.11.a. Alzheimer´s disease (AD):

Alzheimer´s disease is a progressive neurodegenerative disorder that begins as mild short term memory deficits and culminates in total loss of cognition and executive functions. It is characterized by neuronal loss, mainly in hippocampus and cerebral cortex associated with two hallmark pathological lesions: extracellular deposition of amyloid plaques (mainly formed by Aβ) and intracellular deposition of neurofibrillary tangles (mainly formed by paired helical filaments of hyperphosphorylated tau) [309].

The majority of AD patients present the late onset form of the disease and have no familial history; the etiology is unknown although some genetic and/or environmental factors have been postulated as pathogenic.

The *APOE* gene (*AD2*) is nowadays the strongest and most highly replicated genetic risk factor for non-familial AD, concretely the ε4 allele. In humans, there are mainly three isoforms (Figure 31) [310]:

[30] It may be caused by various types of vascular pathology in the brain, such as infarctions.

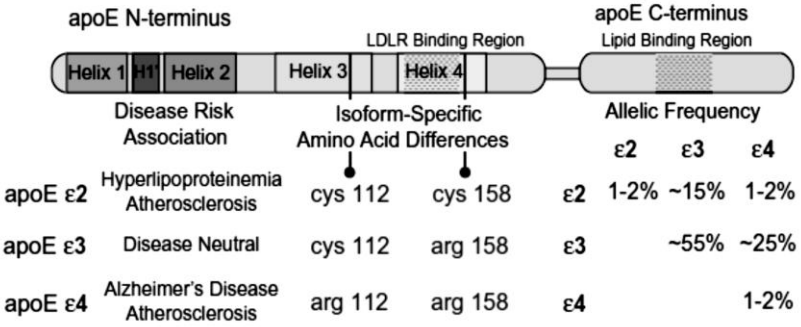


Fig. 31. ApoE isoform-specific differences.
Amino acids in positions 130 and 176, which are traditionally named as 112 and 158, respectively, determine the charge and structural properties of the protein, which may ultimately influence the functional properties of the isoform: ε2 and ε3 bind preferentially to HDLs (high-density lipoproteins) whereas ε4 to VLDLs (very-low-density lipoproteins).

Modified from [310]

APOE (apolipoprotein E) regulates the metabolism of lipids by directing their transport, delivery and distribution from one tissue or cell type to another: for this purpose, APOE binds to lipids and forms lipoprotein particles that bind to specific cell surface receptors [311]. It is expressed by several cell types, but with highest expression in the liver and in the central nervous system. APOE also works in synaptic function, immune regulation and intracellular signaling.

The pathogenic mechanism of APOE in AD is not clear: there are controversial results about the influence of APOE in Aβ aggregation, accumulation and clearance. Those processes could be different between the three isoforms.

The influence of APOE in Parkinson´s disease has also been studied but there are opposite and not definitive conclusions: ε4 allele has been mostly associated with higher PD risk ([312] in familial PD, [313] in Mexican), but also

with lower PD risk ([314] in Caucasians non-Hispanics), decreased age at onset ([315] in Caucasian, [312, 316] in familial PD, [317] in Australian) and risk of dementia in PD ([318] sporadic and familial PD, [316] in familial PD). However, ε2 allele has also been considered a PD risk factor ([319] a meta-analysis, [320] in Thai) and some studies have concluded that any APOE allele is related to Parkinson´s disease ([321] in Irish, [322], [323] in Norwegian, [324] in a large case-control study with thousands of people).

Factors like the ethnicity, the sample size and the objective of the study might have originated those divergences. Future studies should be longitudinal and conducted in large populations.

Around 1% of the cases is clustered in families with mutations in the genes *APP*, *PSEN1* or *PSEN2* and develop the early onset autosomal dominant form of the disease. Familial and sporadic cases have the same clinical features and there is no cure for them.

APP (AD1) is ubiquitously expressed. Its physiological function is unknown [325]. APP encodes for a single-pass transmembrane protein with a large extracellular domain that resembles a signal-transduction receptor [326]. It is cleaved by the γ-secretase complex, a proteolytic complex formed by PSEN1 or PSEN2, Aph1, PEN2 and Nicastrin; the last 3 proteins are necessary for the assembly of the complex [327] (Figure 32).

Presenilins (PSEN1 -AD3- and PSEN2 -AD4-) are ubiquitous. These intramembrane cleaving proteases are the catalytic members of the γ-

secretase complex. They also work in other processes like protein degradation or calcium homeostasis regulation: they are passive ER Ca^{2+} leak channels but when mutated they lead to supranormal Ca^{2+} release from the ER [328].

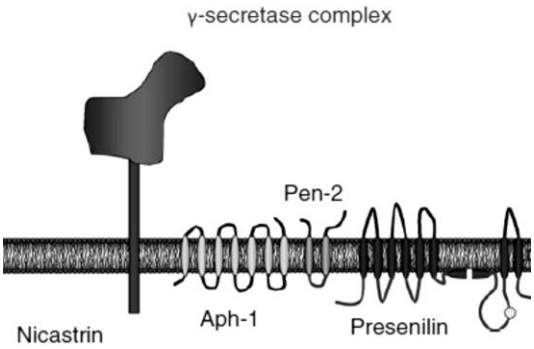


Fig. 32. The γ-secretase complex.
The gap in presenilin reperesents the endoproteolytical cleavage necessary to activate it. The circle represents an early onset Alzheimer´s disease mutation.

Modified from [329]

The main supposed pathological process that originates AD is related to APP processing (Figure 33). The 90% of mutations described in AD familial cases are in presenilins. They cause a change in the cleavage point in the amyloidogenic pathway: in normal conditions, $A\beta_{40}$ is the majority fragment whereas when PSEN1 or PSEN2 are mutated, $A\beta_{42}$ is the most common fragment. $A\beta_{42}$ is more amyloidogenic and more prone to aggregate than $A\beta_{40}$. APP duplications are also pathogenic [330, 331].

Amyloid is transported and released to the surface. In AD patients, Aβ monomers aggregate into fibrils and form extracellular plaques. Amyloid can be removed by autophagy but during the onset of AD, although autophagy is increased, the transfer of autophagic vesicles to the lysosomes is impeded and

this may contribute to the accumulation of Aβ. The subtle effects of the oligomers on synapses progressively cause neural injury, neuritic dysfunction and finally death.

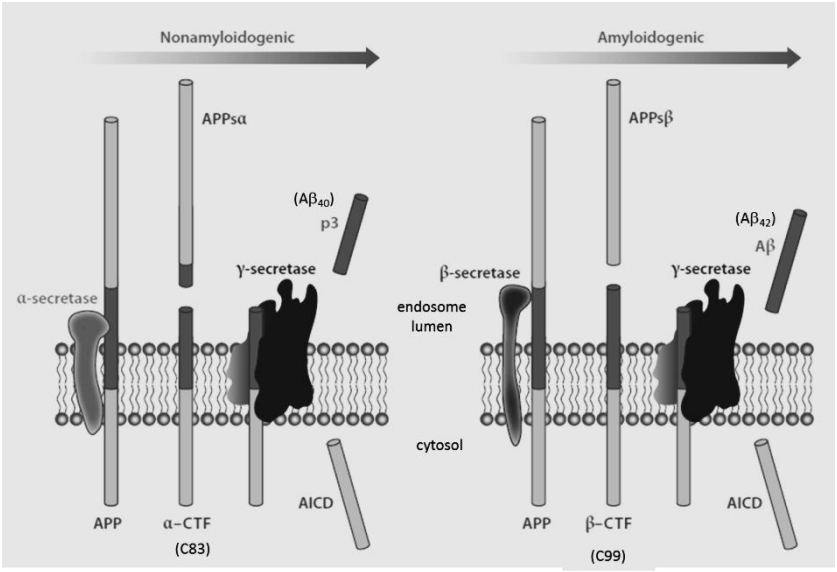


Fig. 33. Cleavage of the amyloid precursor protein (APP).
It is sequential and occurs by two pathways. The nonamyloidogenic processing of APP involves α-secretase and γ-secretase and generates the $A\beta_{40}$ fragment whereas the amyloidogenic processing involves the β-secretase and the γ-secretase and can generate the $A\beta_{40}$ or the $A\beta_{42}$ fragments. Both processes also generate soluble ectodomains (sAPPα and sAPPβ) and identical intracellular C-terminal fragments (AICD). AICD act as transcription factors in the nucleus.
The dark gray fragment in APP is the Aβ peptide.

Modified from [332]

Mutations in any of those 3 genes may act by increasing the steady state level of Aβ, altering the $A\beta_{42/40}$ ratio or altering the amyloidogenic potential of Aβ. More Aβ monomers and less clearance can trigger their aggregation and initiate the pathogenic cascade. That would explain the pathological process in less than 1% of the affected people so there have to be

other undescribed mechanisms involved in the pathogenesis because Aβ deposits are present in all AD cases.

At the present time, the calcium hypothesis is being explored. Aβ oligomers can insert into the plasma membrane and form Ca^{2+}-permeable pores. Calcium signaling is utilized by neurons to control a variety of functions, including membrane excitability, neurotransmitter release, gene expression, cellular growth, differentiation, free radical species formation and cell death. The hypothesis supposes that there could be a remodeling in calcium signaling that would result in the learning and memory deficits that occur early during the onset of Alzheimer´s disease.

[333] described for the first time the gene *CALHM1* (*calcium homeostasis modulator 1*) that encodes for a multipass transmembrane glycoprotein. CALHM1 is a cell surface protein of neuronal origin, although it is also present in ER membrane, that shares sequence similarities with NMDAR. It homomultimerizes and generates a calcium-selective cation current at the plasma membrane. CALHM1 controls cytosolic calcium concentrations, a mechanism that may lead to ERK1/2 activation [334]. Cytosolic Ca^{2+} is critical for the regulation of APP processing, and the p.P86L change decreases the concentration of calcium and that originates an increase in Aβ levels as there is an inhibition of the control of APP processing by CALHM1 (Figure 34).

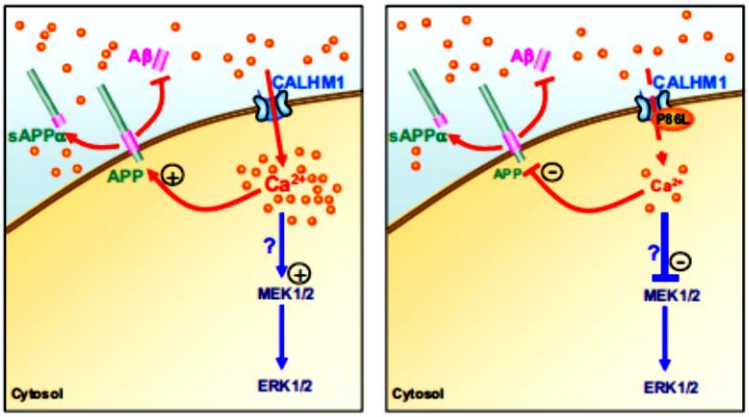


Fig. 34. Influence of the p.86L mutation in the role of CALHM1 in calcium signaling: relevance for APP metabolism and Alzheimer´s disease pathogenesis.

Taken from [334]

CALHM1 also increases Ca^{2+} leak from the ER and, more importantly, reduces ER Ca^{2+} uptake. As a result, the calcium content of the endoplasmic reticulum is drastically decreased and that triggers the ER stress which can evolve to cell damage [335].

[333] conducted a case-control study that showed that the p.P86L mutation was associated with Alzheimer´s disease. However, other case-control studies have not found that association for p.P86L and AD either in Caucasians or in Asians ([336], [337] in Japanese, [338] in Chinese Han). It is noteworthy that the L allele is less frequent in non-Caucasian populations than in Caucasian.

Moreover, there are studies that have analyzed the influence of this mutation in the Aβ levels in cerebrospinal fluid: [339] found that this relation exists in healthy people, whereas [340] found that there is no relation between p.P86L and $A\beta_{42}$ levels in AD patients.

I.11.b. Frontotemporal dementia (FTD):

Frontotemporal dementia (FTD) is the second most common cause of dementia in people younger than 65 years. Men and women are equally affected. There is focal atrophy of the frontal and temporal lobes that leads to progressive behavioral changes, cognitive decline and language difficulties and, finally, to severe dementia.

There are three clinical subtypes (one behavioral and two language variants):

- bvFTD (behavioral variant of FTD) is characterized by loss of empathy, apathy, selfishness, neglect of personal hygiene, disinhibition, irritability, gluttony…
- in SD (semantic dementia) there is anomia[31] and impaired comprehension of words, objects or faces, although there is a relative preservation of grammar and pronunciation and the speech is fluent. Patients use generic terms like "thing" and are unable to understand less frequent words.
- in PNFA (progressive non-fluent aphasia[32]), patients have problems with pronunciation; their speech is non-fluent, agrammatical, and poorly articulated with phonological errors. However, word comprehension and object recognition are well preserved [341, 342].

[31] Anomia: language disorder that prevents call things by their name.
[32] Aphasia: loss of the ability to produce or understand the language.

FTD forms part of a neuropathologically heterogeneous disorder called frontotemporal lobar degeneration (FTLD). FTLD also includes amyotrophic lateral sclerosis (ALS), progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD). There is no cure for FTLD patients although there are some symptomatic treatments.

- PSP is characterized by postural instability and falls within the first year of disease, supranuclear gaze palsy and symmetric parkinsonism minimally or unresponsive to levodopa (there is brainstem-predominant atrophy).
- CBD is characterized by markedly asymmetric parkinsonism with dystonic posture, myoclonus[33], apraxia[34], alien limb syndrome and sensory or visual neglect (there is cortical-predominant atrophy) [343].
- ALS is a progressive neurodegenerative disease that affects motor neurons: the death of upper and lower motor neurons in the brainstem, motor cortex and spinal cord leads to progressive muscle weakening, paralysis and eventually death due to respiratory failure. Around 90% of cases are sporadic whereas the remaining 10% are familial caused by mutations in some genes: *SOD1*, *TARDBP*, *FUS, VCP* and *CHMP2B* amongst others [344].

FTLD can be histologically divided into two major subtypes: FTLD with tau-positive inclusions (FTLD-tau) and FTLD with ubiquitin-positive and TDP-43-positive but tau-negative inclusions (FTLD-TDP) (Figure 35) [345].

[33] Myoclonus: brief, shock-like, involuntary muscle jerks.
[34] Apraxia: the patient wants to do a movement, and there is no physical constraint, but it is not possible.

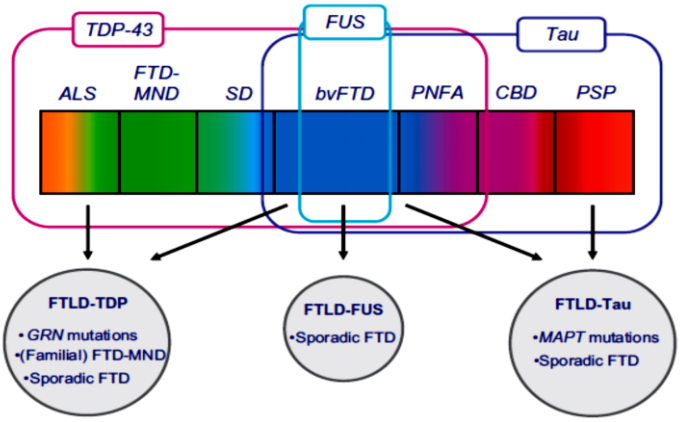


Fig. 35. Clinical, genetic and pathological spectrum of frontotemporal lobar degeneration.
FTD overlaps with motor neuron disease (FTD-MND or FTD-ALS) as well as with the parkinsonian syndromes PSP and CBD. MND comprises a group of conditions with progressive motor neuronal loss where ALS is the most frequent presentation (>75%).
It can be viewed as a motor-dementia-parkinsonism continuum: premortem diagnosis is often a challenge due to the overlapping clinical features.

Taken from [345]

Tau inclusions accumulate in neurons and are formed by hyperphosphorylated and aggregated tau proteins. Although they are present in PSP and CBD patients only some FTD cases carry mutations in *MAPT* gene: missense mutations, that modify tau interaction with microtubules and that can change its potential to aggregate into filaments, and splicing mutations, that affect the splicing of exon 10 and thus alter the ratio 4R/3R. Those proteins are more easily abnormally hyperphosphorylated and self-aggregate into filaments more readily [255].

MAPT mutations are mainly present in families with dominant forms of FTD. Nevertheless, other families with similar clinical features that show linkage to the same chromosomal region (17q21) have been described. They

carry mutations in a different gene: *PGRN*. Those patients have FTLD-TDP inclusions.

PGRN encodes for a 593 amino acid extracellular glycoprotein widely expressed. It is composed of seven and a half granulin peptides (6kDa molecules rich in cysteine) separated by interlinked spacer regions that are cleaved by some proteases (elastase, proteinase 3…) into individual granulins (Figure 36). The levels of PGRN and granulins are controlled by some proteins, like SLPI, and both act as multifunctional secreted growth factors playing key roles in cell division, survival, migration, embryogenesis, transcriptional repression, inflammation and others [346, 347].

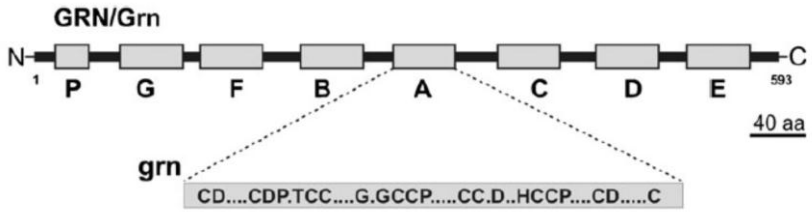


Fig. 36. Progranulin structure.
The progranulin protein (top) and the consensus sequence of the processed granulin peptides (grn) (bottom).

Modified from [347]

PGRN mutations lead to a reduction in protein activity or level: they are mostly nonsense mutations or frameshifts that originate defective mRNAs that are degraded by the nonsense-mediated RNA decay pathway (NMD). However complete or near-complete deletions of the gene, splicing mutations, signal peptide mutations and coding variants that might cause loss of function by structural changes or by reduced protein production or secretion have also

been described. Consequently, [348, 349] concluded that haploinsufficiency is pathogenic and leads to neurodegeneration.

It is necessary to elucidate the connection between *PGRN* mutations, the affected function and TDP-43 accumulation.

TAR DNA binding protein-43 (TDP-43) is encoded by *TARDBP* gene. It is a nuclear DNA/RNA binding protein of 414 amino acids, highly conserved and ubiquitously expressed that acts as a shuttle between the nucleus and the cytoplasm (Figure 37). TDP-43 regulates transcription, pre-mRNA splicing (by recruiting splicing factors) and microRNA processing [350]. TDP inclusions are composed of ubiquitinated and hyperphosphorylated C-terminal fragments of TDP-43 and accumulate in neuronal cytoplasm. *TARDBP* is mutated in familial and sporadic ALS but not frequently in FTD although FTD patients also have TDP-43 inclusions. It is unclear whether those aggregates are a cause or a consequence of the disease pathogenesis.

Recently a new subgroup has been described: a small number of FTLD cases with ubiquitin-positive TDP-43-negative inclusions have shown FUS-positive inclusions but no abnormal posttranslational modifications in this protein have yet been identified. *FUS* gene is not usually mutated; mutations have only been described in some familial ALS patients.

FUS is a ubiquitously expressed nuclear DNA/RNA binding protein of 526 amino acids that regulates transcription, pre-mRNA splicing and mRNA transport. It can act as a shuttle between the nucleus and the cytoplasm (Figure 37). Its functions and its structure are very similar to those described for TDP-43 [350].

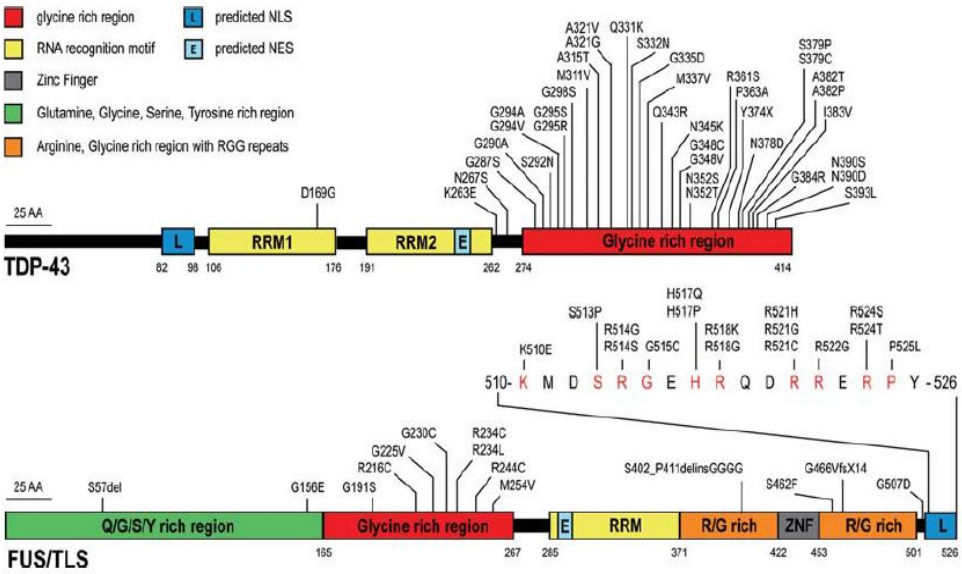


Fig. 37. TDP-43 and FUS structure.

TDP-43 contains two RNA binding domains, a glycine-rich C-terminal domain involved in protein-protein interactions, a nuclear localization signal (NLS) and a nuclear export signal (NES). The majority of mutations described are located on the Gly-rich region and that could mean that the ability to interact with other proteins is essential for TDP-43.

FUS contains a glycine-rich region, an RNA binding domain, a NLS, a NES, and other regions to interact with proteins like a Zinc finger. Most of the mutations described in familial ALS are located on the NLS. Therefore, its accumulation in the cytoplasm seems to be pathogenic.

Taken from [350]

FTLD is a genetically complex disorder, with multiple genetic factors contributing to the disease. The majority of FTLD cases are sporadic (60-90%).

Mutations in *MAPT* and *PGRN* account for the 20% of the familial cases, whereas mutations in *TARDBP*, *FUS*, *VCP* and *CHMP2B* are rare. Patients with *VCP* mutations present FTLD-TDP inclusions whereas those with *CHMP2B* present ubiquitin-positive TDP-43-negative inclusions.

Some important genetic, functional and structural aspects of FTLD pathogenesis remain unknown: the influence of other genes that have not been described yet, the relation between protein inclusions and gene mutations, especially in those cases where the mutated protein and the aggregated protein are different, the initial cause that triggers the pathogenic mechanism or the cellular processes that are affected initially or subsequently.

I.11.c. Huntington´s disease (HD):

Huntington´s disease is an autosomal dominant neurodegenerative disorder characterized by chorea[35], cognitive decline, intellectual impairment and emotional disturbances. With the progression of the disease, motor rigidity and dementia predominate. There is massive GABAergic striatal neuronal death and also atrophy of the cerebral cortex, thalamus and other brain regions though with a less severe extent [351].

It is progressive, fatal (with death in the 15-20 years after diagnosis) and caused by an abnormal CAG repeat expansion in exon 1 of the *HTT* gene that

[35] Chorea: sudden jerking movements which are entirely random in their timing and in their distribution about the body.

originates a protein (htt or huntingtin) with abnormally long polyglutamine (polyQ) fragments in their N-terminal extreme. As with most of the diseases caused by expansion of triplet repeats, there are three main groups of expansions: the normal range, that comprises expansions with less than 27 repeats, the preclinical range, that includes expansions between 27 and 35 repeats, and the pathogenic range, with those expansions of 36 or more repeats [352].

HD is monogenic, fully penetrant and marked by the anticipation effect, i.e. the onset is earlier and the progression of the disease faster in successive generations of an affected family. The 90% of HD cases are familial although there are "de novo" patients too: cases that originate from asymptomatic parents with normal repeat lengths that have expanded to the symptomatic range.

Htt is ubiquitously expressed in the central nervous system and is mainly a cytosolic protein although it has also been identified in nucleus, plasma membrane, mitochondria, lysosomes and endoplasmic reticulum. It can act as a shuttle between the cytoplasm and the nucleus but the expanded polyQ disrupts the balance of htt distribution and leads to its accumulation in the nucleus. Those large intranuclear toxic inclusion bodies formed by mutant htt are the pathological hallmark of HD.

The precise function of huntingtin remains elusive; although it is required for normal development, deletion in mice is letal. It interacts, mainly by its N-terminus, with many proteins with very different functions: endocytosis and vesicle transport, cell signaling, apoptosis or transcriptional

regulation amongst others. However, that changes due to the polyQ repeat because the structure of the protein is altered, it undergoes extensive posttranslational modifications such as cleavage, that forms toxic N-terminal fragments, and due to the polyQ fragment, the protein is more prone to aggregate [353].

The mutant htt has some conformational states: a soluble (nontoxic) conformer in equilibrium with a soluble toxic conformer. The ubiquitin-proteasome system can clear them. Nevertheless, this mechanism is impaired in HD patients and, therefore, the protein changes its structure to the misfolded conformer that cannot be degraded and forms the insoluble aggregates.

All these structural changes are supposed to trigger the pathogenic cascade that affects some cellular processes in HD patients. There is transcriptional dysregulation (the aggregates bind and sequester proteins, including factors that regulate transcription; in addition, the soluble mutated htt has different affinity to its usual interactors; as a consequence, some genes have different expression, for example, BDNF expression is decreased) [354], caspase activation, NMDAR activation, calcium dyshomeostasis, defective axonal trafficking (mutant proteins interact with proteins localized in axons and block the transport of organelles along them) and abnormal mitochondrial dynamics (fission is increased and fusion is decreased and that originates an increased mitochondrial fragmentation; there is also abnormal mitochondrial bioenergetics) (Figure 38) [355].

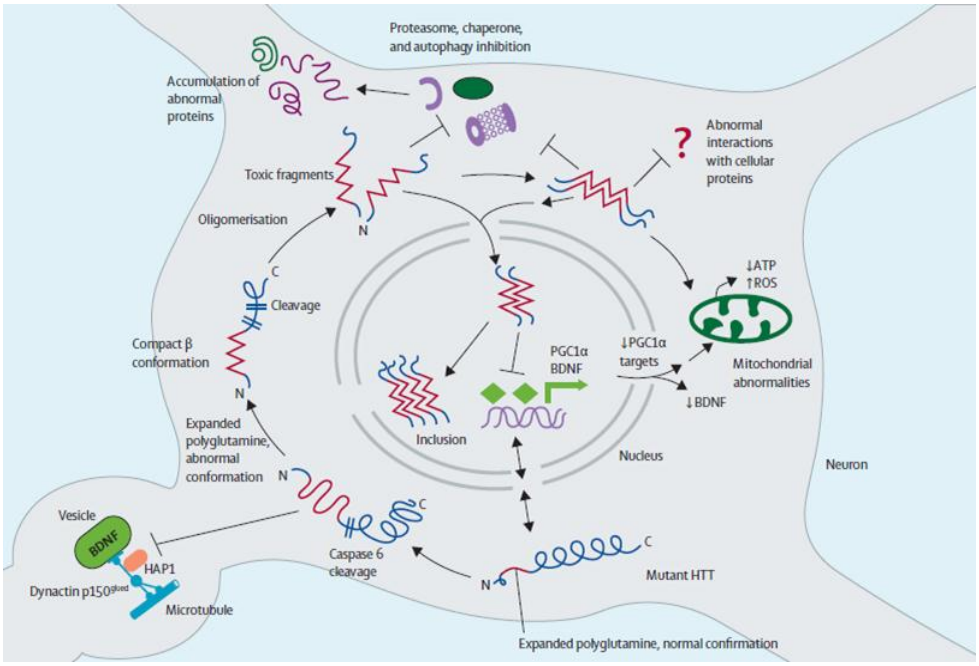


Fig. 38. Intracellular pathogenesis in Huntington´s disease.
Mutant htt (blue helical structure) with an expanded polyglutamine repeat (in red) undergoes a conformational change and interferes with cellular trafficking, especially of BDNF. Mutant htt is cleaved at several points to generate toxic fragments with abnormal compact β conformation. A major action of mutant htt is interference with gene transcription, in part via PGC1α, leading to decreased transcription of BDNF and nuclear-encoded mitochondrial proteins. Mutant htt can also lead to increased transglutaminase activity thus producing an abnormal covalent link between proteins.

Taken from [356]

I.11.d. Creutzfeldt-Jakob disease (CJD)

Although some recent studies, that need further validation, have concluded that other genes are associated, at present, all prion diseases are caused by, and only, the prion protein.

The human prion protein (PrP) is encoded by the *PRNP* gene. This protein is N-glycosylated and mainly attached to the plasma membrane by a C-

terminally linked glycosyl phosphatidylinositol anchor. PrP is highly expressed in brain, hearth, lungs… and has an unknown function.

The pathogenic mechanism that causes prion diseases starts when PrP changes from its normal conformation (PrP^{c}, or cellular PrP) to the pathogenic conformer (PrP^{Sc}, or scrapie PrP). PrP^{c} changes its structure by the direct interaction with the scrapie form: this is the only requisite. The spontaneous conversion of PrP^{c} to PrP^{Sc} is very slow but this conversion is accelerated when there are PrP^{Sc} "seeds". This is just a conformational change, neither splicing nor postranslational modifications are involved. PrP^{c} has a majority of α-helix in its structure, whereas PrP^{Sc} is mostly folded in β-sheets, is more prone to aggregate and also protease-resistant. PrP^{Sc} polymerizes into amyloid aggregates and spreads between cells and even between organisms (Figure 39).

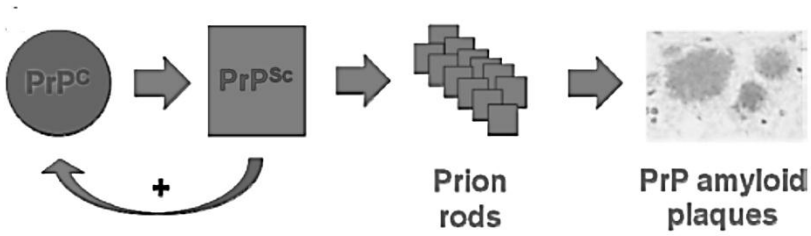


Fig. 39. The conformational change in PrP triggers the pathogenic pathway.

Modified from [357]

The human prion[36] diseases (also called transmissible spongiform encephalopathies or TSE) can be (Table 3):

- Inherited: familial Creutzfeldt-Jakob disease, or CJD, is characterized by dementia with fast evolution; Gerstmann-Sträussler-Scheinker disease,

[36] Prion means "protein infectious agent" and they represent a new type of infectious agent that do not have nucleic acids.

or GSS, is characterized by ataxia[37] and unpredictable clinical course; and fatal familial insomnia, or FFI, is characterized by insomnia and autonomic nervous system dysfunction that progresses to dementia. They are caused by mutations in the *PRNP* gene. Mutations are sufficient to cause the conformational conversion and, therefore, the disease. Only about the 10% of cases are familial. They have an autosomal dominant mode of inheritance and are nearly 100% penetrant.

- Acquired: iatrogenic CJD and kuru, that is mainly due to cannibalism and generally presented as an ataxic disorder although in the terminal phases of the disease dementia is frequent. The transmission of prions from one specie to another with non-identical sequences is possible. Animals like sheep or cattle are also affected. Depending on the host, the type of inoculum and the route of inoculation, the incubation time lasts for weeks, months or years.
- Sporadic: sCJD; although recently sFFI has also been described. The majority of prion disease cases are sporadic. Their ethiology is unknown. There are no mutations in the *PRNP* gene but it has been postulated that horizontal transmission, somatic mutations that affect the exonic region in the *PRNP* gene or spontaneous events that trigger the conformational change in PrP can be involved.

[37] Ataxia: decreased ability to coordinate movements.

Table 3. Human prion diseases.

Disease	Cause
Kuru	Infection
Creutzfeldt-Jakob disease	
Iatrogenic	Infection
Sporadic	Unknown
Familial	Prion protein mutation
Gerstmann-Sträussler-Scheinker disease	Prion protein mutation
Fatal familial insomnia	Prion protein mutation

Taken from [358]

The neurophatological features that describe human prion diseases are spongiform degeneration, gliosis and neuronal loss in the absence of an inflammatory reaction. The damage is mainly located on cerebral cortex, basal ganglia, thalamus and cerebellum. There is no cure and men and women are equally affected [359, 360].

Although most of the mutations described in *PRNP* gene cause one concrete prion disease, in some cases, the same mutation generates different phenotypes. That is the case for the p.D178N mutation that has been described in FFI and familial CJD cases. [361] observed that the disease was determined by the co-segregation of this mutation and the polymorphism located on position 129 (p.M129V): FFI for 129M carriers or CJD for 129V carriers.

The p.M129V polymorphism also increases the risk for sCJD: homozygous individuals (MM or VV) have more risk to develop sCJD than heterozygous (MV). And homozygosity for either allele is also related with faster progression in all familial and sporadic prion diseases.

This polymorphism has been widely studied in other neurodegenerative diseases (FTLD, ALS or AD). However, no definite conclusion has been obtained. For example, in Alzheimer´s disease, [362] found that homozygous carriers (129VV) were significantly more frequent in early onset Dutch AD cases; nevertheless, [363] found that in early onset German AD cases 129MM genotype was the most frequent whereas [364-366] could not associate any of the genotypes of this polymorphism with sporadic Alzheimer´s disease (independently of the onset) in any of the populations studied (Spanish, Italian and Japanese, respectively).

I.12. Others.

There are other proteins considered potential candidates to explain PD pathogenesis:

* **I.12.a. NFE2L2 and KEAP1 (or Nrf2 and INrf2):**

Cells have mechanisms to control the levels of ROS and electrophiles to avoid the damage that oxidative stress can have on their survival, development and evolution. One of them involves NFE2L2 and KEAP1.

Although some aspects are still on debate, at the present time, it is thought that, in basal conditions, NFE2L2 forms a complex with KEAP1 and other proteins (Figure 40) and is constantly ubiquitinated and degraded via the proteasome.

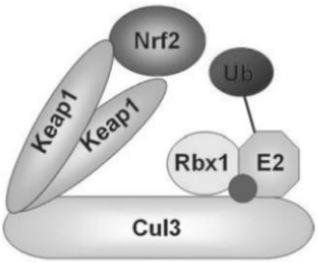


Fig. 40. Structure of the complex.
A KEAP1 homodimer acts as the substrate adaptor and binds, NFE2L2 via its C-terminal domain, and Cul3, via its N-terminal domain. Cul3 serves as the scaffold protein that forms the E3 ligase complex with Rbx1 which recruits the E2 enzyme.
Cul3 is neddylated (small circle). Neddylation, is the process by which Nedd8, a small ubiquitin-like protein (there are structural and sequence similarities between both proteins), is conjugated onto a conserved lysine residue. Neddylation occurs by a similar process as ubiquitination (E1, E2 and E3) and is reversible. Nedd8 stimulates the transfer of ubiquitin from E2 onto the substrate (NFE2L2 in this case) and also facilitates the elongation of the polyubiquitin chain.

Modified from [367]

However, when there is oxidative stress, the sensor cysteines of KEAP1 are oxidized and NFE2L2 is released from the complex and translocates to the nucleus where it heterodimerizes with other proteins, such as small Maf and Jun, and triggers the transcription of genes involved in the protective cascade against oxidative conditions such as enzymes involved in drug metabolism -NQO1- and glutathione homeostasis -GST-, chaperones, proteasome subunits, neurotrophic factors -BDNF- or transporters -Mrp1- amongst others which have antioxidant or electrophile responsive elements (ARE/EpRE) on their promoters [368, 369]. In addition, NFE2L2 regulates the expression of KEAP1, Cul3 and Rbx1, what can be considered a feedback auto-regulatory loop.

* **I.12.b. LAMP-2A and hsc70:**

There are two main cellular mechanisms involved in protein degradation: the ubiquitin-proteasome system and autophagy.

There are three types of autophagic processes:

- Macroautophagy and microautophagy are not selective processes. Both degrade cytosolic regions that contain proteins or even organelles present in a wide range of eukaryotes, mammals included. In macroautophagy, a *de novo* formed isolation membrane sequesters a cytosolic region and then, this membrane seals into a double membrane vesicle (autophagosome) that fusions with lysosomes to acquire the necessary enzymes to degrade its inner content (autophagolysosome).

However, in microautophagy, the lysosomal membrane directly engulfes the cytosolic region that will be degraded [370].

- Chaperone-mediated autophagy (CMA) only occurs in mammals and consists in the degradation of individual cytosolic proteins. The target proteins have a motif, KFERQ, which is recognized by the hsc70 chaperone. Hsc70 plus cochaperones, that regulate its activity or act as cochaperones themselves (hop, hip, bag-1, hsp40 and hsp90), and the target protein bind to the CMA receptor, LAMP-2A, at the lysosomal membrane. After unfolding the protein, it crosses the lysosomal membrane, supposedly through a pore formed by oligomerized LAMP-2A, assisted by a lysosomal form of hsc70 (lys-hsc70) present in the lumen. There, the protein is rapidly degraded and the hsc70 chaperone complex is released from the lysosomal membrane and can bind to other proteins to start the process again [371].

The gene *LAMP-2* codifies for three different isoforms (LAMP-2A, LAMP-2B and LAMP-2C) that have the same heavily glycosylated luminal domain but differ in their short cytoplasmic tail and the single transmembrane domain (those domains are partially codified by exon 9 which is alternatively spliced). Only LAMP-2A is involved in CMA.

The HSP70 multigene family of chaperones is composed of constitutively expressed members such as hsc70 and stress-inducible members such as hsp70. Hsc70, also called HSPA8, works in the proper folding of newly synthesized proteins and those subject to stress-induced denaturation, in the refolding of misfolded or aggregated proteins,

preventing protein aggregation, promoting ubiquitination and degradation of misfolded proteins, in the translocation of proteins into cellular compartments such as mitochondria and chloroplasts and also in CMA.

CMA is maximally activated during stresses that damage proteins such as starvation or oxidative stress. Its activity is decreased in aging, because the levels of LAMP-2A at the lysosomal membrane decrease with age, and in familial PD, although wild type αSyn is degraded by CMA, the mutants p.A53T and p.A30P bind strongly to the CMA-receptor at the lysosomal membrane but do not translocate into the lysosomal lumen and even block the process[38] [372]. As a consequence, long-lived damaged cytoplasmic proteins are less efficiently degraded and the accumulation of damaged proteins increases cellular susceptibility to stressors. Although macroautophagy, that occurs constitutively in cells, is activated to compensate this deficiency, that could not be enough.

In addition, [373] found that the expression of LAMP-2A and hsc70 was significantly reduced in *substantia nigra* and amygdala of PD brains when comparing to healthy controls, and that autophagy-related proteins seemed to accumulate in Lewy bodies.

Therefore, dysfunctional autophagy is considered one of the possible causes of Parkinson´s disease.

[38] Some frequent covalent modifications of αSyn such as oxidation, nitration, formation of dopamine adducts… also have the same effect on CMA activity.

* **I.12.c. Neurotrophic factors (CDNF, MANF, BDNF):**

Neurotrophic factors are small secreted proteins that regulate the number, development, maturation and survival of neurons by binding to their receptors [374].

CDNF (cerebral dopamine neurotrophic factor) and MANF (mesencephalic astrocyte-derived neurotrophic factor) belong to the fourth and most recently discovered family of neurotrophic factors. Their expression in the human body is widespread and, if they follow the same expression pattern observed in other animals, both are expressed at all developmental stages and in adults.

CDNF and MANF are highly homologous and evolutionary conserved in vertebrates; in invertebrates, however, there is just one similar neurotrophic factor that resembles more to MANF than to CDNF:

- in *Drosophila*, it is required for the maturation of the embryonic nervous system and the maintenance of neuronal connectivity. When it is abolished, there is high dopaminergic cellular death, embryonic lethality and a rough decrease of dopamine levels [374].
- in zebrafish, there is no apparent abnormal phenotype when MANF is abolished but there is a reduction of dopamine-containing neurons during embryogenesis and a lower dopamine level. The phenotype is not as lethal as in fly but the different outcomes may be due to significant differences in the organization of neurotransmitter systems between invertebrates and vertebrates [375].

In any case they/it are/is necessary for the proper neuronal development and survival.

Nowadays, due to the fact that the current PD treatments are just symptomatic and cannot stop the degenerative process, CDNF and MANF are considered promising therapeutic agents in PD, even more than GDNF, for their effectiveness, neuroprotective and neurorestorative role observed in animal models:

- [376] observed that, in rats, CDNF treatment, before or after the striatal injection of 6-OHDA[39], had a protective or reparative effect, respectively, on dopamine-containing neurons in the SN. In mice, CDNF also protected the nigrostriatal system and promoted its recovery after MPTP treatment [377].
- MANF selectively promoted *in vitro*, in embryonic rat neurons, the survival of dopaminergic nigral neurons. This effect was not observed in either GABAergic or serotonergic cells. Moreover, its effect was more powerful than that observed for GDNF and BDNF [378].

Nevertheless, the receptors and signaling pathways where they are involved remain unclear. There is just some more information about MANF obtained from in *in vitro* studies that have reported that

[39] 6-OHDA does not cross the blood-brain barrier. When it is injected in the striatum to reproduce the PD phenotype in an animal, it produces lesions on nerve endings and progresses retrogradely towards the *substantia nigra* to damage the cell bodies. When it is injected in the SN, the neurons lose their phenotype within 24h and the striatal dopamine depletion starts 2 to 3 days later. Its damaging mechanism is unknown but it is supposed to involve ROS generation (although inhibition of mitochondrial complex I or induction of ER stress have been proposed too).

- it is upregulated in non-neuronal cultures when there is an excess of unfold proteins and the unfold protein response (UPR) is induced to restore homeostasis. It also inhibits ER stress-induced cell death [379],
- it protects against apoptosis in neuronal cultures; it is supposed that MANF mimics Ku70 mechanism, i.e. the inhibition of the proapoptotic Bax and the prevention of mitochondrial cell death signaling, because both share high structural homology at their C-terminal domains [380].

BDNF is highly and widely expressed in the central nervous system and supports the differentiation, maturation and survival of dopaminergic neurons, at least *in vitro*: it is supposed that inhibits apoptosis and stimulates sprouting and neuronal reorganization [381].

BDNF is translated as a preprotein. There are many different functional promoters that are tissue and brain-region specific and generate many isoforms with the same 3′ end but with different 5′ extremes resulting in signal peptides with different lengths that are cleaved to obtain the mature protein [382]. The precursor protein and the mature protein are both active via two different receptor systems.

The most studied polymorphism, p.V66M, is localized in the signal peptide and, consequently, is cleaved during the maturation of BDNF. Its relevance is due to the study conducted by [383]. They concluded that the M allele was associated with poorer verbal episodic memory and hippocampal function *in vivo*. *In vitro* assays demonstrated that this allele

impaired intracellular trafficking and secretion of BDNF and also changed the cellular localization: the V allele was mainly localized in dendrites whereas the M allele was mainly localized in cell bodies.

The M allele is not observed in lower primates so, if it is conserved, it may confer to humans some compensatory advantage in other biological processes.

BDNF has been associated with varied neurological (AD, PD, HD, multiple sclerosis…) and psychiatric disorders (schizophrenia, depression, bipolar disorder, eating disorders…) with controversial results.

With regard to Parkinson´s disease:

- In the rat PD model, BDNF is neuroprotective against MPTP and 6-OHDA treatment. It is effective when administered before those toxic agents but not after, that is, it is not neurorestorative.
- BDNF concentration was significantly lower, i.e. decreased expression, in the brain of PD patients when comparing to healthy controls, especially in the nigrostriatal dopaminergic regions [384]. This could explain the massive neuronal death observed in PD cases.
- Some association studies have been conducted to analyze the influence of the p.V66M change in PD pathogenesis.
  - The majority of them have concluded that there is no significant difference in the allelic or genotypic frequency between PD cases and controls:

- in sporadic PD in Chinese populations [385] and [386]. [386] also analyzed the effect of p.G2385R in *LRRK2* and observed that when both polymorphisms were present in the same person, the PD risk was increased, particularly in patients with an onset age older than 60.
- in a population from USA [387], in a Greek population [388] and in a Caucasian population, where, moreover, the polymorphism does not seem to modify clinical features in PD cases [389],
- in familial PD in a worldwide population [390],
- and also in a meta-analysis based on 6 studies in sporadic PD, 4 in Asians, Chinese and Japanese, and 2 in Caucasians, from UK and Sweden [391].

- However, [392] observed that in an Italian population, the allelic and genotypic frequencies for this variant were different between sporadic PD cases and controls. The M allele was more frequent in aged PD patients with cognitive impairment and more severe disease progression.

In conclusion, the p.V66M polymorphism does not seem to play a major role in the pathogenesis of PD.

* **I.12.d. ARMCX family:**

After the completion of the Human Genome Project there are plenty of new genes with unknown features. The components of the ARMCX cluster are some of them.

These 6 short genes, ≈4.5 to 8kb, are localized in the X chromosome (Xq22.1) and are exclusive to Eutherian mammals. In humans, they evolved from a single ancestor gene, *ARMC10*, in chromosome 7.

All the members of the family share some characteristics that were predicted *in silico*:

- their entire coding region is in a single exon. The absence of alternative splicing might point out the importance of their correct expression;
- have six armadillo or armadillo-like repeats in tandem and a DUF634 mammal domain with unknown function. The arm repeat is a degenerate protein sequence motif of about 42 amino acids that forms a conserved three-dimensional structure composed of 3 α-helices that allows proteins to have many functions and to interact with many proteins [393];
- have a signal peptide and a mitochondrial targeting signal. Four of the members of this gene family also have a nuclear localization signal.

At the present time just some information has been obtained by *in vitro* studies:

All the murine members of the cluster (ARMCX1-6) are expressed in the nervous system. In addition, murine ARMCX3, which is an outer

mitochondrial membrane protein, interacts with Sox10, that is a transcription factor involved in a wide variety of developmental processes, including sex determination and neurogenesis. This interaction might reveal that there is a signal transduction cascade that connects the nucleus and the mitochondria [394].

In humans, ARMCX1 is transcriptionally regulated by CREB and Wnt/β-catenin signaling [395]. Moreover, ARMCX1 and ARMCX2 expression, which is widespread in the human body, is lost in cell lines established from different human carcinomas [396]. For those reasons, it has been related to tumorigenesis.

Human ARMCX1, 2, 3 and 6 are mitochondrial proteins. And, at least ARMCX3, is a member of the KIF5/Miro/Trak2[40] protein complex responsible for the mitochondrial transport along exons. ARMCX3 interaction is Ca^{2+}-dependent and is not related with the regulation of the motor activity of the kinesin [397].

Nowadays, just some details are known and not for the six ARMCX. Therefore, plenty of work is still needed to clarify the essential features of all the members of the cluster.

However, they are good candidates to explain PD pathogenesis due to their localization in chromosome X, in the hypothetical *PARK12 locus*, and

[40] Mitochondrial trafficking in neurons is mediated by kinesin motors (KIF5). Rho GTPases (Miro1 and Miro2) and kinesin adaptors (Trak2) are also necessary to link the mitochondria to the kinesin motors.

their possible connection with the mitochondria. *In silico* predictions and biochemical assays seem to confirm their cellular localization; in addition, ARMCX members possibly regulate mitochondrial dynamics and trafficking, which are essential to supply appropriate energy to distal neuronal branches, and thus for correct neurotransmission and neuronal viability.

* **I.12.e. TOR1A:**

Dystonia is a movement disorder characterized by repetitive or sustained involuntary muscle contractions, writhing and torsion are conspicuous, that affect one or more body parts. Patients present abnormal interneuronal signaling but not neurodegeneration.

Dystonia can be primary, when is the only or major symptom, or secondary, when is just one of several symptoms within another, frequently neurological, oncological or metabolic disorder, after intoxication or after trauma.

Although some primary dystonias have an unknown cause, they are frequently inherited as Mendelian traits. Due to their wide phenotypical spectrum, their classification followed a clinical criterion for long time. Nevertheless, recently it has been replaced by a genetic criterion (Table 4) [398].

DYT1 is the most common and severe form of hereditary primary dystonia. The disease typically begins in childhood (age at onset ≈12), in an

extremity, and frequently generalizes. It is due to mutations in the *TOR1A* gene.

Table 4. The genetic dystonias.

Subgroup	Designation	Type	Inheritance, chromosomal *locus*	Gene
Primary dystonia	Dystonia 1	Early onset primary torsion dystonia	AD 9q34	*TOR1A*
	Dystonia 2	Generalized dystonia	AR unknown	
	Dystonia 4	Hereditary whispering dystonia	AD unknown	
	Dystonia 6	Dystonia of mixed type	AD 8p11.21	*THAP1*
	Dystonia 7	Focal adult onset dystonia	AD 18p	
	Dystonia 13	Dystonia of mixed type	AD 1p36.13-p36.32	
Dystonia-plus syndromes	Dystonia 5	Dopa-responsive dystonia	AD 14q22.1-q22.2	*GCH1*
	Dystonia 11	Myoclonus dystonia	AD 7q21	*SGCE*
	Dystonia 12	Rapid-onset dystonia parkinsonism	AD 19q12-q13.2	*ATP1A3*
	Dystonia 15	Myoclonus dystonia	AD 18p11	
Paroxysmal dystonia	Dystonia 8	PDC	AD 2q35	*MR1*
	Dystonia 9	PDC with ataxia and spasticity	AD 1p21-p13.3	
	Dystonia 10	Paroxysmal kinisigencic choreoathetosis	AD 16p11.2-q12.1	
	Dystonia 18	Paroxysmal exercise-induced dystonia	AD 1p31.3-p35	*SLC2A1*
Others	Dystonia 3	X-linked dystonia parkinsonism	X-linked Xq13.1	*TAF1*
	Dystonia 16	Dystonia parkinsonism	AR 2q31.3	*PRKRA*

PDC: paroxysmal dystonic choreoathetosis; AD: autosomal dominant; AR: autosomal recessive. Dystonia plus syndromes are monogenic dystonias without detectable neuroanatomical abnormalities but with additional neurological manifestations such as myoclonus and parkinsonism. Paroxysmal dystonias or dyskinesias include episodes of dystonic involvement and no gross neuropathological anomalies.

Taken from [398]

TOR1A, or torsinA, is widely distributed across the body and the central nervous system [399]. Nevertheless, DYT1 manifests as a tissue-specific (neuronal) disorder despite affecting a widely expressed protein. It is noteworthy that its highest expression is in dopaminergic neurons. That

observation raised questions about a potential function of TOR1A in dopaminergic cells and/or a role for a dopamine-related defect in DYT1. Subsequent neurochemical analyses of postmortem brain tissue did not clearly resolve these questions. In addition, no pathological lesions have been detected neither in SNpc neurons nor in any other central nervous system region from DYT1 patients [400].

At the present time, the only confirmed connection between dopamine and dystonia is that dystonia affects patients with mutations in *PARK2* or *PARK1*4 and is also a common secondary motor symptom in PD. Moreover, [401] reported that TOR1A accumulated in Lewy bodies in brains from sporadic Parkinson´s disease cases. The pattern of staining was different in SNpc, where TOR1A was preferentially distributed around the halo rather than the central core, than in cortex, where it was uniformly distributed throughout the Lewy body.

TOR1A is a lumenal glycoprotein, without transmembrane domain, localized inside the ER and the nuclear envelope (NE) [402] that belongs to the AAA family of ATPases [403, 404]. Due to its activity and cellular localization, it has been postulated to work as a molecular chaperone assisting in the proper folding of secreted and/or membrane proteins.

TOR1A interacts with LULL1, an ER transmembrane protein, LAP1, an integral protein of the NE inner membrane, and also with SUN1, another NE protein component of the LINC complex that couples the nuclear interior to cytoskeletal networks [405]. For that reason it has also been proposed to

have a role in the structure and/or function of the endoplasmic reticulum and the nuclear envelope.

There are some other proposals about its function:

- It may play a role in the secretory pathway because the overexpression of TOR1A suppresses the delivery to the plasma membrane of membrane proteins such as the dopamine transporter; or it may regulate the synaptic vesicle recycling;
- It can play a role in the cytoskeletal network because it interacts with vimentin in the cytoplasm, which is important for cellular motility and adhesion; and also binds tau and kinesin light chain in the cytoplasm so it may affect microtubule stabilization, neurite outgrowth and polarity.

In any case, its real function is unknown.

A 3 bp deletion in *TOR1A* gene causes the majority of cases (≈80%) although this mutation has low penetrance (30-40%). The in frame deletion results in the loss of a glutamic acid residue at position 302 or 303 in the carboxy terminal region of the protein.

The nuclear envelope is the primary site of dysfunction in DYT1: mutated proteins accumulate in perinuclear inclusions. Moreover, the mutated proteins alter their normal cellular distribution and are more present in the NE than in the ER. They also sequester wild type TOR1A in NE [406].

The deletion does not alter the solubility of TOR1A neither its ability to multimerize nor its ATPase activity. However, its stability is altered: the mutated protein is degraded via macroautophagy, like other short-lived proteins, whereas the wild type protein is degraded by the proteasome, as the long-lived proteins. Its influence in TOR1A function is unclear: [407, 408] postulated that mutations in *TOR1A* could cause changes in the interaction of cytoskeletal components with the nuclear envelope; due to the mutation, this interaction could be prolonged changing from transient to permanent.

The p.D216H polymorphism has been proposed to modify the penetrance of dystonia in p.delE302/303 carriers: the H allele was less frequent in patients that carried the deletion compared to nonmanifesting carriers of the deletion [409]. Moreover, [410] observed that the frequency of the H allele *in trans* (i.e., on the non-GAG deleted chromosome) was significantly increased in nonmanifesting GAG-deletion carriers without dystonia compared to those with dystonia. They concluded that the H allele is protective *in trans* and that the D allele *in cis* is required for the disease to be penetrant. However, the H allele is rare (frequency ≈12%) so this modifying effect only affects a small subset of deletion carriers.

Nevertheless, how this could happen remains unknown: in cell cultures, the H allele caused protein inclusions similar to those observed in ΔGAG-TOR1A. Furthermore, the presence of both alleles *in cis* cancelled the effect: when both were present, the protein had a reduced tendency to form inclusions [411].

* **I.12.f. GSK3β:**

GSK3β is a serine-threonine kinase expressed in the brain and some other tissues, that works in multiple signaling pathways.

Among other functions, it is a potential negative regulator of BDNF and is, in turn, regulated by BDNF-mediated signaling: to induce microtubule assembly and axonal outgrowth, DOCK3 promotes GSK3β recruitment to the membrane, where it is phosphorylated and thus inactivated. GSK3β is a master regulator of microtubule dynamics in growth cones

It is associated with neurodegenerative diseases, mainly with tauopathies, where tau is hyperphosphorylated and assembles into neurofibrillary tangles, because GSK3β is one of the 20 kinases that can phosphorylate tau. However, its exact implication is not clear. Moreover, there are some evidences that point out to an involvement of GSK3β in PD pathogenesis, although Parkinson´s disease is a synucleinopathy:

- αSyn stimulates the GSK3β kinase activity on tau. The increased phosphorylation of tau results in destabilization of the microtubule associated network, cytoskeletal dysfunction and modification of synaptic plasticity [412].
- [413] observed that, *in vitro*, two SNPs altered the transcription and the splicing of the gene:
  - the T allele of rs334558, located in the promoter, had greater transcriptional activity (1.4-fold increase) than the C allele;
  - the intronic polymorphism rs6438552 had an effect on splicing: there was a difference between both alleles on the use of splice acceptor sites in downstream introns resulting thus in different

proportion of protein isoforms depending on the allele although the mechanism for this splicing modulation is unknown. The T allele originated more splicing variants with higher kinase activity on tau.

They also conducted a genetic study in Australian and Chinese PD patients and controls that highlighted that there was no significant difference neither in the allelic nor in the genotypic frequency between both groups for any of the two polymorphisms. Nevertheless, they observed that in H1H1 haplotype carriers (*MAPT* gene) there was an underrepresentation of the TT genotype for any of the SNPs, suggesting that the T allele, in rs334558 or in rs6438552, was protective. That was just a trend, not a significant result.

Subsequent genetic analyses in PD patients, [95] in a British population, [93] in a Caucasian population and [412] in an Indian population, have also reported that there is no significant difference neither in the allelic nor in the genotypic frequency between both groups for any of the two polymorphisms (rs334558, rs6438552). Despite all this, there are other additional controversial conclusions:

- [95] observed that there was no interaction between any of the 2 SNPs in *GSK3*β and *MAPT* H1/H2 haplotype.
- [93] concluded that there was no pairwise interaction between any of them and H1/H2 haplotype or rs356219 (*SNCA* gene).
- [412] reported that there was no significant difference even considering haplotypes. However, when considering age at onset,

the CC haplotype (rs334558-6438552) was moderately associated with increased risk of late onset PD (LOPD; defined by an onset after 40) whereas the TC haplotype showed a protective effect against LOPD.

Nevertheless, other studies did find significant differences:

- [414] observed that the CC genotype in rs334558 was protective against PD in a Greek population. They also observed that the frequency of this genotype was decreased in PD cases carrying the H1H1 haplotype. There was no significant result for rs6438552. In addition, the TT haplotype (rs334558-6438552) was overrepresented in PD cases compared to controls, independently of the *MAPT* haplotype, so it could be considered a risk factor.
- [415] only analyzed rs334558 in Han Chinese sporadic PD patients and controls and found that the T allele was a protective PD factor.

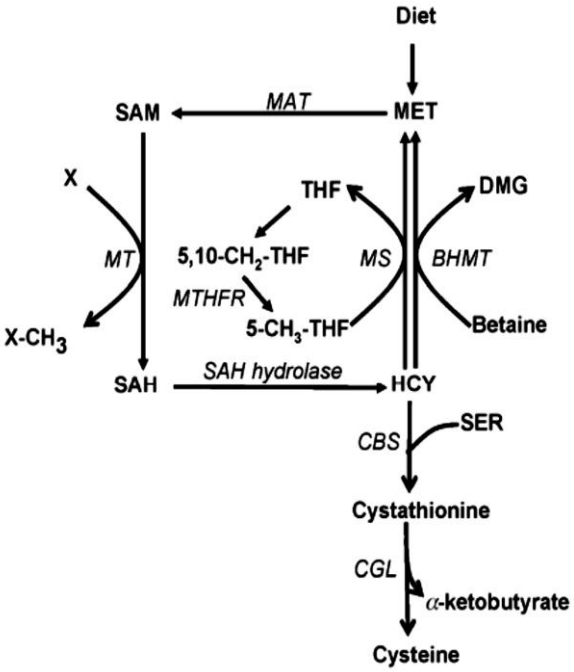
Therefore, the influence of both SNPs on PD pathogenesis is confusing.

* **I.12.g. Homocysteine (MTHFR, MTR and CBS):**

Homocysteine (Hcy) is a non-protein amino acid that arises during methionine metabolism (Figure 41).

When there is a metabolic problem, homocysteine accumulates inside the cells. This excess of Hcy is toxic for them and, consequently, it is

exported into the circulation even when that exposes all tissues to the potential toxicity of hyperhomocysteinemia (HHcy)[41].



Taken from [416]

Fig. 41. Methionine metabolism.
Methionine (MET) is activated to S-adenosylmethionine (SAM) by the methionine adenosyltransferase (MAT). SAM is the major methyl group donor in the cell (it is used by methyltransferases -MT- such as the catechol-O-methyltransferase, COMT, that catalyzes the conversion of L-dopa -X- into 3-OMD -X-CH_3-). The resultant product is SAH (S-adenosylhomocysteine) which is subsequently hydrolyzed to homocysteine (HCY) and adenosine. These reactions form the transmethylation pathway.
Homocysteine is removed either by its irreversible conversion to cysteine (cystathionine β-synthase, CBS, and cystathionine γ-lyase, CGL, are involved in the transulfuration pathway which has a limited distribution: it is only found in the liver, kidney, small intestine and pancreas) or by remethylation to methionine (by the betaine:homocysteine methyltransferase -BHMT- or by the methionine synthase -MS or MTR-; BHMT has a limited tissue distribution and is mainly present in the liver whereas MS is widely distributed).
SER; serine;
MTHFR:methylenetetrahydrofolate reductase;
5-CH_3-THF: 5-methyltetrahydrofolate;
5,10-CH_2-THF: 5,10-methylenetetrahydrofolate;
THF: tetrahydrofolate;
DMG: dimethyl glycine.

HHcy occurs when there is a deficiency in vitamins, that are essential cofactors for MTR -vitamin B12-, CBS and CGL -vitamin B6- (moreover THF derives from

[41] HHcy: abnormally high level of homocysteine in blood.

folic acid -vitamin B9-), or due to genetic mutations in *MTHFR*, *MTR* or *CBS*, that are rare disorders.

However, some polymorphisms have also been associated with altered Hcy levels in plasma [417]. Hyperhomocysteinemia is more pronounced and has more toxic and serious effects in those individuals who carry mutations. They represent a minority of severe cases.

- **MTHFR:** due to mutations that abrogate the enzyme activity, patients in infancy or adolescence present developmental delay, motor and gait dysfunction, seizures and other neurological abnormalities. There are two frequent polymorphisms:
  - In the heterozygous or homozygous state, c.C677T, which is located in the predicted catalytic domain of the protein, correlates with reduced enzymatic activity and increased thermolability of the protein [417]. The majority of studies that have analyzed this polymorphism conclude that homozygous TT carriers have significantly elevated plasma Hcy levels. However, [418] observed that, in North-Irish men, the insertion c.844ins68 in *CBS* seems to "normalize" homocysteine levels in c.677TT individuals.
  - c.A1298C, located in the predicted regulatory domain, decreases MTHFR activity although to a lesser extent. Only compound heterozygous carriers, that is c.A1298C and c.C677T, show elevated Hcy levels [419].
- **CBS**: missense mutations in this gene also cause homocystinuria[42]. The major clinical manifestations involve the eyes, and the central nervous, skeletal and vascular systems.
  - c.T833C is a frequent polymorphism. [420], working in fibroblasts, observed that some of the carriers also have an insertion of 68bp *in cis* with

[42] Homocystinuria: increased excretion of homocysteine in the urine.

it. This insertion, c.844ins68, is an almost exact duplication of the previous 68bp and includes intronic and exonic sequence as well as the wild-type T allele at position 833. The insertion creates a new acceptor splicing site that allows the skipping of the C allele: splicing using this new splicing acceptor site generates mRNAs with the normal size and without the C whereas splicing using the canonical acceptor site generates aberrant mRNAs that include the insertion with its 2 nonsense codons and that are degraded in the nucleus. [421] found mRNAs with the C allele in liver and proposed that there could be a third putative noncanonical splicing site that originates mRNAs with the normal size but including the C variant. The insertion is only present in individuals who carry the c.T833C polymorphism. [422] reported that the insertion is not pathogenic and does not alter either the enzymatic activity or the level of plasma homocysteine.

In the major part of individuals with HHcy, the concentration of homocysteine reached in plasma is not high enough to cause direct toxic effects, i.e. Hcy does not act as a direct toxin, but there is an increase in the susceptibility to those effects. High homocysteine levels accelerate cell ageing and promote DNA damage; in addition, there are some evidences that point out that homocysteine facilitates the aggregation of Aβ and the phosphorylation of tau [423, 424].

Hyperhomocysteinemia is a confirmed risk factor for the development of vascular disease [425, 426]. In addition, it seems that there is a link between HHcy and risk of developing dementia but the influence of vitamin

levels and genetic factors and the identification of which is the cause and which the consequence are controversial:

- [427] conducted a follow-up study in mostly Caucasian people from the Framingham (USA) cohort and observed that there was a strong association between plasma homocysteine levels and the risk of developing dementia and Alzheimer's disease. That conclusion was independent of age, sex, *APOE* genotype, plasma vitamin levels, and other putative risk factors for dementia and Alzheimer's disease.
- Other cross-sectional studies also reported the same conclusion, i.e. elevated plasma homocysteine levels were associated with poor cognition and dementia [428-430].
- In two studies conducted only in Australian men[43], one cross-sectional [431] and one longitudinal [424], increased levels of plasma homocysteine were related to higher risk of cognitive impairment or dementia, respectively. In both analyses, for the c.C677T polymorphism, the T allele carriers, that have higher Hcy levels, had higher risk than the C allele carriers. However, [432], in a longitudinal study conducted in a British population, observed that the polymorphisms c.C677T and c.A1298C (*MTHFR*) were not associated with changes in cognitive performance (verbal fluency, verbal memory and abstract reasoning). The influence of vitamin levels was not analyzed in any of those three studies.

[43] There is no suggestion to date that Hcy may have a gender-specific effect with regards to cognition.

In Parkinson´s disease patients, hyperhomocysteinemia has been repeatedly reported. However, there are experimental evidences that point out that L-dopa treatment increases the concentration of plasma homocysteine rather than vitamin deficiency (normal vitamin levels were described in PD cases with HHcy; moreover, higher intake of folate or vitamins B6 and B12 did not lower the risk of developing PD) or the disease *per se* [433]:

- In a Polish population, [434] observed that levodopa-treated PD patients, PD-L, had Hcy levels higher than controls. That conclusion was independent of c.C677T genotype and L-dopa doses.
- [435] reported that homocysteine plasma levels were significantly increased in a Taiwanese population when comparing PD-L with non-treated PD cases or controls. Only in PD-L, Hcy levels were related to c.C677T genotype being the TT genotype the most frequent and associated with the highest Hcy concentration, as previously reported. Homocysteine levels were not correlated with folate and vitamin B12 concentrations, disease duration or L-dopa dose. There were no significant results for c.A1298C.
- Hcy levels analyzed in Spanish controls and PD cases, classified as cognitively normal, with mild cognitive impairment or demented, were only correlated with age and inversely associated with vitamin B12, B6 and folic acid. PD patients (all had received levodopa therapy) presented homocysteine levels higher than controls although the level did not predict the cognitive status of the cases. Homocysteine levels were independent of sex, disease duration, L-dopa doses and genetic

variants in *MTHFR* (c.C677T, c.A1298C), *MTR* (c.A2756G) or *CBS* (c.844ins68). Moreover, any of those variants was associated neither with PD risk nor with cognitive status [436]. Nonetheless, a previous study concluded that PD patients with HHcy had worse mood and cognitive function [437].

More studies are required to obtain definite conclusions and to elucidate the pathogenic mechanisms, if any. Moreover, all those neurological disorders are more frequent in elder people, where vitamin deficiencies, which can develop because of low dietary intake, disturbed absorption in the gastrointestinal tract or by interaction with some medications, are the most common cause for HHcy and it is not clear whether dietary vitamin supplies, which decrease homocysteine plasma levels, can be considered protective factors against those disorders being that elevated Hcy levels have been described even in people without vitamin deficiency and homocysteine lowering treatments for older people without dementia have not been able to demonstrate an improvement in cognitive performance.

* **I.12.h. TOMM40:**

The most repeated statistically significant results obtained when analyzing the genetic influences in late onset Alzheimer´s disease (LOAD) are located in the APOE region (*APOE*, *TOMM40* and *APOC1* genes). It is

currently accepted that the ε4 allele (*APOE*) is a risk factor for LOAD and that it is also associated with lower age at onset. However, genetic markers in the three genes have shown association and, although that has been considered the effect of the high linkage disequilibrium in the region, where the accurate risk marker is located remains unknown.

To solve this question, [438] decided to deeply analyze the region by phylogenetic studies in Caucasian AD patients and healthy controls. They obtained very interesting results for a poly-T polymorphism located in *TOMM40* gene, rs10524523. The length of the homopolymer was correlated with the genotype in *APOE*: for ε4 alleles, the poly-T was relatively long -21 to 30T- with a unimodal distribution of lengths, whereas for ε3 alleles, a bimodal distribution of lengths was evident -12 to 16T and 28 to 37T-; ε2 alleles had a similar distribution than ε3.

[439] analyzed the length of the homopolymer in different populations residing in USA and compared their results with previously published studies conducted in other ethnicities. Whites and Hispanics shared a similar allelic distribution, and the same was reported for Asian (Japanese, Han Chinese and Korean) and for African Americans and Ghanaians. Long alleles (l; 20≤T≤29) were the minority in all the populations and had similar values; the differences were observed in the percentage of short (s; T≤19) and very long (vl; T≥30) alleles: African Americans and Ghanaians had a higher frequency of short alleles than the others and carried very very long alleles, even 54T, whereas Asian had a higher frequency of very long alleles.

Surprisingly, in Japanese, and especially in Ghanaians and African Americans, there were s and vl alleles associated with the ε4 allele.

[438] also observed that alleles ≥27T were associated with earlier age at onset in LOAD.

This finding could explain the current heterogeneity observed in the age at onset of Alzheimer´s disease patients that carry the ε3 allele, which was supposed to be a neutral factor for AD (Figure 42).

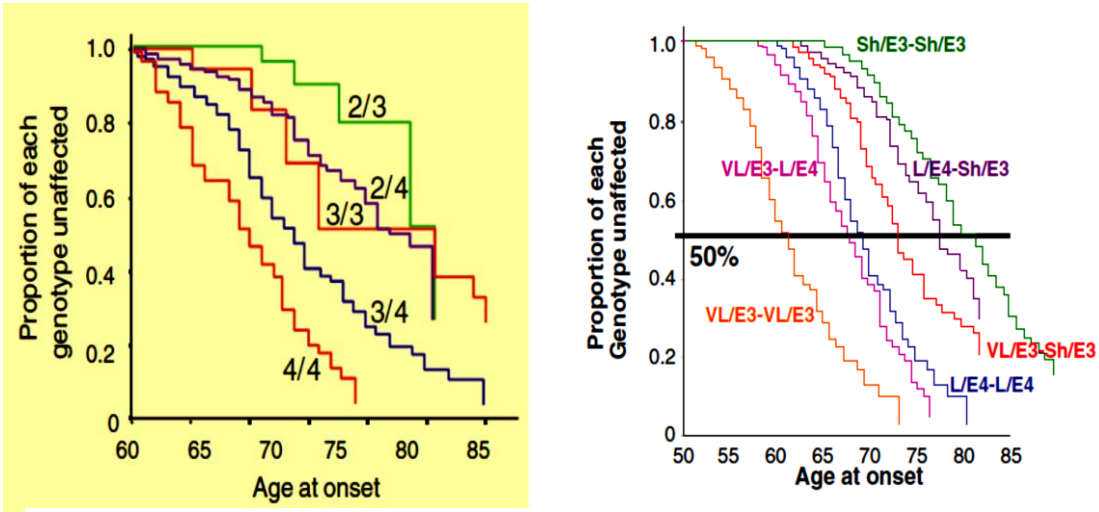


Fig. 42. Left: Alzheimer´s disease age at onset curves by *APOE* genotype.
Right: Hypothetical Alzheimer´s disease age at onset curves by *TOMM40-APOE* haplotype.
rs10524523 marker, when modeled together with APOE subtype and age will translate into a more precise prediction of risk at a particular age than predictions based on APOE subtype alone.
ε3: short, Sh, (T≤19) and very long, VL, poly-T (T≥30); ε4: long, L, (20≤T≤29).

Taken from [440]

However, other studies that analyzed more LOAD cases than [438] obtained opposite conclusions: [441], only in LOAD cases, and [442] in

LOAD cases and controls, both in Caucasian population, did not find any correlation between the length of the poly-T and the age at onset.

Surprisingly, [442] also observed that there was a trend for very long alleles and older age at onset. Moreover, they found no association between the homopolymer length and the mRNA expression of APOE or TOMM40 in the parietal cortex and also no different expression of those genes in LOAD cases when comparing to controls.

Nevertheless, [443] observed that the expression level of TOMM40 was decreased in blood in patients when comparing to age, sex and ethnicity matched Asian controls. They did not analyze the length of the polymorphism and its relation with gene expression.

TOMM40 is the channel-forming subunit of the translocase located in the mitochondrial outer membrane. It forms a complex with other proteins that is essential for protein import into the mitochondria. Almost all the mitochondrial proteins employ the TOM machinery for their import because the majority of them are encoded by nuclear genes [444].

For its function, it has been postulated that it could influence APOE transcription depending on the poly-T length or that it could alter the interaction between APOE and the mitochondria.

Although TOMM40 has not been studied in PD cases, its probable relation with other neurodegenerative disorder such as AD, and even with the age at onset, plus its function in mitochondria, make this gene a good candidate for further analysis.

I.13. Epigenetics in PD.

Epigenetics is the study of mitotically and/or meiotically heritable changes that cannot be explained by changes in DNA sequence. The term epigenetics literally means ‘‘above genetics’’ and describes mechanisms layered on top of the DNA sequence information that are perpetuated. Some broader definitions do not consider the requisite of heritability, thus defining epigenetics as the information not encoded in the DNA sequence, i.e. stable and long-term alterations not present in the DNA sequence.

Although all somatic cells in a given individual are genetically identical, with the exception of T- and B- cells, there are different cell types that form highly distinct anatomic structures with disparate physiologic functions. This is due to the epigenetic control of gene expression that plays a critical role in cellular differentiation [445].

Therefore, one person carries one genome but hundreds or even thousands of epigenomes: even each cell can have its own epigenome and change it depending on time and environmental factors (the environment especially influences epigenome during developmental periods) (Figure 43) [446].

For example:

[447] conducted an epigenetic study in Spanish monozygotic twin pairs and concluded that some epigenetic marks such as DNA methylation and H3 and H4 acetylation, and, consequently, gene expression profiles, were more different between the components of monozygotic twin pairs as they aged, and

the difference was more pronounced in those that had spent less time together and/or had different lifestyles, pointing out the influence of environmental factors. These differential marks were distributed throughout their genomes, affecting repeat DNA sequences and single-copy genes.

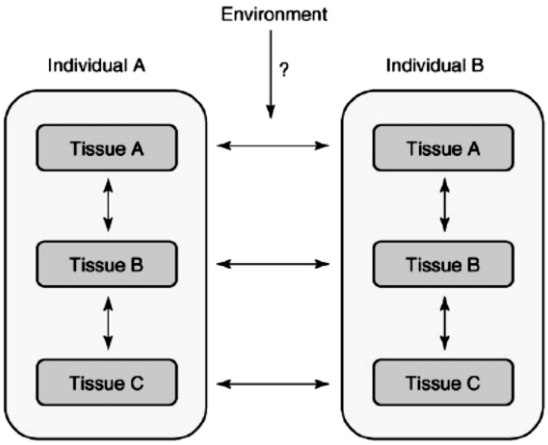


Fig. 43. Intra- and interindividual variation in epigenetics.
Epigenetics is dynamic and there are evidences for variation in it between tissues and individuals although the influence of environment needs further investigation.
For the same genotype, different epigenotypes could explain different phenotypes.

Modified from [446]

[448] postulated that differences in the DNA methylation patterns observed between some parts of the brain (cerebral cortex, cerebellum and pons) could explain their different expression patterns. They also concluded that the DNA methylation pattern correlated much more strongly within a brain region across individuals than within an individual across brain regions.

Histone modifications, DNA methylation and some RNA-mediated processes[44] are part of epigenetics.

A. Histone modifications:

Chromatin is mainly composed of DNA, histones and non-histone chromatin proteins that facilitate the packing of the DNA into higher order structures thus allowing its storage in the nucleus.

The basic structural unit of the chromatin is the nucleosome, which consists of 147bp of DNA wrapped around a histone octamer made of two copies each of the core histones H2A, H2B, H3 and H4 (Figure 44).

Fig. 44. Chromatin structure.
Schematic representation of three nucleosomes and three histone modifications (small circles).

Modified from [449]

[44] RNA-mediated processes involve non-coding RNAs: small (shorter than 200 nucleotides; miRNAs -microRNAs-, siRNAs -short interfering RNAs-, piRNAs -PIWI-interacting RNAs-, rasiRNAs -repeat-associated RNAs- and others less well characterized) and long (lncRNA: longer than 200 nucleotides). They can epigenetically transmit regulatory information thus controlling gene expression. Although some examples of those processes have been described, more extensive research is necessary to clarify and define all their characteristics.

Core histones are evolutionarily highly conserved basic proteins that tightly bind with the negatively charged DNA. Each histone, which is predominantly globular, contains a flexible domain called the histone tail, which remains outside of the nucleosome and that is the target for reversible post-translational modifications [450].

The histone modifications[45] that have been described in humans are:

- methylation (R-me1, R-me2a, R-me2s or K-me1, K-me2, K-me3) is carried out by lysine methyltransferases and arginine methyltransferases, whereas demethylation by lysine demethylases;
- acetylation (K-ac): acetyltransferases (HATs) and deacetylases (HDACs) are responsible for this modification;
- phosphorylation (S-ph or T-ph) is due to serine/threonine kinases;
- ubiquitination (K-ub) is carried out by ubiquitilases;
- sumoylation (K-su): sumoylases are responsible for this modification;
- ADP ribosylation (E-ar) is due to ADP ribosyltransferases;
- deamination (R>citrulline) is carried out by deaminases;
- and proline isomerization (P-cis>P-trans): proline isomerases are responsible for this modification.

All the modifications are dynamic except methylation of arginines [451].

Histone modifications can alter chromatin structure, and therefore gene expression, by recruiting non-histone proteins or by changing the

[45] R, K, S, T, E and P refer to the corresponding amino acids; 1: mono; 2: di; 3: tri; a: asymmetric; S: symmetric.

interaction between histones and DNA (changes in internucleosomes interactions). Nevertheless, many questions remain unclear: its individual contribution to this regulation, the possible combined effect when some are present in the same or different histone tail or its time-dependent variation amongst others, making the histone code an extremely complex and unknown mechanism.

B. DNA methylation:

Mammalian DNA methylation has been implicated in a diverse range of cellular functions and pathologies, including tissue-specific gene expression, development, cell differentiation, silencing of mobile elements, genomic imprinting, X chromosome inactivation, regulation of chromatin structure, carcinogenesis and aging [449].

DNA methylation is generally associated with a repressed chromatin state and inhibition of promoter activity, i.e. transcriptional repression. Two models of repression have been proposed (Figure 45):

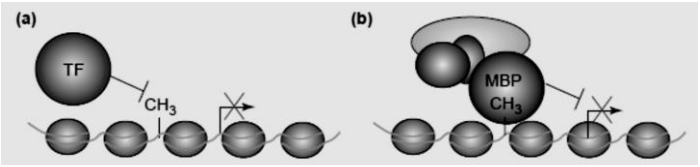


Fig. 45. Mechanisms of DNA-methylation-mediated repression.

a. DNA methylation in the binding sequences of some transcription factors can prevent their binding thus inhibiting transcriptional activation.
b. Methyl-CpG-binding proteins (MBPs) recognize and bind to methylated DNA and recruit chromatin modifiers to establish a repressive chromatin environment.

Modified from [452]

DNA methylation is a covalent modification of DNA catalyzed by DNA methyltransferases (DNMTs).

There are different DNMTs:

- DNMT1 is the maintenance methylase. It works on hemimethylated DNA to methylate cytosines and thus restore the symmetrical DNA methylation pattern on daughter DNA strands generated during replication.
- DNMT3A and DNMT3B are the *de novo* methyltransferases that methylate previous unmethylated DNA.
- DNMT2 has a very weak methylating activity. Its structure suggests that this enzyme is actually involved in the recognition of DNA damage, DNA recombination and mutation repair.
- DNMT3L does not contain intrinsic DNA methyltransferase activity but physically associates with DNMT3A and DNMT3B and modulates their catalytic activity.

DNMT1, DNMT3A, DNMT3B and DNMT3L interact with HDAC and other proteins and participate in the regulation of gene transcription [453, 454].

Some of those other proteins include the two families of MBPs that exist in mammals (•) and proteins involved in chromatin remodeling complexes, that use the energy associated with ATP hydrolysis to effect changes in nucleosome arrangement or composition (□):

- MBDs: MeCP2, MBD1, MBD2, MBD3 and MBD4 share a MBD (methyl-CpG-binding domain) domain to bind to methylated DNA. However,

MBD3 harbors a critical mutation in its MBD domain that abrogates its binding to methylated DNA. MBD1 can also bind unmethylated DNA via its CxxC zinc-finger motif.

All MBDs, except MBD4, repress gene transcription by interacting with histone deacetylases, other histone-modifying enzymes and chromatin remodeling complexes. In contrast, MBD4 is a thymine glycosylase which acts as a DNA repair protein and targets sites of cytosine deamination: it is likely to have a role in limiting the mutagenicity of 5-methylcytosine (5mC) which converts to thymine (T). This enzyme is also able to repair and excise non-methylated CpG when mutates to UpG.

- Kaiso-like: Kaiso, ZBTB4 and ZBTB38 share a triple zinc-finger domain, to bind methylated DNA, and a BTB/POZ domain, involved in protein-protein interactions to repress transcription. Kaiso and ZBTB4 can also bind to unmethylated DNA [455].

- The SNF2H or ISWI family, that act by mobilizing nucleosomes along the DNA.
- The Brahma or SWI/SNF family, that transiently alter the structure of the nucleosomes, thus exposing the DNA-histone contacts. Some of the remodelling complexes that belong to this family also promote the replacement of conventional core histones with variant forms thus acting as “exchanger complexes” [456].

Gene expression and chromatin structure are tightly correlated and dependent on the interaction of numerous proteins: DNMTs, MBPs, chromatin remodeling complexes, histone-modifiers, transcription factors or polymerases amongst others.

This is reflected in the two types of chromatin that can be distinguished (Figure 46):

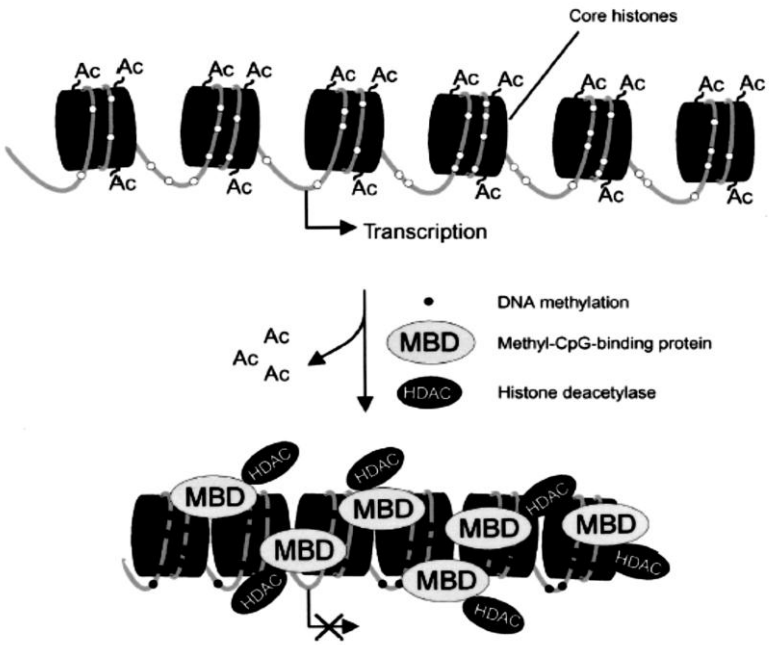


Fig. 46. Schematic representation of silent heterochromatin and active euchromatin.
On the top, open chromatin or euchromatin, where transcriptional machinery easily works: acetylation of the histones causes an open chromatin configuration that is associated with transcriptional activity.
On the bottom, closed chromatin or heterochromatin where there is gene silencing: methylated cytosines are recognized by MBPs, which in turn recruit histone deacetylases to the site of methylation, converting the chromatin into a closed structure that can no longer be accessed by the transcription machinery.

Taken from [457]

- Euchromatin is more loosely packed, which allows the access of transcription factors and other components to promoter regions, thereby enabling genes to be transcribed. There are low levels of DNA methylation and high levels of histone acetylation and trimethylation at H3K4, H3K36 and H3K79.
- Conversely, heterochromatin is more compactly arranged, leading to transcriptional inhibition due to the inaccessibility of promoter elements. It is associated with low levels of acetylation, high levels of methylation at H3K9, H3K27 and H4K20 and high DNA methylation [451].

In addition to the previously described interactions associated with DNA methylation, this modification has some specific characteristics:

In eukaryotes, methylation occurs almost exclusively in cytosines and, in mammals, more concretely in CpG[46] dinucleotides (Figure 47).

Modified from [456]

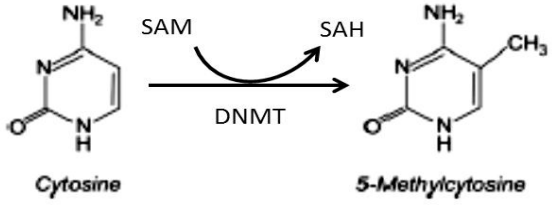


Fig. 47. Conversion of cytosine to 5-methylcytosine by a DNA methyltransferase (DNMT). DNMT catalyzes the transfer of a methyl group from SAM to the C5 position of the cytosine.

[46] The "p" represents the phosphodiester bond that links both nucleotides.

It is noteworthy that the frequency of CpG dinucleotides in the human genome is lower than the expected due to cytosine deamination[47]. Nevertheless, there are some genomic regions where the observed frequency is closer to the expected. Those intergenic and intragenic regions across the genome are called **CpG islands**. CpGs abound in CpG islands but they are generally not methylated or have relatively low levels of methylation. Paradoxically, mostly of CpG dinucleotides are scattered across the genome, not located in the CpG islands, and methylated.

In overall, the 80% of cytosines located in CpG dinucleotides are methylated [446].

CpG islands are defined by three characteristics: G+C content of 0.5 or greater; observed/expected CpG dinucleotide ratio of 0.6 or greater; and all these characteristics occurring within a sequence longer than 200 nucleotides [458].

Depending on their presence, genes are classified as:

I. CG rich: the 60%; all the housekeeping genes and one half of the tissue-specific. There is a CpG island in their promoter but with low methylation level.

[47] Unmethylated cytosine deaminates and converts to uracil (U). Nevertheless, when it occurs, the mutation is rapidly corrected and almost never fixes in the genome. However, the deamination of 5-methylcytosine (it converts to T) is slowly corrected and, as a consequence, it is more often fixed in the genome.

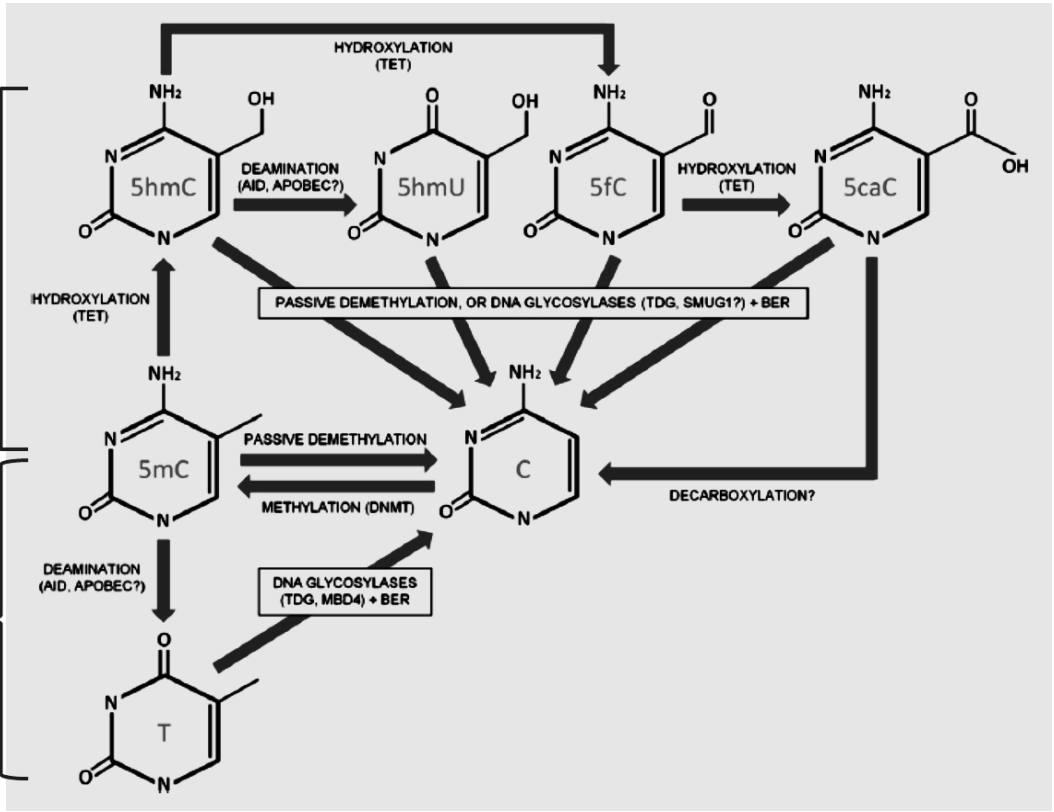
II. CG poor: the remaining 40%; the other half of tissue-specific genes. There is no CpG island in their promoter but the CpG dinucleotides are highly methylated [459, 460].

The promoters in housekeeping genes are poorly methylated and gene transcription is not repressed thus allowing their expression in all the cells. But the differences observed in the promoters of the tissue-specific genes generate rather than explain questions about the regulation of genic expression: how are the promoters of the CG rich genes organized in expressing and non-expressing tissues, considering that CpG islands remain unmethylated in both situations?

Although DNA methylation has been viewed as a stable epigenetic mark, studies in the past decade have revealed that this modification is not static. **DNA demethylation** has been observed in specific contexts and can occur through active or passive mechanisms:

- Passive demethylation occurs when there is no maintenance of the methylation pattern after DNA replication. If DNMT1 is inhibited or absent, the newly synthesized DNA will not be methylated and successive rounds of cell division will result in passive demethylation.
- Active demethylation: a large variety of mechanisms have been proposed to explain how 5mC is replaced by C, but still need to be verified experimentally *in vivo*. There could be different pathways of

demethylation depending on the cell type, the moment and the genomic target (Figure 48).



Modified from [461]

Fig. 48. Putative pathways for cytosine demethylation.

1. Deamination of 5mC into T by cytosine deaminases such as AID or APOBEC, followed by excision of the T (generation of an abasic site) by the DNA glycosylases TDG or MBD4 and repair by the base excision repair (BER) pathway, which removes the rest of the nucleotide and fills the gap by DNA polymerases and ligases.
2. Oxidative demethylation of 5mC which is hydroxylated into 5-hydroxymethylcytosine (5hmC) by TET enzymes. It can be further oxidized into 5-formylcytosine (5fC), and finally into 5-carboxylcytosine (5caC) or deaminated into 5-hydroxymethyluracil (5hmU). All these modified bases could lead to replication-dependent passive demethylation or to active demethylation via excision by DNA glycosylases (TDG, SMUG1) followed by BER. In addition, a putative decarboxylase could lead to decarboxylation of 5caC directly back to cytosine.

5hmC is the only modified base that has been detected but [462] could not correlate the levels of 5hmC and 5mC as one could expect considering that 5hmC is exclusively generated from 5mC.

The presence of 5mC in genomic DNA is constant (≈4-5%). However, the levels of 5hmC are strongly tissue-specific [463] with the highest levels in the central nervous system, where it is present in all the brain regions, and the lower levels in liver, muscle, heart... The amount of 5hmC increases with age until it reaches a final stable plateau but, in any case, it represents less than the 1% [464].

These details point out the possibility that active demethylation via 5hmC is an inaccurate mechanism. Although there are other possibilities to explain these controversial data such as short lives of 5fC, 5caC and 5hmU or inefficient detection limits in current analytical methods or even that 5hmC is more than an intermediate of oxidative demethylation and have a specific function in development or regulation acting as a new epigenetic mark.

Moreover, previous methylation datasets need careful reevaluation because the standard methods to detect 5mC do not differentiate it from 5hmC or do not recognize this new base:

- methylation-sensitive restriction enzymes do not distinguish 5mC and 5hmC [465];

- the anti-5mC antibody has a high selective affinity to 5mC but not to 5hmC (anti-5hmC are in development) [466];
- known 5mC binding proteins such as MBD1, MBD2 and MBD4, do not to recognize 5hmC;
- bisulfite treatment, which is the gold standard, converts unmethylated cytosines into uracils whereas 5mCs remain unaltered and 5hmCs form an adduct (cytosine-5-methylenesulfonate or CMS). CMS and 5mC are resistant to deamination and therefore both are read as C whereas Us are replaced by Ts in the subsequent PCR amplification [467]. Nevertheless, it has been postulated that 5hmC and 5mC could be differentiated by the efficiency of the PCR process, because the formation of the CMS adduct could inhibit the amplification; different studies, however, obtained opposite conclusions about that [466, 467].

Nowadays some techniques have been developed to overpass this limitation. Most of the methods use glucosyltransferases to transfer a glucose group, that can be modified, onto 5hmC followed by mass spectroscopy, enzyme digestion, affinity purification [468] or even by radioactivity [469].

However, these assays determine the genomic position of specific cytosines or quantify their level of hydroxymethylation. But do not combine both possibilities.

That has recently changed, because [470] developed the first method for quantitative mapping of 5hmC in genomic DNA at single-nucleotide resolution called oxidative bisulfite sequencing (oxBS-Seq) which compares two

sequences obtained from the same sample that has been treated with bisulfite (unmethylated Cs convert into Us) or chemically oxidized (5hmC converts into 5fC which is bisulfite treated and converted into U). Nevertheless, more methods need to be developed to determine 5hmC in a cheap, easy, sensitive, efficient, reproducible and precise way as it is for 5mC.

Could epigenetic processes lead to Parkinson´s disease?

Only few studies have focused on this new hypothesis. Some of them have analyzed the levels of 5-methylcytosine in genes related to PD pathogenesis (Table 5).

Table 5. Comparison of the studies that have carried out DNA methylation anaylisis in PD.

Work	[471]	[472]	[473]	[474]
Region analyzed	293bp in *TNF*α promoter	469bp in *UCHL1* promoter and exon 1; 346bp in *MAPT* promoter; 464 bp and 292bp in *APP* promoter; 329bp in *PSEN1* promoter and exon1	443bp in *SNCA* intron1	153bp in *SNCA* intron1
Brain regions	SNpc and cortex	Cortex	SNpc, cortex and *putamen*	SNpc, cortex and *putamen*
Healthy controls/PD patients	8/7	26/8	6/6	7 people affected by other non-neurodegenerative diseases

				where used as controls /12PD+DLB
Method of analysis	Bisulfite treatment, PCR amplification, cloning and sequencing	Bisulfite treatment, PCR amplification and mass spectroscopy	Bisulfite treatment, PCR amplification, cloning and sequencing	Bisulfite treatment, PCR amplification, cloning and sequencing
Conclusions	DNA methylation levels were not different when comparing cases and controls: in both groups, SNpc was less methylated than cortex	No differences in DNA methylation levels were detected in any of the regions analyzed when comparing controls and cases	Hypomethylation was observed across the three regions in PD cases. In addition, the differentially methylated CpG sites were associated with predicted transcription factor binding sites, suggesting that reduced methylation could promote αSyn expression in PD brain	There were no differences in cortex and *putamen*. However, PD cases presented lower DNA methylation levels than the controls in SNpc

The conclusions are promising: although there were not significant differences in all the studies, these were observed in *SNCA*, one of the five genes responsible for familial PD.

Low DNA methylation levels in its intron 1 could lead to increased gene expression and thus be a pathogenic risk factor.

These results highlight the possible influence that epigenetic marks could have on PD pathogenesis. Although more studies are needed to confirm this hypothesis, these reports point out that not only genetic marks should be considered in the future to analyze susceptibility factors in Parkinson´s disease.

II. OBJECTIVES.

II. OBJECTIVES.

Genetics in Parkinson´s disease:

1. Analysis of genetic susceptibility factors in Spanish population.

2. Analysis of genetic determinants in cognitive status in Parkinson´s disease.

3. Analysis of genetic variability in the *ARMCX* gene family in PD.

Epigenetics in Parkinson´s disease:

4. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

II. OBJECTIVES.

III. MATERIAL AND METHODS.

Subjects:

For all the individuals, their personal data were obtained from clinical histories. In addition, all provided an informed written consent[48].

PD cases were diagnosed by neurologists that evaluated the motor impairment and disability following the UPDRS scale. Healthy controls were not related to PD cases and belonged to DNA/brain banks present in the respective institutions.

Individuals from different Spanish regions were analyzed in our **genetic studies**:

A. Basque country: Hospital Universitario Donostia, San Sebastián.

Table 6. Demographic characteristics of the entire Basque population.

	Healthy controls (n = 96)	**PD cases (n = 170)**
Sex	women: 57 (59.38%)	women: 76 (44.71%)
Age (years)*	mean: 73.41 (11.45)**	mean: 63.83 (10.15)**
Evolution (years)	-	mean: 6.18 (4.64)**
Familial history	-	yes: 72 (42.35%)

n = number of individuals.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

[48] For patients with dementia from Navarra, this was obtained from a relative.

The sex distribution between PD cases and healthy controls was statistically different (Pearson´s X^2 = 5.281; p<0.05, p=0.022). Furthermore, their mean age was statistically different too (Mann-Whitney U test; p<0.05, p=0.000).

Therefore, we eliminated individuals with extreme values for age. The groups that we finally analyzed were (Table 7):

Table 7. Demographic characteristics of the Basque population analyzed.

	Healthy controls (n = 86)	**PD cases (n = 151)**
Sex	women: 50 (58.14%)	women: 68 (45.03%)
Age (years)*	mean: 66.91 (8.28)**	mean: 65.98 (7.67)**
Evolution (years)	-	mean: 6.40 (4.77)**
Familial history	-	yes: 64 (42.38%)

n = number of individuals used in the calculation.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

There was no statistically significant difference in sex distribution between healthy controls and PD cases (Pearson´s X^2 = 3.765; p>0.05, p=0.052). Their mean age was not statistically different either (Mann-Whitney U test; p>0.05, p=0.308).

B. Andalusia: Hospital Virgen del Rocío, Sevilla.

Catalonia: Hospital de la Santa Creu i Sant Pau, Barcelona.

Region of Valencia: Hospital Universitari i Politècnic la Fe, València; Hospital Clínic Universitari, València; Hospital Universitari Doctor Peset, València; Hospital Universitari Arnau de Vilanova, València; Hospital Universitari General, València; Hospital Universitari de la Ribera, Alzira and Hospital de Sagunt, Sagunt.

Table 8. Demographic characteristics of the entire non-Basque population.

	Healthy controls (n = 311)	**PD cases (n = 419)**
Sex	women: 119 (38.26%)	women: 181 (43.20%)
Age (years)*	mean: 61.83 (12.25)**	mean: 56.43 (11.47)**
Evolution (years)	-	mean: 10.92 (6.64)**
Familial history	-	yes: 78 (28.57%)***

n = number of individuals. All the healthy controls were from Sevilla whereas PD cases were from Valencia (113), Sevilla (180) and Barcelona (126).

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

***There was no information about familial history for all the individuals from Valencia and for 33 from Sevilla. Therefore, familial history was calculated considering 273 people.

The sex distribution between PD cases and healthy controls was not statistically different (Pearson´s X^2 = 1.795; p>0.05, p=0.180). Nevertheless, their mean age was statistically different (Mann-Whitney U test; p<0.05, p=0.000).

Therefore, we eliminated the individuals with extreme values for age. The groups that we finally analyzed were (Table 9):

Table 9. Demographic characteristics of the non-Basque population analyzed.

	Healthy controls (n = 304)	**PD cases (n = 349)**
Sex	women: 117 (38.49%)	women: 146 (41.83%)
Age (years)*	mean: 60.48 (10.84)**	mean: 59.35 (9.47)**
Evolution (years)	-	mean: 9.62 (5.37)**
Familial history	-	yes: 59 (26.82%)***

n = number of individuals. There were 100 PD patients from Valencia, 163 from Sevilla and 86 from Barcelona. All the healthy controls were from Sevilla.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

***There was no information about familial history for all the individuals from Valencia and for 29 from Sevilla. Therefore, the calculation was done considering 220 people.

The sex distribution between PD cases and healthy controls was not statistically different (Pearson´s X^2 = 0.757; p>0.05, p=0.384). Their mean age was not statistically different either (Mann-Whitney U test; p>0.05, p=0.081).

- To carry out our third objective, we selected some individuals from the non-Basque population previously described in Tables 8 and 9. More concretely (Table 10):

Table 10. Demographic characteristics of the individuals selected from the non-Basque population.

	Healthy controls (n = 160)	PD cases (n = 113)
Sex	Women: 57 (35.63%)	Women: 44 (38.94%)
Age (years)*	mean: 61.81 (12.40)**	mean: 57.55 (9.12)**
Evolution (years)	-	mean: 11.08 (5.89)**

n = number of individuals used in the calculation. Healthy controls were from Sevilla; PD cases were from Valencia.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

The distribution of sex between PD cases and healthy controls was not statistically different (Pearson´s X^2; p>0.05, p=0.577). However, their mean age was statistically different (t-student test; p<0.05, p=0.01).

Therefore, we had to conduct a paired case-control study. PD cases and healthy controls with the same sex and age (even with one year of difference) were matched. For those new groups (Table 11):

Table 11. Demographic characteristics of the individuals included in the paired case-control study.

	Healthy controls (n = 95)	PD cases (n = 95)
Sex	Women: 38 (40%)	Women: 38 (40%)
Age (years)*	mean: 58.96 (8.63)**	mean: 58.92 (8.61)**
Evolution (years)	-	mean: 10.45 (5.32)**

n = number of individuals used in the calculation.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

C. Navarre: Clínica Universitaria de Navarra, Pamplona.

As previously said, our second objective was the analysis of genetic determinants in cognitive status in Parkinson´s disease. However, as we had the possibility to analyze healthy controls from Navarra, and not only PD cases, we decided to also search for genetic susceptibility factors on PD pathogenesis on Navarrese individuals (Table 12):

Table 12. Demographic characteristics of the entire Navarrese population.

	Healthy controls (n = 59)	**PD cases (n = 119)**
Sex	women: 32 (54.24%)	women: 45 (37.82%)
Age (years)*	mean: 68.53 (6.40)**	mean: 59.61 (9.05)**
Evolution (years)	-	mean: 14.35 (4.67)**

n = number of individuals.

* For PD cases, that is age at onset.

**In parenthesis, the standard deviation.

The sex distribution between PD cases and healthy controls was statistically different (Pearson´s X^2 = 4.334; $p<0.05$, $p=0.037$). Furthermore, their mean age was statistically different too (Mann-Whitney U test; $p<0.05$, $p=0.000$).

Due to the differences observed in the demographic characteristics between both groups, the relevance of the results obtained could be compromised.

PD patients were classified into three categories depending on their results after the cognitive assessment (Table 13):

Table 13. Demographic characteristics of PD patients.

	Normal cognition (n = 50)	**Mild cognitive impairment (n = 36)**	**Dementia (PDD) (n = 25)**
Sex	women: 16 (32%)	women: 13 (36.11%)	women: 12 (48%)
Age (years)*	mean: 57.24 (7.34)**	mean: 62.51 (8.66)**	mean: 62.96 (9.59)**
Evolution (years)	mean: 14.55 (3.81)**	mean: 13.83 (4.76)**	mean: 14.61 (4.87)**

n = number of individuals.

* As they are PD cases, that is age at onset.

**In parenthesis, the standard deviation.

Sex distribution between them was not statistically different (Pearson´s X^2 = 1.847 (2df); p>0.05, p=0.397). Their period of evolution was not statistically different either (Kruskal-Wallis test; p=0.794). Nevertheless, their mean age was statistically different (ANOVA test; p<0.05, p=0.003): the group with normal cognition had a younger onset than the others.

As the individuals with cognitive impairment (mild or severe) had similar mean age at onset, period of evolution and proportion of males/females (and small genetic differences -see Results IV.2-) these individuals were included in a unique group named cognitive impairment (C.I.; 61 PD cases, i.e. 36 with M.C.I.

and 25 with dementia) that was compared against the individuals with normal cognition but Parkinson´s disease (N.C.; 50 PD cases) to find genetic susceptibility factors on cognitive impairment.

It is noteworthy that both groups, N.C. and C.I., were statistically different for mean age at onset and that could influence the results obtained.

With regard to the **epigenetic study**, the analyses were conducted in two different Spanish groups:

- In blood from 5 randomly selected male-female pairs of Parkinson´s disease patients, age and period of evolution-matched, from the Region of Valencia.
  - Mean age at onset, years: 58.40 (standard deviation: 10.54).
  - Mean evolution, years: 12.10 (standard deviation: 8.32).

- In *substantia nigra*, parietal cortex and occipital cortex from 5 PD cases and 5 controls from the Biobanc HCB-IDIBAPS, Barcelona, Catalonia (Table 14). Their clinical status was confirmed by *post-mortem* brain analysis. Controls, unlike PD patients, did not present Lewy bodies. However, 4 had other neurological injuries such as vascular encephalopathy and/or AD related pathology and, therefore, 3 of them had suffered vascular dementia.

Table 14. Demographic characteristics of the brain donors.

	Healthy controls (n = 5)	**PD cases (n = 5)**
Sex	3 men and 2 women	3 men and 2 women
Age (years)*	mean: 77.80 (6.80)**	mean: 81.00 (3.81)**

n = number of individuals.

* That is age at the moment of death.

**In parenthesis, the standard deviation.

For one of the controls, we could not obtain DNA from SN, so for this brain region our results are based on the values of only 4 individuals.

Cognitive evaluation:

The cognitive state was determined for all the subjects from Navarra with the Mini Mental State Examination (MMSE) and the Blessed Dementia Scale (BDS). In addition, to precisely define the rate of cognitive impairment, the different cognitive domains were evaluated by the following tests: verbal episodic memory using the Free and Cue Selective Reminding test of Buschke and Cerad word list; visual episodic memory by copy and delayed recall of two simple figures [Massachusetts General Hospital Boston]; the Boston naming test to check language; attention and executive functions with the Raven Progressive matrices, semantic (animals), and phonetic (words starting with ''p'') verbal fluency, Trail Making Test parts A and B, the Stroop test, and digit span forward and backwards. Furthermore, activities of daily living were evaluated with the Interview for Deterioration in Daily Living in Dementia scale (IDDD) and depression was rated using the Yesavage Geriatric Depression Rating Scale (GDS). All these tests were blindly applied by the same person [436].

The methodological process was different between genetic and epigenetic studies and, therefore, it will be explained separately.

III.1. Genetics in PD.

DNA extraction:

The DNA we used was extracted and quantified in the respective hospitals from whole blood using standardized protocols.

Primers design:

Primers, except for the pyrosequencing assays, were designed using Gene Runner (Michael Spruyt and Frank Buquicchio, version 3.01, 1994, Hastings Software Inc.; **http://www.generunner.net/**) based on genomic DNA sequences obtained from the UCSC genome browser (**http://genome.ucsc.edu/**) [475].

To enhance their ability to amplify the target DNA during PCR, we designed them, whenever possible, with the following characteristics: length between 18 and 25 nucleotides, absence of secondary structures (neither intramolecular nor intermolecular, including with the other member of the pair), melting temperature lower than 65ºC, melting temperature difference between the members of the pair lower than 2ºC and cytosine or guanine as the last nucleotides at the 3´ end.

Genotyping:

All the genotyping methods carried out were based on PCR amplification: PCR, allele-specific PCR (AS-PCR), restriction fragment length polymorphism (RFLP), pyrosequencing and fragment analysis. As a consequence, each time, an aliquot of the PCR was loaded in an agarose gel

(1.5%) to confirm the presence of a band of the correct size, and thus the success of the reaction, by electrophoresis.

1. PCR: to determine the presence of the insertion in *CBS* and the deletion in *MAPT*, the previously described steps were sufficient (PCR reaction followed by electrophoresis in a 1.5% agarose gel).

   PCR mix per one reaction (1X) for a final volume of 15µL: 10.8µL Milli-Q water + 1.5µL 10X buffer + 0.5µL dNTPs 5mM each + 0.5µL primers 10µM (F+R) + 0.7µL TAQ DNA polymerase 1U/µL (Biotools) + 1µL DNA 50ng/µL.

2. AS-PCR: two primers, that only differed in their last 3´nucleotide, which corresponded to the two alleles of the SNP to analyze, plus a common primer at the opposite strand were necessary; two PCRs were conducted per individual, each one with a different primer combination. The band pattern observed after the 1.5% agarose electrophoresis was enough to determine the genotype.

   PCR mix per one reaction (1X) for a final volume of 20µL: 15µL Milli-Q water + 2µL 10X buffer + 0.5µL dNTPs 5mM each + 0.5µL primerF 10 µM + 0.5µL primerR 10 µM + 0.5µL TAQ DNA polymerase 1U/µL (Biotools) + 1µL DNA 50ng/µL.

   Exception: in *UCHL1* assays, 1µL of DMSO was also added adjusting the final volume with 14µL of Milli-Q water.

3. RFLP: these assays are based in that differences in the nucleotide sequence at the position analyzed create or destroy restriction endonuclease recognition sequences.

Once the region of interest had been amplified, a restriction reaction was carried out and the genotype determined by the restriction fragments pattern obtained visualized in an electrophoresis.

PCR mix per one reaction (1X) for a final volume of 25µL: 20µL Milli-Q water + 2.5µL 10X buffer + 0.5µL dNTPs 5mM each + 0.5µL primers (F+R) 10 µM + 0.5µL TAQ DNA polymerase 1U/µL (Biotools) + 1µL DNA 50ng/µL.

For *APOE* the PCR mix per one 50µL reaction (1X) was: 36.5µL Milli-Q water + 5µL 10X buffer + 1.1µL dNTPs 5mM each + 1.1µL primer (F+R) 10 µM + 2.5µL DMSO + 0.8µL TAQ DNA polymerase (Biotools) + 3µL DNA 50ng/µL.

Exception: when additives were needed to improve specificity, the final volume was kept constant by adjusting the amount of water in the mix. The SNPs requiring additives were 1.25µL of DMSO for rs6684770 in *ATP13A2* and rs2298969 in *HTT*.

Restriction mix per one 15µL reaction (1X): 1.5µL 10X buffer + 0.2µL restriction enzyme 5-10U/µL (Fermentas or NEB) + 13.3µL PCR.

Manufacturer´s instructions were followed to obtain successful reactions: mostly of them were incubated overnight at 37ºC but some required specific conditions such as different incubation temperature, addition of BSA and/or SAM[49].

In these reactions, the presence of a second recognition site for the same restriction enzyme, not affected by the DNA change analyzed, which was thus always used by the enzyme, served as an internal control, in addition of the use of positive controls with known genotype. This was a useful strategy to ensure that there had not been partial restrictions.

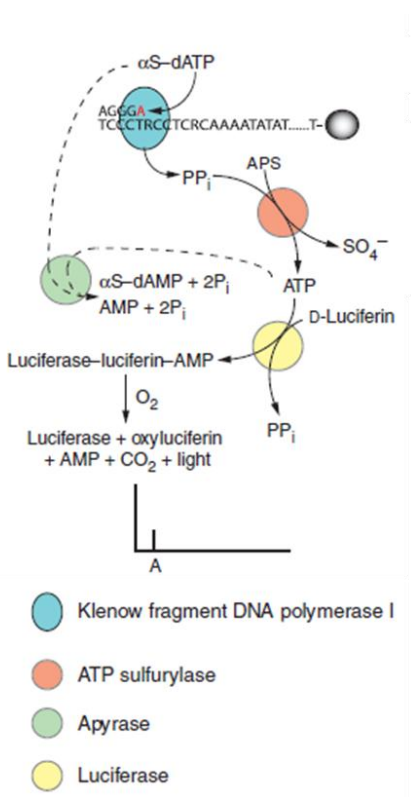
[49] BSA: bovine serum albumin. SAM: S-adenosyl methionine.

For the design of these assays, the NEBcutter V2.0 tool (**http://tools.neb.com/NEBcutter2/**) from NEB was used.

4. Pyrosequencing: it is a sequencing-by-synthesis method. The first step of the process involves a PCR amplification with one of the primers biotin-labeled at its 5´-terminus. Once it has finished, the incorporated biotinylated primer is immobilized on streptavidin-coated beads and, thus, a single-stranded PCR product is purified. Then, the pyrosequencing primer, which has been designed in the opposite strand to the biotin-labeled primer, and, moreover, with its 3´ end around 1bp before the variable nucleotide to interrogate, is hybridized and the sequencing reaction starts.

   Only if the added nucleotide is complementary to the template DNA it is incorporated by a DNA polymerase. The release of pyrophosphate molecules, PPi, during the iterative incorporation of nucleotides, can be monitored in real time because they are quantitatively converted into a proportional light signal by the consecutive activity of three enzymes (Figure 49) [476].

   The pyrosequencing assays, i.e. the primer design, the dispensation order and other features necessary for the proper development of the reaction, were designed using the PyroMark Assay Design Software 2.0 from Qiagen. The reaction was carried out in a PyroMark MD sequencer using NDTS (nucleotide dispensing tips). Results were analyzed by the program PyroMark MD 1.0. These and all the other necessary products for the pyrosequencing were from Qiagen.



Modified from [476]

Fig. 49. Enzymatic cascade of the pyrosequencing reaction. αS-dATP is added as nucleotide for extension. The incorporation of the complementary αS-dATP by the Klenow fragment of DNA polymerase I at the 3′ end of the pyrosequencing primer results in the release of PPi, which is in turn used to convert adenosine phosphosulfate (APS) into ATP. The ATP provides the energy to form an unstable luciferase–luciferin–AMP complex, which in the presence of oxygen results in the release of light in a proportional amount to the available ATP and thus PPi. Unincorporated αS-dATP as well as to a small extent ATP is degraded to the mononucleotides by apyrase before adding the next nucleotide. Carefully balanced proportions of the respective enzymes ensure the preferential incorporation of the nucleotide instead of degradation.
When a non-complementary nucleotide is added, it is degraded by apyrase, there is no generation of ATP and, therefore, no release of light.
This modified nucleotide, αS-dATP, is used instead of dATP, which would act as a direct substrate of the luciferase and thereby uncouple the generation of a bioluminometric signal from the release of PPi. In all the other cases, unmodified nucleotides are used: dCTP, dTTP and dGTP.
To prohibit primer degradation and thus out-of-phase signals, an exonuclease-deficient DNA polymerase is used for nucleotide incorporation.

PCR mix per one reaction (1X) for a final volume of 25µL: 19.4µL Milli-Q water + 2.5µL 10X buffer + 0.5µL dNTPs 5mM each + 0.5µL primerF 10 µM + 0.5µL primerR 10 µM + 0.6µL TAQ DNA polymerase 1U/µL (Biotools) + 1µL DNA 50ng/µL.

5. Fragment analysis: one of the primers used in the PCR amplification had to be labeled with a fluorescent tag. In our case, 6FAM.

   PCR mix per one reaction (1X) for a final volume of 10µL: 7.2µL Milli-Q water + 1µL 10X buffer + 0.25µL dNTPs 5mM each + 0.25µL primers (F+R) 10 µM + 0.5µL TAQ DNA polymerase 1U/µL (Biotools) + 0.8µL DNA 50ng/µL.

1µL of PCR was diluted with Milli-Q water to a final volume of 20µL and then processed following the appropriate protocol.

These assays analyzed the length of the fragment, rather than its sequence. For that purpose, amplicons were subjected to capillary electrophoresis using an automated DNA sequencer ABI Prism 3130XL (Life Technologies) with the GeneScan-500Liz Size Standard as the size marker. Fragment sizes were obtained by analyzing the data with the program Gene Mapper. These and all the other necessary products were from Life Technologies.

There is a frequent problem when using PCR to analyze homopolymers such as those found in *TOMM40*: there is a significant amount of 'slippage' during each DNA amplification cycle, causing the newly polymerized strand to have either fewer or more nucleotides than the original template strand. After several cycles of PCR, the amplification product contains a complex mixture of PCR amplicons that vary in length at the poly-T *locus*, some of them with the true poly-T length. As a consequence, it is difficult to analyze the electropherograms produced by capillary electrophoresis of the PCR products in order to determine precisely the original template length. However, the lengths of the PCR amplicons are normally distributed, i.e. each electropherogram shows a cluster of peaks with a normal distribution of peak heights (Figure 50).

We assumed that the true amplicon length had the highest frequency in the mixture of PCR-product lengths, reflected by the highest intensity peak

within the cluster: the fragment length of the highest peak, or the Mode value, indicated the original (pre-PCR) fragment length [439].

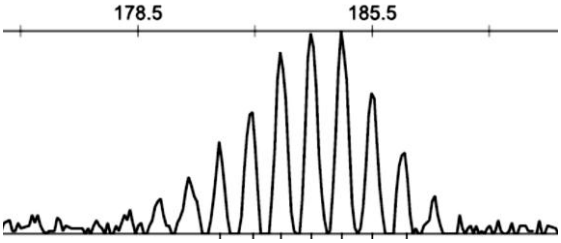


Fig. 50. Example of an electropherogram showing the slippage and the "stutter" peaks.

[Obtained from our own experiments]

Genotyping controls.

To control the quality and reproducibility of the genotyping results, we included, in each reaction, individuals with known genotypes obtained by Sanger sequencing[50]. Moreover, 20% of the assays were independently repeated in order to check for consistency.

[50] We used the same primers, PCR mix and PCR conditions for the sequencing than for the genotyping except in the AS-PCR assays where it was necessary to design the second common primer (forward or reverse depending on the case) and thus different PCR conditions were followed. PCR mix in those cases was the same than for RFLP assays.

All the sequencing reactions to obtain the positive controls as well as those to genotype PD cases and healthy controls by fragment analysis and pyrosequencing assays were carried out at the DNA Analysis Facility of the Instituto de Biomedicina de Valencia by Beatriz Águeda Gómez and Silvia Aparicio Domingo.

We would like to thank to Dr. Elvira V. De Marco (Institute of Neurological Sciences, National Research Council, Mangone, Cosenza, Italy) for sending us DNA from two heterozygous individuals for *GBA* mutations (one for p.N370S and other for p.L444P) that we used as positive controls.

All the primers used in this section are shown, alphabetically ordered, in Annex I, where the mutations and polymorphisms studied are described as well as the criteria to select them, the primer sequences, the size of the fragment amplified, the PCR conditions and the method of DNA analysis. For those which are RFLP, there are two more features, the restriction enzymes used and the percentage of agarose necessary to differentiate the restriction fragments by electrophoresis.

III.2. Epigenetics in PD.

DNA extraction:

The DNA we used from the Valencian patients was extracted and quantified in the respective hospitals from whole blood using standardized protocols.

We followed the Maxwell 16 Mouse Tail DNA Purification Kit instructions to extract DNA from each individual´s brain region: around 30mg of tissue were dissected and afterwards introduced in the Maxwell 16 Instrument (Promega). Finally, it was quantified by using the Qubit dsDNA BR Assay kit and the Qubit Fluorometer (Life Technologies).

Bisulfite treatment:

Although there are other options, bisulfite treatment is the gold standard technique to analyze 5-methylcytosine (5mC). In this process, unmethylated C converts into U whereas 5mC remains unaltered. Although 5mC can also react with bisulfite, the reaction is extremely slow and the equilibrium favors 5mC rather than the deaminated product T.

[477] were the first who used sodium bisulfite to differentiate C and 5mC, and later, [478] optimized the technique (Figure 51).

At the present time, commercial kits are available for this purpose. The original procedure has been replaced by easy, short, effective, exhaustively

tested and optimized alternative protocols which minimize DNA degradation and conversion failure, i.e. Cs that do not change to Us or 5mCs that change to Ts.

We opted for the EZ DNA Methylation-Gold kit (Zymo Research) and treated 1μg of DNA per individual and region. Finally, it was quantified by using NanoDrop 2000 (Thermo Scientific) as if it was RNA, because bisulfite treated DNA has an absorption coefficient at 260nm that resembles that of RNA.

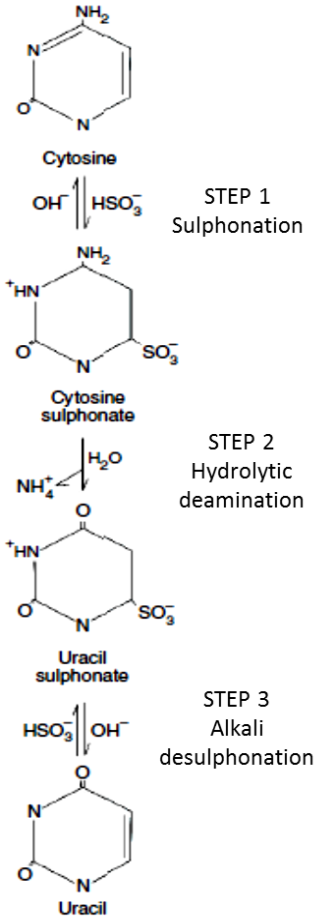


Fig. 51. Chemical conversion of cytosine into uracil.
The reaction is highly single-strand specific and cannot be performed on double-stranded DNA.
The first step, the formation of the sulfonated cytosine derivative (cytosine-SO_3) is reversible. The extent of adduct formation is controlled by pH, bisulfite concentration and temperature. The forward reaction is favoured by low pH and the reverse reaction by high pH.
In the second step of the reaction, cytosine-SO_3 undergoes hydrolytic deamination to give uracil-SO_3. This step is catalyzed by basic substances, such as sulfite, bisulfite and acetate anions. Since sulfonation is favoured by acidic pH, the reversible sulfonation reaction and the subsequent irreversible deamination step are both carried out at pH below 7.
The third step involves alkali treatment to remove the bisulfite adduct.

Modified from [479]

<u>CpG island prediction:</u>

The 5 genes responsible for the familial forms of Parkinson´s disease were analyzed: *DJ-1*, *LRRK2*, *PINK1*, *PRKN* and *SNCA*.

For each one, we considered ≈3000bp upstream the transcription starting site (TSS) and ≈3500bp downstream it to include the promoter and the first exons, where CpG islands are frequently present.

It was sufficient to observe the distribution of CG dinucleotides in those sequences to confirm that there was an accumulation of CG dinucleotides around the TSS. Nevertheless, to precisely delimit the CpG island, we uploaded the regions to 5 prediction programs:

- Softberry: **http://linux1.softberry.com/berry.phtml?topic=cpgfinder&group=programs&subgroup=promoter**
- CpG cluster: **http://bioinfo2.ugr.es/CpGcluster/** [480]
- Zeus2: **http://zeus2.itb.cnr.it/cgi-bin/wwwcpg.pl**
- Bioinformatics: **http://www.bioinformatics.org/sms2/cpg_islands.html**
- Emboss: **http://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/**

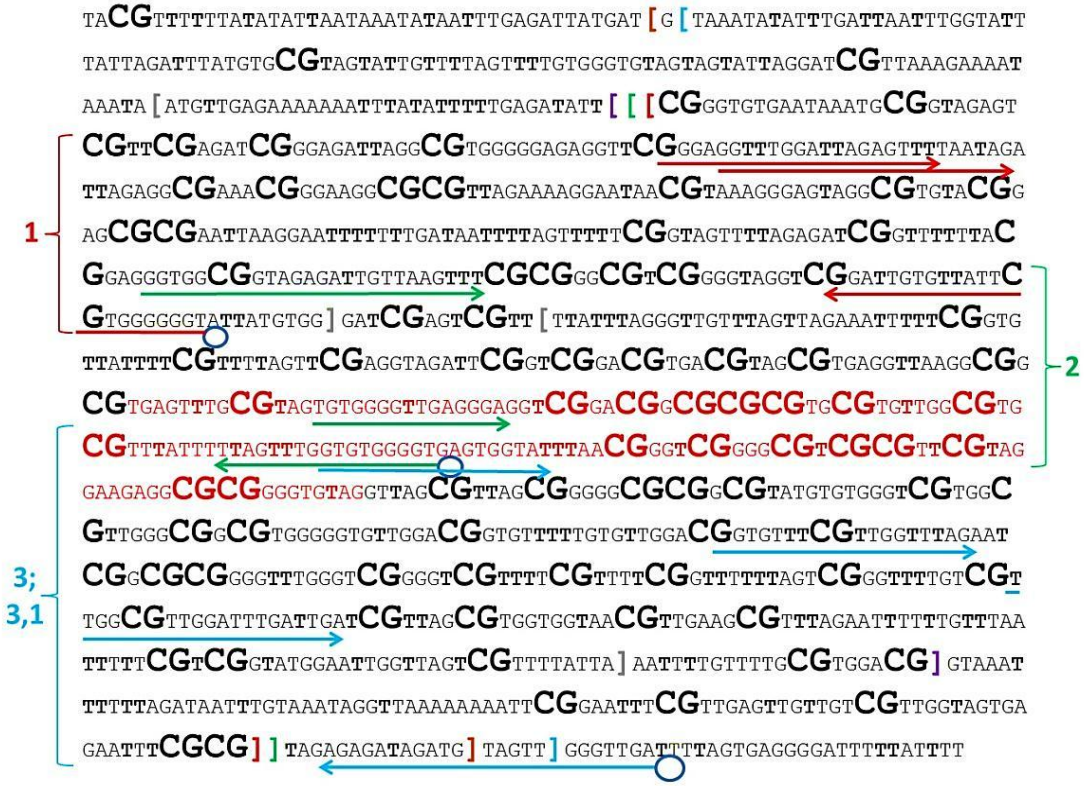
All the programs except *CpG cluster* define CpG islands as [458] did and therefore calculate G+C content and CpG fraction using a 200bp window moving across the sequence at 1bp intervals. However, *CpG cluster* considers that there is no objective standard for defining CpG islands and thus do not calculate any of those three parameters: it is based on the physical distance between neighboring CGs and can find shorter CpG islands. Furthermore it points out that the algorithms that use the moving-window approach cannot accurately define the island boundaries to single-pair resolution.

Furthermore, we annotated the CpG island that UCSC genome browser considered.

Below are the results we obtained for the **CpG island prediction**. In each case, the table indicates the different predicted sizes whereas the image shows their localization in the converted sequence and, in addition, the position of the designed primers (arrows; the white circle indicates which is biotin-labeled). Bold thymines correspond to unmethylated cytosines. The numbers on the sides of the sequence indicate the assays designed to analyze the CpG island.

1. ***DJ-1***: NM_007262 (it has a longer exon 1 than NM_01123377).

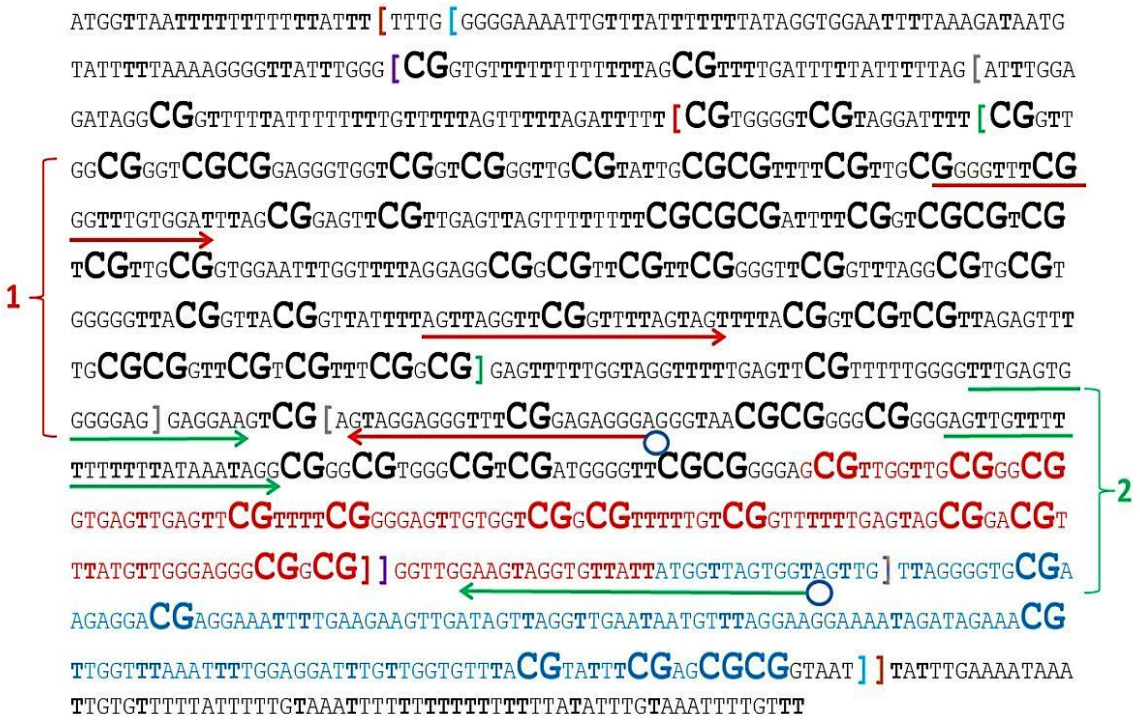
[UCSC]	[Softberry]	[CpG cluster]	[Zeus 2]	[Bioinformatics]	[Emboss]
925	925	840	1079	1075	335+507



Non-coding exon 1 is colored red.

2. ***LRRK2***: NM_198578.

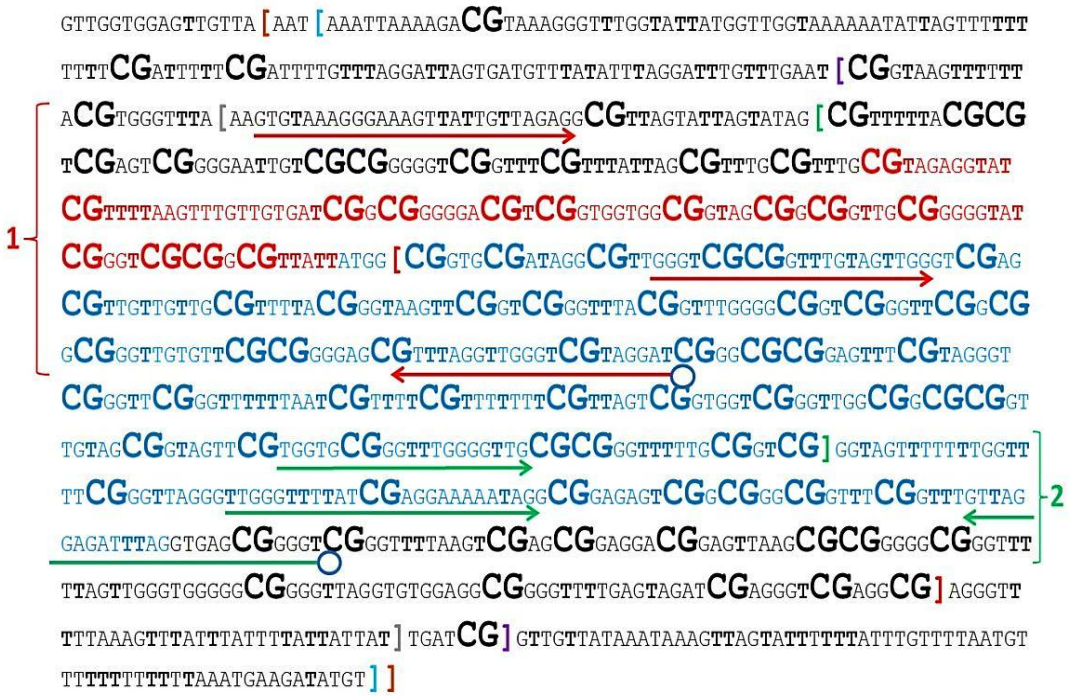
[UCSC]	[Softberry]	[CpG cluster]	[Zeus 2]	[Bioinformatics]	[Emboss]
558	282	649	895	899	403+235



Exon 1 is colored in red (non-coding) and in blue (coding).

3. ***PINK1***: NM_032409.

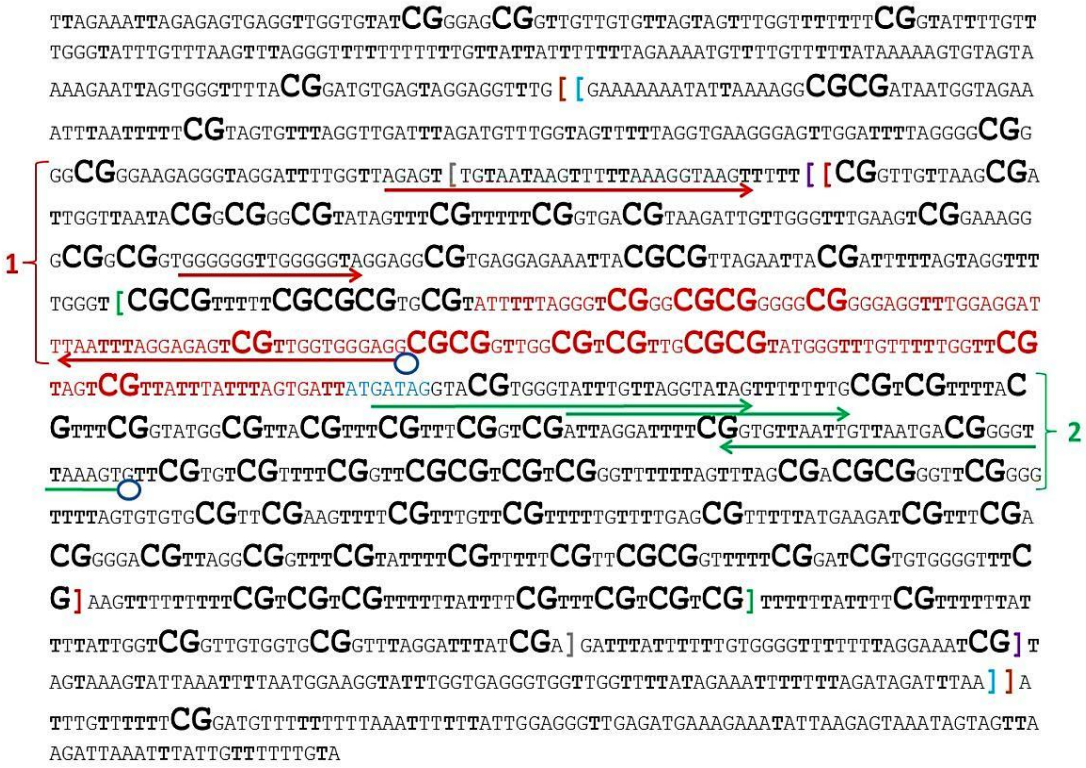
[UCSC]	[Softberry]	[CpG cluster]	[Zeus 2]	[Bioinformatics]	[Emboss]
506	435	779	965	969	749



Exon 1 is colored in red (non-coding) and in blue (coding).

4. ***PRKN***: NM_013988 (although the three isoforms share the same exon 1).

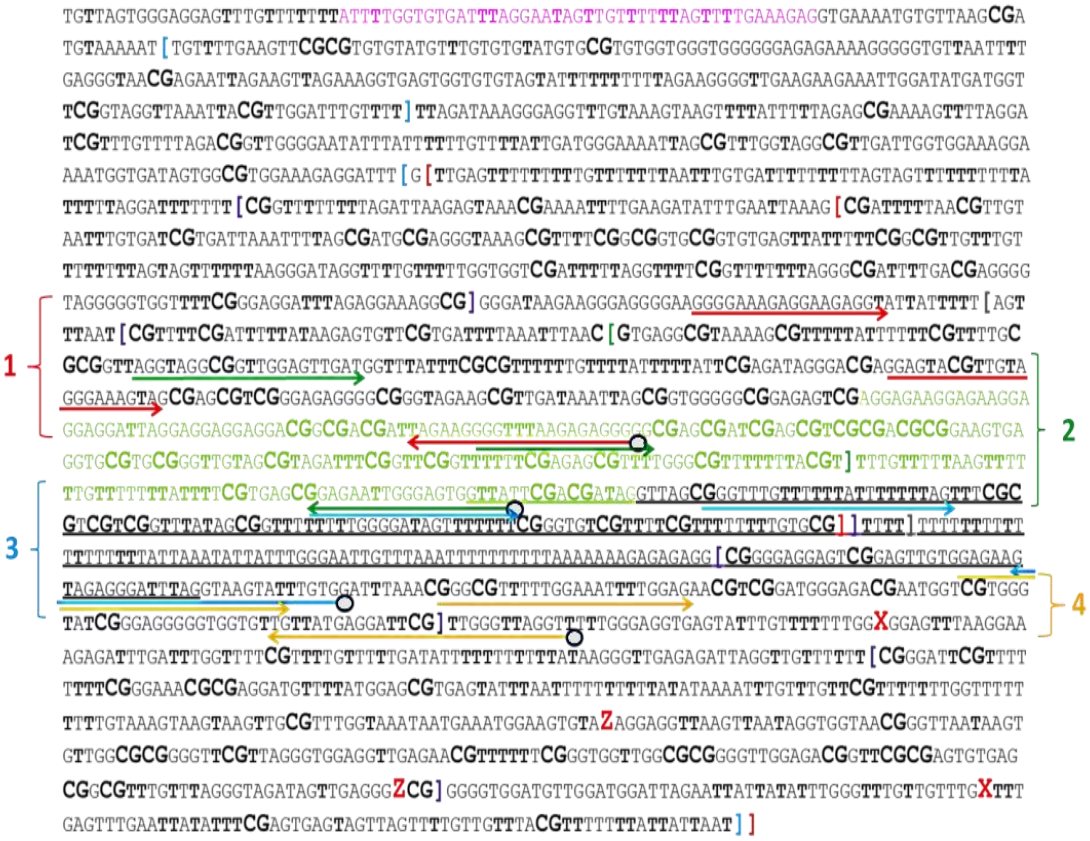
[UCSC]	[Softberry]	[CpG cluster]	[Zeus 2]	[Bioinformatics]	[Emboss]
641	522	778	1027	1027	772



Exon 1 is colored in red (non-coding) and in blue (coding).

5. ***SNCA***: exon 1 is non-coding in all four isoforms.

[UCSC]	[Softberry]	[CpG cluster]	[Zeus 2]	[Bioinformatics]	[Emboss]
862	364	282+579+149+306	200+1762	1761	591



Exon 1 in NM_001146055 is colored in pink. Exon 1 in NM_000345 is colored in green. Exon 1 in NM_007308 and NM_001146054 is underlined.

X and **Z** delimit the regions analyzed by [473] and [474], respectively.

It is noteworthy that the limits of the CpG islands were mostly different between programs even when they were based on the same principles. At least, in all the 5 cases, there was a "core" island shared by all of them.

Primers design:

The methylation analysis was carried out by pyrosequencing. Therefore, three primers were necessary: two for the PCR reaction (forward plus reverse; one of them biotin-labeled) and one more, the pyrosequencing primer, in the opposite strand to the biotin-labeled primer for the sequencing reaction.

The pyrosequencing assays, i.e. the primer design, the dispensation order and other features necessary for the proper development of the reaction, were designed using the PyroMark Assay Design Software 2.0 from Qiagen based on genomic DNA sequences obtained from UCSC genome browser (**http://genome.ucsc.edu/**) [475].

Prior to performing the experiment, we analyzed the presence of other frequent SNPs in the sequences to be analyzed by using the Single Nucleotide Polymorphism database, dbSNP (**http://www.ncbi.nlm.nih.gov/projects/SNP/**) because their presence could alter the assay. Only for DJ-1 3 we found that there was a polymorphism (rs35675666) that was taken into account in the design and thus in the dispensation order to overpass failures.

Primers satisfied some specific features to enhance the success of the amplification reaction:

- length between 18 and 30 nucleotides;
- absence of secondary structures or primer dimer formation;
- melting temperature difference between the members of the pair lower than 2ºC;
- amplicon length between 100 and 500 nucleotides (optimal size around 300);
- inclusion of a limited number of CG dinucleotides in the primer sequence, no more than 2, which should be located as far as possible from the 3′ end. Otherwise the primers would have been entirely selective for methylated templates and thus will only amplify methylated sequences;
- inclusion of some Ts originated from non-CpG Cs at, or near, the 3′ end of the primer to ensure amplification of only bisulfite modified DNA [481].

It should be noted that, after the bisulfite reaction, the two DNA strands are no longer complementary and therefore can be independently amplified (Figure 52).

Our first intention was to analyze the whole predicted CpG islands. Nevertheless, pyrosequencing has its own limitations and reactions longer than 50 nucleotides and/or sequences that include homopolymers longer than 4 nucleotides do not work properly. Moreover, the distribution of CG

dinucleotides in the target sequence determined where primers could be designed following the previously explained necessary characteristics.

Therefore, we designed some assays for each gene but only covered specific regions (specific CpG sites). We tried to, at least, study areas before, in and after the TSS, included in the "core" CpG island.

In addition we made a trial prediction about the position of the promoter in those genes with FirstEF (**http://rulai.cshl.edu/tools/FirstEF/**) and WWW Promoter Scan (**http://www-bimas.cit.nih.gov/molbio/proscan/**) to check that, anyhow, analyzed regions partially overlapped with it (see Annex II).

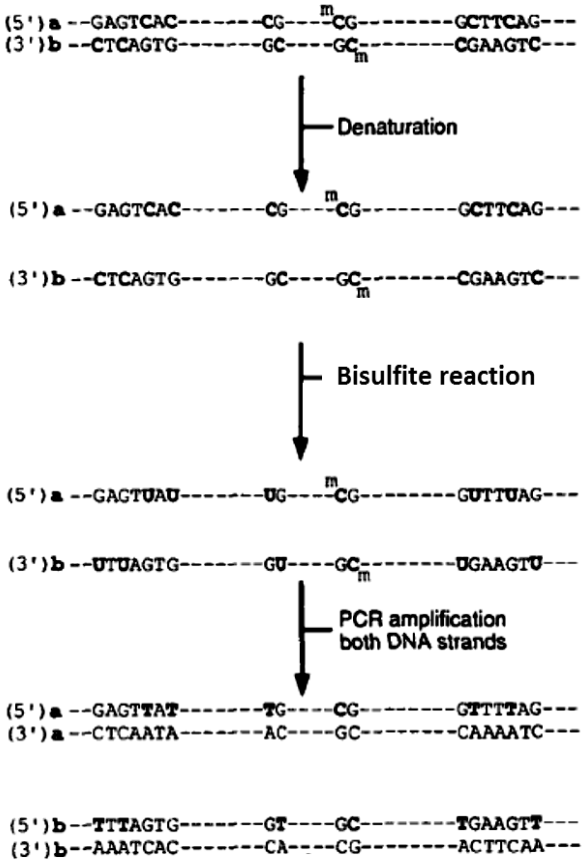


Fig. 52. Bisulfite genomic sequencing procedure.
The two complementary strands in the original DNA are labeled a and b. Cytosine residues and their corresponding uracil and thymine conversion products are shown in bold type. a and b are no longer complementary after the bisulfite treatment.

Modified from [478]

Pyrosequencing:

Previously to pyrosequencing, we confirmed by sequencing[51] that all the PCR amplifications corresponded to our target sequences. In addition, all the primer sets were tested with non-bisulfite treated DNA as a template to eliminate the possibility that they amplified PCR products from unconverted DNA.

Furthermore, other important detail was checked: PCR bias. It is a sequence-dependent and often strand-specific bias that originates because the primers might not proportionally amplify methylated and unmethylated sequences, which are different after bisulfite conversion, thus leading to an inaccurate estimation of methylation [482]. For this purpose, PCRs were carried out with DNA with known methylation percentages (0, 50 and 100) and then DNA methylation levels were analyzed by pyrosequencing to confirm that the observed and the expected methylation levels matched and, thus, there was no bias. We used the EpiTect PCR Control DNA Set from Qiagen which includes methylated and unmethylated bisulfite treated human DNA.

Each time, an aliquot of the PCR was loaded in an agarose gel (1.5%) to confirm the presence of a band, and thus the success of the reaction, by electrophoresis[52]. Then, they were pyrosequenced or, when appropriate, Sanger sequenced.

[51] We used the same PCR mix and PCR conditions in sequencing than in pyrosequencing.
[52] Except when using non-bisulfite treated DNA where the success was the absence of amplification.

For each condition tested, individual and region or tissue analyzed, we obtained three independent replicates.

The reactions were carried out in a PyroMark MD sequencer using NDTS (nucleotide dispensing tips). Results were analyzed by the program PyroMark Q-CpG 1.0.9. These and all the other necessary products for the pyrosequencing were from Qiagen. [For more details about pyrosequencing, see section III.1. Genetics in PD, Genotyping, Pyrosequencing].

All pyrosequencing (and sequencing) reactions were carried out at the DNA Analysis Facility of the Instituto de Biomedicina de Valencia by Beatriz Águeda Gómez and Silvia Aparicio Domingo.

In Annex II, and alphabetically ordered, all the pyrosequencing assays conducted are described as well as the primer sequences, the size of the fragment amplified and the PCR conditions.

Statistics:

The X^2 test was used to compare the proportion of males/females between groups. The t-student test (ANOVA test) or the Mann-Whitney test (Kruskal-Wallis test), depending on the number of groups to compare and their distribution, were used to compare the mean age and the period of evolution.

The asymptotic Pearson's chi-square goodness-of-fit test was used to check that the distribution of the polymorphisms at control groups followed the HWE. There were only three deviations, i.e. p-values lower than 0.05, that are indicated when corresponds at the discussion section.

Most of genotype and allele frequencies between the PD cases and controls were compared by using the X^2 test. Nevertheless, Fisher´s exact test was used for mutations at *GBA*, *NR4A2* and *APP*, which are less frequent.

To estimate the effects of genotypes and demographic factors on the probability to develop PD, a binary logistic regression model was calculated when appropriate, thus obtaining sex- and age-correlated OR.

Haploview [483] was used to determine haplotypes when necessary. Those with frequency lower than 1% were not considered. The colors and values indicated at the linkage disequilibrium maps have been calculated using the parameter D´.

Regarding to the epigenetic study, due to the low number of data, a non-parametric test (the Mann-Whitney test) was used to compare means.

To estimate the statistical power of our studies we used the tool in **http://statpages.org/proppowr.html**. Results are described at the following table and below there is an explanation:

Table 15. Results for the calculation of the proportion difference power.

Population	Difference in proportions (X)	1-β (Y)	α → Z
Basque[53]	0.19	80	0.05 → 5%
non-Basque	0.11	80	0.05 → 5%
non-Basque paired case-control	0.20	80	0.05 → 5%
Navarrese cases vs controls	0.20	80	0.10 → 10%
Navarrese 2 groups cases	0.23	80	0.10 → 10%

"Taking into account the number of individuals included, if, between the two groups, there is a true difference in proportions of X or more, it could be detected in the Y% (power, i.e. 1-β) of cases with an error of Z% (α)[54]."

[53] Basque: Table 7; Results IV.1. non-Basque: Table 9; Results IV.1. non-Basque paired case-control: Table 11; Results IV.3. Navarrese cases vs controls: Table 12; Results IV.2. Navarrese 2 groups cases, i.e. considering N.C. and C.I.: Table 13; Results IV.2.

[54] Type I error or α error is related to the rate of false positives. In our case that is the error of accepting that there is a difference in proportions between the two groups when there is not such difference. Frequently used values for α are 0.05 and 0.1.
Type II error or β error is related to the rate of false negatives. In our case, the error of accepting that there is no difference in proportions when there is a difference. 1-β, the power, is the capacity of detecting a true difference; a value of 80% is the most frequently used.

The statistical analysis was conducted by using IBM SPSS Statistics, version 20.

Vicent Giner Bosch, assistant professor at the Department of applied statistics and operational research and quality at the Universitat Politècnica de València, guided us and solved our statistical doubts during the analysis of the results.

III. MATERIAL AND METHODS.

IV. RESULTS.

Genetics in Parkinson´s disease:

IV.1. Analysis of genetic susceptibility factors in Spanish population.

To determine possible genetic susceptibility factors in Parkinson´s disease in Spanish population, Basque and non-Basque, the genes and mutations/polymorphisms analyzed were[55]:

- the three genes that have been considered genetic susceptibility factors in PD*:*
  - *MAPT*: H1/H2 haplotype;
  - *GBA*: p.N370S and p.L444P;
  - *NR4A2*: c.864+246C>T;
- a controversial *PARK locus*: *UCHL1*, rs5030732 (p.S18Y);
- *TOR1A*, which is related to dystonia, a secondary motor symptom in PD: rs1801968 (p.D216H) and p.delE302/303[56];
- the gene that is considered a genetic susceptibility factor in AD: *APOE*, genotype;
- a new one that could replace it: *TOMM40*, rs10524523 (poly-T homopolymer).

[55] The criteria to select those mutations and polymorphisms is explained in Annex I.

[56] There was no individual, neither PD case nor healthy control, that presented the p.delE302/303. Therefore, we did not include it in the analysis.

We studied two different Spanish populations: one from the Basque country and another one of non-Basque origin (Sevilla, Barcelona and Valencia). Although both are Spanish, we analyzed them by separate because individuals from the Basque country may be genetically different to the rest of Spanish people due to their higher rate of endogamy.

1. Basque population:

The individuals analyzed were described in Table 7 (see section III. Material and methods, Subjects).

First, we analyzed the presence of mutations in *GBA* (p.N370S and p.L444P) (Table 16):

Table 16. Frequency of mutations p.N370S and p.L444P (*GBA*).

GBA	p.N370S	p.L444P	p.N370S + p.L444P
Healthy controls	0 heterozygous	0 heterozygous	0 heterozygous
PD cases	1 heterozygous	0 heterozygous	1 heterozygous
Fisher´s exact test	p=1	p= -	p=1

Although there was no statistically significant difference, the individual that presented the mutation was eliminated from the study on genetic susceptibility factors in PD because mutations in *GBA* are considered some of those factors.

The relevant results, that is, those that reached statistical significance ($p<0.05$) or a strong tendency towards this ($p<0.1$) are listed below. The results obtained for the rest of polymorphisms (p-value higher than 0.1) are detailed in Annex III.

Frequencies are indicated in parenthesis. p-values lower than 0.05, i.e. statistically significant results, are in bold and shadowed. p-values lower than 0.1 are highlighted in bold. We calculated odds ratio for the distribution of alleles in such polymorphisms to know their effect on PD risk or protection.

(1): *MAPT*.

Table 17. Allelic frequency of H1/H2 haplotype (*MAPT*).

Gene and polymorphism	**Alleles**	
MAPT, H1/H2 haplotype	H1	H2
Controls	113 (65.70)	59 (34.30)
PD cases	220 (73.33)	80 (26.67)
	X^2 = 3.068; 1df; **p=0.080** OR:1.44, $CI_{95\%}$=[0.96-2.16]	

The distribution of alleles between healthy controls and PD cases was almost statistically different (p=0.080), being the H1 haplotype the potential susceptibility factor (Table 17). Nevertheless, there was no relevant result when considering genotypes (Table D, Annex III). Only when grouping them in H1H1 and H1H2+H2H2, the H1H1 genotype showed a trend towards increasing the risk of developing PD (p=0.073).

(**2**): *APOE*.

Table 18. Frequency of *APOE* genotype when considering the number of ε4 alleles.

Gene and polymorphism	Genotypes		
APOE, genotype	ε4 +/+	ε4 +/-	ε4 -/-
Controls	0 (0)	12 (13.95)	74 (86.05)
PD cases	0 (0)	38 (25.33)	112 (74.67)
	X^2 = 4.239; 1df; **p=0.039** OR: 2.09, $CI_{95\%}$=[1.03-4.26]		

The only statistically significant result was obtained when individuals were classified according to the number of ε4 alleles that they carried (Table 18; see also Tables F and G, Annex III). The proportion of ε4 +/- individuals was higher for PD cases than for healthy controls. Moreover, bearing a copy of the ε4 allele doubles the risk for PD (OR:2.09).

We could not, however, analyze the effect of carrying two ε4 alleles as no homozygous, ε4 +/+, were found in our populations.

2. Non-Basque population.

The individuals analyzed were described in Table 9 (see section III. Material and methods, Subjects).

First, we analyzed the presence of mutations in *GBA* (p.N370S and p.L444P) (Table 19):

Table 19. Frequency of mutations p.N370S and p.L444P (*GBA*).

GBA	p.N370S	p.L444P	p.N370S + p.L444P
Healthy controls	1 heterozygous	2 heterozygous	3 heterozygous
PD cases	7 heterozygous	10 heterozygous	17 heterozygous
Fisher´s exact test	**p=0.074**	**p=0.042**	**p=0.005**

We observed that the presence of mutations in *GBA* was a genetic susceptibility factor that increased the risk to develop PD.

The individuals that presented the mutations were, consequently, eliminated from the study on genetic susceptibility factors in PD because they were already carrying those factors; even healthy controls were eliminated because they might develop PD in the future due to the presence of this risk factor.

The relevant results, that is, those that reached statistical significance ($p<0.05$) or a strong tendency towards this ($p<0.1$) are listed below. The results obtained for the rest of polymorphisms (p-value higher than 0.1) are detailed in Annex III.

Frequencies are indicated in parenthesis. p-values lower than 0.05, i.e. statistically significant results, are in bold and shadowed. p-values lower than

0.1 are highlighted in bold. We calculated odds ratio for the distribution of alleles in such polymorphisms to know their effect on PD risk or protection.

In addition, we also include the sex and age-correlated odds ratios calculated by using a binary logistic regression model that estimated the effect of their genotypes and demographic factors on the probability to develop PD.

(1): *MAPT*.

Table 20. Genotypic and allelic frequency of H1/H2 haplotype (*MAPT*).

Gene and polymorphism	Genotypes			Alleles	
MAPT, H1/H2 haplotype	H1H1	H1H2	H2H2	H1	H2
Controls	138 (45.85)	124 (41.20)	39 (12.96)	400 (66.45)	202 (33.55)
PD cases	199 (59.94)	105 (31.63)	28 (8.43)	503 (75.75)	161 (24.25)
	X^2 = 12.937; 2df; **p=0.002**			X^2 = 13.375; 1df; **p=0.000** OR:1.58 , $CI_{95\%}$=[1.24-2.02]	

Table 21. Results of the binary logistic regression model for the H1/H2 haplotype.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
H1H1 vs H2H2	1	**0.007**	2.133	1.230	3.699
H1H2 vs H2H2	1	0.497	1.219	0.689	2.156

The H2H2 genotype was taken as a reference, i.e. OR:1.

The distribution of genotypes and alleles between healthy controls and PD cases was strongly statistically different (p=0.002 and p=0.000, respectively; Table 20), being the H1 haplotype an important risk factor that increases the risk to develop Parkinson´s disease in 1.58 times ($CI_{95\%}$=[1.24-2.02]). Under the binary logistic regression model, H1H1 carriers presented a significant increased risk to develop PD, with an age and sex-correlated odds ratio of 2.133 ($CI_{95\%}$=[1.230-3.699]). Although H1H2 carriers seemed to have more risk than H2H2, there was no statistically significant difference (p=0.497) (Table 21).

(2): *TOR1A*.

Table 22. Genotypic frequency of p.D216H (*TOR1A*).

Gene and polymorphism	**Genotypes**		
TOR1A, rs1801968 (p.D216H)	GG	GC	CC
Controls	249 (82.72)	52 (17.28)	0 (0)
PD cases	269 (81.02)	57 (17.17)	6 (1.81)
	X^2 = 5.497; 2df; **p=0.064**		

There was a trend towards significance when considering genotypes in p.D216H (p=0.064). We consider that this result is due to the absence of homozygous CC controls, as this genotype has a very low frequency. Due to that fact, we could not calculate its effect on PD pathogenesis under the binary logistic regression model (Table 22).

(3): *APOE*.

The distribution of genotypes between healthy controls and PD cases was statistically different (p=0.040) as well as the distribution of alleles (p=0.010) (Table 23). ε2 allele was the main responsible (p=0.005/0.001) (Table 24).

Table 23. Genotypic and allelic frequency of *APOE* genotype.

	Genotypes						**Alleles**		
APOE, genotype	ε2ε2	ε2ε3	ε3ε3	ε3ε4	ε4ε4	ε2ε4	ε2	ε3	ε4
Controls	9 (2.99)	28 (9.30)	189 (62.79)	60 (19.93)	2 (0.66)	13 (4.32)	59 (9.80)	466 (77.41)	77 (12.79)
PD cases	12 (3.61)	52 (15.66)	185 (55.72)	53 (15.96)	4 (1.20)	26 (7.83)	102 (15.36)	475 (71.54)	87 (13.10)
	X^2 = 11.615; 5df; **p=0.040**						X^2 = 9.166; 2df; **p=0.010**		

Table 24. Frequency of *APOE* genotype when considering the number of ε2 alleles.

	ε2 +/+	ε2 +/-	ε2 -/-	ε2 +	ε2 -
Controls	9 (2.99)	41 (13.62)	251 (83.39)	50 (16.61)	251 (83.39)
PD cases	12 (3.61)	78 (23.49)	242 (72.89)	90 (27.11)	242 (72.89)
	X^2 = 10.604; 2df; **p=0.005**			X^2 = 10.099; 1df; **p=0.001** OR: 1.87, $CI_{95\%}$=[1.27-2.76]	

In our Spanish non-Basque population, carrying the ε2 allele increased the risk of developing PD in almost two times (OR:1.87), but most of this

effect was due to ε2 +/- carriers, i.e. the main risk factor is the presence of only one ε2 allele, as it was observed for the results obtained under the binary logistic regression models (Tables 25 and 26): only ε2ε3 and ε2ε4 genotypes reached statistical significance (p=0.011 and 0.043, respectively) and were related to an increased risk to develop PD (OR:1.928 and 2.055, respectively). When considering the number of ε2 alleles, the age- and sex-correlated odds ratio again reached statistical significance (p=0.001) for those bearing one allele (OR:2.005; $CI_{95\%}$=[1.319-3.048]).

All the values for OR are similar and indicate that the probability of developing PD doubles for ε2 +/- carriers.

Table 25. Results of the binary logistic regression model for *APOE* genotypes.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
ε2ε2 vs ε3ε3	1	0.471	1.387	0.570	3.376
ε2ε3 vs ε3ε3	1	**0.011**	1.928	1.165	3.191
ε2ε4 vs ε3ε3	1	**0.043**	2.055	1.024	4.125
ε4ε4 vs ε3ε3	1	0.423	2.014	0.364	11.141
ε3ε4 vs ε3ε3	1	0.588	0.890	0.583	1.358

The ε3ε3 genotype was taken as a reference, i.e. OR:1.

Table 26. Results of the binary logistic regression model considering the number of ε2 alleles.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
ε2 +/+ vs ε2 -/-	1	0.444	1.413	0.584	3.420
ε2 +/- vs ε2 -/-	1	**0.001**	2.005	1.319	3.048

Absence of ε2 alleles was taken as a reference, i.e. OR:1.

As we observed that H1H1 carriers (*MAPT*) and ε2 +/- carriers (*APOE*) have an increased risk to develop PD, we calculated if there was a combined effect of both factors in non-Basque individuals (Tables 27 and 28):

Table 27. Distribution of non-Basque cases and controls depending on their genotype for *MAPT* and *APOE*.

H1H1	ε2 +/-	Healthy controls	PD cases
+	+	20 (6.64)	44 (13.25)
+	-	118 (39.20)	155 (46.69)
-	+	21 (6.98)	34 (10.24)
-	-	142 (47.18)	99 (29.82)

+: carrier of H1H1 genotype or one ε2 allele in *APOE*; -: absence.

Table 28. Results of the binary logistic regression model for the combined effect.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
H1H1 and ε2 +/- vs no H1H1 and no ε2 allele	1	**0.000**	3.229	1.790	5.824
H1H1 and no ε2 allele vs no H1H1 and no ε2 allele	1	**0.000**	1.886	1.326	2.681
no H1H1 and ε2 +/- vs no H1H1 and no ε2 allele	1	**0.006**	2.336	1.279	4.266

Individuals that do not carry any ε2 allele and are not H1H1 were taken as a reference, i.e. OR:1.

There was statistical significance in all three comparisons ($p<0.05$):

- H1H1 carriers without any ε2 allele showed an increased risk of developing PD (almost two times; OR:1.886) when compared to those that were ε2 -/- and H1H2 or H2H2 (effect of H1/H2 haplotype).

- The risk was higher, OR: 2.336, for those that were H1H2 or H2H2 but differed in the number of ε2 alleles that carried (effect of ε2 allele).

[In both cases, OR values are around 2 as previously obtained: Tables 21, 25 and 26.]

- However, for those individuals that differed in both factors, there was a difference in the risk of developing PD of more than three times: OR:3.229 ($CI_{95\%}$=[1.790-5.824]). Therefore, the presence of both factors is even riskier.

IV.2. Analysis of genetic determinants in cognitive status in Parkinson´s disease.

Some PD patients develop cognitive impairment (from mild to severe, that is, dementia) during the evolution of the disease. This study pretended to find the genetic susceptibility factors that could explain the appearance of this non-motor symptom.

To that purpose we analyzed the frequency of mutations/polymorphisms located in

- genes related to other diseases where dementia is consubstantial such as
  - AD: *APP* (rs463946), *PSEN1* (rs165932), *PSEN2* (rs6426554) and *CALHM1* (rs2986017 -p.P86L-);
  - HD: *HTT* (rs10015979, rs363066, rs363096, rs2298969 and rs110501);
  - CJD: *PRNP* (rs4815729 and rs1799990 -p.M129V-);
  - FTD: *GRN* (rs11547442 -p.L46P-, rs34975779 -p.V452V- and rs1141754 -p.Q479K-) and *TARDBP* (rs11547736 -p.S92X-);
  - Kufor-Rakeb syndrome, that is also characterized by parkinsonism: *ATP13A2* (rs2871776, rs6684770 and rs4920608);
- genes that encode neurotrophic factors such as *CDNF* (rs7094179 and rs7099185), *MANF* (rs11538558 and *DOCK3* -rs4441646-) and *BDNF* (rs6265 -p.V66M-);
- genes involved in processes proposed to be related to PD:

- homocysteine metabolism, *CBS* (c.844ins68), *MTR* (rs1805087 -c.A2756G-) and *MTHFR* (rs1801133 -c.C677T- and rs1801131 -c.A1298C-);
- protection against oxidative stress, *NFE2L2* (rs1806649 and rs10183914) and *KEAP1* (rs8113472);
- chaperone-mediated autophagy, *HSPA8* (rs1461496 and rs4936770) and *LAMP-2A* (rs7057652, rs42897, rs42895 and rs42890);
- and, due to its relation with tau homeostasis, *GSK3β* (rs334558 and rs6438552).

In addition, some possible genetic susceptibility factors to PD pathogenesis were studied: p.N370S and p.L444P mutations (*GBA*), Rep1 microsatellite and rs356219 polymorphism (*SNCA*), H1/H2 haplotype (*MAPT*), as well as three pathogenic mutations: rs33939927 -p.R1441G and p.R1441C- and rs34637584 -p.G2019S- (*LRRK2*). Furthermore, a genetic susceptibility factor in AD (*APOE* genotype) and a new one related to it (poly-T homopolymer length, rs10524523, *TOMM40*) were considered[57].

[57] The criteria to select those mutations and polymorphisms is explained in Annex I.
The following mutations were not present in any of the individuals. Therefore, they were not included at the subsequent analysis:
- *GRN*: rs11547442 (p.L46P), rs34975779 (p.V452V) and rs1141754 (p.Q479K);
- *TARDBP*: rs11547736 (p.S92X);
- *MANF*: rs11538558;
- *LRRK2*: rs33939927 (p.R1441C).

The analysis was divided in two phases:

1. initially, all these mutations/polymorphisms were studied in the individuals described in Table 12 (see section III. Material and methods, Subjects), i.e. healthy controls and PD cases, to determine if any of them could be considered a genetic susceptibility factor on PD risk in Navarrese population.

First, we analyzed the presence of mutations in

- *LRRK2*: 2 PD cases (1.68%) carried the p.R1441G mutation (both were cognitively normal) and 3 (2.52%) carried the p.G2019S (two presented mild cognitive impairment and the third was cognitively normal). Obviously, no control carried them.
- *GBA*: two patients presented the p.L444P mutation (both were cognitively normal), whereas the PD case that carried the p.N370S variant had mild cognitive impairment. One of the control individuals also carried this mutation. There was no statistically significant difference when comparing the frequency of those mutations between controls and cases (Fisher´s exact test, p=1).

It is noteworthy that the individuals that presented mutations in *GBA* or *LRRK2* were excluded from the subsequent analysis because those factors could involve specific pathological pathways not shared by the other non-

carriers: they were already carrying pathogenic mutations (*LRRK2*) or high risk susceptibility factors (*GBA*).

The relevant results, that is, those that reached statistical significance ($p<0.05$) or a strong tendency towards this ($p<0.1$) are listed below. Frequencies are indicated in parenthesis. p-values lower than 0.05, i.e. statistically significant results, are in bold and shadowed. p-values lower than 0.1 are highlighted in bold. We calculated odds ratio for the distribution of alleles in such polymorphisms to know their effect on PD risk or protection.

When it was possible, we estimated haplotypes that could modify the risk to develop PD. We did not calculate sex and age-correlated odds ratios for the genotypes by using a binary logistic regression model as controls and cases had statistically significant different mean age and sex distribution and, moreover, both groups were composed by a significant different number of individuals. Therefore, caution should be taken when interpreting these results.

The results obtained for the rest of polymorphisms, i.e. those without statistical significance (p-value higher than 0.1), are detailed in Annex IVa.

(1): *SNCA*.

Table 29. Genotypic and allelic frequency of rs356219 (*SNCA*).

Gene and polymorphism	Genotypes			Alleles	
SNCA, rs356219	AA	AG	GG	A	G
Controls	30 (51.72)	25 (43.10)	3 (5.17)	85 (73.28)	31 (26.72)
PD cases	38 (34.23)	58 (52.25)	15 (13.51)	134 (60.36)	88 (39.64)
	X^2 = 6.034; 2df; **p=0.049**			X^2 = 5.571; 1df; **p=0.018** OR: 0.56, $CI_{95\%}$=[0.34-0.92]	

The distribution of genotypes and alleles between healthy controls and PD cases was statistically different (p=0.049 and p=0.018, respectively). The A allele was protective against PD (OR: 0.56; $CI_{95\%}$=[0.34-0.92]), especially in AA genotypes, which were more frequent amongst controls (51.72%) than cases (34.23%) (Table 29). When classifying the individuals as AA vs AG+GG, there was a statistically significant difference (p=0.028): AA carriers had a decreased risk to develop PD (OR:0.49; $CI_{95\%}$=[0.26-0.94]).

(2): *MAPT*.

Table 30. Genotypic and allelic frequency of H1/H2 haplotype (*MAPT*).

Gene and polymorphism	Genotypes			Alleles	
MAPT, H1/H2 haplotype	H1H1	H1H2	H2H2	H1	H2
Controls	18 (31.03)	25 (43.10)	15 (25.86)	61 (52.59)	55 (47.41)
PD cases	69 (62.16)	26 (23.42)	16 (14.41)	164 (73.87)	58 (26.13)
	X^2 = 14.781; 2df; **p=0.001**			X^2 = 15.514; 1df; **p=0.000** OR: 2.55, $CI_{95\%}$=[1.59-4.09]	

The distribution of genotypes and alleles between healthy controls and PD cases was strongly statistically different (p=0.001 and p=0.000, respectively), being the H1 haplotype an important risk factor that increased the risk to develop Parkinson´s disease in 2.55 times ($CI_{95\%}$=[1.59-4.09]); H1H1 carriers were overrepresented amongst PD patients (62.16% vs 31.03% in controls) and had an increased risk to develop the disease (Table 30). When classifying the individuals as H1H1 vs H1H2+H2H2, there was a statistical significance (p=0.0001): H1H1 carriers had an increased risk to develop PD (OR:3.65; $CI_{95\%}$=[1.86-7.17]).

(3): *HTT*.

Table 31. Genotypic and allelic frequency of rs363096 (*HTT*).

Gene and polymorphism	Genotypes			Alleles	
HTT, rs363096	TT	TC	CC	T	C
Controls	5 (8.62)	30 (51.72)	23 (39.66)	40 (34.48)	76 (65.52)
PD cases	22 (19.82)	54 (48.65)	35 (31.53)	98 (44.14)	124 (55.86)
	X^2 = 3.796; 2df; p=0.150			X^2 = 2.944; 1df; **p=0.086** OR: 1.50, $CI_{95\%}$=[0.94-2.39]	

Only the polymorphism rs363096 in the gene *HTT* showed an almost statistically significant result (p=0.086): the T allele tended to increase the risk to develop PD (OR: 1.50; $CI_{95\%}$=[0.94-2.39]) (Table 31). When classifying the individuals as TT vs TC+CC, the same tendency was observed (p=0.059).

This effect was also observed when considering the haplotypes formed by the 5 markers analyzed along the gene, which were in high linkage disequilibrium as observed at Figure 53: the haplotypes ATCAT and ATTAT, that only differentiate in the third position, i.e. the correspondent to rs363096, were protective against PD (p=0.002) and almost risky (p=0.087), respectively (Table 32).

Table 32. Haplotypes in *HTT*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GTTAT	0.328	0.346, 0.293	0.983	0.321	1.28 [0.71-2.32]
ATCAT	0.266	0.211, 0.371	9.995	**0.002**	0.45 [0.24-0.84]
AGCGC	0.151	0.163, 0.129	0.681	0.409	1.31 [0.59-2.89]
ATCGC	0.147	0.156, 0.129	0.437	0.508	1.25 [0.56-2.77]
ATTGT	0.033	0.037, 0.026	0.300	0.584	1.44 [0.29-7.21]
ATTAT	0.017	0.025, 0.000	2.930	**0.087**	-
GTTAC	0.012	0.009, 0.017	0.433	0.510	0.53 [0.04-6.93]
GTTGT	0.011	0.017, 0.000	2.007	0.157	-

The order of the markers at the haplotype is: rs10015979 – rs363066 – rs363096 – rs2298969 – rs110501. Missing ORs correspond to those haplotypes absent in one of the populations.

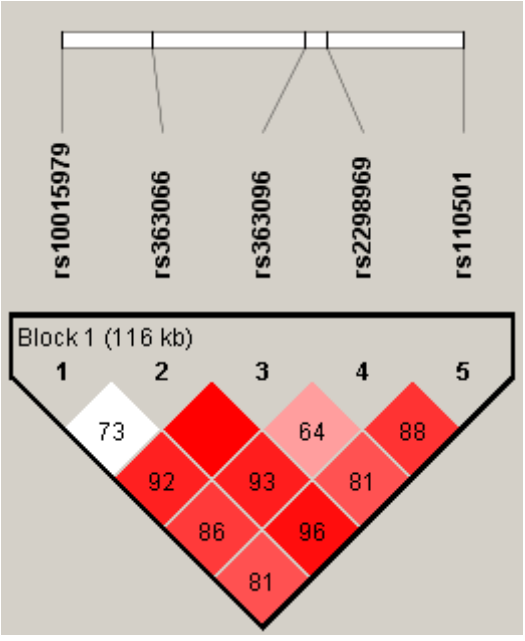


Fig. 53. Linkage disequilibrium map in *HTT*.

(4): *ATP13A2* (Kufor-Rakeb syndrome; dementia and parkinsonism).

Table 33. Genotypic and allelic frequency of rs6684770 (*ATP13A2*).

Gene and polymorphism	Genotypes			Alleles	
ATP13A2, rs6684770	TT	TC	CC	T	C
Controls	2 (3.45)	21 (36.21)	35 (60.34)	25 (21.55)	91 (78.45)
PD cases	11 (9.91)	54 (48.65)	46 (41.44)	76 (34.23)	146 (65.77)
	X^2 = 6.237; 2df; **p=0.044**			X^2 = 5.849; 1df; **p=0.016** OR: 1.89, $CI_{95\%}$=[1.12-3.19]	

Only the polymorphism rs6684770 gave relevant results: the distribution of genotypes and alleles between healthy controls and PD cases was statistically different (p=0.044 and p=0.016, respectively), being the T allele the responsible for the 1.89 times ($CI_{95\%}$=[1.12-3.19]) increased risk to develop Parkinson´s disease (Table 33). When classifying the individuals as TT+TC vs CC, a statistical significance was observed (p=0.020). TT and TC carriers were affected by an increased risk (OR:2.15; $CI_{95\%}$=[1.13-4.11]).

Furthermore, we studied the effect of the haplotypes formed by this and the other two polymorphisms analyzed in *ATP13A2*, rs4920608 and rs2871776, on PD pathogenesis:

Table 34. Haplotypes in *ATP13A2*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
ACA	0.314	0.307, 0.327	0.149	0.700	0.91 [0.50-1.65]
GCG	0.206	0.179, 0.258	2.927	**0.087**	0.63 [0.32-1.24]
ATG	0.148	0.179, 0.089	4.890	**0.027**	2.23 [0.95-5.26]
ACG	0.147	0.130, 0.180	1.552	0.213	0.68 [0.31-1.48]
GTG	0.114	0.111, 0.119	0.044	0.834	0.92 [0.39-2.19]
GCA	0.034	0.042, 0.018	1.290	0.256	2.39 [0.41-14.01]
ATA	0.027	0.037, 0.007	2.746	**0.098**	5.45 [0.42-71.21]
GTA	0.010	0.015, 0.001	1.434	0.231	-

The order of the markers at the haplotype is: rs4920608 – rs6684770 – rs2871776.

One of them, ATG, was statistically risky ($p<0.05$) and two more, GCG and ATA, were almost protective and risky, respectively ($p<0.1$) (Table 34). Nevertheless, other factors, besides the allele present in rs6684770, determine the protective or risky effect of the haplotypes, maybe due to the low linkage disequilibrium observed along the region (Figure 54).

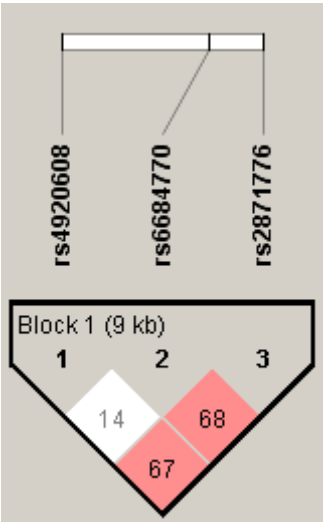


Fig. 54. Linkage disequilibrium map in *ATP13A2*.

(5): *CDNF* (neurotrophic factor).

Table 35. Genotypic and allelic frequency of rs7099185 (*CDNF*).

Gene and polymorphism	Genotypes			Alleles	
CDNF, rs7099185	TT	TC	CC	T	C
Controls	33 (56.90)	21 (36.21)	4 (6.90)	87 (75.00)	29 (25.00)
PD cases	73 (65.77)	38 (34.23)	0 (0)	184 (82.88)	38 (17.12)
	X^2 = 8.175; 2df; **p=0.017**			X^2 = 2.979; 1df; **p=0.084** OR: 1.61, $CI_{95\%}$=[0.93-2.78]	

The distribution of genotypes between cases and controls was statistically different (p=0.017), mainly due to the absence of homozygous CC patients and the higher frequency of TT carriers in PD cases (65.77% vs 56.90% in controls). The same tendency was observed when analyzing the alleles: the T allele showed a trend towards increasing the risk to develop PD (OR: 1.61; $CI_{95\%}$=[0.93-2.78]) (Table 35).

The combined effect of this polymorphism and rs7094179, also located at the *CDNF* gene, revealed two trends in accordance with the previous observation: carriers of the haplotype TT presented a higher risk to develop PD (p=0.055), whereas those carrying the GC were more protected against Parkinson´s disease (p=0.077) (Table 36).

Table 36. Haplotypes in *CDNF*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GT	0.572	0.567, 0.581	0.060	0.807	0.94 [0.54-1.65]
TT	0.230	0.262, 0.169	3.692	**0.055**	1.75 [0.88-3.48]
GC	0.103	0.082, 0.143	3.137	**0.077**	0.54 [0.22-1.34]
TC	0.095	0.089, 0.107	0.261	0.610	0.82 [0.32-2.09]

The order of the markers at the haplotype is: rs7094179 – rs7099185.

(6): *DOCK3* ("neurotrophic factor").

Table 37. Genotypic and allelic frequency of rs4441646 (*DOCK3*).

Gene and polymorphism	Genotypes			Alleles	
DOCK3, rs4441646	AA	AC	CC	A	C
Controls	37 (63.79)	21 (36.21)	0 (0)	95 (81.90)	21 (18.10)
PD cases	75 (67.57)	27 (24.32)	9 (8.11)	177 (79.73)	45 (20.27)
	X^2 = 6.678; 2df; **p=0.035**			X^2 = 0.228; 1df; p=0.633	

The distribution of genotypes between cases and controls was statistically different (p=0.035), mainly due to the absence of homozygous CC controls. Nevertheless, the proportion of alleles A and C was almost the same when comparing healthy controls and PD cases (p=0.633) (Table 37).

(7): *MTHFR* (homocysteine metabolism).

Table 38. Genotypic and allelic frequencies of rs1801133 and rs1801131 (*MTHFR*).

Gene and polymorphism	Genotypes			Alleles	
MTHFR, rs1801133 (c.C677T)	CC	CT	TT	C	T
Controls	18 (31.03)	29 (50.00)	11 (18.97)	65 (56.03)	51 (43.97)
PD cases	46 (41.44)	55 (49.55)	10 (9.01)	147 (66.22)	75 (33.78)
	X^2 = 4.130; 2df; p=0.127			X^2 = 3.378; 1df; **p=0.066** OR: 1.54, $CI_{95\%}$=[0.97-2.44]	
MTHFR, rs1801131 (c.A1298C)	AA	AC	CC	A	C
Controls	38 (65.52)	18 (31.03)	2 (3.45)	94 (81.03)	22 (18.97)
PD cases	53 (47.75)	46 (41.44)	12 (10.81)	152 (68.47)	70 (31.53)
	X^2 = 5.816; 2df; **p=0.055**			X^2 = 6.073; 1df; **p=0.014** OR: 0.51, $CI_{95\%}$=[0.30-0.88]	

The C allele in rs1801133 (c.C677T) showed a trend (p=0.066) towards increasing the risk to develop PD. However, there were no relevant results when considering genotypes: only when grouping the individuals in CC+CT and TT carriers, the same trend was observed (p=0.062).

Nevertheless, in rs1801131 (c.A1298C), the distribution of alleles between healthy controls and PD cases was statistically different (p=0.014), being the A allele protective against PD (OR: 0.51; $CI_{95\%}$=[0.30-0.88]): it was more frequent in controls (81.03%) than in cases (68.47%). This effect was only observed when this allele was found in homozygosity: AA carriers represented the 65.52% of controls whereas the 47.75% of cases (Table 38). When classifying the individuals as AA vs AC+CC, there was a statistically significant difference (p=0.028): AA carriers had a decreased risk to develop PD (OR:0.48; $CI_{95\%}$=[0.25-0.93]).

As both are in the same gene, we analyzed the possible effect that a haplotype could have on PD risk:

Table 39. Haplotypes in *MTHFR*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
TA	0.368	0.332, 0.438	3.728	**0.054**	0.64 [0.36-1.14]
CA	0.360	0.353, 0.372	0.119	0.730	0.92 [0.52-1.64]
CC	0.268	0.309, 0.188	5.673	**0.017**	1.93 [1.00-3.72]

The order of the markers at the haplotype is: rs1801133 (c.C677T) – rs1801131 (c.A1298C).

The results were the expected for the possible combinations of the almost risky C allele (rs1801133) and the protective A allele (rs1801131): they compensated the effect of the other when together (CA; p=0.730), and gave an almost protective haplotype when the almost risky C was absent (TA; 0.054) and a risky haplotype when the protective A was absent (CC; p=0.017) (Table 39).

(8): *HSPA8* (chaperone-mediated autophagy).

Table 40. Genotypic and allelic frequency of rs1461496 (*HSPA8*).

Gene and polymorphism	Genotypes			Alleles	
HSPA8, rs1461496	TT	TC	CC	T	C
Controls	11 (18.97)	26 (44.83)	21 (36.21)	48 (41.38)	68 (58.62)
PD cases	20 (18.02)	30 (27.03)	61 (54.95)	70 (31.53)	152 (68.47)
	X^2 = 6.421; 2df; **p=0.040**			X^2 = 3.252; 1df; **p=0.071** OR: 0.65, $CI_{95\%}$=[0.41-1.04]	

The distribution of alleles between healthy controls and PD patients was almost statistically different (p=0.071) and the T allele tended to be protective (OR: 0.65; $CI_{95\%}$=[0.41-1.04]). The frequency of homozygous TT was almost the same in cases and controls, whereas TC carriers were more frequent in controls (44.83%) than in cases (27.03%) (Table 40). When classifying the individuals as TT+TC vs CC, a statistical significance was observed (p=0.021): carrying at least one T allele was protective against PD (OR:0.47; $CI_{95\%}$=[0.24-0.90]).

When combining these results and the obtained for the other polymorphism analyzed in *HSPA8*, rs4936770, it was observed that the presence of the T allele was always protective although it only reached an almost statistically significant effect at the haplotype TG (Table 41).

Table 41. Haplotypes in *HSPA8*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
CG	0491	0.517, 0.443	1.673	0.196	1.35 [0.77-2.35]
TG	0.293	0.262, 0.350	2.844	**0.092**	0.66 [0.36-1.21]
CA	0.160	0.168, 0.143	0.338	0.561	1.21 [0.56-2.61]
TA	0.056	0.053, 0.063	0.160	0.689	0.83 [0.25-2.73]

The order of the markers at the haplotype is: rs1461496 – rs4936770.

(9): *LAMP-2A* (chaperone-mediated autophagy).

rs7057652: There was a statistically significant difference between both groups for the distribution of genotypes and alleles. The T allele increased the risk to develop PD in 2.40 times ($CI_{95\%}$=[1.38-4.17]), especially in homozygosis: there were more patients (43.24%) that carried the TT genotype than controls (20.69%). The effect was the same independently of sex, although in males it did not reach statistically significant difference (Table 42).

rs42890: There was a statistically significant difference between both groups for the distribution of alleles, and almost for genotypes. The G allele increased the risk to develop PD in 1.94 times ($CI_{95\%}$=[1.14-3.29]), especially in homozygosis: there were more patients (54.95%) that carried the GG genotype than controls (36.21%). When analyzing the results by sex, the G allele was more frequent in PD cases than in controls; however, in men the difference did not reach statistically significant difference. Moreover, the distribution of genotypes in women was different than in the overall group, although the GG genotype was the most frequent (Table 43).

Table 42. Genotypic and allelic frequency of rs7057652 (*LAMP-2A*).

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs7057652	TT	TG	GG	T	G
Controls, overall	12 (20.69)	12 (20.69)	34 (58.62)	27 (30.00)	63 (70.00)
PD cases, overall	48 (43.24)	17 (15.32)	46 (41.44)	77 (50.66)	75 (49.34)
	X^2 = 8.474; 2df; **p=0.014** Z_{mfG}^2 = 7.431; 1df; **p=0.006** Z_C^2 = 8.230; 2df; **p=0.016**			$X^2 = Z_A^2$ = 9.844; 1df; **p=0.002** Z_{mfA}^2 = 8.999; 1df; **p=0.003** OR: 2.40, $CI_{95\%}$=[1.38-4.17]	
Controls, women	3 (9.38)	12 (37.50)	17 (53.13)	18 (28.13)	46 (71.88)
PD cases, women	12 (29.27)	17 (41.46)	12 (29.27)	41 (50.00)	41 (50.00)
	X^2 = 6.107; 2df; **p=0.047** Z_{fG}^2 = 6.078; 1df; **p=0.014**			$X^2 = Z_{fA}^2$ = 7.143; 1df; **p=0.008** OR: 2.56, $CI_{95\%}$=[1.28-5.13]	
Controls, men	-	-	-	9 (34.62)	17 (65.38)
PD cases, men	-	-	-	36 (51.43)	34 (48.57)
	-			$X^2 = Z_m^2$ = 2.152; 1df; p=0.142	

Table 43. Genotypic and allelic frequency of rs42890 (*LAMP-2A*).

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs42890	GG	GT	TT	G	T
Controls, overall	21 (36.21)	13 (22.41)	24 (41.38)	41 (45.56)	49 (54.44)
PD cases, overall	61 (54.95)	16 (14.41)	34 (30.63)	94 (61.84)	58 (38.16)
	X^2 = 5.463; 2df; **p=0.065** Z_{mfG}^2 = 3.828; 1df; **p=0.051** Z_C^2 = 4.875; 2df; **p=0.087**			$X^2 = Z_A^2$ = 6.079; 1df; **p=0.014** Z_{mfA}^2 = 5.098; 1df; **p=0.024** OR: 1.94, $CI_{95\%}$=[1.14-3.29]	
Controls, women	7 (21.88)	13 (40.63)	12 (37.50)	27 (42.19)	37 (57.81)
PD cases, women	17 (41.46)	16 (39.02)	8 (19.51)	50 (60.98)	32 (39.02)
	X^2 = 4.232; 2df; p=0.121 Z_{fG}^2 = 4.231; 1df; **p=0.040**			$X^2 = Z_{fA}^2$ = 5.091; 1df; **p=0.024** OR: 2.14, $CI_{95\%}$=[1.10-4.16]	
Controls, men	-	-	-	14 (53.85)	12 (46.15)
PD cases, men	-	-	-	44 (62.86)	26 (37.14)
	-			$X^2 = Z_m^2$ = 0.644; 1df; p=0.422	

It is noteworthy that the gene *LAMP-2A* is located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [484][58].

The high linkage disequilibrium pattern observed along this gene (Figure 55) revealed that there were two haplotypes with relevance on PD pathogenesis: GAAT was more frequent in PD cases than in controls whereas TAAG was more frequent in controls than in patients. Both differentiate in their first and fourth position, i.e. rs42890 and rs7057652, respectively: when the G allele (rs42890) and the T allele (rs7057652) are present, haplotype GAAT, as both increase the risk to develop PD, carriers have an increased risk too, whereas their absence, haplotype TAAG, protects against PD (Table 44).

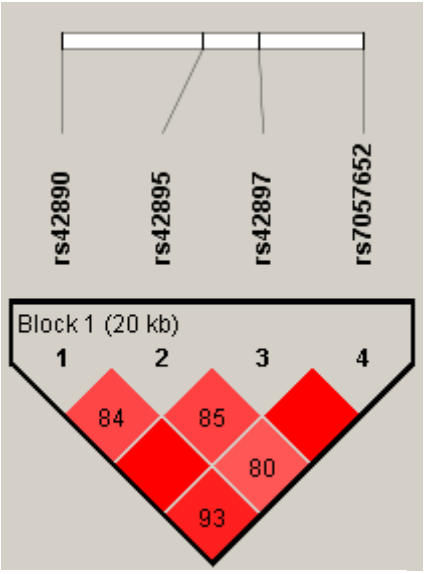


Fig. 55. Linkage disequilibrium map in *LAMP-2A*.

[58] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

Table 44. Haplotypes in *LAMP-2A*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GAAT	0.388	0.461, 0.265	9.191	**0.002**	2.37 [1.31-4.29]
TCAG	0.349	0.347, 0.354	0.012	0.913	0.97 [0.54-1.73]
GAGG	0.091	0.085, 0.100	0.145	0.703	0.84 [0.32-2.19]
TAAG	0.080	0.021, 0.179	19.059	**0.00001**	0.10 [0.02-0.43]
GAAG	0.045	0.031, 0.067	1.682	0.195	0.45 [0.11-1.78]
GCAT	0.029	0.032, 0.024	0.142	0.707	1.34 [0.25-7.31]

The order of the markers at the haplotype is rs42890 – rs42895 – rs42897 – rs7057652.

❖ Modelling genetic risk against PD.

We have tried to find a mathematical model to quantify the contribution of the variants on the risk to develop Parkinson´s disease[59]. To simplify the process, we considered the genotypes, neither the alleles nor the haplotypes, that reported relevant results, i.e. p-value < 0.05 (X^2 test)[60]:

1. Risk of developing Parkinson´s disease (58 controls and 111 PD patients)[61]:

I. rs356219 (*SNCA*): AA vs AG+GG, p=0.028.

II. haplotype (*MAPT*): H1H2+H2H2 vs H1H1, p=0.0001.

III. rs6684770 (*ATP13A2*): CC vs TT+CT, p=0.020.

IV. rs7099185 (*CDNF*): CC vs TT+TC, p=0.005.

[59] Due to the number of individuals analyzed and their demographic and cognitive characteristics, the relevance of this model is compromised.

[60] Again, individuals carrying mutations in *GBA* or *LRRK2* were not considered.

[61] The risk variants are underlined.

V. rs4441646 (*DOCK3*): AA+AC vs CC, p=0.026.

VI. rs1801131 (c.A1298C, *MTHFR*): AA vs AC+CC, p=0.028.

VII. rs1461496 (*HSPA8*): TT+CT vs CC, p=0.021.

VIII. rs7057652 (*LAMP-2A*): TG+GG vs TT, p=0.004.

IX. rs42890 (*LAMP-2A*): GT+TT vs GG, p=0.021.

We employed a binary logistic regression model to quantify the relevance of these nine variables on PD risk.

The following equation explains what a binary logistic regression is:

$$\frac{P}{1-P} = e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4+\cdots}$$

$$= e^{a} \cdot e^{b_1X_1} \cdot e^{b_2X_2} \cdot e^{b_3X_3} \cdot e^{b_4X_4} \cdot \ldots$$

In our case,

- P is the probability that an individual presents PD (to be a case), whereas 1 – P is the probability that an individual is healthy (to be a control).
- The term P / (1-P) is the odds ratio, i.e. the increase or decrease in the risk to develop Parkinson´s disease.
- Each X_i represents a variable (genotype) that influences on the risk (value 1 for the underlined risk variants and 0 for the others) and a, b_1, b_2… are numerical coefficients.

Table 45 shows how many cases and controls were correctly assigned by the model.

Table 45. Classification of the individuals analyzed under the binary logistic regression model.

Observed		**Predicted**		
		health		Percentage Correct
		Control	PD case	
health	Control	32	26	55.2 (**specificity**)[62]
	PD case	18	93	83.8 (**sensitivity**)
Overall Percentage				74.0

The results we obtained were[63] (Table 46):

Table 46. Results of the binary logistic regression model to determine PD risk.

	b	df	p-value	OR (e^b)	$CI_{95\%}$	
					Lower	Upper
haplotype (*MAPT*) $\mathbf{X_1}$	1.402 $\mathbf{b_1}$	1	**0.000**	4.062 $\mathbf{e^{b_1}}$	1.886	8.750
rs6684770 (*ATP13A2*) $\mathbf{X_2}$	1.034 $\mathbf{b_2}$	1	**0.009**	2.812 $\mathbf{e^{b_2}}$	1.295	6.106
rs7099185 (*CDNF*) $\mathbf{X_3}$	21.878 $\mathbf{b_3}$	1	0.999	$3.17*10^9$ $\mathbf{e^{b_3}}$	0.000	∞
rs4441646 (*DOCK3*) $\mathbf{X_4}$	20.090 $\mathbf{b_4}$	1	0.999	$5.31*10^8$ $\mathbf{e^{b_4}}$	0.000	∞
rs1461496 (*HSPA8*) $\mathbf{X_5}$	0.988 $\mathbf{b_5}$	1	**0.011**	2.685 $\mathbf{e^{b_5}}$	1.259	5.724
rs7057652 (*LAMP-2A*) $\mathbf{X_6}$	0.954 $\mathbf{b_6}$	1	**0.023**	2.595 $\mathbf{e^{b_6}}$	1.138	5.918
constant	-23.131 **a**	1	0.999	0.000		

Although we introduced nine variables, I, VI and IX were not included in the model[64]. The model fitted best including variables IV and V, despite that they did not show statistical significance in the model (p-value ≈1)[65].

[62] Sensitivity can be defined as the proportion of PD cases that are correctly assigned by the model, that is, capacity to identify positives. Specificity is defined as the proportion of controls that are correctly assigned by the model, i.e. capacity to identify negatives.

[63] Due to the differences between controls and cases regarding age at onset/age at collection and sex distribution, these variables were not considered.

In all cases, individual OR>1. Nevertheless, these values were different to the previously reported as they were obtained under other mathematical assumptions. For example, an individual that carried the H1H1 genotype in *MAPT* presented 4.062 times ($CI_{95\%}$=[1.886-8.750]), instead of 3.65 ($CI_{95\%}$=[1.86-7.17]), more risk to develop PD than other person that carried the H1H2 or the H2H2.

What would happen if a person presents the risk variants in variables II (X_1), III (X_2), IV (X_3), VII (X_5) and VIII (X_6)?

$$\text{odds ratio} = e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4+b_5X_5+b_6X_6} =$$

$$e^{-23.131+1.402\cdot1+1.034\cdot1+21.878\cdot1+20.090\cdot0+0.988\cdot1+0.954\cdot1} = e^{3.125} =$$

$$22.76$$

$$\frac{P}{1-P} = 22.76;\ \text{P} = \ 0.9579$$

This person would have 22.76 times more possibilities to develop the disease or would develop PD with a probability of 95.79%.

Finally, we analyzed the distribution of the nine variables in cases and controls (Figure 56 and Table 47):

---

[64] These three variables were not included because they did not present statistical significance and, furthermore, did not help to differentiate controls and cases. As they were redundant and not useful, were excluded.

[65] In addition, for their particular distribution, as the CC genotype in rs4441646 (*DOCK3*) is only present in PD cases and only a few controls carry the genotype CC in rs7099185 (*CDNF*), b_3 and b_4 show an extreme value and a as well to compensate it.

Fig. 56. Distribution of the number of risk variants carried by healthy controls and PD patients.

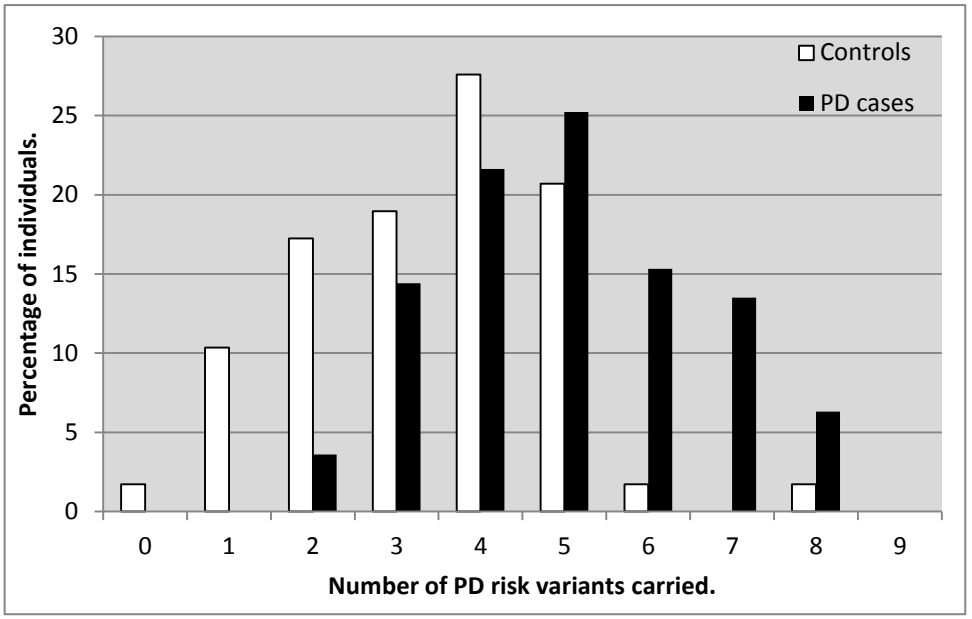


Table 47. Percentage of controls and cases that carry each number of risk variants.
Number of individuals are in parenthesis.

Number of risk variants carried	0	1	2	3	4	5	6	7	8	9
Healthy controls	1.72 (1)	10.34 (6)	17.24 (10)	18.97 (11)	27.59 (16)	20.69 (12)	1.72 (1)	0 (0)	1.72 (1)	0 (0)
PD patients	0 (0)	0 (0)	3.60 (4)	14.41 (16)	21.62 (24)	25.23 (28)	15.32 (17)	13.51 (15)	6.31 (7)	0 (0)

The mean of PD risk variants carried by healthy controls was 3.39, whereas for PD patients was 5. Medians were 4 and 5, respectively.

We also tried to determine if there was some correlation between them as this could shed some light on the biological pathogenic mechanisms. Only rs7057652 and rs42890, both in *LAMP-2A*, showed some (intermediate for controls and high for cases) as expected due to the high linkage disequilibrium observed along the region.

2. subsequently, the frequency of these mutations/polymorphisms was analyzed only in PD patients to determine if any of them could be considered a genetic determinant of cognitive status in Parkinson´s disease. Although cases had been classified into three categories depending on their results after the cognitive assessment (Table 13; see section III. Material and methods, Subjects), considering the demographic similarities and that there were almost no differences in genotype nor in allele distributions (-see **"Below…"** at the final part of this section IV.2.-) between PD patients with cognitive impairment (mild or severe)[66], we grouped them into a unique group named cognitive impairment (C.I.; 61 PD cases, i.e. 36 with M.C.I. and 25 with dementia) that was compared against the individuals with normal cognition but Parkinson´s disease (N.C.; 50 PD cases). Thereby, we increased the power of our analysis.

[66] M.C.I. and dementia do not represent the same process but are somehow connected.

The relevant results, that is, those that reached statistical significance ($p<0.05$) or a strong tendency towards this ($p<0.1$) are listed below. The results obtained for the rest of polymorphisms (p-value higher than 0.1) are detailed in Annex IVb.

Frequencies are indicated in parenthesis. p-values lower than 0.05, i.e. statistically significant results, are in bold and shadowed. p-values lower than 0.1 are highlighted in bold. We calculated odds ratio for the distribution of alleles in such polymorphisms to know their effect on PD risk or protection.

When it was possible, we calculated haplotypes that could increase/decrease the risk to develop cognitive impairment during PD. When appropriate, we calculated sex, but not age-correlated odds ratios for the genotypes by using a binary logistic regression model because the cases from the groups N.C. and C.I. had similar distribution of sex but statistically different mean age at onset.

(1): *APOE*.

Table 48. Frequency of *APOE* genotype when considering the number of ε2 alleles.

	ε2 +/+	ε2 +/-	ε2 -/-	ε2 +	ε2 -
N.C.	1 (2.00)	3 (6.00)	46 (92.00)	4 (8.00)	46 (92.00)
C.I.	0 (0)	12 (19.67)	49 (80.33)	12 (19.67)	49 (80.33)
	X^2 = 5.458; 2df; **p=0.065**			X^2 = 3.034; 1df; **p=0.082** OR: 2.82, $CI_{95\%}$=[0.85-9.37]	

There were not statistically or almost statistically significant results when considering genotypes or alleles (Table AH) or even the number of ε4 alleles (Table AI). Nevertheless, when individuals where classified depending on the number of ε2 alleles they carry, we observed that the ε2 allele was more frequent in C.I. (19.67%) than in N.C. (8.00%) and that there was a trend ($p<0.1$) that pointed out to the involvement of this allele in the increased risk to develop cognitive impairment during the evolution of PD (OR: 2.82; $CI_{95\%}$=[0.85-9.37]). This effect was mainly observed for ε2+/- carriers (ε2+/+ individuals were almost non-existent) (Table 48).

(2): *CDNF* (neurotrophic factor).

Table 49. Genotypic and allelic frequency of rs7094179 (*CDNF*).

Gene and polymorphism	Genotypes			Alleles	
CDNF, rs7094179	GG	GT	TT	G	T
N.C.	25 (50.00)	21 (42.00)	4 (8.00)	71 (71.00)	29 (29.00)
C.I.	21 (34.43)	31 (50.82)	9 (14.75)	73 (59.84)	49 (40.16)
	X^2 = 3.135; 2df; p=0.209			X^2 = 3.005; 1df; **p=0.083** OR: 0.61, $CI_{95\%}$=[0.35-1.07]	

There was only a trend: the G allele in rs7094179 showed a tendency to protect against cognitive impairment in PD patients (p=0.083; OR: 0.61; $CI_{95\%}$=[0.35-1.07]). It was more frequent at the N.C. group (71.00%) than at the C.I. (59.84%). However, there was no effect when considering genotypes (Table 49). When classifying the individuals as GG vs GT+TT, GG carriers showed a tendency to protect against cognitive impairment in PD (p=0.097).

When this polymorphism and rs7099185, also located at the *CDNF* gene, were analyzed together, the haplotype GT (that included the G allele from rs7094179) also showed a tendency to protect against cognitive impairment (p=0.092). It was more frequent at the N.C. group (0.630) than at the C.I. (0.517) (Table 50).

Table 50. Haplotypes in *CDNF*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GT	0.568	0.517, 0.630	2.835	**0.092**	0.63 [0.36-1.11]
TT	0.261	0.294, 0.220	1.559	0.212	1.48 [0.78-2.81]
TC	0.090	0.107, 0.070	0.949	0.330	1.59 [0.59-4.31]
GC	0.081	0.081, 0.080	0.001	0.981	1.01 [0.36-2.80]

The order of the markers at the haplotype is: rs7094179 – rs7099185.

(3): *CBS* (homocysteine metabolism).

Table 51. Genotypic and allelic frequency of c.844ins68 (*CBS*).

Gene and polymorphism	Genotypes			Alleles	
CBS, c.844ins68	ins/ins	ins/-	-/-	ins	-
N.C.	0 (0)	5 (10.00)	45 (90.00)	5 (5.00)	95 (95.00)
C.I.	1 (1.64)	12 (19.67)	48 (78.69)	14 (11.48)	108 (88.52)
	X^2 = 2.918; 2df; p=0.233			X^2 = 2.944; 1df; **p=0.086** OR: 2.46, $CI_{95\%}$=[0.85-7.08]	

The insertion c.844ins68 was more frequent in C.I. (11.48%) than in N.C. (5.00%). However, that was just a trend (p=0.086). The supposed increased risk to develop cognitive impairment for carriers of the insertion (OR: 2.46; $CI_{95\%}$=[0.85-7.08]) was not observed in genotypes (p=0.233) even though ins/ins and ins/- were more frequent in C.I. than in N.C. (Table 51).

(4): *MTR* (homocysteine metabolism).

Table 52. Genotypic and allelic frequency of rs1805087 (*MTR*).

Gene and polymorphism	Genotypes			Alleles	
MTR, rs1805087 (c.A2756G)	AA	AG	GG	A	G
N.C.	29 (58.00)	17 (34.00)	4 (8.00)	75 (75.00)	25 (25.00)
C.I.	50 (81.97)	11 (18.03)	0 (0)	111 (90.98)	11 (9.02)
	X^2 = 9.875; 2df; **p=0.007**			X^2 = 10.333; 1df; **p=0.001** OR: 3.36, $CI_{95\%}$=[1.56-7.24]	

The A allele increased the risk to develop cognitive impairment during PD in 3.36 times ($CI_{95\%}$=[1.56-7.24]). It had a statistically significant effect especially for AA carriers (81.97% C.I. vs 58.00% N.C.) (Table 52).

(5): *NFE2L2* (protection against oxidative stress).

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table AU). Nevertheless, there was a haplotype more frequent in PD patients with cognitive impairment (0.617) than in those with normal cognition (0.502): GG. That was just a trend (p=0.086) (Table 53).

Table 53. Haplotypes in *NFE2L2*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GG	0.565	0.617, 0.502	2.942	**0.086**	1.60 [0.91-2.81]
AA	0.250	0.215, 0.292	1.735	0.188	0.66 [0.35-1.26]
GA	0.142	0.129, 0.158	0.374	0.541	0.79 [0.36-1.75]
AG	0.043	0.039, 0.048	0.108	0.743	0.80 [0.20-3.14]

The order of the markers at the haplotype is: rs1806649 – rs10183914.

(6): *KEAP1* (protection against oxidative stress).

Table 54. Genotypic and allelic frequency of rs8113472 (*KEAP1*).

Gene and polymorphism	Genotypes			Alleles	
KEAP1, rs8113472	GG	GT	TT	G	T
N.C.	47 (94.00)	2 (4.00)	1 (2.00)	96 (96.00)	4 (4.00)
C.I.	49 (80.33)	12 (19.67)	0 (0)	110 (90.16)	12 (9.84)
	X^2 = 7.165; 2df; **p=0.028**			X^2 = 2.799; 1df; **p=0.094** OR: 0.38, $CI_{95\%}$=[0.12-1.22]	

The distribution of alleles between C.I. and N.C. was almost statistically different (p=0.094) and the G allele tended to be protective against mental deterioration (OR: 0.38; $CI_{95\%}$=[0.12-1.22]). Its effect reached statistical significance in genotypes (p=0.028), concretely for the homozygous GG cases,

which were more frequent at the N.C. group (94.00%) than at the C.I. (80.33%) (Table 54).

(7): *HSPA8* (chaperone-mediated autophagy).

Table 55. Genotypic and allelic frequency of rs4936770 (*HSPA8*).

Gene and polymorphism	Genotypes			Alleles	
HSPA8, rs4936770	AA	AG	GG	A	G
N.C.	5 (10.00)	20 (40.00)	25 (50.00)	30 (30.00)	70 (70.00)
C.I.	2 (3.28)	15 (24.59)	44 (72.13)	19 (15.57)	103 (84.43)
	X^2 = 6.203; 2df; **p=0.045**			X^2 = 6.649; 1df; **p=0.010** OR: 0.43, $CI_{95\%}$=[0.22-0.82]	

The distribution of alleles between C.I. and N.C. was statistically different, being the A allele protective against cognitive impairment in PD (OR: 0.43; $CI_{95\%}$=[0.22-0.82]). The distribution of genotypes between both groups was statistically different too (p=0.045) and the protective effect of the A variant was important in homozygous AA carriers (p=0.094) but more in heterozygous AG (p=0.036; OR: 0.407; $CI_{95\%}$=[0.175-0.943]), as results from the binary logistic regression model conclude (Table 56).

Table 56. Results of the binary logistic regression model for rs4936770 (*HSPA8*) genotypes.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
AA vs GG	1	**0.094**	0.230	0.041	1.285
AG vs GG	1	**0.036**	0.407	0.175	0.943

The GG genotype was taken as a reference, i.e. OR:1.

When combining these results and the obtained for the other polymorphism analyzed in *HSPA8*, rs1461496, it was observed that the presence or absence of the A allele, only in the presence of the C allele at rs1461496, determined the protective or risky effect of two haplotypes: CA and CG, respectively (Table 57).

Table 57. Haplotypes in *HSPA8*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
CG	0.517	0.588, 0.431	5.448	**0.020**	1.88 [1.07-3.30]
TG	0.262	0.256, 0.269	0.049	0.826	0.94 [0.50-1.77]
CA	0.167	0.117, 0.229	4.987	**0.026**	0.45 [0.21-0.97]
TA	0.053	0.039, 0.071	1.101	0.294	0.53 [0.15-1.88]

The order of the markers at the haplotype is: rs1461496 – rs4936770.

(8): *LAMP-2A* (chaperone-mediated autophagy).

There were only statistically and almost statistically significant results at the polymorphism rs42897. However, the effect of alleles and genotypes on cognitive impairment was opposite depending on the sex of the cases: in women, the A allele increased the risk to develop cognitive impairment whereas in men it was protective against the development of cognitive impairment on PD (Table 59).

The study of the haplotypes composed by the four markers analyzed along the gene revealed that one, GAAG, was statistically more frequent in patients with normal cognition (0.064) than in those with cognitive impairment (none of them carried it): p=0.018. In this case, the A allele in rs42897 showed a protective effect against cognitive impairment (Table 58). There was an irregular linkage disequilibrium pattern along the region although it was mainly high (Figure 57).

Table 58. Haplotypes in *LAMP-2A*.

Haplotype	Frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GAAT	0.465	0.476, 0.451	0.098	0.755	1.11 [0.64-1.94]
TCAG	0.348	0.360, 0.333	0.123	0.726	1.13 [0.63-2.02]
GAGG	0.085	0.081, 0.091	0.041	0.839	0.88 [0.33-2.37]
GAAG	0.028	0.000, 0.064	5.559	**0.018**	-
GCAT	0.028	0.035, 0.019	0.348	0.555	1.87 [0.31-11.18]
TAAG	0.020	0.035, 0.000	2.248	0.134	-
TAAT	0.013	0.000, 0.030	2.507	0.113	-

The order of the markers at the haplotype is rs42890 – rs42895 – rs42897 – rs7057652. Missing ORs correspond to those haplotypes absent in one of the populations.

Table 59. Genotypic and allelic frequency of rs42897 (*LAMP-2A*).

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs42897	AA	AG	GG	A	G
N.C., overall	44 (88.00)	5 (10.00)	1 (2.00)	60 (90.91)	6 (9.09)
C.I., overall	53 (86.89)	2 (3.28)	6 (9.84)	78 (90.70)	8 (9.30)
	X^2 = 4.648; 2df; **p=0.098** Z_{mfG}^2 = 7.248; 1df; **p=0.007** Z_C^2 = 7.385; 2df; **p=0.025**			$X^2 = Z_A^2$ = 0.002; 1df; p=0.964 Z_{mfA}^2 = 7.012; 1df; **p=0.008** OR: 0.98, $CI_{95\%}$=[0.32-2.98]	
N.C., women	11 (68.75)	5 (31.25)	0 (0)	27 (84.38)	5 (15.63)
C.I., women	23 (92.00)	2 (8.00)	0 (0)	48 (96.00)	2 (4.00)
	X^2 = 3.725; 1df; **p=0.054** Z_{fG}^2 = 3.725; 1df; **p=0.054**			$X^2 = Z_{fA}^2$ = 3.377; 1df; **p=0.066** OR: 4.44, $CI_{95\%}$=[0.81-24.46]	
N.C., men	-	-	-	33 (97.06)	1 (2.94)
C.I., men	-	-	-	30 (83.33)	6 (16.67)
	-			$X^2 = Z_m^2$ = 3.660; 1df; **p=0.056** OR: 0.15, $CI_{95\%}$=[0.02-1.32]	

It is noteworthy that *LAMP-2A* is located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [484][67].

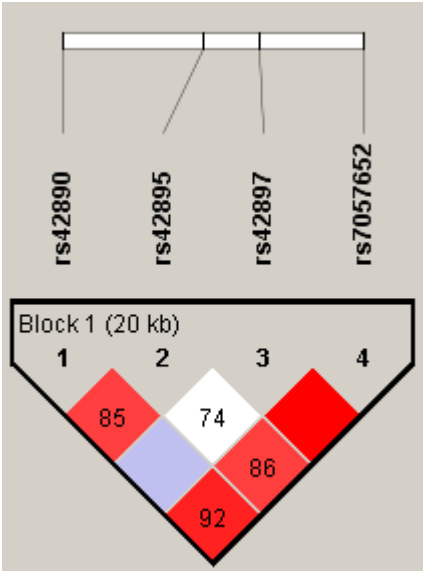


Fig. 57. Linkage disequilibrium map in *LAMP-2A*.

[67] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

❖ **Modelling the genetic risk for cognitive decline in PD.**

We have tried to find a mathematical model to quantify the contribution of the variants on the risk to develop cognitive impairment during the disease[68].

To simplify the process, we considered the genotypes, neither the alleles nor the haplotypes, that reported significant results, i.e. p-value < 0.05 (X^2 test):

2. **Risk of developing cognitive impairment during PD (111 PD patients: 50 normal cognition (N.C.) and 61 cognitive impairment (C.I.))[69]:**
   - X. rs1805087 (c.A2756G, *MTR*): AG+GG vs AA, p=0.006.
   - XI. rs8113472 (*KEAP1*): GG vs GT+TT, p=0.036.
   - XII. rs4936770 (*HSPA8*): AA+AG vs GG, p=0.017.

Again, we employed a binary logistic regression model to quantify the influence of these three variables on the risk to develop cognitive impairment during PD.

The following equation explains what a binary logistic regression is:

[68] Due to the number of individuals analyzed and their demographic and cognitive characteristics, the relevance of this model is compromised.
[69] The risk variants are underlined.

$$\frac{P}{1-P} = e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4+\cdots}$$

$$= e^{a} \cdot e^{b_1X_1} \cdot e^{b_2X_2} \cdot e^{b_3X_3} \cdot e^{b_4X_4} \cdot \ldots$$

In our case,

- P is the probability that an individual presents cognitive impairment (to belong to the C.I. group), whereas 1-P is the probability that an individual is cognitively normal (to be part of the N.C. group).
- The term P / (1-P) is the odds ratio, i.e. the increase or decrease in the risk to develop cognitive impairment during PD.
- Each X_i represents a variable (genotype) that influences on the risk (value 1 for the underlined risk variants and 0 for the others) and a, b_1, b_2… are numerical coefficients.

Table 60 shows how many subjects were correctly assigned by the model.

Table 60. Classification of the subjects analyzed under the binary logistic regression model.

Observed		**Predicted**		
		cognitive status		Percentage Correct
		N.C.	C.I.	
cognitive status	N.C.	35	15	70.0 (**specificity**)[70]
	C.I.	18	43	70.5 (**sensitivity**)
Overall Percentage				70.3

[70] Sensitivity can be defined as the proportion of C.I. that are correctly assigned by the model, that is, capacity to identify positives. Specificity is defined as the proportion of N.C. that are correctly assigned by the model, i.e. capacity to identify negatives.

The results we obtained were[71] (Table 61):

Table 61. Results of the binary logistic regression model to determine the risk of developing cognitive impairment.

	b	df	p-value	OR (e^b)	$CI_{95\%}$	
					Lower	Upper
rs1805087 (c.A2756G, *MTR*) $\mathbf{X_1}$	1.141 $\mathbf{b_1}$	1	**0.013**	3.131 $\mathbf{e^{b1}}$	1.271	7.715
rs8113472 (*KEAP1*) $\mathbf{X_2}$	1.382 $\mathbf{b_2}$	1	0.051	3.985 $\mathbf{e^{b2}}$	0.996	15.942
rs4936770 (*HSPA8*) $\mathbf{X_3}$	0.989 $\mathbf{b_3}$	1	**0.021**	2.689 $\mathbf{e^{b3}}$	1.163	6.220
constant	-1.384 **a**	1	0.004	0.250		

In all cases, individual OR > 1. That means that all the variables increased the risk to develop cognitive impairment during Parkinson´s disease, although only for variables X and XII it reached statistical significance. For example, an individual that carried the AA genotype in *MTR* presented 3.131 times ($CI_{95\%}$=[1.271-7.715]) more risk to show an altered cognitive status than other person that carried the AG or the GG genotypes.

What would happen if a person carries the risk variants in variables X (X_1), XI (X_2) and XII (X_3)?

[71] The mean age at onset when comparing N.C. and C.I. was statistically different. Therefore, it was not considered. The percentage of males and females was similar between groups and it was considered in the model but the model did not included it because sex did not reach statistical significance.
It is noteworthy that the number of individuals in both groups was small and that could be considered a limitation for the calculations.

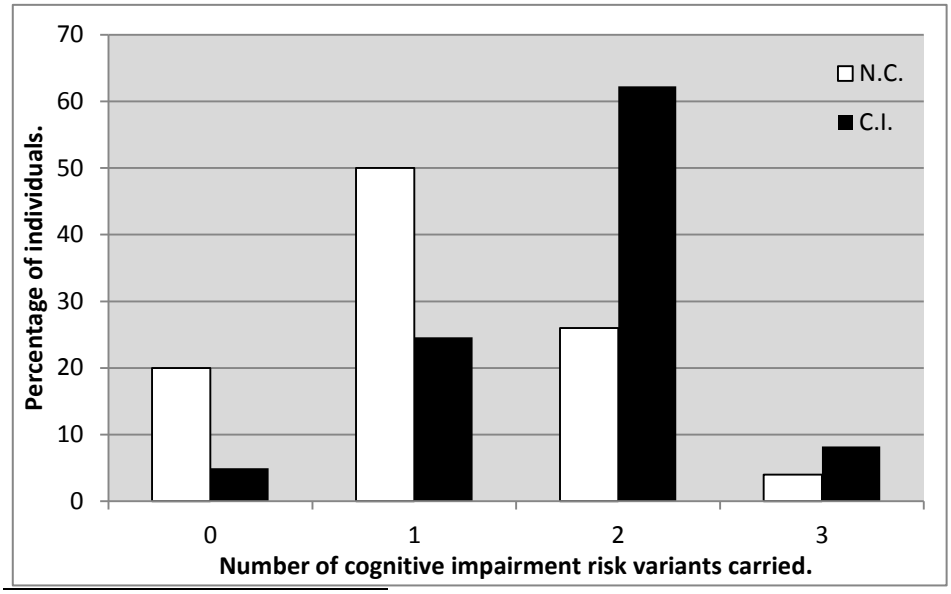
$$\text{odds ratio} = e^{a+b_1X_1+b_2X_2+b_3X_3} = e^{-1.384+1.141\cdot1+1.382\cdot1+0.989\cdot1} = e^{2.128} = 8.40$$

$$\frac{P}{1-P} = 8.40;\ \text{P} = 0.8936$$

This person would have 8.40 times more possibilities to develop cognitive impairment during PD or would develop it with a probability of 89.36%.

Finally, we analyzed the distribution of the three variables in patients with normal cognition (N.C.) and with cognitive impairment (C.I.)[72] (Figure 58 and Table 62):

Fig. 58. Distribution of the number of risk variants carried by N.C. and C.I. groups.



[72] We also tried to determine if there was some correlation between them, but we observed that there was not any. There was not correlation between these three and the previous nine either.

Table 62. Percentage of N.C. and C.I. subjects that carry each number of risk variants.
Number of individuals are in parenthesis.

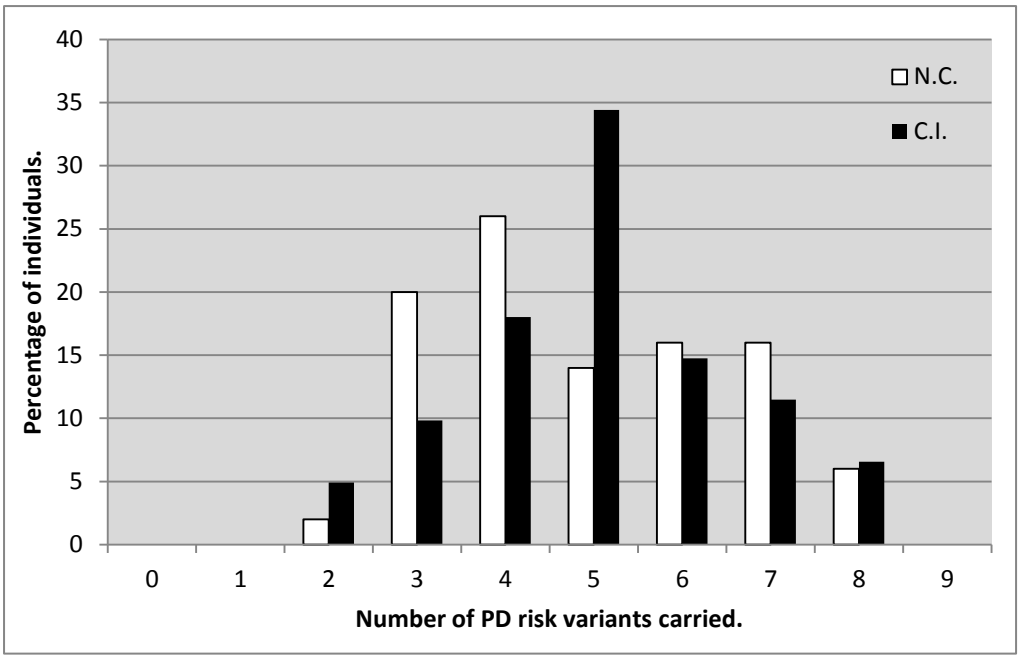
Number of cognitive impairment risk variants carried	0	1	2	3
N.C.	20.00 (10)	50.00 (25)	26.00 (13)	4.00 (2)
C.I.	4.92 (3)	24.59 (15)	62.30 (38)	8.20 (5)

Furthermore, we studied how many variants affecting PD presented individuals of both groups (Table 63 and Figure 59).

Table 63. Percentage of N.C. and C.I. cases that carry each number of PD risk variants.
Number of individuals are in parenthesis.

Number of PD risk variants carried	0	1	2	3	4	5	6	7	8	9
N.C.	0 (0)	0 (0)	2.00 (1)	20.00 (10)	26.00 (13)	14.00 (7)	16.00 (8)	16.00 (8)	6.00 (3)	0 (0)
C.I.	0 (0)	0 (0)	4.92 (3)	9.84 (6)	18.03 (11)	34.43 (21)	14.75 (9)	11.48 (7)	6.56 (4)	0 (0)

Fig. 59. Distribution of the number of PD risk variants carried by N.C. and C.I. groups.



The mean of cognitive impairment risk variants carried by the cases with normal cognition was 1.14, whereas for those with cognitive impairment was 1.74. Medians were 1 and 2, respectively.

The mean of PD risk variants carried by the subjects from the N.C. group was 4.94, whereas for C.I. was 5.28. Medians were 5 and 5, respectively.

Below are the relevant results (p<0.1) we obtained when comparing PD patients with mild cognitive impairment to those that had developed dementia (PDD). Frequencies are in parenthesis[73].

(1): The A allele at rs6426554 in *PSEN2* was more frequently found in individuals with M.C.I. than in PD cases suffering from dementia although without reaching statistical significance (p=0.075) pointing out to a protective tendency against dementia (Table 64). This trend reached statistical significance (p=0.039) when classifying the cases in AA and AG+GG, being the AA genotype protective against dementia (OR:0.33; $CI_{95\%}$=[0.11-0.95]).

Table 64. Genotypic and allelic frequency of rs6426554 (*PSEN2*).

Gene and polymorphism	**Genotypes**			**Alleles**	
PSEN2, rs6426554	AA	AG	GG	A	G
M.C.I.	24 (66.67)	11 (30.56)	1 (2.78)	59 (81.94)	13 (18.06)
Dementia	10 (40.00)	14 (56.00)	1 (4.00)	34 (68.00)	16 (32.00)
	X^2 = 4.280; 2df; p=0.118			X^2 = 3.167; 1df; **p=0.075** OR: 0.47, $CI_{95\%}$=[0.20-1.09]	

[73] A binary logistic regression model and some descriptive graphics and tables comparing the groups M.C.I. and PDD are detailed in Annex IVc.

(2): The T allele in rs2986017, *CALHM1*, increased the risk to develop dementia in PD in 2.81 times ($CI_{95\%}$=[1.07-7.41]). Although there was no statistically significant difference in the frequency of genotypes (p=0.124) (Table 65). When they were grouped depending on the presence of the T allele, i.e. TT+TC vs CC, there was a statistical significance (p=0.042) pointing out that the T allele increased the risk to develop dementia (OR:3.33; $CI_{95\%}$=[1.02-10.91]).

Table 65. Genotypic and allelic frequency of rs2986017 (*CALHM1*).

Gene and polymorphism	**Genotypes**			**Alleles**	
CALHM1, rs2986017 (p.P86L)	TT	TC	CC	T	C
M.C.I.	2 (5.56)	4 (11.11)	30 (83.33)	8 (11.11)	64 (88.89)
Dementia	3 (12.00)	7 (28.00)	15 (60.00)	13 (26.00)	37 (74.00)
	X^2 = 4.170; 2df; p=0.124			X^2 = 4.590; 1df; **p=0.032** OR: 2.81, $CI_{95\%}$=[1.07-7.41]	

(3): In the gene *HTT*, the ATCAT haplotype was protective against dementia in PD: it was significantly more present in cases with M.C.I. than in those with dementia (Table 66, Figure 60).

Table 66. Haplotypes in *HTT*.

Haplotype	Overall frequency	Frequencies: Dementia, M.C.I.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GTTAT	0.376	0.439, 0.331	1.470	0.225	1.58 [0.89-2.81]
ATCAT	0.193	0.092, 0.263	5.502	**0.019**	0.28 [0.12-0.63]
ATCGC	0.189	0.200, 0.181	0.073	0.787	1.13 [0.56-2.29]
AGCGC	0.141	0.144, 0.138	0.007	0.932	1.05 [0.47-2.33]
ATTGT	0.026	0.003, 0.042	1.777	0.183	0.07 [0.00-2.87]
ATTAT	0.017	0.040, 0.001	2.654	0.103	-
GTTGT	0.015	0.017, 0.014	0.020	0.889	1.22 [0.13-11.62]
GGCGC	0.015	0.016, 0.014	0.006	0.938	1.15 [0.12-11.30]
GTCAT	0.012	0.028, 0.002	1.608	0.205	-

The order of the markers at the haplotype is: rs10015979 – rs363066 – rs363096 – rs2298969 – rs110501.

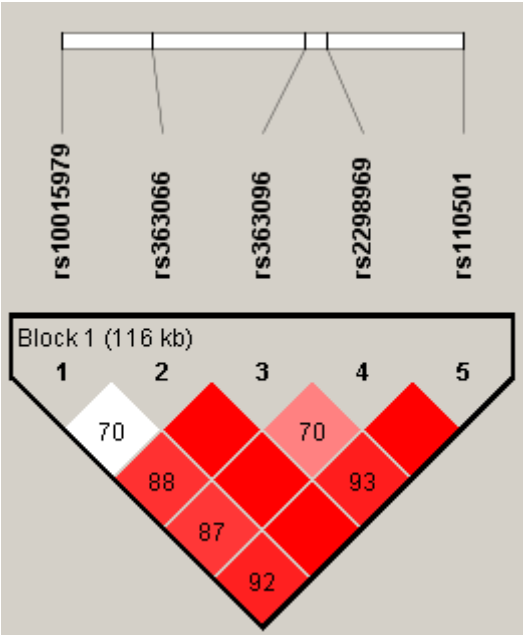


Fig. 60. Linkage disequilibrium map in *HTT*.

(4): The distribution of alleles in SNP rs2871776 in *ATP13A2*, was very similar between cases with M.C.I. and those with dementia. However, the difference in genotypes was almost statistically significant (p=0.053) with a higher frequency of heterozygous individuals in M.C.I. cases when compared to those with dementia (Table 67).

Table 67. Genotypic and allelic frequency of rs2871776 (*ATP13A2*)

Gene and polymorphism	Genotypes			Alleles	
ATP13A2, rs2871776	AA	AG	GG	A	G
M.C.I.	7 (19.44)	12 (33.33)	17 (47.22)	26 (36.11)	46 (63.89)
Dementia	9 (36.00)	2 (8.00)	14 (56.00)	20 (40.00)	30 (60.00)
	X^2 = 5.891; 2df; **p=0.053**			X^2 = 0.190; 1df; p=0.663	

(5): In *NFE2L2*, there was a protective haplotype, GA, that was significantly more present in cases with M.C.I. than in those with dementia (Table 68).

Table 68. Haplotypes in *NFE2L2*.

Haplotype	Overall frequency	Frequencies: Dementia, M.C.I.	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GG	0.618	0.673, 0.579	1.093	0.296	1.50 [0.84-2.67]
AA	0.216	0.233, 0.204	0.142	0.706	1.19 [0.61-2.33]
GA	0.128	0.047, 0.184	4.989	**0.026**	0.22 [0.08-0.63]
AG	0.038	0.047, 0.031	0.190	0.663	1.54 [0.36-6.64]

The order of the markers at the haplotype is: rs1806649 – rs10183914.

IV.3. Analysis of genetic variability in the *ARMCX* gene family in PD.

The *ARMCX* gene family is comprised by 6 genes. We analyzed polymorphisms located along them:

- *ARMCX1*: rs6616255 and rs1044275;
- *ARMCX2*: rs5951282;
- *ARMCX3*: rs6995;
- *ARMCX4*: rs2179670 and rs6523506;
- *ARMCX5*: rs2235827;
- *ARMCX6*: rs2858162[74].

The individuals analyzed were described in Table 11 (see section III. Material and methods, Subjects).

As these individuals were selected from those used for the analysis of genetic susceptibility factors in Spanish populations (see Results IV.1), we also had, for them, genetic information about the *APOE* genotype, the poly-T homopolymer length (rs10524523, *TOMM40*), the H1/H2 haplotype (*MAPT*), mutations p.N370S and p.L444P (*GBA*), mutation c.864+246C>T (*NR4A2*), polymorphism p.S18Y (rs5030732, *UCHL1*), and rs1801968 (p.D216H) and p.delE302/303 (*TOR1A*).

[74] The criteria to select those polymorphisms is explained in Annex I.

First, we determined the presence of mutations in *GBA* in healthy controls and PD cases and found two heterozygous carriers of the p.N370S and one heterozygous carrier of the p.L444P mutation in the PD group (no control carried the mutations). There were not significant differences between groups[75].

The individuals that presented the mutations were eliminated from the study on genetic susceptibility factors in PD as their disease was most likely caused by those same mutations.

The relevant results, that is, those that reached statistical significance ($p<0.05$) or a strong tendency towards this ($p<0.1$) are listed below. The results obtained for the rest of polymorphisms (p-value higher than 0.1) are detailed in Annex V.

Frequencies are indicated in parenthesis. p-values lower than 0.05, i.e. statistically significant results, are in bold and shadowed. p-values lower than 0.1 are highlighted in bold. We calculated odds ratio for the distribution of alleles in such polymorphisms to know their effect on PD risk or protection.

In addition, we show the sex and age-correlated odds ratios calculated by using a binary logistic regression model that estimated the effect of their genotypes and demographic factors on the probability to develop PD.

[75] Fisher´s exact test: p=0.497 (p.N370S); p=1 (p.L444P); p=0.246 (p.N370S+p.L444P).

It is noteworthy that the *ARMCX* genes are located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [484][76].

(1): *APOE*.

Table 69. Genotypic and allelic frequency of *APOE* genotype.

	Genotypes						**Alleles**		
APOE, genotype	ε2ε2	ε2ε3	ε3ε3	ε3ε4	ε4ε4	ε2ε4	ε2	ε3	ε4
Controls	0 (0)	7 (7.37)	65 (68.42)	22 (23.16)	0 (0)	1 (1.05)	8 (4.21)	159 (83.68)	23 (12.11)
PD cases	3 (3.26)	8 (8.70)	69 (75.00)	10 (10.87)	0 (0)	2 (2.17)	16 (8.70)	156 (84.78)	12 (6.52)
	X^2 = 7.973; 4df; **p=0.093**						X^2 = 6.058; 2df; **p=0.048**		

The distribution of alleles between healthy controls and PD cases was statistically different (p=0.048). However, that was not the case for genotypes (p=0.093) (Table 69). When considering the number of ε4 alleles, it was observed that ε4 +/- carriers (there were not ε4ε4 carriers) were almost

[76] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

protected against PD (p=0.050; OR: 0.47; $CI_{95\%}$=[0.22-1.01]) (Table 70). Under the binary logistic regression model, the protective effect reached statistical significance (p=0.044; OR:0.428; $CI_{95\%}$=[0.188-0.976]) only for ε3ε4 carriers (Table 71).

Table 70. Frequency of *APOE* genotype when considering the number of ε4 alleles.

	ε4 +/+	ε4 +/-	ε4 -/-
Controls	0 (0)	23 (24.21)	72 (75.79)
PD cases	0 (0)	12 (13.04)	80 (86.96)
	X^2 = 3.831; 1df; **p=0.050** OR: 0.47, $CI_{95\%}$=[0.22-1.01]		

Table 71. Results of the binary logistic regression model for *APOE* genotypes.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
ε2ε3 vs ε3ε3	1	0.891	1.078	0.369	3.150
ε2ε4 vs ε3ε3	1	0.606	1.893	0.167	21.480
ε3ε4 vs ε3ε3	1	**0.044**	0.428	0.188	0.976

The ε3ε3 genotype was taken as a reference, i.e. OR:1. As there were not individuals in both groups carrying genotypes ε2ε2 and ε4ε4, these calculations could not been conducted for them.

(2): *MAPT*.

The distribution of alleles between healthy controls and PD cases was strongly statistically different (p=0.005), being H1 haplotype an important risk factor to develop Parkinson´s disease (OR: 1.99; $CI_{95\%}$=[1.23-3.22]). Furthermore, this effect was also observed when considering genotypes:

there was a statistically significant difference between both groups (p=0.029) (Table 72) and carriers of the H1H1 genotype had around 4 times more risk to develop PD than those that carried the H2H2 genotype. On its turn, heterozygous carriers seemed to have higher risk than the H2 homoygous individuals, but the result did not reach statistical significance (Table 73).

Table 72. Genotypic and allelic frequency of H1/H2 haplotype (*MAPT*).

Gene and polymorphism	Genotypes			Alleles	
MAPT, H1/H2 haplotype	H1H1	H1H2	H2H2	H1	H2
Controls	48 (50.53)	35 (36.84)	12 (12.63)	131 (68.95)	59 (31.05)
PD cases	62 (67.39)	26 (28.26)	4 (4.35)	150 (81.52)	34 (18.48)
	X^2 = 7.063; 2df; **p=0.029**			X^2 = 7.991; 1df; **p=0.005** OR: 1.99, $CI_{95\%}$=[1.23-3.22]	

Table 73. Results of the binary logistic regression model for the H1/H2 haplotype.

	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
H1H1 vs H2H2	1	**0.025**	3.922	1.184	12.991
H1H2 vs H2H2	1	0.200	2.260	0.650	7.859

The H2H2 genotype was taken as a reference, i.e. OR:1.

(3): *ARMCX6*.

The only *ARMCX* gene that showed some relevant result was *ARMCX6* where the distribution of alleles and genotypes between healthy controls and

PD cases were close to be statistically different (p≈0.070) (Table 74). However, there was no significant result under the binary logistic regression model, i.e. there was no specific risk or protective genotype (Table 75). Moreover, sex seemed to have no influence.

Table 74. Genotypic and allelic frequency of rs2858162 (*ARMCX6*).

Gene and polymorphism	**Genotypes**			**Alleles**	
ARMCX6, rs2858162	CC	CT	TT	C	T
Controls, overall	10 (10.53)	12 (12.63)	73 (76.84)	23 (17.29)	110 (82.71)
PD cases, overall	16 (17.39)	16 (17.39)	60 (65.22)	34 (26.56)	94 (73.44)
	X^2 = 3.179; 2df; p=0.204 Z_{mfG}^2 = 3.338; 1df; **p=0.068** Z_C^2 = 3.443; 2df; p=0.179			$X^2 = Z_A^2$ = 3.283; 1df; **p=0.070** Z_{mfA}^2 = 3.307; 1df; **p=0.069** OR: 1.73, $CI_{95\%}$=[0.95-3.14]	
Controls, women	1 (2.63)	12 (31.58)	25 (65.79)	14 (18.42)	62 (81.58)
PD cases, women	2 (5.56)	16 (44.44)	18 (50)	20 (27.78)	52 (72.22)
	X^2 = 1.992; 2df; p=0.369 Z_{fG}^2 = 1.965; 1df; p=0.161			$X^2 = Z_{fA}^2$ = 1.829; 1df; p=0.176	
Controls, men	-	-	-	9 (15.79)	48 (84.21)
PD cases, men	-	-	-	14 (25.00)	42 (75.00)
	-			$X^2 = Z_m^2$ = 1.478; 1df; p=0.224	

Table 75. Results of the binary logistic regression model for rs2858162 (*ARMCX6*).

ARMCX6	df	p-value	OR	$CI_{95\%}$	
				Lower	Upper
rs2858162 CC vs TT	1	0.325	1.597	0.629	4.058
rs2858162 CT vs TT	1	0.309	1.690	0.615	4.646

The TT genotype was taken as a reference, i.e. OR:1.

As members of the *ARMCX* gene family are relatively close at the X chromosome, we decided to analyze the relevance of haplotypes on PD pathogenesis (Figure 61).

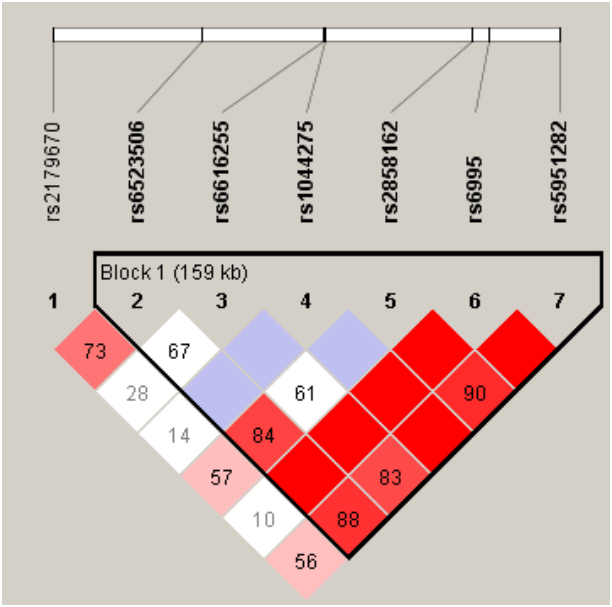


Fig. 61. Linkage disequilibrium map.
The polymorphisms analyzed in *ARMCX* gene family are ordered by their location at chromosome X. rs2235827 (*ARMCX5*) is located ≈943kb downstream rs5951282 and shows no LD with any of the other polymorphisms. Therefore, it is not present. The intensity of the red color is directly correlated to the strength of the linkage disequilibrium between the markers.

Based on the pattern observed, we considered that linkage disequilibrium was strong in one block that included six polymorphisms (showed in bold at Figure 61) and, thus, we analyzed whether its frequency was different between controls and cases and could be considered a PD risk or protective factor (Table 76):

Table 76. Haplotypes in *ARMCX* gene family.

Haplotype	Overall frequency	Frequencies: Case,Control	Chi square (X^2)	p-value	OR [$CI_{95\%}$]
GACTAC	0.260	0.248, 0.271	0.171	0.679	0.89 [0.47-1.68]
GACTAT	0.178	0.150, 0.205	1.371	0.247	0.68 [0.33-1.42]
TACCAC	0.176	0.218, 0.135	3.120	**0.077**	1.79 [0.85-3.77]
GGCTAT	0.134	0.123, 0.145	0.269	0.604	0.83 [0.37-1.88]
GATTGC	0.092	0.094, 0.090	0.010	0.922	1.05 [0.40-2.74]
GACTGC	0.080	0.094, 0.068	0.600	0.439	1.42 [0.51-3.97]
GACCAC	0.027	0.032, 0.023	0.190	0.663	1.40 [0.25-7.80]
TACTAC	0.018	0.000, 0.035	4.490	**0.034**	-
GGCTAC	0.013	0.017, 0.010	0.254	0.614	1.71 [0.14-20.54]

The order of the markers at the haplotype is: rs6523506(*ARMCX4*) - rs6616255(*ARMCX1*) - rs1044275(*ARMCX1*) - rs2858162(*ARMCX6*) - rs6995(*ARMCX3*) - rs5951282(*ARMCX2*).

Missing ORs correspond to those haplotypes absent in one of the populations.

The TAC**C**AC haplotype appeared more frequently in PD cases (21.8%) than in healthy controls (13.5%) although without reaching statistical significance, so it could not be considered a risk haplotype. However, the TAC**T**AC haplotype only appeared in healthy controls (3.5%) and the statistical analysis pointed out to a protective effect against PD pathogenesis. Both just differ in the fourth position, that corresponds to rs2858162 (*ARMCX6*).

Epigenetics in Parkinson´s disease:

IV.4. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

There are 5 genes responsible for the familial forms of PD: *SNCA*, *PRKN*, *PINK1*, *DJ-1* and *LRRK2*. We wanted to study if their influence on pathogenesis was not only genetic. For that purpose, we analyzed the levels of DNA methylation around their transcription start site where variations in this epigenetic mark could influence their expression.

Initially, we conducted a pilot trial in blood, frontal cortex, occipital cortex, hippocampus and *substantia nigra* from one individual from Hospital 12 de Octubre, Madrid, to test the methodological process (as described in section III.2. Epigenetics in PD), to better know the range of values we would obtain lately and to explore the concordance in methylation levels between different areas of the same tissue as well as with other tissues.

As expected for housekeeping genes, the promoters[77] of *DJ-1*, *LRRK2*, *PINK1*, *PRKN* and *SNCA* were poorly methylated in all the tissues to allow their ubiquitous expression.

[77] There was only one CpG island in each gene, located around the TSS, as expected for CG rich genes.

Then, we carried out a second trial with blood from 5 randomly selected male-female pairs of Parkinson´s disease patients, age and period of evolution-matched, from the Region of Valencia (they were described in section III.Material and methods, Subjects).

As previously observed for the individual from the initial trial, there were low methylation levels in blood[78].

Finally, we conducted our study in *substantia nigra*, parietal cortex and occipital cortex from the individuals described in Table 14 (see section III.Material and methods, Subjects).

We compared the levels of DNA methylation in *substantia nigra*, parietal cortex and occipital cortex between healthy controls and PD cases, with special interest for the results in SN because, unlike parietal cortex and occipital cortex, that have not been described as affected by PD pathogenesis and could be considered “control brain regions”, it is extensevely affected in PD [17].

Furthermore, we did not compare the values between brain regions because, as [448] concluded, the DNA methylation pattern correlates much more strongly within a brain region across individuals than within an individual across brain regions.

[78] The results are shown in Annex VI.

The results are shown in Annex VI. Per each assay and brain region, there is a table and a graphic with the mean methylation percentages for each group, healthy controls and PD cases, per position and overall. The number in parenthesis at the tables, as well as the error bars at the graphics, represent the standard deviation.

Statistically significant results, i.e. p-values lower than 0.05, are highlighted in bold and shadowed at the tables and indicated with an ∗ at the graphics. In addition, p-values lower than 0.1 are highlighted in bold and marked with a ○.

Due to the characteristics of our study, i.e. low number of individuals analyzed, low levels of DNA methylation observed, small differences and high standard deviation values, our results could be considered as trends that should be replicated in a larger study.

In addition, as observed in the two previous trials, the promoters of *DJ-1*, *LRRK2*, *PINK1*, *PRKN* and *SNCA* were poorly methylated in all brain regions to allow their ubiquitous expression and there was not an overall tendency in differences of DNA methylation levels between controls and cases.

Thus, we will focus on the most relevant results (Tables 77 and 78):

Table 77. Results with 0.05<p<0.1.

Assay	Target[79]	Brain region	Position	p-value	Methylation percentage
DJ-1 2	Exon 1 •	Occipital cortex	2	0.095	Healthy controls > PD cases
DJ-1 2 **(1)**	Exon 1 •	*Substantia nigra*	3	0.063	Healthy controls > PD cases
LRRK2 2	Exon 1	Parietal cortex	9	0.095	Healthy controls < PD cases
PRKN 2 **(3b)**	Intron 1 (• partially overlaps)	Occipital cortex	6	0.056	Healthy controls > PD cases
PRKN 2 **(3b)**	Intron 1 (• partially overlaps)	*Substantia nigra*	2	0.063	Healthy controls > PD cases
PRKN 2 **(3b)**	Intron 1 (• partially overlaps)	*Substantia nigra*	3	0.063	Healthy controls > PD cases
PRKN 2	Intron 1 (• partially overlaps)	*Substantia nigra*	Overall	0.063	Healthy controls > PD cases
SNCA 1	"Promoter"	Parietal cortex	Overall	0.095	Healthy controls < PD cases
SNCA 2	Exon 1[80] (• partially overlaps)	Occipital cortex	3	0.095	Healthy controls > PD cases
SNCA 2	Exon 1[81]	Occipital	Overall	0.095	Healthy controls >

[79] "Promoter" indicates that our trial predictions located it in this area. However, in some assays, the predicted promoter overlapped with exons or introns (marked with a •).
[80] In isoform NM_000345.
[81] In isoform NM_000345.

	(• partially overlaps)	cortex			PD cases
SNCA 3	Exon 1[82]	Occipital cortex	1	0.095	Healthy controls < PD cases

Table 78. Results with p<0.05.

Assay	Target[83]	Brain region	Position	p-value	Methylation percentage
PINK1 2 **(2)**	Exon 1 •	*Substantia nigra*	2	0.016	Healthy controls > PD cases
PRKN 1 **(3a)**	"Promoter"	*Substantia nigra*	3	0.016	Healthy controls > PD cases
PRKN 2 **(3b)**	Intron 1 (• partially overlaps)	Parietal cortex	6	0.032	Healthy controls > PD cases
PRKN 2 **(3b)**	Intron 1 (• partially overlaps)	Occipital cortex	2	0.032	Healthy controls > PD cases
SNCA 1 **(4a)**	"Promoter"	Parietal cortex	3	0.008	Healthy controls < PD cases
SNCA 1 **(4a)**	"Promoter"	Parietal cortex	6	0.016	Healthy controls < PD cases
SNCA 1 **(4a)**	"Promoter"	*Substantia nigra*	6	0.016	Healthy controls > PD cases

[82] In isoforms NM_007308 and NM_001146054.

[83] "Promoter" indicates that our trial predictions located it in this area. However, in some assays, the predicted promoter overlapped with exons or introns (marked with a •).

SNCA 2 **(4b)**	Exon 1[84] (• partially overlaps)	Occipital cortex	2	0.032	Healthy controls > PD cases
SNCA 2 **(4b)**	Exon 1[85] (• partially overlaps)	Occipital cortex	7	0.008	Healthy controls > PD cases
SNCA 2 **(4b)**	Exon 1[86] (• partially overlaps)	*Substantia nigra*	8	0.016	Healthy controls > PD cases

We predicted, *in silico*, for the positions with statistically significant differences between PD cases and healthy controls and, moreover, for the almost statistically significant positions related to them or located in *substantia nigra*, that where all located at the "promoter", if they where transcription factor binding sites. This could correlate the differences in DNA methylation with possible pathogenic differences in transcription and, thus, in protein expression.

We uploaded fragments of ≈100 nucleotides centered in our target position in each case to the following prediction programs, selecting vertebrates and/or human when it was possible:

- TFSEARCH: **http://www.cbrc.jp/research/db/TFSEARCH.html** [485]
- JASPAR CORE: **http://jaspar.genereg.net/** [486]
- AliBaBa 2.1 and PATCH: **http://www.gene-regulation.com/pub/programs.html**

[84] In isoform NM_000345.
[85] In isoform NM_000345.
[86] In isoform NM_000345.

Per each transcription factor, we obtained information from UniProt (**http://www.uniprot.org/**): the [X] represents its UniProt ID. They are more deeply analyzed in the Discussion.

Below are the results we obtained[87].

(1) *DJ-1* (NM_007262); assay 2.

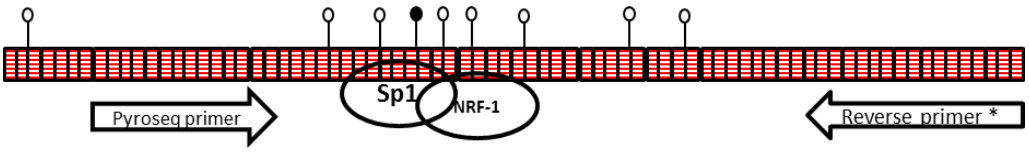


Fig. 62. Schematic representation of the DNA region from -1126 to -1048 in *DJ-1* (NM_007262).
The DNA sequence is striped red because it was located in the non-coding exon 1 and, moreover, the predicted promoter overlapped with it. There was an almost statistical difference (p=0.063) for DNA methylation levels between healthy controls and PD cases in position 3 (black lollipop, -1094) in *substantia nigra*.
The * at the reverse primer indicates that it was biotinylated in 5´.

Sp1 [P08047] binds to GGACGG**C**GCG whereas NRF-1 [Q16656] binds to **C**GCGCGTGCG.[88]

[87] In all figures, lollipops represent cytosines in CpG dinucleotides.
For the numeration, the +1 was assigned to the A from the first codon translated, i.e. ATG. Therefore, negative positions are located upstream to it, whereas positive positions are downstream.
These representations are derived from those present in section III.2. Epigenetics in PD, CpG island prediction.
[88] **C** corresponds to the black lollipop.

(2) *PINK1* (NM_032409); assay 2.

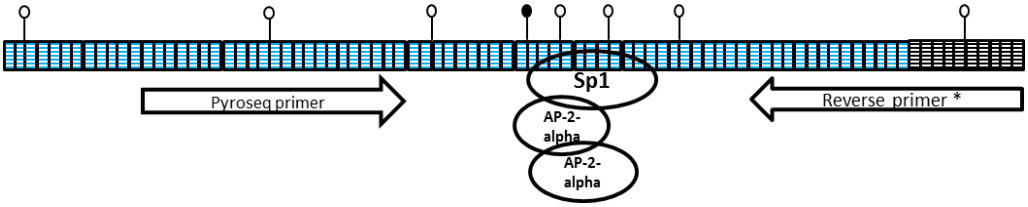


Fig. 63. Schematic representation of the DNA region from +310 to +398 in *PINK1* (NM_032409). The DNA sequence is striped because the predicted promoter overlapped with it. In addition, it is coloured blue when represents the coding part of exon 1 and black for intron 1. There was a statistical difference (p=0.016) for DNA methylation levels between healthy controls and PD cases in position 2 (black lollipop, +355) in *substantia nigra*.
The * at the reverse primer indicates that it was biotinylated in 5´.

Sp1 [P08047] binds to **C**GGCGGGCGGTC whereas AP-2-alpha [P05549] binds to C**C**GGCGGGC or **C**GGCGGGCGG, depending on the prediction program.

(3) *PRKN* (NM_013988);

(3a) assay 1.

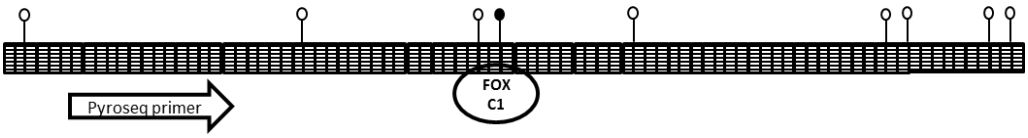


Fig. 64. Schematic representation of the DNA region from -230 to -142 in *PRKN* (NM_013988). The DNA sequence is striped black because it was located 5´ upstream the transcription start site and, moreover, the predicted promoter overlapped with it. There was statistical difference (p=0.016) for DNA methylation levels between healthy controls and PD cases in position 3 (black lollipop, -187) in *substantia nigra*.

FOXC1 [Q12948] binds to AA**C**GCGTA (- strand).

(3b) assay 2.

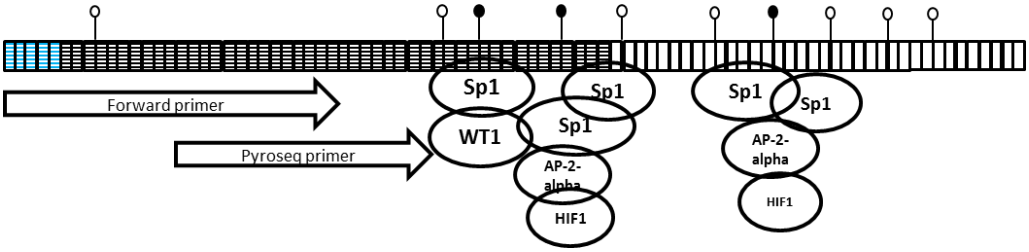


Fig. 65. Schematic representation of the DNA region from +3 to +91 in *PRKN* (NM_013988).
The DNA sequence is striped where the promoter was predicted to be located. In addition, it is blue for the coding part of exon 1 and black (or white) for intron 1. When comparing DNA methylation levels between healthy controls and PD cases, there were relevant results for positions 2 (in SN and occipital cortex), 3 (in SN) and 6 (in occipital and parietal cortex) (black lollipops, +44, +51 and +69, respectively).

- position 2: Sp1 [P08047] and WT1 [P19544] bind to GCGC**C**GCCCC.
- position 3: Sp1 [P08047] binds to **C**GCCCCGGT and CCCA**C**GCCCCG; AP-2-alpha [P05549] binds to CCCA**C**GCCC; HIF1 [Q16665] binds to GGG**C**GTGG (- strand).
- position 6: Sp1 [P08047] binds to **C**GCCCCGCC and GGCGCCA**C**GC; AP-2-alpha [P05549] binds to GCCA**C**GCCC; HIF1 [Q16665] binds to GGG**C**GTGG (- strand).

(4) *SNCA*;

(4a) assay 1.

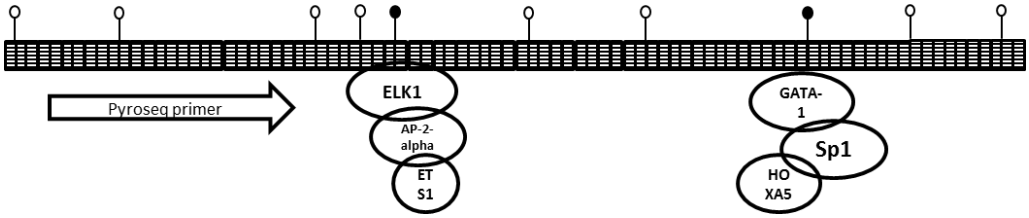


Fig. 66. Schematic representation of the DNA region from -1620 to -1532 in *SNCA*.
The DNA sequence is striped black because it was located 5´ upstream to the transcription start site, or in intron 1, depending on the isoform considered, and, moreover, the predicted promoter overlapped with it. When comparing DNA methylation levels between healthy controls and PD cases, there were relevant results for positions 3 (in parietal cortex) and 6 (in parietal cortex and SN) (black lollipops, -1586 and -1551, respectively).

- position 3: ELK1 [P19419] binds to GCGC**C**GGGAG; AP-2-alpha [P05549] binds to GC**C**GGGAGA; ETS1 [P14921] binds to CTCC**C**G (- strand).
- position 6: GATA-1 [P15976] binds to ATCAG**C**GGTG; Sp1 [P08047] binds to CCCCAC**C**GCT (- strand); HOXA5 [P20719] binds to **C**GCTGATT (- strand).

(4b) assay 2.

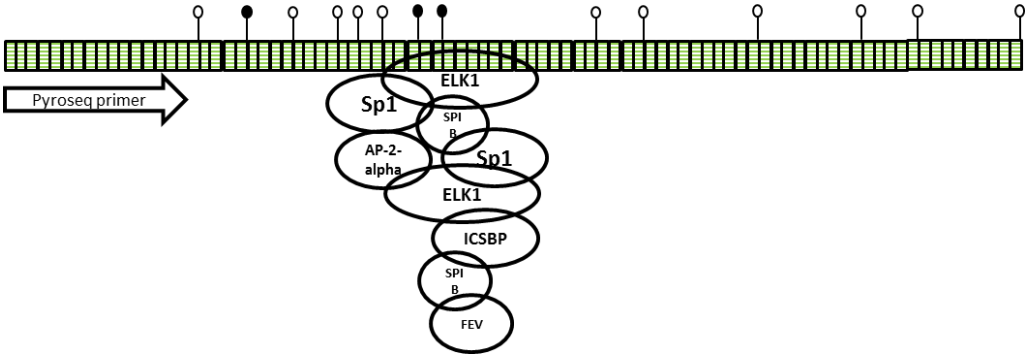


Fig. 67. Schematic representation of the DNA region from -1479 to -1390 in *SNCA*.
The DNA sequence is striped green because it was located in exon 1 (isoform NM_000345) and, moreover, the predicted promoter overlapped with it. When comparing DNA methylation levels between healthy controls and PD cases, there were relevant results for positions 2 (in occipital cortex), 7 (in occipital cortex) and 8 (in SN) (black lollipops, -1458, -1442 and -1440, respectively).

- position 2: there was no transcription factor binding site predicted for it.
- position 7: ELK1 [P19419] binds to CGA**C**GCGGAAGTGA; Sp1 [P08047] binds to CGCCGCGA**C**G; AP-2-alpha [P05549] binds to GCCGCGA**C**G; SPIB [Q01892] binds to **C**GCGGAA.
- position 8: Sp1 [P08047] binds to **C**GGAAGTGAG; ELK1 [P19419] binds to CGACG**C**GGAAGTGA; ICSBP [Q02556] binds to G**C**GGAAGTGA; SPIB [Q01892] binds to CG**C**GGAA; FEV [Q99581] binds to G**C**GGAAGT.

V. DISCUSSION.

V. DISCUSSION.

❖ Our **first objective** was to determine the effect of known or suspected genetic susceptibility factors in two Spanish populations, one of Basque descent and a second of non-Basques. The candidates were: three genes that have been considered genetic risk factors in PD (*MAPT*, *GBA* and *NR4A2*), a controversial *PARK locus* (*UCHL1*), *TOR1A*, which is related to dystonia, the gene that is considered a genetic susceptibility factor in AD (*APOE*) and a new one related to it (*TOMM40*).

UCHL1 **(*PARK5*)** encodes for a protein involved in the ubiquitin proteasome system that is highly specific for neurons [191]. Both characteristics could be relevant in PD pathogenesis. It was considered a cause of familial forms of Parkinson´s disease when [193] found that members of one German family carried the p.I93M mutation, although penetrance was incomplete. However, it could not be replicated in subsequent studies [196, 197] and at present, this mutation is considered either a rare cause of PD or with no influence on the disease (its presence in the two German siblings could be coincidental).

Despite this result, the intensive analysis of the gene reported other possible candidate to explain PD pathogenesis: the p.S18Y polymorphism [197]. The Y allele has been considered a protective factor in some studies [196, 198, 199], whereas for other groups it has no influence on PD risk [200-203].

The Y allele is more frequent in Asians than in Caucasians. Its protective effect has been more frequently observed in Asians, but also in

Caucasians. However, in overall, the studies point out that this allele has no influence on PD risk. [199], based on a meta-analysis of thousands of individuals, only observed some effect depending on the model employed, whereas [201], in a larger population, concluded that there was no effect.

In our case, our statistical power to detect some significance was the same (in Basque population) or higher (in non-Basque population) than the described in most of the previous studies and we observed that the Y allele in p.S18Y showed neither a tendency for protection nor for risk in any of the two Spanish populations (Annex III, Tables E and I). Therefore, we do not consider it a genetic susceptibility factor in our population.

It could be possible that this allele was related to protection against PD pathogenesis at early age [198, 200, 203], but we did not analyze this possibility due to the average age of our populations.

Although dystonia is secondary in Parkinson´s disease, we studied the influence of mutations in ***TOR1A***, which is responsible for the most common and severe form of hereditary primary dystonia -DYT1-, on PD pathogenesis. More concretely, the presence of p.delE302/303, that causes the majority of DYT1 cases (≈80%).

This protein has its highest expression in dopaminergic neurons and it accumulates in Lewy bodies in brains from sporadic PD cases [401].

Nevertheless, no pathological lesions have been detected neither in SNpc neurons nor in any other central nervous system region from DYT1 patients [400]. Despite that, we thought that the presence of *TOR1A* mutations in PD patients could follow a similar pattern to the observed for *GBA*. However, none of the individuals analyzed carried the deletion. This stresses the difference that exists between dystonia and GD in their relation with PD.

Moreover, we studied the p.D216H polymorphism, that has been proposed to modify the penetrance of dystonia in p.delE302/303 carriers [409, 410]. As there were no p.delE302/303 carriers, it was, therefore, not possible to analyze the relation between both variants. We did not find any relevant result (Annex III, Tables E and J): only for non-Basques the distribution of genotypes for this polymorphism showed an almost statistically significant result but it was due to the absence of healthy CC carriers (Table 22). As a consequence, our results do not point out that *TOR1A* has an influence on PD pathogenesis. At most, it could be a rare PD risk factor.

NR4A2, also known as Nurr1, is essential for the development and survival of dopaminergic neurons. It is important not only during development but also in adulthood. It is highly expressed in the *substantia nigra* but also in other parts of the brain and the body. For all these reasons, some genetic studies have been conducted to determine if it is involved in PD but with controversial results:

- two mutations were described in heterozygosis in the noncoding exon 1 (c.-291delT and c.-245T>G) by [290] when analyzed German PD patients (familial and sporadic cases) and controls. Nevertheless, [291-293] could not replicate this conclusion in other European populations;
- [294] concluded that the 7048G7049 variant in intron 6 (c.1361+16insG) was a PD risk factor in homozygosis, but [295] observed that only heterozygous carriers had an increased risk to develop PD and [296, 297] found that it was not a genetic susceptibility factor

Therefore, we opted for analyzing a different mutation that we found after sequencing intron 3, which is highly conserved in mammals. The c.864+246C>T variant was not present in the Basque population and its frequency in non-Basque population was not statistically different between healthy controls and PD cases (Annex III).

As the previously described mutations, it is not a genetic susceptibility factor on PD pathogenesis in our populations.

GBA is a genetic susceptibility factor on PD: some GD patients develop parkinsonism and present LB with mutated glucocerebrosidase in their brains. Moreover, relatives of those patients (usually carriers of *GBA* mutations in heterozygosis) develop PD with higher frequency than general population.

Furthermore, there is a higher frequency of *GBA* mutations in Parkinson´s disease patients when compared to healthy population and [267] has also reported that there is a significant deficiency of glucocerebrosidase activity in *substantia nigra* and cerebellum in PD cases without *GBA* mutations.

Some models have been proposed to explain the possible connection between *GBA* mutations and PD. None can explain why only some GD patients develop parkinsonism and why only some PD patients are carriers of mutations in *GBA* gene. It is supposed that *GBA* just contributes to, but not initiates, the development of *SNCA* pathology: the mutations just exacerbate and accelerate the process.

Most of the studies that have analyzed the frequency of *GBA* mutations in PD patients and in controls have concluded that *GBA* mutations can be considered as a PD risk factor (even for familial and early onset Parkinson´s disease): in Ashkenazi Jews [268], in Caucasians from Canada [269], in people from different ethnicities from USA [270], in Portuguese [271], in Italian [272], in Brazilian [273], in Chinese [274, 275] and in Korean [276].

Although some studies (in Tunisian [277] and in Norwegian [278] for example) did not found significant association for *GBA* mutations and Parkinson´s disease, the large, collaborative, international multicenter study, with thousands of PD patients and controls, conducted by [279] observed that there is an increased probability to develop PD for carriers of *GBA* mutations, and that is not exclusive for a specific ethnicity or a specific mutation.

V. DISCUSSION.

Even though we only analyzed the two most frequent mutations (p.N370S, p.L444P) from the almost 300 that have been described in the gene, we observed that they represent a genetic susceptibility factor that increases the risk to develop PD, at least on non-Basque population (Table 19). It is noteworthy that in Basque population, despite the reduced size of the groups, which affects the power to detect statistical significance, the relative frequency of mutations was significantly smaller than in non-Basque individuals: only one patient presented a mutation (p.N370S) (Table 16).

PD is a synucleinopathy, i.e. it is characterized by αSyn deposits. Nevertheless, ***MAPT*** is an important genetic susceptibility factor. There is a ≈2Mb region in 17q21.31, centered in *MAPT*, with strong linkage disequilibrium that includes other genes like *CRHR1*, *IMP5* and *Saitohin* [256]. Inside this region there is an inverted fragment of 900kb that defines two different haplotypes called H1 (direct orientation) and H2 (inverted orientation; almost exclusive of Caucasian population [257]). Homozygous H1H1 carriers have an increased risk to develop PD [260-263] in both, familial and sporadic PD cases, independently of sex, age at onset and even ethnicity.

In our non-Basque population, we confirmed the effect of the H1H1 haplotype: the distribution of genotypes and alleles between healthy controls and PD cases was strongly statistically different (p=0.002 and p=0.000, respectively), being the H1 haplotype an important risk factor that increased

the risk to develop Parkinson´s disease in 1.58 times ($CI_{95\%}$=[1.24-2.02]; Table 20). Under a binary logistic regression model, H1H1 carriers presented a significant increased risk, with an age and sex-correlated odds ratio of 2.133 ($CI_{95\%}$=[1.23-3.699]; Table 21). However, in the Basques, no significant results for this *locus* were observed (Table 17).

APOE, and concretely the ε4 allele, is nowadays the strongest and most highly replicated genetic risk factor for non-familial AD. Actually, most of the statistically significant results obtained when analyzing the genetic influences in late onset AD are located in a region that comprises this gene but also *TOMM40* and *APOC1*.

The precise risk marker could remain unknown. To further study this, some groups analyzed ***TOMM40*** and [438] observed that the length of a poly-T homopolymer present in its intron 6 was correlated to the genotype in *APOE* in the majority of individuals: ε3 carriers presented short (T≤19) and very long (T≥30) alleles whereas ε4 carriers presented long (20≤T≤29) alleles with ε2 carriers having a similar distribution than ε3 carriers.

We studied the effect of *APOE* genotype and *TOMM40* poly-T length on PD pathogenesis.

Despite that AD and PD are very different in, for example, symptoms, brain regions affected or proteic accumulations, the effect of APOE on PD has

been deeply analyzed. Nonetheless, results are controversial. We hypothesize that, as there are frequent but opposed statistically significant results when analyzing the *APOE* genotype [312-324], this is not the factor directly involved on PD pathogenesis but, in turn, is closely related to the real pathogenic factor by either location or function.

In our Basque population, we found that carrying at least one copy of the ε4 allele increased the risk for PD (p=0.039; Table 18). This result is in agreement with those of [312], in familial PD, and [313], in Mexican.

Nevertheless, a different result was obtained in the non-Basques: the distribution of alleles between healthy controls and PD cases was statistically different (p=0.010) as well as the distribution of genotypes (p=0.040) (Table 23). ε2 allele was the main responsible for this differences (Table 24) and more concretely, the presence of only one ε2 allele: only ε2ε3 and ε2ε4 genotypes reached statistical significance under the binary logistic regression models (Tables 25 and 26) and were related to an increased risk to develop PD (OR:1.928 and 2.055, respectively). This result is similar to the obtained by [319], a meta-analysis, and [320], in Thai (with lower statistical power than us).

Furthermore, for non-Basque population we observed that there was a combined effect of H1H1 genotype (*MAPT*) and the presence of one ε2 allele in *APOE*: carriers of both presented three times more tendency to develop PD than those that did not present any of these variants (OR:3.229; $CI_{95\%}$=[1.790-5.824]; Tables 27 and 28).

V. DISCUSSION.

Our results do not shed light into the controversy on *APOE* alleles and their relation with PD pathogenesis: each Spanish population points out in a different direction and, moreover, this direction is different to the observed by other studies conducted in sporadic Caucasian cases (with similar statistical power) that mostly concluded that any of the alleles influenced PD pathogenesis [321-323].

With regard to the poly-T homopolymer in *TOMM40*, mostly of the healthy controls and PD patients presented the previously described pattern for *APOE* genotype – *TOMM40* length, being the long allele the less frequent in both groups, as expected. We observed no influence of the poly-T length on PD pathogenesis neither in Basque nor in non-Basque population (Annex III, Tables H and L). The same result was observed by [487], that analyzed, in Polish controls and PD cases, the *APOE* genotype and the poly-T length and concluded that there was no significant association at the single allele, genotype or haplotype (*TOMM40-APOE*) level for any of the genes[89]. Only those two studies have studied the influence of the homopolymer on PD pathogenesis but both point out in the same direction: this polymorphism on *TOMM40* is not a genetic susceptibility factor, at least in our populations.

Furthermore, there is an important detail to point out about these results in *APOE* and *TOMM40*:

[89] There is a difference between their study and ours because they divided the long alleles in two types: L_a (T: 20-22) and L_b (T: 26-30).

the Hardy–Weinberg principle states that, in the absence of natural selection, mutation, migration, non-random mating, random genetic drift[90], gene flow, and meiotic drive, the genotypic frequencies and the allele frequencies of a population remain constant from one generation to the next [488, 489][91].

When conducting a genetic association study, the control group is analyzed to determine if there are deviations from the Hardy-Weinberg equilibrium (HWE), that is, if any of those processes is acting on it and, therefore, if it is appropriate to be compared with the case group[92].

For the assessment of deviations from the Hardy–Weinberg equilibrium in the data, the most popular approaches include the asymptotic Pearson's chi-square goodness-of-fit test, which is simple and straightforward although very sensitive to small sample size or rare allele frequency, and the exact test, which is valid for any sample size and minor allele frequency. The exact test can be performed through complete enumeration of heterozygote genotypes or on the basis of the Markov chain Monte Carlo procedure [489].

In our genetic association study, both control groups (Basque and non-Basque) presented a deviation from the HWE for the polymorphism in *TOMM40* and also for the genotype in *APOE* in the non-Basque controls.

[90] When a population is small, the allele frequencies can drift from generation to generation. This process is known as genetic drift.
[91] This is an ideal situation because a population will never be exactly in HWE.
[92] Cases do not need to be in HWE. In fact, screening with HWE of data sets of affected individuals has been proposed as a relatively efficient method for detecting gene–disease associations.

As this result was obtained by using the Pearson's X^2 goodness-of-fit test, we decided to use the exact test (GENEPOP program on the web: **http://genepop.curtin.edu.au/genepop_op1.html** [490]). The program calculated the exact test for HWE by the two previously described methods and confirmed the previously obtained result: there was a deviation from the HWE in all three cases.

Although traditionally it has been considered that deviations from the HWE can also be due to genotyping errors and thus could highlight methodological errors, see [491] for example, [492] demonstrated that testing for HWE should not be used as a criterion for identifying SNPs with genotyping errors at unmatched case-control association studies and that this practice was unjustified. Furthermore, [493] remarked that only gross genotyping errors could cause deviations from the HWE as could happen in GWAS where the number of markers genotyped can be greater than tens of thousands, and genotypes are assigned through automated procedures which analyze hybridization intensities.

Moreover, in our case, the methodological process we followed to determine the genotypes in *APOE* and *TOMM40* was robust: standardized protocols that had been widely checked were used and, in addition, the results passed the respective quality controls. As a consequence, genotyping errors do not represent a feasible reason to explain the deviation.

A priori, those control populations were not affected by any of the selective processes that explain deviations from HWE and the influence of both polymorphisms on population selection is not strong enough to explain

that fact. Even population stratification seems unlikely because populations from Sevilla or Donosti are mostly Caucasian, i.e. they are not compossed by different ethnical groups. Therefore, there is no obvious apparent reason to explain the deviation beyond that it is occurring by chance.

Anyway, as a consequence of this deviations from HWE for Basque and non-Basque control groups, we have to reformulate two of our explanations:

- it is possible that, actually, the polymorphism in *TOMM40* has an effect on PD pathogenesis that we have not detected either in Basque or in non-Basque population.
- the effect of the ε2 allele in *APOE* on PD pathogenesis in non-Basques might be an artifact due to the HWE deviation observed that distorts the real genetic effect.

In summary, in Basques and non-Basques, relevant results have been obtained for the same mutations/polymorphisms, i.e. *GBA*, *MAPT* and *APOE*. Nevertheless, the differences we have observed, especially in *MAPT* and *GBA*, which are clear genetic susceptibility factors in PD (but even in *APOE*) may involve that despite Basque and non-Basque populations are Spanish, due to the historical isolation that Basques have maintained and thus their higher rate of endogamy, the genetic susceptibility factors that affect them in respect to PD pathogenesis are slightly different than those that affect other Spanish populations. Therefore, our results should be maintained as for

Basque and non-Basque population and not be extrapolated to the overall Spanish population.

Parkinson´s disease is a complex maladie with unknown etiology. Nowadays, there still are plenty of unanswered questions about it: could there be more than one pathological mechanism but only one disease? What explains the different symptoms that PD patients present? Why are there differences in how cases respond to treatment?...

There are five genes responsible for familial PD, a minority of cases (≈10%) [50]. However, even in those familiar cases there are remaining questions to solve such as, for example, the mechanisms that explain how they act in a pathological way. Why they affect dopamine-containing neurons in SNpc if their expression is widespread? In addition, not all those cells and not only those cells are affected, and the spatial and temporal pattern of cell death is not always the same…

We decided to focus on the sporadic forms of the disease, that is, on the majority of patients (≈90%), the individuals where there is not a known cause that explains why they are affected.

The common disease-common variant hypothesis [494] postulates that these patients could carry genetic susceptibility factors that trigger the onset of the disease even in combination with a relatively common genetic

background. These genetic factors are ancient variants, i.e. they are frequent at the population, but have an inherent pathogenic capacity that can be enhanced when combined with some environmental factors.

Nevertheless, how could those pathogenic variants not be eliminated by the natural selection over the years if they have a pathogenic and active potential? Maybe it could be explained by the fact that PD is mostly a late onset disease, and therefore relatively immune from natural selection, whose prevalence has recently increased as the human life expectancy has done it. These variants could confer some competitive advantage in young/reproductive age even at the price of increasing the chances of an unhealthy ageing. Moreover, in overall population the pathogenic effect seems to be weak.

- Based on the common disease-common variant hypothesis, we propose an **additive model** to explain **Parkinson´s disease pathogenesis**:

  Let´s assume that there are

  - genetic variants frequent in the population that are susceptibility factors in PD, for example, the H1 haplotype[93]. We could also include here even the APOE allele, although this is more controversial. Those genetic susceptibility factors

[93] H1 haplotype could be considered not risky *per se* although less optimum than the H2 haplotype, which has suffered positive selection. For H1/H2 haplotype (*MAPT*) and *APOE* genotype, as it could be for other variants, it remains unclear if they are the risk factors or are closely related to the real pathogenic factors by either location or function.

originate elements, i.e. proteins/RNAs, with standard or low quality.

- other minority factors, such as *GBA* mutations, which are less frequent but have higher pathogenic potential: they originate defective proteins or RNAs (extremely low quality elements).[94]

Each person carries a unique combination of these factors[95] that has a specific pathogenic potential.

Each cell works employing the elements it has: therefore, in individuals that carry a high-risk load, cells work at suboptimal conditions due to the low quality of their components and are, as a consequence, more prone to suffer stress. Some people carry a very high-pathogenic combination and cross the threshold: there will be a period of time while their cells can return to the homeostatic situation, but in a specific moment (due to some environmental factor or to the ageing process…) they are overpassed, and thus, there is cellular dysfunction that finally leads to cellular death and PD pathogenesis.

It is noteworthy that, for example, glucocerebrosidase, tau and APOE, which are the proteins that could be affected for those variants, do not work in processes exclusive for dopamine-containing neurons. In addition, their

[94] There could be even more types of variants with different degrees of pathogenicity and inversely related frequencies.
The variants could affect expression, function, interactions… of proteins/RNAs and thus decrease their quality. This aspect is the most difficult to determine: even for the already accepted genetic susceptibility factors in PD, almost all those questions remain unclear.

[95] People from the same population share more similarity: that would explain the genetic differences observed for Basques and non-Basques. Moreover, each combination of factors could originate different pathological mechanisms and thus, different evolutions and symptoms.

expression is widespread[96]. Therefore, to explain why only some tissues are affected in PD, there should be some kind of *tissue specificity*, i.e. in these brain regions, the affected elements (for example, glucocerebrosidase, tau and APOE) would participate in critical pathways or would work in combination with tissue-specific proteins/RNAs...

❖ Our **second objective** was to study possible genetic risk factors for the development of cognitive impairment during the evolution of Parkinson´s disease.

To that purpose, using a case-control approach in PD patients with different cognitive status (N.C. and C.I.), we compared the frequency of mutations or polymorphisms located in:

- genes related to other diseases where dementia is consubstantial such as AD (*APP*, *PSEN1*, *PSEN2* and *CALHM1*), HD (*HTT*), CJD (*PRNP*), FTD (*GRN* and *TARDBP*) and Kufor-Rakeb syndrome, also characterized by parkinsonism (*ATP13A2*);
- genes that encode neurotrophic factors such as *CDNF*, *MANF* (and *DOCK3*) and *BDNF*;
- and genes involved in homocysteine metabolism, *CBS*, *MTR* and *MTHFR*, protection against oxidative stress, *NFE2L2* and *KEAP1*, chaperone-mediated autophagy, *HSPA8* and *LAMP-2A*, and *GSK3β*,

[96] Furthermore, their pathogenic mechanism remains unknown, although all have been related to αSyn, and more concretely, to changes in its proneness to aggregate.

due to its relation with tau homeostasis, that is, processes and proteins proposed to be related to PD.

Furthermore, as healthy controls were also included in the study, some possible genetic susceptibility factors to PD pathogenesis were studied in *GBA*, *SNCA* and *MAPT*. Furthermore, a genetic susceptibility factor in AD (*APOE* genotype) and a new one related to it (poly-T homopolymer length, *TOMM40*) were considered.

We observed that there was no influence of poly-T length (*TOMM40*) (Annex IVa, Table P) on PD pathogenesis as previously determined in Basque and non-Basque population. As some studies previously concluded, [321] in Irish, [322], [323] in Norwegian and [324] in a large case-control study with thousands of people, we observed no risk for PD associated to *APOE* (Annex IVa, Tables N and O). We did not obtain this conclusion either in Basque or in non-Basque population: again, our results about *APOE* genotype and PD were controversial. At least we observed that the long allele in *TOMM40* was the most frequent, as expected, and that mostly of the individuals presented the previously described pattern for *APOE* genotype – *TOMM40* length.

Although there were more PD cases carrying mutations in *GBA* than controls, the difference had no statistical significance. The frequency was higher than the observed in Basque population but not as high as the observed in non-Basque population. It is possible that, in this case, as

happened in other studies, the analysis of only the two most frequent mutations (p.N370S, p.L444P), instead of all the existing variants that have been described in the gene, and the size of the sample analyzed precluded the proper quantification of the risk.

For the Rep1 microsatellite (*SNCA*), some alleles have been considered as risk factors: 273 by [77] in a German population although other articles, [71] in a Italian population or [78] in a Singaporean population, did not obtain the same conclusion (we do not even observed this allele); 271 by [79] in a Greek population, [80] a meta-analysis and [81, 82] in a population from USA; moreover, [83] in Australians with European ancestries, [80] a meta-analysis and [81] in a population from USA[97] observed that allele 267 could be neuroprotective. However, we did not obtain any relevant result (Annex IVa, Table M): the allele 269 was the most frequent, as expected, but the distribution of the alleles and genotypes was not different between groups. The length of this polymorphism has been related to PD susceptibility by its influence in αSyn expression, although this is still controversial [74, 75].

It is noteworthy that the demographic characteristics of controls and cases, that is, the percentage of males and females and their mean age, were statistically different. Therefore, the relevance of our results is compromised.

[97] Depending on the article used as a reference, allele name may change: [76] described alleles 0 (267), 1 (269), 2 (271), 3 (273)… Nevertheless, alleles in [77] are 10 bp shorter than those. These are the most common nomenclatures, although some articles ([80] and [82] for example) mix both: their allele 263 is actually 271 in [76].

Nevertheless, and despite that, there were two statistically significant results that confirmed that two polymorphisms were important genetic susceptibility factors in PD, as other groups have previously concluded:

- rs356219 (*SNCA*) (Table 29); the distribution of genotypes and alleles between healthy controls and PD cases was statistically different (p=0.049 and p=0.018, respectively). The A allele was protective against PD (OR: 0.56; $CI_{95\%}$=[0.34-0.92]), especially in AA genotypes, (OR:0.49; $CI_{95\%}$=[0.26-0.94]). [84] in a Norwegian population, [85] in a Caucasian population from the USA, [86] in an Italian population and [87] in Han Chinese, also concluded that the G variant was a PD risk factor, that is, that the A variant was protective. Nevertheless, [88] in Swedish observed the opposite.
- H1/H2 haplotype (*MAPT*) (Table 30); the distribution of genotypes and alleles between healthy controls and PD cases was strongly statistically different (p=0.001 and p=0.000, respectively), being the H1 haplotype an important risk factor that increased the risk to develop Parkinson´s disease in 2.55 times ($CI_{95\%}$=[1.59-4.09]). The same effect was observed for the H1H1 carriers. We, in non-Basque population (the result in Basque population was slightly different) as well as [260-263], obtained the same conclusion.

These variants have been mainly studied for their relation with PD pathogenesis.

Any of them, except the *APOE* genotype, showed any statistically or almost statistically significant result when comparing the PD patients depending on their cognitive status.

For *APOE*, there were not statistically or almost statistically significant results when considering genotypes or alleles or even the number of ε4 alleles (Annex IVb, Tables AH and AI). Nevertheless, when individuals where classified depending on the number of ε2 alleles they carry, we observed that the ε2 allele was more frequent in C.I. (19.67%) than in N.C. (8.00%) and that there was a trend ($p<0.1$) that pointed out to the involvement of this allele in the increased risk to develop cognitive impairment during the evolution of PD (OR: 2.82; $CI_{95\%}$=[0.85-9.37]). This effect was mainly observed for ε2+/- carriers (Table 48).

Previous studies reported that the ε4 allele, not the ε2, could be associated with risk of dementia in PD ([318] sporadic and familial PD, [316] in familial PD). In addition, [495] concluded that carrying at least one ε4 allele was associated with more rapid cognitive decline in British PD patients. The same study observed that H1 haplotype was correlated with lower performance in memory tasks but not with the rate of general cognitive decline but [496] obtained that the H1 haplotype was a genetic risk factor for PD and also was associated with an increased risk to develop dementia in Spanish PD patients.

We did not obtain any similar result for *APOE*, *MAPT* or even *GBA*: [497], in white Europeans, concluded that mutations in *GBA*, not only p.N370S and p.L444P, were not only more frequent in PD patients than in controls, but

also more frequent in PD patients with dementia. The presence of those mutations had no influence on age at onset or motor symptoms.

It is noteworthy that we studied cognitive impairment[98] and not only dementia as the other studies did, but, M.C.I. and dementia groups were not different for the distribution of *APOE, MAPT* or *GBA*. It is possible that the characteristics of the two groups that were compared, controls and cases, have influenced this result.

For the other polymorphisms and mutations analyzed (see the first lines of this second objective), the major part of differences were observed when we compared PD patients and healthy controls, which was not our main objective, especially because, as previously said, the relevance of our results was compromised due to the differences present between both groups in their demographic characteristics and number of individuals. Nevertheless, as observed with rs356219 (*SNCA*) and H1/H2 haplotype (*MAPT*), those inconvenients might not be enough to cover their real risk or protection potential:

[98] Mild cognitive impairment is present since the earliest stages of PD, but not all PD patients will develop dementia because M.C.I. can stabilize, evolve to dementia or recover to normal cognition. Despite that M.C.I. and dementia do not represent the same process as are somehow connected, the size of M.C.I. and dementia groups was small, and both did not present statistically different percentage of males and females, mean age at onset and period of evolution, and were genetically similar (see Results IV.2.), we opted for considering individuals with dementia and mild cognitive impairment together, i.e. C.I. group, to increase the statistical power of our analysis.

- *HTT* (Huntington´s disease): the T allele in rs363096, tended to increase the risk to develop PD (p=0.086; OR: 1.50; $CI_{95\%}$=[0.94-2.39]) (Table 31). This effect was also observed for the haplotypes ATCAT, protective against PD (p=0.002), and ATTAT, almost risky (p=0.087) (Table 32).
- *ATP13A2* (Kufor-Rakeb syndrome; dementia and parkinsonism): for rs6684770, the distribution of genotypes and alleles was statistically different (p=0.044 and p=0.016, respectively), being the T allele the responsible for the 1.89 times ($CI_{95\%}$=[1.12-3.19]) increased risk to develop Parkinson´s disease (Table 33). Moreover, the haplotype ATG was significantly risky (p<0.05) and two more, GCG and ATA, were almost protective and risky, respectively (p<0.1) (Table 34).

Previous studies have analyzed mutations, not polymorphisms, in *ATP13A2*, and have concluded that its connection with PD is minimal [207, 208]. Nevertheless, there were discordant opinions [209].

- *CDNF* (neurotrophic factor): the distribution of genotypes in rs7099185 was statistically different (p=0.017) and the T allele showed a trend towards increasing the risk to develop PD (p=0.084; OR: 1.61; $CI_{95\%}$=[0.93-2.78]) (Table 35). There were trends for two haplotypes (rs7094179-rs7099185): TT was almost risky (p=0.055), whereas GC was almost protective (p=0.077) (Table 36).
- *DOCK3* ("neurotrophic factor"): the distribution of genotypes in rs4441646 between cases and controls was statistically different (p=0.035) (Table 37).

- *MTHFR* (homocysteine metabolism): the C allele in rs1801133 (c.C677T) seemed to increase the risk to develop PD in 1.54 times (p=0.066). Moreover, in rs1801131 (c.A1298C), the distribution of alleles was statistically different (p=0.014), being the A allele protective against PD (OR: 0.51; $CI_{95\%}$=[0.30-0.88]). Its effect almost reached statistically significant results when considering genotypes (p=0.055) (Table 38). There were also relevant results when considering haplotypes (Table 39).
- *HSPA8* (chaperone-mediated autophagy): in rs1461496, the distribution of alleles was almost statistically different and the T allele tended to be protective (p=0.071; OR: 0.65; $CI_{95\%}$=[0.41-1.04]). The difference in genotypes reached statistical significance (p=0.040) (Table 40). The TG haplotype was almost protective against PD pathogenesis (p=0.092) (Table 41).
- *LAMP-2A* (chaperone-mediated autophagy): there was a statistically significant difference between both groups for the distribution of genotypes and alleles in rs7057652. The T allele increased the risk to develop PD in 2.40 times ($CI_{95\%}$=[1.38-4.17]). The effect was the same independently of sex, although in males it did not reach statistically significant difference (Table 42). In rs42890, there was a statistically significant difference for the distribution of alleles. The G allele increased the risk to develop PD in 1.94 times ($CI_{95\%}$=[1.14-3.29]), especially in homozygosis (Table 43). In addition, there were two haplotypes with relevance on PD pathogenesis: GAAT (p=0.002) and TAAG (p=0.00001) (Table 44).

However, the differences observed between the groups N.C. and C.I., which only differed in their mean age at onset, were slightly more numerous than the observed between the individuals with M.C.I. and dementia:

- *CDNF* (neurotrophic factor): the G allele in rs7094179 showed a tendency to protect against cognitive impairment in PD patients (p=0.083; OR: 0.61; $CI_{95\%}$=[0.35-1.07]) (Table 49). The haplotype GT (rs7094179-rs7099185) also showed the same tendency (p=0.092) (Table 50).

[498] observed that the G allele, rs7094179, showed a trend towards increased risk to develop PD in a Korean population, although the genotypes did not affected mRNA expression levels in peripheral lymphoblasts. However, as previously described (Tables 35 and 36), we only found some relevant results, when comparing PD cases and healthy controls, for rs7099185 and for some haplotypes where this G allele was not risky.

- *CBS* (homocysteine metabolism): the difference in alleles was almost statistically significant (p=0.086) being the insertion c.844ins68 the risk factor (Table 51).
- *MTR* (homocysteine metabolism): the A allele increased the risk to develop cognitive impairment during PD in 3.36 times (p=0.001; $CI_{95\%}$=[1.56-7.24]). The difference was statistically significant even in genotypes (p=0.007) (Table 52).
- *NFE2L2* (protection against oxidative stress): the haplotype GG (rs1806649-rs10183914) was more frequent in PD patients with cognitive impairment (p=0.086) (Table 53). This was the only relevant

result we observed. Nevertheles, [499] observed that in rs10183914, frequencies were not different when comparing either Polish or Swedish PD patients and healthy controls, whereas the distribution of genotypes in rs1806649 was statistically different, being the GG genotype more frequent in Polish PD cases than in controls (in Swedish population there was no difference). In addition, some of the haplotypes formed by these two plus other polymorphisms influenced the risk to develop PD and/or the age at onset in both populations.

- *KEAP1* (protection against oxidative stress): the distribution of alleles was almost statistically different (p=0.094) and the G allele tended to be protective against mental deterioration (OR: 0.38; $CI_{95\%}$=[0.12-1.22]). Its effect reached statistical significance in genotypes (p=0.028), concretely for the homozygous GG cases (Table 54).
- *HSPA8* (chaperone-mediated autophagy): in rs4936770, the distribution of alleles was statistically different, being the A allele protective against cognitive impairment in PD (p=0.010; OR: 0.43; $CI_{95\%}$=[0.22-0.82]). The distribution of genotypes between both groups was statistically different too (p=0.045) (Table 55). Relevant results were also observed for some haplotypes (Table 57).
- *LAMP-2A* (chaperone-mediated autophagy): there were only statistically and almost statistically significant results at the polymorphism rs42897. However, the effect of alleles and genotypes on cognitive impairment was confusing and opposite depending on the sex of the cases (Table 59). One haplotype, GAAG, was statistically

more frequent in patients with normal cognition than in those with cognitive impairment (p=0.018) (Table 58).

We did not observe relevant results either for *BDNF* (neurotrophic factor) or for *GSK3β* in any of the comparisons we carried out. These conclusions have also been obtained by other groups:

- p.V66M (*BDNF*): the majority of the studies except one, [392] in an Italian population, have concluded that there is no significant difference in the allelic or genotypic frequency between PD cases and controls: [385] and [386] in sporadic PD in Chinese populations, [387] in a population from USA, [388] in a Greek population, [389] in a Caucasian population, [390] in familial PD in a worldwide population and also [391] in a meta-analysis based on 6 studies in sporadic PD, 4 in Asians, Chinese and Japanese, and 2 in Caucasians, from UK and Sweden. It does not seem to play a major role in the pathogenesis of PD.
- rs334558 and rs6438552 (*GSK3β*): [413] conducted a genetic study in Australian and Chinese PD patients and controls that highlighted that there was not significant difference either in the allelic or in the genotypic frequency between both groups for any of the two polymorphisms. Subsequent genetic analyses in PD patients, [95] in a British population, [93] in a Caucasian population, [412] in an Indian population and [500] a meta-analysis, have also reported the same conclusion. Nevertheless, other studies did find significant differences:

[414] observed that the CC genotype in rs334558 was protective against PD in a Greek population whereas [415] only analyzed rs334558 in Han Chinese sporadic PD patients and controls and found that the T allele was a protective PD factor. Its role in PD pathogenesis seems to be minor.

In our study, none of the individuals carried the mutations we analyzed in *TARDBP* or *GRN* (frontotemporal dementia) pointing out that they had no influence on PD pathogenesis. Other groups have studied which is the influence of carrying mutations on these genes on PD pathogenesis:

- [501] sequenced the entire *GRN* gene in a Belgian population and concluded that mutations on it had no role in the genetic etiology of PD.
- Mutations in *TARDBP* do not appear to contribute to the pathogenesis of PD either: [502] in a population from USA, [503] in a French-Canadian population and [504] in Dutch patients and controls obtained this conclusion after sequencing the entire coding region or only the exon 6, where almost all the mutations that have been described at the present time are located.

A priori, the genetic determinants selected and analyzed could be divided in two groups[99]:

[99] Nevertheless, the results we obtained pointed out that both categories, i.e. PD risk and cognitive impairment, could be related.

- For those that have been related to PD risk, we replicated some previously described results (also observed in our Spanish population – see Results IV.1-): the H1 haplotype (*MAPT*) and the G allele in rs356219 (*SNCA*) were PD risk factors whereas the poly-T length in *TOMM40* had no influence and the effect of the *APOE* genotype on PD risk was again controversial. Nevertheless, mutations in *GBA* were not genetic susceptibility factors on PD pathogenesis.

  With regard to the effect of those variants on the development of dementia in Parkinson´s disease, we only observed that the ε2 allele in *APOE* showed some tendency to increase the risk to develop cognitive impairment. However, neither this result nor the others for *MAPT* and *GBA* agreed with the few previous conclusions reported by others.

- For those that we considered that could determine the appearance of dementia in PD:
  - For some of these variants, *TARDBP*, *GRN*, *BDNF* and *GSK3β*, studies have already been conducted, mainly to determine their effect on PD risk, and most of the results are negative, as was our case. Therefore, they do not seem to have an important role either on PD risk or in the development of cognitive impairment.
  - For others, we observed differences not only for their influence on cognitive status, but also when analyzing the PD risk. Polymorphisms present in the same gene (*ATP13A2, HTT, CDNF, HSPA8, LAMP-2A, NFE2L2*) as well as genes that work in the same pathway (homocysteine metabolism, chaperone-mediated

autophagy) pointed out in different directions. Those, as well as *KEAP1*, *DOCK3* and the genes related to AD, which mainly gave differences when comparing cases with M.C.I. and dementia, should be more deeply analyzed.

In any case, it is hasty to postulate which could be their influence on Parkinson´s disease. We selected them for their hypothetical relation to PD or dementia but there are no pathogenic mechanisms described that connect any of them with the beginning of the disease or its evolution. To that purpose, the influence of the variants on protein function, expression, interactions… should be studied as well as the combined effect that some of them could have, as our additive model proposes.

Our mathematical models have served as drafts of the results we could obtain once the study has been replicated with larger and more homogeneous groups. Their values of specificity and sensitivity were low, i.e. an important percentage of influence on the fact of developing PD/cognitive impairment was not explained by the variables analyzed.

We also analyzed the distribution of the

- nine PD risk variants in cases and controls (Figure 56 and Table 47) and observed that the 90% of healthy controls carried 1 to 5 whereas 90% of patients carried 3 to 7, pointing out that these were genetic

susceptibility factors, not pathogenic factors, as they were observed in both groups, although, on average, cases carried more risk variants than controls (5 vs 3.39, respectively).

- three cognitive impairment risk variants in cases with different cognitive status (N.C. and C.I.) (Figure 58 and Table 62) and observed that the mean of cognitive impairment risk variants carried by the cases with normal cognition was 1.14, whereas for those with cognitive impairment was 1.74. As previously said, they could be genetic susceptibility factors, not pathogenic factors.

We could suppose that the variants have an additive effect (**additive model; common disease - common variant hypothesis**) and act in conjunction but also with other genetic and even environmental factors, as the sensitivity and specificity values indicate, to influence on the cause or evolution of Parkinson´s disease, respectively.

Nevertheless, there was no correlation between the variables (9 and 3, respectively and separately) that could have shed some light on the affected biological mechanisms that contribute to the development of Parkinson´s disease or cognitive dysfunction.

Moreover, our results did not indicate that both processes were genetically related as relevant variants were different in both processes, the distribution of the nine PD risk variants on cases was not different when comparing N.C. and C.I. (Table 63 and Figure 59) and there was no correlation between these 9 and the other 3.

V. DISCUSSION.

It is noteworthy that in our analysis, there were demographic differences between groups, the statistical power was low and, furthermore, we grouped cases with mild cognitive impairment and dementia together.

However, despite that, we replicated some previously reported results, i.e. *MAPT* and *SNCA*, and obtained some relevant results for the variants that we considered that could determine the appearance of dementia in PD. We should replicate our study with larger and more homogeneous groups but, at least, these results represent some trends to be considered in the future.

❖ Our **third objective** was to determine the relation between genetic variability in the *ARMCX* gene family and PD pathogenesis in Spanish population.

To that purpose, we used a non-Basque subpopulation of healthy controls and PD patients that had been previously analyzed for possible genetic susceptibility factors in Parkinson´s disease (see Results IV.1. and Discussion, first objective): therefore, as observed for the overall non-Basque population there was no influence of p.S18Y (***UCHL1***), p.delE302/303 and p.D216H (***TOR1A***), c.864+246C>T (***NR4A2***) and poly-T length (***TOMM40***) on PD pathogenesis (Annex V) and H1H1 carriers (***MAPT***) had an increased risk to develop the disease (p=0.025; OR: 3.922; $CI_{95\%}$=[1.184-12.991]). (Tables 72 and 73).

However, there were two differences:

- although p.N370S and p.L444P mutations (***GBA***) were more frequent in PD patients than in healthy controls there was not statistical significance. In this, like in other studies that have analyzed *GBA* gene, methodology, i.e. we only considered two mutations from the almost 300 that have been described, and sample size, that is, the number of individuals analyzed in this case was ≈3-4 smaller than before (Results IV.1.), might be the reason why there were not relevant results when considering the presence of mutations in *GBA* as a risk factor.
- furthermore, the results for ***APOE*** genotype revealed that the ε4 allele was protective against the disease as [314], in Caucasians non-Hispanics, concluded. More concretely, the protective effect reached statistical significance (p=0.044; OR:0.428; $CI_{95\%}$=[0.188-0.976]) only for ε3ε4 carriers (Table 71). In this subpopulation, *APOE* genotypes were in HWE. One more time, the results for *APOE* were controversial and it remains unclear if it is a risk factor or if it is closely related to the real pathogenic factors by either location or function.

Mitochondria and Parkinson´s disease are related but the sense of this relation remains unclear: amongst other facts, mitochondrial function is altered in patients although it is not clear whether it is cause or consequence of the malady, there is a close relation between some pathological genes and mitochondrial function or integrity, complex I is impaired in PD patients and, in addition, some PD patients have mutations in their mtDNA.

V. DISCUSSION.

We decided to study the ***ARMCX* gene family** because the 6 members share some interesting characteristics that could shed light into this connection and imply a new pathological mechanism in Parkinson´s disease based on the mitochondria: at least *in silico*, the 6 proteins have a mitochondrial targeting signal[100], *in vitro* studies have reported that human ARMCX1, 2, 3 and 6 are mitochondrial proteins [397], and moreover, ARMCX3 is a member of the KIF5/Miro/Trak2 protein complex responsible for the mitochondrial transport along axons [397]. Thus, these proteins could regulate mitochondrial dynamics and trafficking, which are essential to supply appropriate energy to distal neuronal branches, and thus for the correct neurotransmission and neuronal viability.

In addition, they are located at chromosome X and could be related to *PARK12* (Xq21-25), and thus to the possibility that sex has some influence on PD susceptibility, despite that the studies that defined the *locus* ([215] first, and [216] although only when using broad criteria for inclusion of PD patients[101]) did not include the *ARMCX* cluster inside the region limited by the markers that showed the highest LOD scores.

We obtained that, from the 8 markers located along the cluster that were analyzed, only rs2858162 (*ARMCX6*) gave some relevant results: the

[100] Four members of the family also have a nuclear localization signal.

[101] In both studies, the major part of the association for this region in X chromosome was due to brother-brother pairs, with sister-sister pairs and mixed sex sibships having lower linkage values.

distribution of alleles and genotypes between healthy controls and PD cases was close to be statistically different (p≈0.070). That was observed when comparing cases and controls in overall, being the C allele and C-containing genotypes more frequent in affected individuals than in healthy people. The same tendency was observed when dividing by sex, for women and men, although there was not statistical significance, maybe due to the reduced size of those groups that had not enough statistical power (Table 74).

Nevertheless, when we analyzed the relevance of haplotypes formed by the eight polymorphisms, we obtained a statistically significant result were the role of rs2858162 was decisive: the haplotype TACTAC only appeared in healthy controls, i.e. it was protective (p=0.034). Nevertheless, the haplotype TACCAC was more frequent in PD cases than in controls although the statistical significance was not enough to consider it a risk haplotype (p=0.077) (Table 76). Both haplotypes only differentiate in their fourth position, that corresponds to rs2858162 and, as observed for genotypes and alleles, the C variant was risky whereas the T variant was protective.

Therefore, although *ARMCX* genes were good candidates to represent new genetic susceptibility factors for their cellular localization, their structure, their expression pattern and their predicted function, our results did not point out to an effect of the *ARMCX* gene family, or any of its members, on PD pathogenesis and thus to a relation with the *PARK12 locus*. Moreover, we did not observe any influence of sex on PD pathogenesis, although this result could be influenced by the reduced size of those groups, i.e. the low statistical

power. Only the polymorphism rs2858162 (*ARMCX6*) should be more deeply studied to confirm the trend observed for its alleles and genotypes (and haplotypes).

❖ And for our **fourth objective** we used an epigenetic perspective: we analyzed if the five genes responsible for the familial forms of Parkinson´s disease could be epigenetically related to PD pathogenesis.

Our interest focused on DNA methylation levels around their transcription start site becuase variations of this epigenetic mark in this area can influence gene expression.

For *SNCA* and *LRRK2*, overexpression is pathogenic and thus, lower DNA methylation levels that involve higher transcription and gene expression could be pathogenic too. However, for *PRKN*, *PINK1* and *DJ-1*, the pathogenic factor is the lack of enough active protein; therefore, higher DNA methylation levels that involve transcription silencing could lead to decreased gene expression and thus, be pathogenic.

Previous studies analyzed DNA methylation levels around the TSS after bisulfite treatment and PCR amplification in SNpc and cortex, even in *putamen*, from PD cases and healthy controls [471-474].

There are two differences when comparing to our study: instead of cloning and sequencing or using mass spectroscopy to quantify the

methylation levels, we carried out pyrosequencing. Moreover, we analyzed the five genes mutated in familial PD whereas [471] and [472] studied *loci* that are relatively related to Parkinson´s disease, that is, *MAPT* which is one clear genetic susceptibility factor but also *UCHL1*, *TNF*α, *APP* and *PSEN1* that have no such clear relation with it. Those studies concluded that in any of the brain regions analyzed DNA methylation levels were different between cases and controls.

With regard to [473] and [474], although the region that they analyzed in *SNCA* was not included in our study, it was close to the area in intron 1 that we studied. Those studies concluded that PD cases showed low methylation levels in SNpc, cortex and *putamen* when compared to controls, and that PD cases presented lower DNA methylation levels but only in SNpc (not in cortex or *putamen*), respectively.

In our case, as expected for housekeeping genes, the promoters of *DJ-1*, *LRRK2*, *PINK1*, *PRKN* and *SNCA* were poorly methylated in all the tissues to allow their ubiquitous expression. The range of our values was similar to the observed by [473] (the percentages observed by [474] were completely different). The major part of our most relevant (statistically significant or close to significance) results were obtained for specific CpG sites mainly located in *PRKN* and *SNCA* (for *PINK1, LRRK2* and *DJ-1* there were almost no relevant results). Most of them showed that DNA methylation levels were higher in controls than in cases but there was no overall tendency as [473] and [474] concluded (Tables 77 and 78).

V. DISCUSSION.

We analyzed *substantia nigra*, parietal cortex and occipital cortex with special interest for the results in SN because, unlike parietal cortex and occipital cortex, that have not been described as affected early in PD and could be considered "control brain regions", it is extensevely affected in PD [17]. Nevertheless, those CpG sites with statistically or almost statistically significant differences in DNA methylation levels were observed not only in SN, but also in parietal and occipital cortex and even, sometimes, were shared by more than one brain region (Tables 77 and 78).

Unfortunately, there was no blood from any of the individuals, so we could not analyze if there was some change in methylation levels in this tissue due to the maladie and if, therefore, it could serve as a new biomarker to detect PD[102]. Furthermore, we did not compare the values between brain regions because, as [448] concluded, the DNA methylation pattern correlates much more strongly within a brain region across individuals than within an individual across brain regions. Moreover, our methodological approach was not able to differentiate 5hmC from 5mC but, although 5hmC has its highest levels in the central nervous system and its amount increases with age, the impact of that fact on methylation percentages is expected to be almost imperceptible.

Due to the characteristics of our study, i.e. low number of individuals analyzed, low levels of DNA methylation observed, small differences and high standard deviation values, as well as presence of relevant results not only in

[102] This was one of our objectives in this epigenetic study, so we decided to include in Annex VI the results we observed in our second trial in blood from 10 PD cases to, at least, have an idea of how they could have been.

substantia nigra, our results could be considered as trends that should be replicated in a larger study. In addition, those methylation levels need to be correlated with gene expression levels.

Subsequently, for the differentially methylated CpG sites, we conducted an *in silico* prediction to determine if they were target sites for transcription factor binding and that, consequently, could alter the binding of transcription factors and thus, gene expression.

In each case some candidates were proposed, but due to their expression pattern, that did not include SN in the brain (that was the case for WT1, ELK1, ETS1, GATA-1, HOXA5, SPIB, FEV and ICSBP), and/or their targets, that did not include our genes (as was the case for NRF-1[103], AP-2-alpha[104], FOXC1[105], HIF1[106], GATA-1[107], SPIB[108] and ICSBP[109]), finally only Sp1 could be considered as a possible candidate: it is ubiquitous and binds to GC-rich motifs

[103] NRF-1 is implicated in the control of nuclear genes required for respiration, heme biosynthesis and mitochondrial DNA transcription and replication.
[104] AP-2-alpha is involved in a large spectrum of biological functions including proper eye, face, body wall, limb and neural tube development. Therefore, defects in this gene are a cause of branchiooculofacial syndrome.
[105] FOXC1 is an important regulator of cell viability and resistance to oxidative stress in the eye. Therefore, mutations in this gene cause some glaucoma phenotypes.
[106] HIF1 functions as a master regulator of cellular and systemic homeostatic response to hypoxia.
[107] GATA-1 plays an important role in erythroid development by regulating the switch of fetal hemoglobin to adult hemoglobin. Mutations in this gene have been associated with X-linked dyserythropoietic anemia and thrombocytopenia.
[108] SPIB promotes the development of plasmacytoid dendritic cells, also known as natural interferon-producing cells.
[109] ICSBP regulates the expression of genes stimulated by type I IFNs, namely IFN-alpha and IFN-beta. IFN=interferon.

(for that reason it was obtained as candidate for all our genes). Moreover, it is involved in many cellular processes, including cell differentiation, cell growth, apoptosis, immune responses, response to DNA damage, and chromatin remodeling. It can be an activator or a repressor and its activity is highly regulated by post-translational modifications. Amongst others it interacts with HDAC1 and DNMT1.

In vitro studies should be conducted to confirm that Sp1 binds to the predicted sites and that its binding is altered by changes in methylation levels.

V. DISCUSSION.

VI. CONCLUSIONS.

VI. CONCLUSIONS.

1. We have confirmed that, at least in non-Basque population, mutations in *GBA* and the H1H1 genotype in *MAPT* are clear genetic susceptibility factors that increase the risk to develop PD. Their effect in Basques is less important, pointing out that there might be genetic differences between the two Spanish groups regarding to PD risk.
   APOE remains as a controversial risk factor.

2. Despite the limitations of our study, derived from the characteristics of the individuals analyzed, we have observed some trends for variants do not previously related to PD, or at least, not clearly, with regard to their influence on PD risk and/or cognitive status during the malady. They are important features to consider in future studies.

3. Based on our results, we have concluded that the members of the *ARMCX* gene family cannot be considered genetic susceptibility factors on Parkinson´s disease. Only the polymorphism rs2858162 (*ARMCX6*) showed some tendency that should be more deeply analyzed.

4. And finally, in our pilot epigenetic study, we have reported that there are differences in DNA methylation levels between cases and controls in some specific CpG dinucleotides located in the predicted promoters of genes responsible for familial forms of Parkinson´s disease, although not only in *substantia nigra*. In addition, these CpG are part of predicted binding sites of transcription factor Sp1 and,

consequently, depending on their methylation level, could alter its binding and thus, gene expression.

VII. BIBLIOGRAPHY.

VII. BIBLIOGRAPHY.

1. Parkinson, J., *An essay on the shaking palsy.* 1817, London: Sherwood, Neely and Jones. 66.
2. Jellinger, K.A., *Formation and development of Lewy pathology: a critical update.* J Neurol, 2009. **256 Suppl 3**: p. 270-9.
3. Fahn, S., *Description of Parkinson's disease as a clinical syndrome.* Ann N Y Acad Sci, 2003. **991**: p. 1-14.
4. Fahn, S. and D. Sulzer, *Neurodegeneration and neuroprotection in Parkinson disease.* NeuroRx, 2004. **1**(1): p. 139-54.
5. Murray L. Barr, J.A.K., *The human nervous system. An anatomical viewpoint.* 6th ed. 1993, Philadelphia: J.B. Lippincot Company. 451.
6. Blandini, F., et al., *Functional changes of the basal ganglia circuitry in Parkinson's disease.* Prog Neurobiol, 2000. **62**(1): p. 63-88.
7. Haegelen, C., et al., *The subthalamic nucleus is a key-structure of limbic basal ganglia functions.* Med Hypotheses, 2009. **72**(4): p. 421-6.
8. Fearnley, J.M. and A.J. Lees, *Ageing and Parkinson's disease: substantia nigra regional selectivity.* Brain, 1991. **114 (Pt 5)**: p. 2283-301.
9. Damier, P., et al., *The substantia nigra of the human brain. I. Nigrosomes and the nigral matrix, a compartmental organization based on calbindin D(28K) immunohistochemistry.* Brain, 1999. **122 (Pt 8)**: p. 1421-36.
10. Damier, P., et al., *The substantia nigra of the human brain. II. Patterns of loss of dopamine-containing neurons in Parkinson's disease.* Brain, 1999. **122 (Pt 8)**: p. 1437-48.
11. Double, K.L., et al., *Selective cell death in neurodegeneration: why are some neurons spared in vulnerable regions?* Prog Neurobiol, 2010. **92**(3): p. 316-29.
12. Wilkinson, I.M.S., *Essential neurology.* 3rd ed. 1999, Oxford: Blackwell Science Ltd. 248.
13. Jankovic, J., *Parkinson's disease: clinical features and diagnosis.* J Neurol Neurosurg Psychiatry, 2008. **79**(4): p. 368-76.
14. Pandya, M., C.S. Kubu, and M.L. Giroux, *Parkinson disease: not just a movement disorder.* Cleve Clin J Med, 2008. **75**(12): p. 856-64.
15. Cronin-Golomb, A., *Parkinson's disease as a disconnection syndrome.* Neuropsychol Rev, 2010. **20**(2): p. 191-208.
16. Lees, A.J., J. Hardy, and T. Revesz, *Parkinson's disease.* Lancet, 2009. **373**(9680): p. 2055-66.
17. Braak, H., et al., *Staging of brain pathology related to sporadic Parkinson's disease.* Neurobiol Aging, 2003. **24**(2): p. 197-211.
18. Michell, A.W., et al., *Biomarkers and Parkinson's disease.* Brain, 2004. **127**(Pt 8): p. 1693-705.
19. Molochnikov, L., et al., *A molecular signature in blood identifies early Parkinson's disease.* Mol Neurodegener, 2012. **7**: p. 26.

20. Voon, V., et al., *Chronic dopaminergic stimulation in Parkinson's disease: from dyskinesias to impulse control disorders.* Lancet Neurol, 2009. **8**(12): p. 1140-9.
21. Simpkins, N. and J. Jankovic, *Neuroprotection in Parkinson disease.* Arch Intern Med, 2003. **163**(14): p. 1650-4.
22. Juri, C.C. and C.P. Chana, *[Levodopa for Parkinson's disease: What have we learned?].* Rev Med Chil, 2006. **134**(7): p. 893-901.
23. Singh, N., V. Pillay, and Y.E. Choonara, *Advances in the treatment of Parkinson's disease.* Prog Neurobiol, 2007. **81**(1): p. 29-44.
24. Zesiewicz, T.A., *Parkinson disease: the controversy of levodopa toxicity in Parkinson disease.* Nat Rev Neurol, 2012. **8**(1): p. 8-10.
25. Olanow, C.W., et al., *Levodopa in the treatment of Parkinson's disease: current controversies.* Mov Disord, 2004. **19**(9): p. 997-1005.
26. Fahn, S., et al., *Levodopa and the progression of Parkinson's disease.* N Engl J Med, 2004. **351**(24): p. 2498-508.
27. Nutt, J.G., K.A. Chung, and N.H. Holford, *Dyskinesia and the antiparkinsonian response always temporally coincide: a retrospective study.* Neurology, 2010. **74**(15): p. 1191-7.
28. Langston, J.W., *The promise of stem cells in Parkinson disease.* J Clin Invest, 2005. **115**(1): p. 23-5.
29. Freed, C.R., *Will embryonic stem cells be a useful source of dopamine neurons for transplant into patients with Parkinson's disease?* Proc Natl Acad Sci U S A, 2002. **99**(4): p. 1755-7.
30. Lindvall, O. and A. Bjorklund, *Cell therapy in Parkinson's disease.* NeuroRx, 2004. **1**(4): p. 382-93.
31. Uitti, R.J., *Surgical treatments for Parkinson's disease.* Can Fam Physician, 2000. **46**: p. 368-73.
32. Peterson, A.L. and J.G. Nutt, *Treatment of Parkinson's disease with trophic factors.* Neurotherapeutics, 2008. **5**(2): p. 270-80.
33. Garbayo, E., et al., *[Neuroprotective and neurorestorative therapies for the treatment of Parkinson's disease].* An Sist Sanit Navar, 2006. **29**(3): p. 325-35.
34. Airavaara, M., et al., *Neurorestoration.* Parkinsonism Relat Disord, 2012. **18 Suppl 1**: p. S143-6.
35. Benmoyal-Segal, L. and H. Soreq, *Gene-environment interactions in sporadic Parkinson's disease.* J Neurochem, 2006. **97**(6): p. 1740-55.
36. de Lau, L.M. and M.M. Breteler, *Epidemiology of Parkinson's disease.* Lancet Neurol, 2006. **5**(6): p. 525-35.
37. Van Den Eeden, S.K., et al., *Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity.* Am J Epidemiol, 2003. **157**(11): p. 1015-22.
38. de Lau, L.M., et al., *Incidence of parkinsonism and Parkinson disease in a general population: the Rotterdam Study.* Neurology, 2004. **63**(7): p. 1240-4.

VII. BIBLIOGRAPHY.

39. Langston, J.W., et al., *Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis.* Science, 1983. **219**(4587): p. 979-80.
40. Kamel, F. and J.A. Hoppin, *Association of pesticide exposure with neurologic dysfunction and disease.* Environ Health Perspect, 2004. **112**(9): p. 950-8.
41. Brown, T.P., et al., *Pesticides and Parkinson's disease--is there a link?* Environ Health Perspect, 2006. **114**(2): p. 156-64.
42. Campdelacreu, J., *Parkinson disease and Alzheimer disease: environmental risk factors.* Neurologia, 2012.
43. Schlesinger, I. and N. Schlesinger, *Uric acid in Parkinson's disease.* Mov Disord, 2008. **23**(12): p. 1653-7.
44. Schwarzschild, M.A., et al., *Serum urate as a predictor of clinical and radiographic progression in Parkinson disease.* Arch Neurol, 2008. **65**(6): p. 716-23.
45. Hardy, J., *No definitive evidence for a role for the environment in the etiology of Parkinson's disease.* Mov Disord, 2006. **21**(10): p. 1790-1.
46. Tanner, C.M., et al., *Parkinson disease in twins: an etiologic study.* JAMA, 1999. **281**(4): p. 341-6.
47. Piccini, P., et al., *The role of inheritance in sporadic Parkinson's disease: evidence from a longitudinal study of dopaminergic function in twins.* Ann Neurol, 1999. **45**(5): p. 577-82.
48. Wirdefeldt, K., et al., *Heritability of Parkinson disease in Swedish twins: a longitudinal study.* Neurobiol Aging, 2011. **32**(10): p. 1923 e1-8.
49. Thacker, E.L. and A. Ascherio, *Familial aggregation of Parkinson's disease: a meta-analysis.* Mov Disord, 2008. **23**(8): p. 1174-83.
50. Thomas, B. and M.F. Beal, *Parkinson's disease.* Hum Mol Genet, 2007. **16 Spec No. 2**: p. R183-94.
51. Venda, L.L., et al., *alpha-Synuclein and dopamine at the crossroads of Parkinson's disease.* Trends Neurosci, 2010. **33**(12): p. 559-68.
52. Conway, K.A., J.D. Harper, and P.T. Lansbury, *Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease.* Nat Med, 1998. **4**(11): p. 1318-20.
53. Uversky, V.N., *A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders.* J Biomol Struct Dyn, 2003. **21**(2): p. 211-34.
54. Uversky, V.N., J. Li, and A.L. Fink, *Evidence for a partially folded intermediate in alpha-synuclein fibril formation.* J Biol Chem, 2001. **276**(14): p. 10737-44.
55. Lavedan, C., *The synuclein family.* Genome Res, 1998. **8**(9): p. 871-80.
56. Spillantini, M.G., et al., *Alpha-synuclein in Lewy bodies.* Nature, 1997. **388**(6645): p. 839-40.

57. Desplats, P., et al., *Inclusion formation and neuronal cell death through neuron-to-neuron transmission of alpha-synuclein.* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 13010-5.
58. Dunning, C.J., et al., *Can Parkinson's disease pathology be propagated from one neuron to another?* Prog Neurobiol, 2011. **97**(2): p. 205-19.
59. Li, J.Y., et al., *Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation.* Nat Med, 2008. **14**(5): p. 501-3.
60. Kordower, J.H., et al., *Lewy body-like pathology in long-term embryonic nigral transplants in Parkinson's disease.* Nat Med, 2008. **14**(5): p. 504-6.
61. Polymeropoulos, M.H., et al., *Mutation in the alpha-synuclein gene identified in families with Parkinson's disease.* Science, 1997. **276**(5321): p. 2045-7.
62. Kruger, R., et al., *Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease.* Nat Genet, 1998. **18**(2): p. 106-8.
63. Zarranz, J.J., et al., *The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia.* Ann Neurol, 2004. **55**(2): p. 164-73.
64. Farrer, M., et al., *A chromosome 4p haplotype segregating with Parkinson's disease and postural tremor.* Hum Mol Genet, 1999. **8**(1): p. 81-5.
65. Singleton, A.B., et al., *alpha-Synuclein locus triplication causes Parkinson's disease.* Science, 2003. **302**(5646): p. 841.
66. Miller, D.W., et al., *Alpha-synuclein in blood and brain from familial Parkinson disease with SNCA locus triplication.* Neurology, 2004. **62**(10): p. 1835-8.
67. Fuchs, J., et al., *Phenotypic variation in a large Swedish pedigree due to SNCA duplication and triplication.* Neurology, 2007. **68**(12): p. 916-22.
68. Chartier-Harlin, M.C., et al., *Alpha-synuclein locus duplication as a cause of familial Parkinson's disease.* Lancet, 2004. **364**(9440): p. 1167-9.
69. Ibanez, P., et al., *Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease.* Lancet, 2004. **364**(9440): p. 1169-71.
70. Webb, J.L., et al., *Alpha-Synuclein is degraded by both autophagy and the proteasome.* J Biol Chem, 2003. **278**(27): p. 25009-13.
71. Spadafora, P., et al., *NACP-REP1 polymorphism is not involved in Parkinson's disease: a case-control study in a population sample from southern Italy.* Neurosci Lett, 2003. **351**(2): p. 75-8.
72. Farrer, M., et al., *alpha-Synuclein gene haplotypes are associated with Parkinson's disease.* Hum Mol Genet, 2001. **10**(17): p. 1847-51.
73. Chiba-Falek, O., J.W. Touchman, and R.L. Nussbaum, *Functional analysis of intra-allelic variation at NACP-Rep1 in the alpha-synuclein gene.* Hum Genet, 2003. **113**(5): p. 426-31.
74. Chiba-Falek, O. and R.L. Nussbaum, *Effect of allelic variation at the NACP-Rep1 repeat upstream of the alpha-synuclein gene (SNCA) on transcription in*

a cell culture luciferase reporter system. Hum Mol Genet, 2001. **10**(26): p. 3101-9.

75. Linnertz, C., et al., *Genetic regulation of alpha-synuclein mRNA expression in various human brain tissues.* PLoS One, 2009. **4**(10): p. e7480.
76. Xia, Y., et al., *Genetic studies in Alzheimer's disease with an NACP/alpha-synuclein polymorphism.* Ann Neurol, 1996. **40**(2): p. 207-15.
77. Kruger, R., et al., *Increased susceptibility to sporadic Parkinson's disease by a certain combined alpha-synuclein/apolipoprotein E genotype.* Ann Neurol, 1999. **45**(5): p. 611-7.
78. Tan, E.K., et al., *Alpha synuclein promoter and risk of Parkinson's disease: microsatellite and allelic size variability.* Neurosci Lett, 2003. **336**(1): p. 70-2.
79. Hadjigeorgiou, G.M., et al., *Association of alpha-synuclein Rep1 polymorphism and Parkinson's disease: influence of Rep1 on age at onset.* Mov Disord, 2006. **21**(4): p. 534-9.
80. Maraganore, D.M., et al., *Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease.* JAMA, 2006. **296**(6): p. 661-70.
81. Kay, D.M., et al., *Genetic association between alpha-synuclein and idiopathic Parkinson's disease.* Am J Med Genet B Neuropsychiatr Genet, 2008. **147B**(7): p. 1222-30.
82. Pankratz, N., et al., *Alpha-synuclein and familial Parkinson's disease.* Mov Disord, 2009. **24**(8): p. 1125-31.
83. Mellick, G.D., D.M. Maraganore, and P.A. Silburn, *Australian data and meta-analysis lend support for alpha-synuclein (NACP-Rep1) as a risk factor for Parkinson's disease.* Neurosci Lett, 2005. **375**(2): p. 112-6.
84. Myhre, R., et al., *Multiple alpha-synuclein gene polymorphisms are associated with Parkinson's disease in a Norwegian population.* Acta Neurol Scand, 2008. **118**(5): p. 320-7.
85. Mata, I.F., et al., *SNCA variant associated with Parkinson disease and plasma alpha-synuclein level.* Arch Neurol, 2010. **67**(11): p. 1350-6.
86. Trotta, L., et al., *SNCA and MAPT genes: Independent and joint effects in Parkinson disease in the Italian population.* Parkinsonism Relat Disord, 2012. **18**(3): p. 257-62.
87. Pan, F., et al., *SNP rs356219 of the alpha-synuclein (SNCA) gene is associated with Parkinson's disease in a Chinese Han population.* Parkinsonism Relat Disord, 2012. **18**(5): p. 632-4.
88. Westerlund, M., et al., *Cerebellar alpha-synuclein levels are decreased in Parkinson's disease and do not correlate with SNCA polymorphisms associated with disease in a Swedish material.* FASEB J, 2008. **22**(10): p. 3509-14.
89. McCarthy, J.J., et al., *The effect of SNCA 3' region on the levels of SNCA-112 splicing variant.* Neurogenetics, 2011. **12**(1): p. 59-64.

90. Mamah, C.E., et al., *Interaction of alpha-synuclein and tau genotypes in Parkinson's disease.* Ann Neurol, 2005. **57**(3): p. 439-43.
91. Biernacka, J.M., et al., *Do interactions between SNCA, MAPT, and LRRK2 genes contribute to Parkinson's disease susceptibility?* Parkinsonism Relat Disord, 2011. **17**(10): p. 730-6.
92. Elbaz, A., et al., *Independent and joint effects of the MAPT and SNCA genes in Parkinson disease.* Ann Neurol, 2011. **69**(5): p. 778-92.
93. Wider, C., et al., *SNCA, MAPT, and GSK3B in Parkinson disease: a gene-gene interaction study.* Eur J Neurol, 2011. **18**(6): p. 876-81.
94. Botta-Orfila, T., et al., *Lack of interaction of SNCA and MAPT genotypes in Parkinson's disease.* Eur J Neurol, 2011. **18**(3): p. e32.
95. Goris, A., et al., *Tau and alpha-synuclein in susceptibility to, and dementia in, Parkinson's disease.* Ann Neurol, 2007. **62**(2): p. 145-53.
96. Melrose, H.L., et al., *Parkinson's disease: a rethink of rodent models.* Exp Brain Res, 2006. **173**(2): p. 196-204.
97. Abeliovich, A., et al., *Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system.* Neuron, 2000. **25**(1): p. 239-52.
98. Fleming, S.M., et al., *Early and progressive sensorimotor anomalies in mice overexpressing wild-type human alpha-synuclein.* J Neurosci, 2004. **24**(42): p. 9434-40.
99. Kahle, P.J., et al., *Subcellular localization of wild-type and Parkinson's disease-associated mutant alpha -synuclein in human and transgenic mouse brain.* J Neurosci, 2000. **20**(17): p. 6365-73.
100. Shimura, H., et al., *Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase.* Nat Genet, 2000. **25**(3): p. 302-5.
101. Smith, W.W., et al., *Leucine-rich repeat kinase 2 (LRRK2) interacts with parkin, and mutant LRRK2 induces neuronal degeneration.* Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18676-81.
102. Mata, I.F., P.J. Lockhart, and M.J. Farrer, *Parkin genetics: one model for Parkinson's disease.* Hum Mol Genet, 2004. **13 Spec No 1**: p. R127-33.
103. Shimura, H., et al., *Ubiquitination of a new form of alpha-synuclein by parkin from human brain: implications for Parkinson's disease.* Science, 2001. **293**(5528): p. 263-9.
104. Imai, Y., et al., *CHIP is associated with Parkin, a gene responsible for familial Parkinson's disease, and enhances its ubiquitin ligase activity.* Mol Cell, 2002. **10**(1): p. 55-67.
105. Chung, K.K., et al., *Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease.* Nat Med, 2001. **7**(10): p. 1144-50.

106. Sha, D., L.S. Chin, and L. Li, *Phosphorylation of parkin by Parkinson disease-linked kinase PINK1 activates parkin E3 ligase function and NF-kappaB signaling.* Hum Mol Genet, 2010. **19**(2): p. 352-63.
107. Kitada, T., et al., *Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism.* Nature, 1998. **392**(6676): p. 605-8.
108. Yamamura, Y., et al., *[Early-onset parkinsonism with diurnal fluctuation--clinical and pathological studies].* Rinsho Shinkeigaku, 1993. **33**(5): p. 491-6.
109. Takahashi, H., et al., *Familial juvenile parkinsonism: clinical and pathologic study in a family.* Neurology, 1994. **44**(3 Pt 1): p. 437-41.
110. Mori, H., et al., *Pathologic and biochemical studies of juvenile parkinsonism linked to chromosome 6q.* Neurology, 1998. **51**(3): p. 890-2.
111. Farrer, M., et al., *Lewy bodies and parkinsonism in families with parkin mutations.* Ann Neurol, 2001. **50**(3): p. 293-300.
112. Shimura, H., et al., *Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients.* Ann Neurol, 1999. **45**(5): p. 668-72.
113. Bekris, L.M., I.F. Mata, and C.P. Zabetian, *The genetics of Parkinson disease.* J Geriatr Psychiatry Neurol, 2010. **23**(4): p. 228-42.
114. Lucking, C.B., et al., *Association between early-onset Parkinson's disease and mutations in the parkin gene.* N Engl J Med, 2000. **342**(21): p. 1560-7.
115. Hilker, R., et al., *Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene.* Ann Neurol, 2001. **49**(3): p. 367-76.
116. Kay, D.M., et al., *Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients.* Ann Neurol, 2007. **61**(1): p. 47-54.
117. Kay, D.M., et al., *A comprehensive analysis of deletions, multiplications, and copy number variations in PARK2.* Neurology, 2010. **75**(13): p. 1189-94.
118. Itier, J.M., et al., *Parkin gene inactivation alters behaviour and dopamine neurotransmission in the mouse.* Hum Mol Genet, 2003. **12**(18): p. 2277-91.
119. Kubo, S., N. Hattori, and Y. Mizuno, *Recessive Parkinson's disease.* Mov Disord, 2006. **21**(7): p. 885-93.
120. Valente, E.M., et al., *Hereditary early-onset Parkinson's disease caused by mutations in PINK1.* Science, 2004. **304**(5674): p. 1158-60.
121. Beilina, A., et al., *Mutations in PTEN-induced putative kinase 1 associated with recessive parkinsonism have differential effects on protein stability.* Proc Natl Acad Sci U S A, 2005. **102**(16): p. 5703-8.
122. Healy, D.G., et al., *PINK1 (PARK6) associated Parkinson disease in Ireland.* Neurology, 2004. **63**(8): p. 1486-8.
123. Khan, N.L., et al., *Clinical and subclinical dopaminergic dysfunction in PARK6-linked parkinsonism: an 18F-dopa PET study.* Ann Neurol, 2002. **52**(6): p. 849-53.

124. Bonifati, V., et al., *Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes.* Neurology, 2005. **65**(1): p. 87-95.
125. Jin, S.M. and R.J. Youle, *PINK1- and Parkin-mediated mitophagy at a glance.* J Cell Sci, 2012. **125**(Pt 4): p. 795-9.
126. Plun-Favreau, H., et al., *The mitochondrial protease HtrA2 is regulated by Parkinson's disease-associated kinase PINK1.* Nat Cell Biol, 2007. **9**(11): p. 1243-52.
127. Matsuda, N. and K. Tanaka, *Uncovering the roles of PINK1 and parkin in mitophagy.* Autophagy, 2010. **6**(7): p. 952-4.
128. Narendra, D.P., et al., *PINK1 is selectively stabilized on impaired mitochondria to activate Parkin.* PLoS Biol, 2010. **8**(1): p. e1000298.
129. Gautier, C.A., T. Kitada, and J. Shen, *Loss of PINK1 causes mitochondrial functional defects and increased sensitivity to oxidative stress.* Proc Natl Acad Sci U S A, 2008. **105**(32): p. 11364-9.
130. Park, J., et al., *Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin.* Nature, 2006. **441**(7097): p. 1157-61.
131. Poole, A.C., et al., *The PINK1/Parkin pathway regulates mitochondrial morphology.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1638-43.
132. Cookson, M.R., *Crystallizing ideas about Parkinson's disease.* Proc Natl Acad Sci U S A, 2003. **100**(16): p. 9111-3.
133. Honbou, K., et al., *The crystal structure of DJ-1, a protein related to male fertility and Parkinson's disease.* J Biol Chem, 2003. **278**(33): p. 31380-4.
134. Zhang, L., et al., *Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis.* Hum Mol Genet, 2005. **14**(14): p. 2063-73.
135. Bonifati, V., et al., *Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism.* Science, 2003. **299**(5604): p. 256-9.
136. Taira, T., et al., *DJ-1 has a role in antioxidative stress to prevent cell death.* EMBO Rep, 2004. **5**(2): p. 213-8.
137. Clements, C.M., et al., *DJ-1, a cancer- and Parkinson's disease-associated protein, stabilizes the antioxidant transcriptional master regulator Nrf2.* Proc Natl Acad Sci U S A, 2006. **103**(41): p. 15091-6.
138. Wang, Z., et al., *DJ-1 modulates the expression of Cu/Zn-superoxide dismutase-1 through the Erk1/2-Elk1 pathway in neuroprotection.* Ann Neurol, 2011. **70**(4): p. 591-9.
139. McCoy, M.K. and M.R. Cookson, *DJ-1 regulation of mitochondrial function and autophagy through oxidative stress.* Autophagy, 2011. **7**(5): p. 531-2.
140. Shendelman, S., et al., *DJ-1 is a redox-dependent molecular chaperone that inhibits alpha-synuclein aggregate formation.* PLoS Biol, 2004. **2**(11): p. e362.
141. Tang, B., et al., *Association of PINK1 and DJ-1 confers digenic inheritance of early-onset Parkinson's disease.* Hum Mol Genet, 2006. **15**(11): p. 1816-25.

142. Xiong, H., et al., *Parkin, PINK1, and DJ-1 form a ubiquitin E3 ligase complex promoting unfolded protein degradation.* J Clin Invest, 2009. **119**(3): p. 650-60.
143. Xu, J., et al., *The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis.* Hum Mol Genet, 2005. **14**(9): p. 1231-41.
144. Goldberg, M.S., et al., *Nigrostriatal dopaminergic deficits and hypokinesia caused by inactivation of the familial Parkinsonism-linked gene DJ-1.* Neuron, 2005. **45**(4): p. 489-96.
145. Kim, R.H., et al., *Hypersensitivity of DJ-1-deficient mice to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyrindine (MPTP) and oxidative stress.* Proc Natl Acad Sci U S A, 2005. **102**(14): p. 5215-20.
146. Mata, I.F., et al., *LRRK2 in Parkinson's disease: protein domains and functional insights.* Trends Neurosci, 2006. **29**(5): p. 286-93.
147. Kumar, A. and M.R. Cookson, *Role of LRRK2 kinase dysfunction in Parkinson disease.* Expert Rev Mol Med, 2011. **13**: p. e20.
148. Higashi, S., et al., *Localization of Parkinson's disease-associated LRRK2 in normal and pathological human brain.* Brain Res, 2007. **1155**: p. 208-19.
149. Gloeckner, C.J., et al., *The Parkinson disease causing LRRK2 mutation I2020T is associated with increased kinase activity.* Hum Mol Genet, 2006. **15**(2): p. 223-32.
150. Biskup, S., et al., *Localization of LRRK2 to membranous and vesicular structures in mammalian brain.* Ann Neurol, 2006. **60**(5): p. 557-69.
151. Funayama, M., et al., *A new locus for Parkinson's disease (PARK8) maps to chromosome 12p11.2-q13.1.* Ann Neurol, 2002. **51**(3): p. 296-301.
152. Paisan-Ruiz, C., et al., *Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease.* Neuron, 2004. **44**(4): p. 595-600.
153. Zimprich, A., et al., *Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology.* Neuron, 2004. **44**(4): p. 601-7.
154. Aasly, J.O., et al., *Novel pathogenic LRRK2 p.Asn1437His substitution in familial Parkinson's disease.* Mov Disord, 2010. **25**(13): p. 2156-63.
155. Li, X., et al., *Leucine-rich repeat kinase 2 (LRRK2)/PARK8 possesses GTPase activity that is altered in familial Parkinson's disease R1441C/G mutants.* J Neurochem, 2007. **103**(1): p. 238-47.
156. Lewis, P.A., et al., *The R1441C mutation of LRRK2 disrupts GTP hydrolysis.* Biochem Biophys Res Commun, 2007. **357**(3): p. 668-71.
157. Deng, J., et al., *Structure of the ROC domain from the Parkinson's disease-associated leucine-rich repeat kinase 2 reveals a dimeric GTPase.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1499-504.

158. West, A.B., et al., *Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity.* Proc Natl Acad Sci U S A, 2005. **102**(46): p. 16842-7.
159. Simon-Sanchez, J., et al., *Parkinson's disease due to the R1441G mutation in Dardarin: a founder effect in the Basques.* Mov Disord, 2006. **21**(11): p. 1954-9.
160. Ruiz-Martinez, J., et al., *Penetrance in Parkinson's disease related to the LRRK2 R1441G mutation in the Basque country (Spain).* Mov Disord, 2010. **25**(14): p. 2340-5.
161. Bardien, S., et al., *Genetic characteristics of leucine-rich repeat kinase 2 (LRRK2) associated Parkinson's disease.* Parkinsonism Relat Disord, 2011. **17**(7): p. 501-8.
162. Correia Guedes, L., et al., *Worldwide frequency of G2019S LRRK2 mutation in Parkinson's disease: a systematic review.* Parkinsonism Relat Disord, 2010. **16**(4): p. 237-42.
163. Latourelle, J.C., et al., *The Gly2019Ser mutation in LRRK2 is not fully penetrant in familial Parkinson's disease: the GenePD study.* BMC Med, 2008. **6**: p. 32.
164. Luzon-Toro, B., et al., *Mechanistic insight into the dominant mode of the Parkinson's disease-associated G2019S LRRK2 mutation.* Hum Mol Genet, 2007. **16**(17): p. 2031-9.
165. West, A.B., et al., *Parkinson's disease-associated mutations in LRRK2 link enhanced GTP-binding and kinase activities to neuronal toxicity.* Hum Mol Genet, 2007. **16**(2): p. 223-32.
166. Di Fonzo, A., et al., *A common missense variant in the LRRK2 gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan.* Neurogenetics, 2006. **7**(3): p. 133-8.
167. Tan, E.K., et al., *The LRRK2 Gly2385Arg variant is associated with Parkinson's disease: genetic and functional evidence.* Hum Genet, 2007. **120**(6): p. 857-63.
168. Tan, E.K., et al., *LRRK2 R1628P increases risk of Parkinson's disease: replication evidence.* Hum Genet, 2008. **124**(3): p. 287-8.
169. Guo, L., et al., *The Parkinson's disease-associated protein, leucine-rich repeat kinase 2 (LRRK2), is an authentic GTPase that stimulates kinase activity.* Exp Cell Res, 2007. **313**(16): p. 3658-70.
170. Smith, W.W., et al., *Kinase activity of mutant LRRK2 mediates neuronal toxicity.* Nat Neurosci, 2006. **9**(10): p. 1231-3.
171. Gandhi, P.N., et al., *The Roc domain of leucine-rich repeat kinase 2 is sufficient for interaction with microtubules.* J Neurosci Res, 2008. **86**(8): p. 1711-20.

172. Gillardon, F., *Leucine-rich repeat kinase 2 phosphorylates brain tubulin-beta isoforms and modulates microtubule stability--a point of convergence in parkinsonian neurodegeneration?* J Neurochem, 2009. **110**(5): p. 1514-22.
173. Wang, X., et al., *LRRK2 regulates mitochondrial dynamics and function through direct interaction with DLP1.* Hum Mol Genet, 2012. **21**(9): p. 1931-44.
174. Jorgensen, N.D., et al., *The WD40 domain is required for LRRK2 neurotoxicity.* PLoS One, 2009. **4**(12): p. e8463.
175. Greggio, E., et al., *The Parkinson disease-associated leucine-rich repeat kinase 2 (LRRK2) is a dimer that undergoes intramolecular autophosphorylation.* J Biol Chem, 2008. **283**(24): p. 16906-14.
176. Webber, P.J., et al., *Autophosphorylation in the leucine-rich repeat kinase 2 (LRRK2) GTPase domain modifies kinase and GTP-binding activities.* J Mol Biol, 2011. **412**(1): p. 94-110.
177. Greggio, E., et al., *The Parkinson's disease kinase LRRK2 autophosphorylates its GTPase domain at multiple sites.* Biochem Biophys Res Commun, 2009. **389**(3): p. 449-54.
178. Ito, G. and T. Iwatsubo, *Re-examination of the dimerization state of leucine-rich repeat kinase 2: predominance of the monomeric form.* Biochem J, 2012. **441**(3): p. 987-94.
179. Berger, Z., K.A. Smith, and M.J. Lavoie, *Membrane localization of LRRK2 is associated with increased formation of the highly active LRRK2 dimer and changes in its phosphorylation.* Biochemistry, 2010. **49**(26): p. 5511-23.
180. Sato, T., et al., *Active gamma-secretase complexes contain only one of each component.* J Biol Chem, 2007. **282**(47): p. 33985-93.
181. Andres-Mateos, E., et al., *Unexpected lack of hypersensitivity in LRRK2 knock-out mice to MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine).* J Neurosci, 2009. **29**(50): p. 15846-50.
182. Tong, Y., et al., *Loss of leucine-rich repeat kinase 2 causes impairment of protein degradation pathways, accumulation of alpha-synuclein, and apoptotic cell death in aged mice.* Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9879-84.
183. Melrose, H.L., et al., *Impaired dopaminergic neurotransmission and microtubule-associated protein tau alterations in human LRRK2 transgenic mice.* Neurobiol Dis, 2010. **40**(3): p. 503-17.
184. Lin, X., et al., *Leucine-rich repeat kinase 2 regulates the progression of neuropathology induced by Parkinson's-disease-related mutant alpha-synuclein.* Neuron, 2009. **64**(6): p. 807-27.
185. Li, Y., et al., *Mutant LRRK2(R1441G) BAC transgenic mice recapitulate cardinal features of Parkinson's disease.* Nat Neurosci, 2009. **12**(7): p. 826-8.

186. Sen, S., P.J. Webber, and A.B. West, *Dependence of leucine-rich repeat kinase 2 (LRRK2) kinase activity on dimerization.* J Biol Chem, 2009. **284**(52): p. 36346-56.
187. Gasser, T., et al., *A susceptibility locus for Parkinson's disease maps to chromosome 2p13.* Nat Genet, 1998. **18**(3): p. 262-5.
188. DeStefano, A.L., et al., *PARK3 influences age at onset in Parkinson disease: a genome scan in the GenePD study.* Am J Hum Genet, 2002. **70**(5): p. 1089-95.
189. Pankratz, N., et al., *Genes influencing Parkinson disease onset: replication of PARK3 and identification of novel loci.* Neurology, 2004. **62**(9): p. 1616-8.
190. West, A.B., et al., *Refinement of the PARK3 locus on chromosome 2p13 and the analysis of 14 candidate genes.* Eur J Hum Genet, 2001. **9**(9): p. 659-66.
191. Leroy, E., R. Boyer, and M.H. Polymeropoulos, *Intron-exon structure of ubiquitin c-terminal hydrolase-L1.* DNA Res, 1998. **5**(6): p. 397-400.
192. Liu, Y., et al., *The UCH-L1 gene encodes two opposing enzymatic activities that affect alpha-synuclein degradation and Parkinson's disease susceptibility.* Cell, 2002. **111**(2): p. 209-18.
193. Leroy, E., et al., *The ubiquitin pathway in Parkinson's disease.* Nature, 1998. **395**(6701): p. 451-2.
194. Andersson, F.I., et al., *The effect of Parkinson's-disease-associated mutations on the deubiquitinating enzyme UCH-L1.* J Mol Biol, 2011. **407**(2): p. 261-72.
195. Kabuta, T., et al., *Aberrant interaction between Parkinson disease-associated mutant UCH-L1 and the lysosomal receptor for chaperone-mediated autophagy.* J Biol Chem, 2008. **283**(35): p. 23731-8.
196. Wintermeyer, P., et al., *Mutation analysis and association studies of the UCHL1 gene in German Parkinson's disease patients.* Neuroreport, 2000. **11**(10): p. 2079-82.
197. Lincoln, S., et al., *Low frequency of pathogenic mutations in the ubiquitin carboxy-terminal hydrolase gene in familial Parkinson's disease.* Neuroreport, 1999. **10**(2): p. 427-9.
198. Maraganore, D.M., et al., *UCHL1 is a Parkinson's disease susceptibility gene.* Ann Neurol, 2004. **55**(4): p. 512-21.
199. Ragland, M., et al., *Association between the ubiquitin carboxyl-terminal esterase L1 gene (UCHL1) S18Y variant and Parkinson's Disease: a HuGE review and meta-analysis.* Am J Epidemiol, 2009. **170**(11): p. 1344-57.
200. Elbaz, A., et al., *S18Y polymorphism in the UCH-L1 gene and Parkinson's disease: evidence for an age-dependent relationship.* Mov Disord, 2003. **18**(2): p. 130-7.
201. Healy, D.G., et al., *UCHL-1 is not a Parkinson's disease susceptibility gene.* Ann Neurol, 2006. **59**(4): p. 627-33.

202. Wang, L., et al., *Case-control study of the UCH-L1 S18Y variant in sporadic Parkinson's disease in the Chinese population.* J Clin Neurosci, 2011. **18**(4): p. 541-4.
203. Snapinn, K.W., et al., *The UCHL1 S18Y polymorphism and Parkinson's disease in a Japanese population.* Parkinsonism Relat Disord, 2011. **17**(6): p. 473-5.
204. Hampshire, D.J., et al., *Kufor-Rakeb syndrome, pallido-pyramidal degeneration with supranuclear upgaze paresis and dementia, maps to 1p36.* J Med Genet, 2001. **38**(10): p. 680-2.
205. Williams, D.R., et al., *Kufor Rakeb disease: autosomal recessive, levodopa-responsive parkinsonism with pyramidal degeneration, supranuclear gaze palsy, and dementia.* Mov Disord, 2005. **20**(10): p. 1264-71.
206. Ramirez, A., et al., *Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase.* Nat Genet, 2006. **38**(10): p. 1184-91.
207. Vilarino-Guell, C., et al., *ATP13A2 variability in Parkinson disease.* Hum Mutat, 2009. **30**(3): p. 406-10.
208. Djarmati, A., et al., *ATP13A2 variants in early-onset Parkinson's disease patients and controls.* Mov Disord, 2009. **24**(14): p. 2104-11.
209. Di Fonzo, A., et al., *ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease.* Neurology, 2007. **68**(19): p. 1557-62.
210. Hicks, A.A., et al., *A susceptibility gene for late-onset idiopathic Parkinson's disease.* Ann Neurol, 2002. **52**(5): p. 549-55.
211. Oliveira, S.A., et al., *Identification of risk and age-at-onset genes on chromosome 1p in Parkinson disease.* Am J Hum Genet, 2005. **77**(2): p. 252-64.
212. Maraganore, D.M., et al., *High-resolution whole-genome association study of Parkinson disease.* Am J Hum Genet, 2005. **77**(5): p. 685-93.
213. Li, Y.J., et al., *Investigation of the PARK10 gene in Parkinson disease.* Ann Hum Genet, 2007. **71**(Pt 5): p. 639-47.
214. Haugarvoll, K., et al., *Fine-mapping and candidate gene investigation within the PARK10 locus.* Eur J Hum Genet, 2009. **17**(3): p. 336-43.
215. Pankratz, N., et al., *Genome screen to identify susceptibility genes for Parkinson disease in a sample without parkin mutations.* Am J Hum Genet, 2002. **71**(1): p. 124-35.
216. Pankratz, N., et al., *Genome-wide linkage analysis and evidence of gene-by-gene interactions in a sample of 362 multiplex Parkinson disease families.* Hum Mol Genet, 2003. **12**(20): p. 2599-608.
217. Prestel, J., et al., *PARK11 is not linked with Parkinson's disease in European families.* Eur J Hum Genet, 2005. **13**(2): p. 193-7.
218. Lautier, C., et al., *Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease.* Am J Hum Genet, 2008. **82**(4): p. 822-33.

219. Samaranch, L., et al., *Analysis of the GIGYF2 gene in familial and sporadic Parkinson disease in the Spanish population.* Eur J Neurol, 2009. **17**(2): p. 321-5.
220. Bras, J., et al., *Lack of replication of association between GIGYF2 variants and Parkinson disease.* Hum Mol Genet, 2009. **18**(2): p. 341-6.
221. Li, L., et al., *No evidence for pathogenic role of GIGYF2 mutation in Parkinson disease in Japanese patients.* Neurosci Lett, 2010. **479**(3): p. 245-8.
222. Sutherland, G.T., et al., *Haplotype analysis of the PARK 11 gene, GIGYF2, in sporadic Parkinson's disease.* Mov Disord, 2009. **24**(3): p. 449-52.
223. Jones, J.M., et al., *Loss of Omi mitochondrial protease activity causes the neuromuscular disorder of mnd2 mutant mice.* Nature, 2003. **425**(6959): p. 721-7.
224. Strauss, K.M., et al., *Loss of function mutations in the gene encoding Omi/HtrA2 in Parkinson's disease.* Hum Mol Genet, 2005. **14**(15): p. 2099-111.
225. Simon-Sanchez, J. and A.B. Singleton, *Sequencing analysis of OMI/HTRA2 shows previously reported pathogenic mutations in neurologically normal controls.* Hum Mol Genet, 2008. **17**(13): p. 1988-93.
226. Kruger, R., et al., *A large-scale genetic association study to evaluate the contribution of Omi/HtrA2 (PARK13) to Parkinson's disease.* Neurobiol Aging, 2011. **32**(3): p. 548 e9-18.
227. Yun, J., et al., *Loss-of-function analysis suggests that Omi/HtrA2 is not an essential component of the PINK1/PARKIN pathway in vivo.* J Neurosci, 2008. **28**(53): p. 14500-10.
228. Paisan-Ruiz, C., et al., *Characterization of PLA2G6 as a locus for dystonia-parkinsonism.* Ann Neurol, 2009. **65**(1): p. 19-23.
229. Yoshino, H., et al., *Phenotypic spectrum of patients with PLA2G6 mutation and PARK14-linked parkinsonism.* Neurology, 2010. **75**(15): p. 1356-61.
230. Shi, C.H., et al., *PLA2G6 gene mutation in autosomal recessive early-onset parkinsonism in a Chinese cohort.* Neurology, 2011. **77**(1): p. 75-81.
231. Tian, J.Y., et al., *Analysis of PLA2G6 gene mutation in sporadic early-onset parkinsonism patients from Chinese population.* Neurosci Lett, 2012. **514**(2): p. 156-8.
232. Tomiyama, H., et al., *PLA2G6 variant in Parkinson's disease.* J Hum Genet, 2011. **56**(5): p. 401-3.
233. Shojaee, S., et al., *Genome-wide linkage analysis of a Parkinsonian-pyramidal syndrome pedigree by 500 K SNP arrays.* Am J Hum Genet, 2008. **82**(6): p. 1375-84.
234. Di Fonzo, A., et al., *FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome.* Neurology, 2009. **72**(3): p. 240-5.

235. Zhao, T., et al., *Loss of nuclear activity of the FBXO7 protein in patients with parkinsonian-pyramidal syndrome (PARK15).* PLoS One, 2011. **6**(2): p. e16983.
236. Satake, W., et al., *Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease.* Nat Genet, 2009. **41**(12): p. 1303-7.
237. Simon-Sanchez, J., et al., *Genome-wide association study reveals genetic risk underlying Parkinson's disease.* Nat Genet, 2009. **41**(12): p. 1308-12.
238. Vilarino-Guell, C., et al., *An independent replication of PARK16 in Asian samples.* Neurology, 2010. **75**(24): p. 2248-9.
239. Ramirez, A., et al., *Association of Parkinson disease to PARK16 in a Chilean sample.* Parkinsonism Relat Disord, 2011. **17**(1): p. 70-1.
240. Gan-Or, Z., et al., *Association of sequence alterations in the putative promoter of RAB7L1 with a reduced parkinson disease risk.* Arch Neurol, 2012. **69**(1): p. 105-10.
241. Tan, E.K., et al., *Analysis of GWAS-linked loci in Parkinson disease reaffirms PARK16 as a susceptibility locus.* Neurology, 2010. **75**(6): p. 508-12.
242. Mata, I.F., et al., *Replication of MAPT and SNCA, but not PARK16-18, as susceptibility genes for Parkinson's disease.* Mov Disord, 2010. **26**(5): p. 819-23.
243. Pankratz, N., et al., *Genomewide association study for susceptibility genes contributing to familial Parkinson disease.* Hum Genet, 2009. **124**(6): p. 593-605.
244. Saad, M., et al., *Genome-wide association study confirms BST1 and suggests a locus on 12q24 as the risk loci for Parkinson's disease in the European population.* Hum Mol Genet, 2010. **20**(3): p. 615-27.
245. Hamza, T.H., et al., *Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease.* Nat Genet, 2010. **42**(9): p. 781-5.
246. Simon-Sanchez, J., et al., *Genome-wide association study confirms extant PD risk loci among the Dutch.* Eur J Hum Genet, 2011. **19**(6): p. 655-61.
247. Rhodes, S.L., et al., *Replication of GWAS associations for GAK and MAPT in Parkinson's disease.* Ann Hum Genet, 2010. **75**(2): p. 195-200.
248. Li, N.N., et al., *GWAS-linked GAK locus in Parkinson's disease in Han Chinese and meta-analysis.* Hum Genet, 2012. **131**(7): p. 1089-93.
249. Edwards, T.L., et al., *Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease.* Ann Hum Genet, 2010. **74**(2): p. 97-109.
250. Spencer, C.C., et al., *Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21.* Hum Mol Genet, 2011. **20**(2): p. 345-53.

251. Vilarino-Guell, C., et al., *VPS35 mutations in Parkinson disease.* Am J Hum Genet, 2011. **89**(1): p. 162-7.
252. Zimprich, A., et al., *A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease.* Am J Hum Genet, 2011. **89**(1): p. 168-75.
253. Ando, M., et al., *VPS35 mutation in Japanese patients with typical Parkinson's disease.* Mov Disord, 2012. **27**(11): p. 1413-7.
254. Chartier-Harlin, M.C., et al., *Translation initiator EIF4G1 mutations in familial Parkinson disease.* Am J Hum Genet, 2011. **89**(3): p. 398-406.
255. Liu, F. and C.X. Gong, *Tau exon 10 alternative splicing and tauopathies.* Mol Neurodegener, 2008. **3**: p. 8.
256. Pittman, A.M., et al., *The structure of the tau haplotype in controls and in progressive supranuclear palsy.* Hum Mol Genet, 2004. **13**(12): p. 1267-74.
257. Evans, W., et al., *The tau H2 haplotype is almost exclusively Caucasian in origin.* Neurosci Lett, 2004. **369**(3): p. 183-5.
258. Stefansson, H., et al., *A common inversion under selection in Europeans.* Nat Genet, 2005. **37**(2): p. 129-37.
259. Baker, M., et al., *Association of an extended haplotype in the tau gene with progressive supranuclear palsy.* Hum Mol Genet, 1999. **8**(4): p. 711-5.
260. Farrer, M., et al., *The tau H1 haplotype is associated with Parkinson's disease in the Norwegian population.* Neurosci Lett, 2002. **322**(2): p. 83-6.
261. Healy, D.G., et al., *Tau gene and Parkinson's disease: a case-control study and meta-analysis.* J Neurol Neurosurg Psychiatry, 2004. **75**(7): p. 962-5.
262. Fidani, L., et al., *Association of the Tau haplotype with Parkinson's disease in the Greek population.* Mov Disord, 2006. **21**(7): p. 1036-9.
263. Zabetian, C.P., et al., *Association analysis of MAPT H1 haplotype and subhaplotypes in Parkinson's disease.* Ann Neurol, 2007. **62**(2): p. 137-44.
264. de Silva, R., et al., *The tau locus is not significantly associated with pathologically confirmed sporadic Parkinson's disease.* Neurosci Lett, 2002. **330**(2): p. 201-3.
265. Giasson, B.I., et al., *Initiation and synergistic fibrillization of tau and alpha-synuclein.* Science, 2003. **300**(5619): p. 636-40.
266. Kawakami, F., et al., *Stimulatory effect of alpha-synuclein on the tau-phosphorylation by GSK-3beta.* FEBS J, 2011. **278**(24): p. 4895-904.
267. Gegg, M.E., et al., *Glucocerebrosidase deficiency in substantia nigra of parkinson disease brains.* Ann Neurol, 2012. **72**(3): p. 455-63.
268. Aharon-Peretz, J., H. Rosenbaum, and R. Gershoni-Baruch, *Mutations in the glucocerebrosidase gene and Parkinson's disease in Ashkenazi Jews.* N Engl J Med, 2004. **351**(19): p. 1972-7.
269. Sato, C., et al., *Analysis of the glucocerebrosidase gene in Parkinson's disease.* Mov Disord, 2005. **20**(3): p. 367-70.

270. Lwin, A., et al., *Glucocerebrosidase mutations in subjects with parkinsonism.* Mol Genet Metab, 2004. **81**(1): p. 70-3.
271. Bras, J., et al., *Complete screening for glucocerebrosidase mutations in Parkinson disease patients from Portugal.* Neurobiol Aging, 2007. **30**(9): p. 1515-7.
272. De Marco, E.V., et al., *Glucocerebrosidase gene mutations are associated with Parkinson's disease in southern Italy.* Mov Disord, 2008. **23**(3): p. 460-3.
273. Spitz, M., et al., *Association between Parkinson's disease and glucocerebrosidase mutations in Brazil.* Parkinsonism Relat Disord, 2008. **14**(1): p. 58-62.
274. Sun, Q.Y., et al., *Glucocerebrosidase gene L444P mutation is a risk factor for Parkinson's disease in Chinese population.* Mov Disord, 2010. **25**(8): p. 1005-11.
275. Mao, X.Y., et al., *Association between GBA L444P mutation and sporadic Parkinson's disease from Mainland China.* Neurosci Lett, 2009. **469**(2): p. 256-9.
276. Choi, J.M., et al., *Association of mutations in the glucocerebrosidase gene with Parkinson disease in a Korean population.* Neurosci Lett, 2012. **514**(1): p. 12-5.
277. Nishioka, K., et al., *Glucocerebrosidase mutations are not a common risk factor for Parkinson disease in North Africa.* Neurosci Lett, 2010. **477**(2): p. 57-60.
278. Toft, M., et al., *Glucocerebrosidase gene mutations and Parkinson disease in the Norwegian population.* Neurology, 2006. **66**(3): p. 415-7.
279. Sidransky, E., et al., *Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease.* N Engl J Med, 2009. **361**(17): p. 1651-61.
280. Nichols, W.C., et al., *Mutations in GBA are associated with familial Parkinson disease susceptibility and age at onset.* Neurology, 2009. **72**(4): p. 310-6.
281. Ron, I., D. Rapaport, and M. Horowitz, *Interaction between parkin and mutant glucocerebrosidase variants: a possible link between Parkinson disease and Gaucher disease.* Hum Mol Genet, 2010. **19**(19): p. 3771-81.
282. Westbroek, W., A.M. Gustafson, and E. Sidransky, *Exploring the link between glucocerebrosidase mutations and parkinsonism.* Trends Mol Med, 2011. **17**(9): p. 485-93.
283. Goldin, E., *Gaucher disease and parkinsonism, a molecular link theory.* Mol Genet Metab, 2010. **101**(4): p. 307-10.
284. Yang, Y.X. and D.S. Latchman, *Nurr1 transcriptionally regulates the expression of alpha-synuclein.* Neuroreport, 2008. **19**(8): p. 867-71.
285. Chu, Y., et al., *Nurr1 in Parkinson's disease and related disorders.* J Comp Neurol, 2006. **494**(3): p. 495-514.

286. Lu, L., et al., *DJ-1 upregulates tyrosine hydroxylase gene expression by activating its transcriptional factor Nurr1 via the ERK1/2 pathway.* Int J Biochem Cell Biol, 2011. **44**(1): p. 65-71.
287. Volakakis, N., et al., *NR4A orphan nuclear receptors as mediators of CREB-dependent neuroprotection.* Proc Natl Acad Sci U S A, 2010. **107**(27): p. 12317-22.
288. Zhang, T., et al., *NGFI-B nuclear orphan receptor Nurr1 interacts with p53 and suppresses its transcriptional activity.* Mol Cancer Res, 2009. **7**(8): p. 1408-15.
289. Barneda-Zahonero, B., et al., *Nurr1 protein is required for N-methyl-D-aspartic acid (NMDA) receptor-mediated neuronal survival.* J Biol Chem, 2012. **287**(14): p. 11351-62.
290. Le, W.D., et al., *Mutations in NR4A2 associated with familial Parkinson disease.* Nat Genet, 2003. **33**(1): p. 85-9.
291. Wellenbrock, C., et al., *NR4A2 mutations are rare among European patients with familial Parkinson's disease.* Ann Neurol, 2003. **54**(3): p. 415.
292. Zimprich, A., et al., *Point mutations in exon 1 of the NR4A2 gene are not a major cause of familial Parkinson's disease.* Neurogenetics, 2003. **4**(4): p. 219-20.
293. Ibanez, P., et al., *Absence of NR4A2 exon 1 mutations in 108 families with autosomal dominant Parkinson disease.* Neurology, 2004. **62**(11): p. 2133-4.
294. Xu, P.Y., et al., *Association of homozygous 7048G7049 variant in the intron six of Nurr1 gene with Parkinson's disease.* Neurology, 2002. **58**(6): p. 881-4.
295. Zheng, K., B. Heydari, and D.K. Simon, *A common NURR1 polymorphism associated with Parkinson disease and diffuse Lewy body disease.* Arch Neurol, 2003. **60**(5): p. 722-5.
296. Nichols, W.C., et al., *Evaluation of the role of Nurr1 in a large sample of familial Parkinson's disease.* Mov Disord, 2004. **19**(6): p. 649-55.
297. Hering, R., et al., *Extended mutation analysis and association studies of Nurr1 (NR4A2) in Parkinson disease.* Neurology, 2004. **62**(7): p. 1231-2.
298. Chen, J.H., K.P. Lin, and Y.C. Chen, *Risk factors for dementia.* J Formos Med Assoc, 2009. **108**(10): p. 754-64.
299. Ruitenberg, A., et al., *Incidence of dementia: does gender make a difference?* Neurobiol Aging, 2001. **22**(4): p. 575-80.
300. Aarsland, D., et al., *Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study.* Arch Neurol, 2003. **60**(3): p. 387-92.
301. Meireles, J. and J. Massano, *Cognitive impairment and dementia in Parkinson's disease: clinical features, diagnosis, and management.* Front Neurol, 2012. **3**: p. 88.
302. Aarsland, D., C.G. Ballard, and G. Halliday, *Are Parkinson's disease with dementia and dementia with Lewy bodies the same entity?* J Geriatr Psychiatry Neurol, 2004. **17**(3): p. 137-45.

303. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J Psychiatr Res, 1975. **12**(3): p. 189-98.
304. Nasreddine, Z.S., et al., *The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment.* J Am Geriatr Soc, 2005. **53**(4): p. 695-9.
305. Dong, Y., et al., *The Montreal Cognitive Assessment is superior to the Mini-Mental State Examination in detecting patients at higher risk of dementia.* Int Psychogeriatr, 2012. **24**(11): p. 1749-55.
306. Markwick, A., G. Zamboni, and C.A. de Jager, *Profiles of cognitive subtest impairment in the Montreal Cognitive Assessment (MoCA) in a research cohort with normal Mini-Mental State Examination (MMSE) scores.* J Clin Exp Neuropsychol, 2012. **34**(7): p. 750-7.
307. Hoops, S., et al., *Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease.* Neurology, 2009. **73**(21): p. 1738-45.
308. van der Flier, W.M. and P. Scheltens, *Epidemiology and risk factors of dementia.* J Neurol Neurosurg Psychiatry, 2005. **76 Suppl 5**: p. v2-7.
309. Lee, Y.J., et al., *Inflammation and Alzheimer's disease.* Arch Pharm Res, 2010. **33**(10): p. 1539-56.
310. Hauser, P.S., V. Narayanaswami, and R.O. Ryan, *Apolipoprotein E: from lipid transport to neurobiology.* Prog Lipid Res, 2011. **50**(1): p. 62-74.
311. Verghese, P.B., J.M. Castellano, and D.M. Holtzman, *Apolipoprotein E in Alzheimer's disease and other neurological disorders.* Lancet Neurol, 2011. **10**(3): p. 241-52.
312. Li, Y.J., et al., *Apolipoprotein E controls the risk and age at onset of Parkinson disease.* Neurology, 2004. **62**(11): p. 2005-9.
313. Lopez, M., et al., *Apolipoprotein E epsilon4 allele is associated with Parkinson disease risk in a Mexican Mestizo population.* Mov Disord, 2007. **22**(3): p. 417-20.
314. Gao, J., et al., *Apolipoprotein E genotypes and the risk of Parkinson disease.* Neurobiol Aging, 2011. **32**(11): p. 2106 e1-6.
315. Zareparsi, S., et al., *Age at onset of Parkinson disease and apolipoprotein E genotypes.* Am J Med Genet, 2002. **107**(2): p. 156-61.
316. Pankratz, N., et al., *Presence of an APOE4 allele results in significantly earlier onset of Parkinson's disease and a higher risk with dementia.* Mov Disord, 2006. **21**(1): p. 45-9.
317. Buchanan, D.D., et al., *Association of APOE with Parkinson disease age-at-onset in women.* Neurosci Lett, 2007. **411**(3): p. 185-8.
318. Parsian, A., et al., *Parkinson's disease and apolipoprotein E: possible association with dementia but not age at onset.* Genomics, 2002. **79**(3): p. 458-61.

319. Huang, X., P.C. Chen, and C. Poole, *APOE-[epsilon]2 allele associated with higher prevalence of sporadic Parkinson disease.* Neurology, 2004. **62**(12): p. 2198-202.
320. Pulkes, T., et al., *Association between apolipoprotein E genotypes and Parkinson's disease.* J Clin Neurosci, 2011. **18**(10): p. 1333-5.
321. Whitehead, A.S., et al., *Frequency of the apolipoprotein E epsilon 4 allele in a case-control study of early onset Parkinson's disease.* J Neurol Neurosurg Psychiatry, 1996. **61**(4): p. 347-51.
322. Williams-Gray, C.H., et al., *Apolipoprotein E genotype as a risk factor for susceptibility to and dementia in Parkinson's disease.* J Neurol, 2009. **256**(3): p. 493-8.
323. Vefring, H., et al., *The role of APOE alleles in incident Parkinson's disease. The Norwegian ParkWest Study.* Acta Neurol Scand, 2010. **122**(6): p. 438-41.
324. Federoff, M., et al., *A large study reveals no association between APOE and Parkinson's disease.* Neurobiol Dis, 2012. **46**(2): p. 389-92.
325. Groemer, T.W., et al., *Amyloid precursor protein is trafficked and secreted via synaptic vesicles.* PLoS One, 2011. **6**(4): p. e18754.
326. Bekris, L.M., et al., *Genetics of Alzheimer disease.* J Geriatr Psychiatry Neurol, 2010. **23**(4): p. 213-27.
327. De Strooper, B., *Aph-1, Pen-2, and Nicastrin with Presenilin generate an active gamma-Secretase complex.* Neuron, 2003. **38**(1): p. 9-12.
328. Zhang, H., et al., *Role of presenilins in neuronal calcium homeostasis.* J Neurosci, 2010. **30**(25): p. 8566-80.
329. Schellenberg, G.D. and T.J. Montine, *The genetics and neuropathology of Alzheimer's disease.* Acta Neuropathol, 2012. **124**(3): p. 305-23.
330. Rovelet-Lecrux, A., et al., *APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy.* Nat Genet, 2006. **38**(1): p. 24-6.
331. Sleegers, K., et al., *APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy.* Brain, 2006. **129**(Pt 11): p. 2977-83.
332. O'Brien, R.J. and P.C. Wong, *Amyloid precursor protein processing and Alzheimer's disease.* Annu Rev Neurosci, 2011. **34**: p. 185-204.
333. Dreses-Werringloer, U., et al., *A polymorphism in CALHM1 influences Ca2+ homeostasis, Abeta levels, and Alzheimer's disease risk.* Cell, 2008. **133**(7): p. 1149-61.
334. Marambaud, P., U. Dreses-Werringloer, and V. Vingtdeux, *Calcium signaling in neurodegeneration.* Mol Neurodegener, 2009. **4**: p. 20.
335. Gallego-Sandin, S., M.T. Alonso, and J. Garcia-Sancho, *Calcium homoeostasis modulator 1 (CALHM1) reduces the calcium content of the endoplasmic reticulum (ER) and triggers ER stress.* Biochem J, 2011. **437**(3): p. 469-75.

336. Tan, E.K., et al., *CALHM1 variant is not associated with Alzheimer's disease among Asians.* Neurobiol Aging, 2011. **32**(3): p. 546 e11-2.
337. Inoue, K., et al., *The P86L common allele of CALHM1 does not influence risk for Alzheimer disease in Japanese cohorts.* Am J Med Genet B Neuropsychiatr Genet, 2010. **153B**(2): p. 532-5.
338. Lin, K., et al., *Association between the polymorphisms of CALHM1 and GOLPH2 genes and Alzheimer's disease.* Psychiatr Genet, 2010. **20**(4): p. 190.
339. Koppel, J., et al., *CALHM1 P86L polymorphism modulates CSF Abeta levels in cognitively healthy individuals at risk for Alzheimer's disease.* Mol Med, 2011. **17**(9-10): p. 974-9.
340. Giedraitis, V., et al., *CALHM1 P86L polymorphism does not alter amyloid-beta or tau in cerebrospinal fluid.* Neurosci Lett, 2010. **469**(2): p. 265-7.
341. Pickering-Brown, S.M., *Review: Recent progress in frontotemporal lobar degeneration.* Neuropathol Appl Neurobiol, 2010. **36**(1): p. 4-16.
342. Lillo, P. and J.R. Hodges, *Frontotemporal dementia and motor neurone disease: overlapping clinic-pathological disorders.* J Clin Neurosci, 2009. **16**(9): p. 1131-5.
343. Espay, A.J. and I. Litvan, *Parkinsonism and frontotemporal dementia: the clinical overlap.* J Mol Neurosci, 2011. **45**(3): p. 343-9.
344. Dormann, D. and C. Haass, *TDP-43 and FUS: a nuclear affair.* Trends Neurosci, 2011.
345. Seelaar, H., et al., *Clinical, genetic and pathological heterogeneity of frontotemporal dementia: a review.* J Neurol Neurosurg Psychiatry, 2011. **82**(5): p. 476-86.
346. Gass, J., et al., *Progranulin: an emerging target for FTLD therapies.* Brain Res, 2012. **1462**: p. 118-28.
347. Kumar-Singh, S., *Progranulin and TDP-43: mechanistic links and future directions.* J Mol Neurosci, 2011. **45**(3): p. 561-73.
348. Baker, M., et al., *Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17.* Nature, 2006. **442**(7105): p. 916-9.
349. Cruts, M., et al., *Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21.* Nature, 2006. **442**(7105): p. 920-4.
350. Lagier-Tourenne, C., M. Polymenidou, and D.W. Cleveland, *TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration.* Hum Mol Genet, 2010. **19**(R1): p. R46-64.
351. Damiano, M., et al., *Mitochondria in Huntington's disease.* Biochim Biophys Acta, 2010. **1802**(1): p. 52-61.
352. Warby, S.C., et al., *CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup.* Am J Hum Genet, 2009. **84**(3): p. 351-66.

353. Ehrnhoefer, D.E., L. Sutton, and M.R. Hayden, *Small changes, big impact: posttranslational modifications and function of huntingtin in Huntington disease.* Neuroscientist, 2011. **17**(5): p. 475-92.
354. Seredenina, T. and R. Luthi-Carter, *What have we learned from gene expression profiles in Huntington's disease?* Neurobiol Dis, 2012. **45**(1): p. 83-98.
355. Reddy, P.H. and U.P. Shirendeb, *Mutant huntingtin, abnormal mitochondrial dynamics, defective axonal transport of mitochondria, and selective synaptic degeneration in Huntington's disease.* Biochim Biophys Acta, 2012. **1822**(2): p. 101-10.
356. Ross, C.A. and S.J. Tabrizi, *Huntington's disease: from molecular pathogenesis to clinical treatment.* Lancet Neurol, 2011. **10**(1): p. 83-98.
357. Olanow, C.W. and S.B. Prusiner, *Is Parkinson's disease a prion disorder?* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12571-2.
358. Prusiner, S.B., *Neurodegeneration in humans caused by prions.* West J Med, 1994. **161**(3): p. 264-72.
359. Huang, Z., et al., *Proposed three-dimensional structure for the cellular prion protein.* Proc Natl Acad Sci U S A, 1994. **91**(15): p. 7139-43.
360. Prusiner, S.B., *Inherited prion diseases.* Proc Natl Acad Sci U S A, 1994. **91**(11): p. 4611-4.
361. Goldfarb, L.G., et al., *Fatal familial insomnia and familial Creutzfeldt-Jakob disease: disease phenotype determined by a DNA polymorphism.* Science, 1992. **258**(5083): p. 806-8.
362. Dermaut, B., et al., *PRNP Val129 homozygosity increases risk for early-onset Alzheimer's disease.* Ann Neurol, 2003. **53**(3): p. 409-12.
363. Riemenschneider, M., et al., *Prion protein codon 129 polymorphism and risk of Alzheimer disease.* Neurology, 2004. **63**(2): p. 364-6.
364. Combarros, O., et al., *Polymorphism at codon 129 of the prion protein gene is not associated with sporadic AD.* Neurology, 2000. **55**(4): p. 593-5.
365. Casadei, V.M., et al., *Prion protein gene polymorphism and Alzheimer's disease: one modulatory trait of cognitive decline?* J Neurol Neurosurg Psychiatry, 2001. **71**(2): p. 279-80.
366. Ohkubo, T., et al., *Absence of association between codon 129/219 polymorphisms of the prion protein gene and Alzheimer's disease in Japan.* Ann Neurol, 2003. **54**(4): p. 553-4; author reply 555.
367. Villeneuve, N.F., A. Lau, and D.D. Zhang, *Regulation of the Nrf2-Keap1 antioxidant response by the ubiquitin proteasome system: an insight into cullin-ring ubiquitin ligases.* Antioxid Redox Signal, 2010. **13**(11): p. 1699-712.
368. Kaspar, J.W., S.K. Niture, and A.K. Jaiswal, *Nrf2:INrf2 (Keap1) signaling in oxidative stress.* Free Radic Biol Med, 2009. **47**(9): p. 1304-9.

369. Uruno, A. and H. Motohashi, *The Keap1-Nrf2 system as an in vivo sensor for electrophiles.* Nitric Oxide, 2011. **25**(2): p. 153-60.
370. Massey, A.C., C. Zhang, and A.M. Cuervo, *Chaperone-mediated autophagy in aging and disease.* Curr Top Dev Biol, 2006. **73**: p. 205-35.
371. Dice, J.F., *Chaperone-mediated autophagy.* Autophagy, 2007. **3**(4): p. 295-9.
372. Cuervo, A.M., et al., *Impaired degradation of mutant alpha-synuclein by chaperone-mediated autophagy.* Science, 2004. **305**(5688): p. 1292-5.
373. Alvarez-Erviti, L., et al., *Chaperone-mediated autophagy markers in Parkinson disease brains.* Arch Neurol, 2010. **67**(12): p. 1464-72.
374. Palgi, M., et al., *Evidence that DmMANF is an invertebrate neurotrophic factor supporting dopaminergic neurons.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2429-34.
375. Chen, Y.C., et al., *MANF regulates dopaminergic neuron development in larval zebrafish.* Dev Biol, 2012. **370**(2): p. 237-49.
376. Lindholm, P., et al., *Novel neurotrophic factor CDNF protects and rescues midbrain dopamine neurons in vivo.* Nature, 2007. **448**(7149): p. 73-7.
377. Airavaara, M., et al., *CDNF Protects the Nigrostriatal Dopamine System and Promotes Recovery After MPTP Treatment in Mice.* Cell Transplant, 2011. **21**(6): p. 1213-23.
378. Petrova, P., et al., *MANF: a new mesencephalic, astrocyte-derived neurotrophic factor with selectivity for dopaminergic neurons.* J Mol Neurosci, 2003. **20**(2): p. 173-88.
379. Apostolou, A., et al., *Armet, a UPR-upregulated protein, inhibits cell proliferation and ER stress-induced cell death.* Exp Cell Res, 2008. **314**(13): p. 2454-67.
380. Hellman, M., et al., *Mesencephalic astrocyte-derived neurotrophic factor (MANF) has a unique mechanism to rescue apoptotic neurons.* J Biol Chem, 2010. **286**(4): p. 2675-80.
381. Balaratnasingam, S. and A. Janca, *Brain Derived Neurotrophic Factor: a novel neurotrophin involved in psychiatric and neurological disorders.* Pharmacol Ther, 2012. **134**(1): p. 116-24.
382. Pruunsild, P., et al., *Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters.* Genomics, 2007. **90**(3): p. 397-406.
383. Egan, M.F., et al., *The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function.* Cell, 2003. **112**(2): p. 257-69.
384. Mogi, M., et al., *Brain-derived growth factor and nerve growth factor concentrations are decreased in the substantia nigra in Parkinson's disease.* Neurosci Lett, 1999. **270**(1): p. 45-8.

385. Hong, C.J., et al., *Brain-derived neurotrophic factor (BDNF) Val66Met polymorphisms in Parkinson's disease and age of onset.* Neurosci Lett, 2003. **353**(1): p. 75-7.
386. Liu, J., et al., *Brain-derived neurotrophic factor (BDNF) genetic polymorphism greatly increases risk of leucine-rich repeat kinase 2 (LRRK2) for Parkinson's disease.* Parkinsonism Relat Disord, 2012. **18**(2): p. 140-3.
387. Cramer, S.C., et al., *Increased prevalence of val(66)met BDNF genotype among subjects with cervical dystonia.* Neurosci Lett, 2010. **468**(1): p. 42-5.
388. Karakasis, C., et al., *Association of brain-derived neurotrophic factor (BDNF) Val66Met polymorphism with Parkinson's disease in a Greek population.* J Clin Neurosci, 2011. **18**(12): p. 1744-5.
389. Gao, L., et al., *Brain-derived neurotrophic factor G196A polymorphism and clinical features in Parkinson's disease.* Acta Neurol Scand, 2010. **122**(1): p. 41-5.
390. Karamohamed, S., et al., *BDNF genetic variants are associated with onset age of familial Parkinson disease: GenePD Study.* Neurology, 2005. **65**(11): p. 1823-5.
391. Zintzaras, E. and G.M. Hadjigeorgiou, *The role of G196A polymorphism in the brain-derived neurotrophic factor gene in the cause of Parkinson's disease: a meta-analysis.* J Hum Genet, 2005. **50**(11): p. 560-6.
392. Guerini, F.R., et al., *BDNF Val66Met polymorphism is associated with cognitive impairment in Italian patients with Parkinson's disease.* Eur J Neurol, 2009. **16**(11): p. 1240-5.
393. Coates, J.C., *Armadillo repeat proteins: beyond the animal kingdom.* Trends Cell Biol, 2003. **13**(9): p. 463-71.
394. Mou, Z., A.R. Tapper, and P.D. Gardner, *The armadillo repeat-containing protein, ARMCX3, physically and functionally interacts with the developmental regulatory factor Sox10.* J Biol Chem, 2009. **284**(20): p. 13629-40.
395. Iseki, H., et al., *Human Arm protein lost in epithelial cancers, on chromosome X 1 (ALEX1) gene is transcriptionally regulated by CREB and Wnt/beta-catenin signaling.* Cancer Sci, 2010. **101**(6): p. 1361-6.
396. Kurochkin, I.V., et al., *ALEX1, a novel human armadillo repeat protein that is expressed differentially in normal tissues and carcinomas.* Biochem Biophys Res Commun, 2001. **280**(1): p. 340-7.
397. Lopez-Domenech, G., et al., *The Eutherian Armcx genes regulate mitochondrial trafficking in neurons and interact with Miro and Trak2.* Nat Commun, 2012. **3**: p. 814.
398. Warner, T.T., A. Granata, and G. Schiavo, *TorsinA and DYT1 dystonia: a synaptopathy?* Biochem Soc Trans, 2010. **38**(2): p. 452-6.

399. Konakova, M., et al., *Cellular distribution of torsin A and torsin B in normal human brain.* Arch Neurol, 2001. **58**(6): p. 921-7.
400. Bragg, D.C., et al., *Molecular pathways in dystonia.* Neurobiol Dis, 2011. **42**(2): p. 136-47.
401. Shashidharan, P., et al., *TorsinA accumulation in Lewy bodies in sporadic Parkinson's disease.* Brain Res, 2000. **877**(2): p. 379-81.
402. Kustedjo, K., M.H. Bracey, and B.F. Cravatt, *Torsin A and its torsion dystonia-associated mutant forms are lumenal glycoproteins that exhibit distinct subcellular localizations.* J Biol Chem, 2000. **275**(36): p. 27933-9.
403. Ozelius, L., et al., *Human gene for torsion dystonia located on chromosome 9q32-q34.* Neuron, 1989. **2**(5): p. 1427-34.
404. Ozelius, L.J., et al., *The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein.* Nat Genet, 1997. **17**(1): p. 40-8.
405. Jungwirth, M.T., et al., *The nuclear envelope localization of DYT1 dystonia torsinA-DeltaE requires the SUN1 LINC complex component.* BMC Cell Biol, 2011. **12**: p. 24.
406. Goodchild, R.E. and W.T. Dauer, *Mislocalization to the nuclear envelope: an effect of the dystonia-causing torsinA mutation.* Proc Natl Acad Sci U S A, 2004. **101**(3): p. 847-52.
407. Gerace, L., *TorsinA and torsion dystonia: Unraveling the architecture of the nuclear envelope.* Proc Natl Acad Sci U S A, 2004. **101**(24): p. 8839-40.
408. Naismith, T.V., et al., *TorsinA in the nuclear envelope.* Proc Natl Acad Sci U S A, 2004. **101**(20): p. 7612-7.
409. Kamm, C., et al., *Susceptibility to DYT1 dystonia in European patients is modified by the D216H polymorphism.* Neurology, 2008. **70**(23): p. 2261-2.
410. Risch, N.J., et al., *Intragenic Cis and Trans modification of genetic susceptibility in DYT1 torsion dystonia.* Am J Hum Genet, 2007. **80**(6): p. 1188-93.
411. Kock, N., et al., *Effects of genetic variations in the dystonia protein torsinA: identification of polymorphism at residue 216 as protein modifier.* Hum Mol Genet, 2006. **15**(8): p. 1355-64.
412. Das, G., et al., *Role of tau kinases (CDK5R1 and GSK3B) in Parkinson's disease: a study from India.* Neurobiol Aging, 2010. **33**(7): p. 1485 e9-15.
413. Kwok, J.B., et al., *GSK3B polymorphisms alter transcription and splicing in Parkinson's disease.* Ann Neurol, 2005. **58**(6): p. 829-39.
414. Kalinderi, K., et al., *GSK3beta polymorphisms, MAPT H1 haplotype and Parkinson's disease in a Greek cohort.* Neurobiol Aging, 2009. **32**(3): p. 546 e1-5.
415. Zhao, D.M., et al., *GSK3beta reduces risk of sporadic Parkinson's disease in ethnic Chinese.* Am J Med Genet B Neuropsychiatr Genet, 2012. **159B**(6): p. 718-21.

416. Brosnan, J.T., et al., *Methylation demand: a key determinant of homocysteine metabolism.* Acta Biochim Pol, 2004. **51**(2): p. 405-13.
417. Frosst, P., et al., *A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase.* Nat Genet, 1995. **10**(1): p. 111-3.
418. Summers, C.M., et al., *Influence of the cystathionine beta-synthase 844ins68 and methylenetetrahydrofolate reductase 677C>T polymorphisms on folate and homocysteine concentrations.* Eur J Hum Genet, 2008. **16**(8): p. 1010-3.
419. van der Put, N.M., et al., *A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects?* Am J Hum Genet, 1998. **62**(5): p. 1044-51.
420. Sperandeo, M.P., et al., *A 68-bp insertion found in a homocystinuric patient is a common variant and is skipped by alternative splicing of the cystathionine beta-synthase mRNA.* Am J Hum Genet, 1996. **59**(6): p. 1391-3.
421. Chasse, J.F. and R. Barouki, *Characterisation of a human liver cystathionine beta synthase mRNA sequence corresponding to the c.[833T>C;844_845ins68] mutation in CBS gene.* Mol Cell Biochem, 2009. **332**(1-2): p. 183-7.
422. Tsai, M.Y., et al., *High prevalence of a mutation in the cystathionine beta-synthase gene.* Am J Hum Genet, 1996. **59**(6): p. 1262-7.
423. Martignoni, E., et al., *Homocysteine and Parkinson's disease: a dangerous liaison?* J Neurol Sci, 2007. **257**(1-2): p. 31-7.
424. Ford, A.H., et al., *Plasma homocysteine and MTHFRC677T polymorphism as risk factors for incident dementia.* J Neurol Neurosurg Psychiatry, 2012. **83**(1): p. 70-5.
425. Collaboration, H.S., *Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis.* JAMA, 2002. **288**(16): p. 2015-22.
426. de Ruijter, W., et al., *Use of Framingham risk score and new biomarkers to predict cardiovascular mortality in older people: population based observational cohort study.* BMJ, 2009. **338**: p. a3083.
427. Seshadri, S., et al., *Plasma homocysteine as a risk factor for dementia and Alzheimer's disease.* N Engl J Med, 2002. **346**(7): p. 476-83.
428. Bell, I.R., et al., *Plasma homocysteine in vascular disease and in nonvascular dementia of depressed elderly people.* Acta Psychiatr Scand, 1992. **86**(5): p. 386-90.
429. Lehmann, M., C.G. Gottfries, and B. Regland, *Identification of cognitive impairment in the elderly: homocysteine is an early marker.* Dement Geriatr Cogn Disord, 1999. **10**(1): p. 12-20.
430. Morris, M.S., et al., *Hyperhomocysteinemia associated with poor recall in the third National Health and Nutrition Examination Survey.* Am J Clin Nutr, 2001. **73**(5): p. 927-33.

431. Ford, A.H., et al., *Homocysteine, methylenetetrahydrofolate reductase C677T polymorphism and cognitive impairment: the health in men study.* Mol Psychiatry, 2011. **17**(5): p. 559-66.
432. Schiepers, O.J., et al., *MTHFR polymorphisms and cognitive ageing in the ninth decade: the Lothian Birth Cohort 1921.* Genes Brain Behav, 2011. **10**(3): p. 354-64.
433. Rogers, J.D., et al., *Elevated plasma homocysteine levels in patients treated with levodopa: association with vascular disease.* Arch Neurol, 2003. **60**(1): p. 59-64.
434. Religa, D., et al., *Hyperhomocysteinemia and methylenetetrahydrofolate reductase polymorphism in patients with Parkinson's disease.* Neurosci Lett, 2006. **404**(1-2): p. 56-60.
435. Yuan, R.Y., et al., *Methylenetetrahydrofolate reductase polymorphisms and plasma homocysteine in levodopa-treated and non-treated Parkinson's disease patients.* J Neurol Sci, 2009. **287**(1-2): p. 64-8.
436. Rodriguez-Oroz, M.C., et al., *Homocysteine and cognitive impairment in Parkinson's disease: a biochemical, neuroimaging, and genetic study.* Mov Disord, 2009. **24**(10): p. 1437-44.
437. O'Suilleabhain, P.E., et al., *Elevated plasma homocysteine level in patients with Parkinson disease: motor, affective, and cognitive associations.* Arch Neurol, 2004. **61**(6): p. 865-8.
438. Roses, A.D., et al., *A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease.* Pharmacogenomics J, 2009. **10**(5): p. 375-84.
439. Linnertz, C., et al., *Characterization of the poly-T variant in the TOMM40 gene in diverse populations.* PLoS One, 2012. **7**(2): p. e30994.
440. Grossman, I., et al., *Alzheimer's disease: diagnostics, prognostics and the road to prevention.* EPMA J, 2010. **1**(2): p. 293-303.
441. Chu, S.H., et al., *TOMM40 poly-T repeat lengths, age of onset and psychosis risk in Alzheimer disease.* Neurobiol Aging, 2011. **32**(12): p. 2328 e1-9.
442. Cruchaga, C., et al., *Association and expression analyses with single-nucleotide polymorphisms in TOMM40 in Alzheimer disease.* Arch Neurol, 2011. **68**(8): p. 1013-9.
443. Lee, T.S., et al., *Downregulation of TOMM40 expression in the blood of Alzheimer disease subjects compared with matched controls.* J Psychiatr Res, 2012. **46**(6): p. 828-30.
444. Humphries, A.D., et al., *Dissection of the mitochondrial import and assembly pathway for human Tom40.* J Biol Chem, 2005. **280**(12): p. 11535-43.
445. Christensen, B.C., et al., *Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context.* PLoS Genet, 2009. **5**(8): p. e1000602.

446. Weber, M. and D. Schubeler, *Genomic patterns of DNA methylation: targets and function of an epigenetic mark.* Curr Opin Cell Biol, 2007. **19**(3): p. 273-80.
447. Fraga, M.F., et al., *Epigenetic differences arise during the lifetime of monozygotic twins.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10604-9.
448. Ladd-Acosta, C., et al., *DNA methylation signatures within the human brain.* Am J Hum Genet, 2007. **81**(6): p. 1304-15.
449. Bernstein, B.E., A. Meissner, and E.S. Lander, *The mammalian epigenome.* Cell, 2007. **128**(4): p. 669-81.
450. Stuffrein-Roberts, S., P.R. Joyce, and M.A. Kennedy, *Role of epigenetics in mental disorders.* Aust N Z J Psychiatry, 2008. **42**(2): p. 97-107.
451. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
452. Klose, R.J. and A.P. Bird, *Genomic DNA methylation: the mark and its mediators.* Trends Biochem Sci, 2006. **31**(2): p. 89-97.
453. Chen, T. and E. Li, *Structure and function of eukaryotic DNA methyltransferases.* Curr Top Dev Biol, 2004. **60**: p. 55-89.
454. Turek-Plewa, J. and P.P. Jagodzinski, *The role of mammalian DNA methyltransferases in the regulation of gene expression.* Cell Mol Biol Lett, 2005. **10**(4): p. 631-47.
455. Bogdanovic, O. and G.J. Veenstra, *DNA methylation and methyl-CpG binding proteins: developmental requirements and function.* Chromosoma, 2009. **118**(5): p. 549-65.
456. Gibney, E.R. and C.M. Nolan, *Epigenetics and gene expression.* Heredity (Edinb), 2010. **105**(1): p. 4-13.
457. Worm, J. and P. Guldberg, *DNA methylation: an epigenetic pathway to cancer and a promising target for anticancer therapy.* J Oral Pathol Med, 2002. **31**(8): p. 443-9.
458. Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes.* J Mol Biol, 1987. **196**(2): p. 261-82.
459. Bird, A., *The essentials of DNA methylation.* Cell, 1992. **70**(1): p. 5-8.
460. Antequera, F., *Structure, function and evolution of CpG island promoters.* Cell Mol Life Sci, 2003. **60**(8): p. 1647-58.
461. Auclair, G. and M. Weber, *Mechanisms of DNA methylation and demethylation in mammals.* Biochimie, 2012. **94**(11): p. 2202-11.
462. Munzel, M., et al., *Quantification of the sixth DNA base hydroxymethylcytosine in the brain.* Angew Chem Int Ed Engl, 2010. **49**(31): p. 5375-7.
463. Munzel, M., D. Globisch, and T. Carell, *5-Hydroxymethylcytosine, the sixth base of the genome.* Angew Chem Int Ed Engl, 2011. **50**(29): p. 6460-8.

464. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain.* Science, 2009. **324**(5929): p. 929-30.
465. Nestor, C., et al., *Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA.* Biotechniques, 2010. **48**(4): p. 317-9.
466. Jin, S.G., S. Kadam, and G.P. Pfeifer, *Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine.* Nucleic Acids Res, 2010. **38**(11): p. e125.
467. Huang, Y., et al., *The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing.* PLoS One, 2010. **5**(1): p. e8888.
468. Song, C.X., et al., *Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine.* Nat Biotechnol, 2011. **29**(1): p. 68-72.
469. Szwagierczak, A., et al., *Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA.* Nucleic Acids Res, 2010. **38**(19): p. e181.
470. Booth, M.J., et al., *Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution.* Science, 2012. **336**(6083): p. 934-7.
471. Pieper, H.C., et al., *Different methylation of the TNF-alpha promoter in cortex and substantia nigra: Implications for selective neuronal vulnerability.* Neurobiol Dis, 2008. **32**(3): p. 521-7.
472. Barrachina, M. and I. Ferrer, *DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain.* J Neuropathol Exp Neurol, 2009. **68**(8): p. 880-91.
473. Jowaed, A., et al., *Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains.* J Neurosci, 2010. **30**(18): p. 6355-9.
474. Matsumoto, L., et al., *CpG demethylation enhances alpha-synuclein expression and affects the pathogenesis of Parkinson's disease.* PLoS One, 2010. **5**(11): p. e15522.
475. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.
476. Tost, J. and I.G. Gut, *DNA methylation analysis by pyrosequencing.* Nat Protoc, 2007. **2**(9): p. 2265-75.
477. Frommer, M., et al., *A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1827-31.
478. Clark, S.J., et al., *High sensitivity mapping of methylated cytosines.* Nucleic Acids Res, 1994. **22**(15): p. 2990-7.

479. Clark, S.J., et al., *DNA methylation: bisulphite modification and analysis.* Nat Protoc, 2006. **1**(5): p. 2353-64.
480. Hackenberg, M., et al., *CpGcluster: a distance-based algorithm for CpG-island detection.* BMC Bioinformatics, 2006. **7**: p. 446.
481. Wojdacz, T.K., L.L. Hansen, and A. Dobrovic, *A new approach to primer design for the control of PCR bias in methylation studies.* BMC Res Notes, 2008. **1**: p. 54.
482. Warnecke, P.M., et al., *Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA.* Nucleic Acids Res, 1997. **25**(21): p. 4422-6.
483. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps.* Bioinformatics, 2005. **21**(2): p. 263-5.
484. Zheng, G., et al., *Testing association for markers on the X chromosome.* Genet Epidemiol, 2007. **31**(8): p. 834-43.
485. Heinemeyer, T., et al., *Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL.* Nucleic Acids Res, 1998. **26**(1): p. 362-7.
486. Bryne, J.C., et al., *JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update.* Nucleic Acids Res, 2008. **36**(Database issue): p. D102-6.
487. Peplonska, B., et al., *TOMM40 and APOE common genetic variants are not Parkinson's disease risk factors.* Neurobiol Aging, 2013. **34**(8): p. 2078 e1-2.
488. Hardy, G.H., *Mendelian Proportions in a Mixed Population.* Science, 1908. **28**(706): p. 49-50.
489. Wang, J. and S. Shete, *Testing departure from Hardy-Weinberg proportions.* Methods Mol Biol, 2012. **850**: p. 77-102.
490. Rousset, F., *genepop'007: a complete re-implementation of the genepop software for Windows and Linux.* Mol Ecol Resour, 2008. **8**(1): p. 103-6.
491. Hosking, L., et al., *Detection of genotyping errors by Hardy-Weinberg equilibrium testing.* Eur J Hum Genet, 2004. **12**(5): p. 395-9.
492. Zou, G.Y. and A. Donner, *The merits of testing Hardy-Weinberg equilibrium in the analysis of unmatched case-control data: a cautionary note.* Ann Hum Genet, 2006. **70**(Pt 6): p. 923-33.
493. Teo, Y.Y., et al., *On the usage of HWE for identifying genotyping errors.* Ann Hum Genet, 2007. **71**(Pt 5): p. 701-3; author reply 704.
494. Lander, E.S., *The new genomics: global views of biology.* Science, 1996. **274**(5287): p. 536-9.
495. Morley, J.F., et al., *Genetic influences on cognitive decline in Parkinson's disease.* Mov Disord, 2012. **27**(4): p. 512-8.
496. Seto-Salvia, N., et al., *Dementia risk in Parkinson disease: disentangling the role of MAPT haplotypes.* Arch Neurol, 2011. **68**(3): p. 359-64.

497. Seto-Salvia, N., et al., *Glucocerebrosidase mutations confer a greater risk of dementia during Parkinson's disease course.* Mov Disord, 2011. **27**(3): p. 393-9.
498. Choi, J.M., et al., *Analysis of mutations and the association between polymorphisms in the cerebral dopamine neurotrophic factor (CDNF) gene and Parkinson disease.* Neurosci Lett, 2011. **493**(3): p. 97-101.
499. von Otter, M., et al., *Association of Nrf2-encoding NFE2L2 haplotypes with Parkinson's disease.* BMC Med Genet, 2010. **11**: p. 36.
500. Yuan, Y., et al., *The association between glycogen synthase kinase 3 beta polymorphisms and Parkinson's disease susceptibility: a meta-analysis.* Gene, 2013. **524**(2): p. 133-8.
501. Nuytemans, K., et al., *Progranulin variability has no major role in Parkinson disease genetic etiology.* Neurology, 2008. **71**(15): p. 1147-51.
502. Ticozzi, N., et al., *Mutational analysis of TARDBP in neurodegenerative diseases.* Neurobiol Aging, 2009. **32**(11): p. 2096-9.
503. Kabashi, E., et al., *No TARDBP mutations in a French Canadian population of patients with Parkinson disease.* Arch Neurol, 2009. **66**(2): p. 281-2.
504. van Blitterswijk, M., et al., *Mutational analysis of TARDBP in Parkinson's disease.* Neurobiol Aging, 2012. **34**(5): p. 1517 e1-3.

VII. BIBLIOGRAPHY.

RESUMEN.

RESUMEN.

La enfermedad de Parkinson (PD) es progresiva y neurodegenerativa. Se caracteriza clínicamente por bradiquinesia, temblor en reposo, rigidez e inestabilidad postural y anatomopatológicamente por degeneración nigral y por la presencia de numerosos cuerpos de Lewy (LB) en las neuronas supervivientes.

La degeneración nigral hace referencia a la muerte de neuronas en la *substantia nigra*, concretamente a las que contienen dopamina en la *substantia nigra pars compacta* (SNpc). La *substantia nigra* se encuentra en el mesencéfalo y forma, junto a otros cuatro núcleos (caudado, *putamen*, globo pálido y núcleo subtalámico), los ganglios basales. Ellos, con el córtex y el tálamo, forman un circuito que se encarga de controlar la ejecución de los movimientos voluntarios. Concretamente afectan al inicio y la planificación.

La degeneración nigral podría, por tanto, explicar el enlentecimiento en la ejecución de movimientos, o bradiquinesia, característico de la enfermedad. Sin embargo, no explicaría el resto de síntomas.

Además de los cuatro rasgos motores típicos, los enfermos de PD pueden sufrir otros síntomas motores. Asimismo es muy común que también se vean afectados por síntomas no motores.

Dicho síntomas se dividen en alteraciones sensoriales, cognitivas, disfunciones autónomas y desórdenes del sueño, y podrían deberse a daños en otras regiones del cerebro, que también se ven afectadas como consecuencia del progreso de la enfermedad.

RESUMEN.

Se han postulado factores genéticos y/o ambientales como responsables de la enfermedad de Parkinson. Además, procesos tales como el estrés oxidativo, la agregación de proteínas, la disfunción mitocondrial o la inflamación, entre otros, se encuentran alterados en los enfermos, aunque falta aclarar si ello es causa o consecuencia de la enfermedad.

El 90% de los casos corresponden a formas esporádicas o idiopáticas de PD, de causa desconocida, mientras que el 10% restante, las denominadas formas familiares o monogénicas[110], son causados por mutaciones en alguno de los siguientes cinco genes:

1. *SNCA* (*PARK1, PARK4*), 4q22.1. Este gen codifica para la proteína α-sinucleína, αSyn, que no tiene estructura en su forma nativa y de la cual se desconoce su función. Dicha proteína es uno de los componentes de los LB.
   Cambios en su secuencia, actualmente sólo hay tres mutaciones descritas (p.A30P, p.E46K, p.A53T), o en su cantidad, se han descrito enfermos con duplicaciones o triplicaciones de regiones que contienen el gen, originan PD con herencia autosómica dominante.
   Además, existen dos polimorfismos ampliamente estudiados como posibles factores de susceptibilidad: Rep1, del que se desconoce qué alelo podría estar implicado, y rs356219, cuyo alelo G es considerado de riesgo.

2. *PRKN* (*PARK2*), 6q26. Mutaciones, deleciones, duplicaciones y otros cambios en la secuencia del gen conllevan la pérdida de función de la proteína. Dichas alteraciones, presentes en homocigosis o en heterocigosis combinada, se han descrito en

[110] Las formas familiares de la enfermedad se diferencian de las esporádicas en que su inicio suele ser más temprano.

enfermos de PD siguiendo una herencia autosómica recesiva. Sin embargo, también hay portadores heterocigotos que han desarrollado la enfermedad, lo que sigue siendo un controvertido aspecto de estudio a día de hoy, así como la poca o nula presencia de LB en el cerebro de los enfermos.

La proteína parkina participa en el mecanismo de degradación de proteínas por el proteasoma. Concretamente es una enzima E3 que marca con ubiquitina a la proteína que debe ser degradada (K-48 poliubiquitinación). Además, participa en otros procesos, mediante monoubiquitinación o K-63 poliubiquitinación, como la regulación de la endocitosis o la formación del agregosoma.

3. *PINK1* (*PARK6*), 1p36.12. Los portadores homocigotos o heterocigotos combinados de mutaciones en este gen se ven afectados por la enfermedad de Parkinson, pero también algunos heterocigotos la han desarrollado, lo cual contradice la herencia autosómica recesiva observada en las familias. PINK1 es una quinasa que, junto con parkina, uno de sus sustratos, se encarga de marcar aquellas mitocondrias dañadas para que sean eliminadas por mitofagia. Cuando alguno de los dos genes está mutado, este proceso se ve bloqueado.

4. *DJ-1* (*PARK7*), 1p36.23. Son raras las mutaciones en este gen. A pesar de ello, hay familias en las que se han heredado en homocigosis o en heterocigosis combinada de forma autosómica recesiva. DJ-1 forma homodímeros y participa en mecanismos de protección frente a estrés oxidativo, posiblemente junto a Nrf2 y/o SOD1. Todo ello se pierde cuando está mutado.

5. *LRRK2* (*PARK8*), 12q12. Este gen codifica para una proteína, denominada dardarina, con varios dominios funcionales: uno GTPasa (ROC), otro quinasa y cuatro de interacción con proteínas como son ARM, ANK, LRR y WD40. Se desconoce la función de dardarina, pero se sabe que la actividad GTPasa controla la actividad quinasa, de modo que es necesario que se una GTP en ROC para que la proteína pueda fosforilar sus sustratos. La actividad GTPasa es independiente de la actividad quinasa. Si forma o no dímeros, así como las propiedades que puedan tener, es debatido a día de hoy.

   Solo mutaciones puntuales, ni deleciones ni duplicaciones, se han descrito en este gen. Más de 40 hasta el momento, pero solo 6 se consideran patogénicas y causantes de formas autosómicas dominantes de PD: p.N1437H, p.R1441C, p.R1441G, p.Y1699C, p.G2019S, p.I2020T. Las más comunes, y descritas tanto en casos familiares como en esporádicos, son p.R1441C y p. R1441G, que se localizan en el dominio ROC y disminuyen la actividad GTPasa de la proteína, y p.G2019S, que se localiza en el dominio quinasa y aumenta dicha actividad, tiene penetrancia incompleta y dependiente de la edad y es la mutación causante de PD más común.

   Además, hay dos polimorfismos, p.G2385R y p.R1628P, que son factores de riesgo en población asiática.

Existen otros *loci* que también se han relacionado con la enfermedad de Parkinson, pero hay resultados controvertidos sobre su influencia:

- *PARK3* (2p13) se identificó en familias con posible herencia autosómica dominante pero penetrancia incompleta y fenotipo amplio y no pudo ser replicado en estudios posteriores.

- *PARK5* (*UCHL1*; 4p14) participa en el mecanismo de degradación de proteínas por el proteasoma como enzima desubiquitinante. La mutación p.I93M, que causa una ganancia de función en la proteína, probablemente patogénica aunque rara, y el polimorfismo p.S18Y, posible factor de riesgo, se han estudiado sin llegar a conclusiones concluyentes.
- *PARK9* (*ATP13A2*; 1p36) ocasiona el síndrome de Kufor-Rakeb, de herencia autosómica recesiva y fenotipo solapante con la PD. Por ello se ha estudiado la frecuencia de mutaciones en enfermos de Parkinson familiares y esporádicos, obteniendo mayoritariamente resultados que descartan la relación entre *ATP13A2* y PD.
- *PARK10* (1p32) fue identificado en el estudio de familias islandesas con PD de inicio tardío, pero estudios posteriores no han logrado esclarecer qué gen sería el responsable, si lo hubiera.
- *PARK11* (*GIGYF2*; 2q36-37) y *PARK12* (Xq21-25) fueron postulados tras un extenso estudio de ligamiento, pero, como en el caso anterior, poco más se ha avanzado en su identificación. En el caso de *PARK11*, al obtenerse el valor más alto de ligamiento en el gen *GIGYF2*, éste ha sido estudiado en individuos con formas familiares de PD de diversas poblaciones, pero las conclusiones son contrapuestas y, como mucho, podría hablarse de causa rara de PD.
- *PARK13* (*HtrA2*; 2p13.1) codifica para una serín proteasa mitocondrial. Una mutación en dicho gen origina parkinsonismo y neurodegeneración en un modelo murino y, por ello, se consideró como candidato a explicar la enfermedad de Parkinson. Sin embargo, los estudios genéticos llevados a cabo no apuntan en dicha dirección.
- *PARK14* (*PLA2G6*; 22q13.1) está relacionado con fenotipos más complejos que incluyen parkinsonismo y otros síntomas como

distonía o demencia. Pero también hay descritos individuos portadores de mutaciones en dicho gen asintomáticos o con formas típicas de PD, lo que ha conllevado controversia sobre su implicación en PD.

- *PARK15* (*FBXO7*; 22q12.3) se ha encontrado mutado en personas que presentan parkinsonismo pero también signos piramidales.

El estudio a gran escala del genoma mediante GWAs ha supuesto la descripción de nuevos *PARK loci*: *PARK16* (1q32), *PARK17* (4p16.3; aunque también se propone 16q11.2) y *PARK18* (6p21.3; aunque también se propone 3q27.1). Sin embargo, la replicación de las regiones implicadas así como la identificación de los genes que en ellas se localizan causantes de la enfermedad es complicada.

Además, hay tres factores de susceptibilidad descritos en la enfermedad de Parkinson:

- *MAPT* (17q21.31), codifica para la proteína tau, una proteína de unión a microtúbulos, que promueve su formación, estabiliza su estructura y colabora en el transporte axonal a través de ellos. En las enfermedades clasificadas como taupatías, como la enfermedad de Alzheimer, forma depósitos en el cerebro denominados ovillos neurofibrilares.
  Múltiples y diversos estudios (analizando pacientes con diferentes edades de inicio, etnias, con formas familiares o con formas esporádicas) han observado que en los enfermos de Parkinson, el porcentaje de individuos portadores homocigotos del haplotipo H1 es mayor que en población sana.
- *GBA* (1q22) codifica para la enzima glucocerebrosidasa. Las personas que presentan mutaciones en homocigosis o en

heterocigosis compuesta, desarrollan una enfermedad lisosomal denominada de Gaucher. Algunos de los enfermos además presentan parkinsonismo. Se ha observado que el porcentaje de enfermos de PD entre sus parientes, que suelen ser portadores heterocigotos de mutaciones en el gen, es más elevado que en la población general y que los enfermos de PD tienen, con mayor frecuencia, mutaciones en GBA. De entre las múltiples mutaciones descritas, las más comunes son p.N370S y p.L444P. Por ello, mutaciones en este gen se consideran un factor de riesgo de PD.

- *NR4A2* (2q24.1) es un factor de transcripción esencial para el desarrollo y la supervivencia de las neuronas dopaminérgicas. Dada la imposibilidad de replicar los estudios que inicialmente concluyeron que c.-291delT y c.-245T>G eran causantes de la enfermedad y que 7048G7049 (c.1361+16insG) era un factor de susceptibilidad, en la actualidad, se duda de su relación con PD.

RESUMEN.

Nuestro objetivo principal en este trabajo ha sido la identificación de factores genéticos de susceptibilidad para la enfermedad de Parkinson en población española. Además de ello, analizamos la posible relación de una familia de genes poco estudiada (*ARMCX*) con el riesgo de padecer la enfermedad así como la posible influencia genética en la aparición de demencia en los enfermos de PD.

Cabe destacar que en todo el proceso experimental llevamos a cabo controles de calidad y comprobaciones para confirmar la validez de los resultados.

Para llevar a cabo estos tres estudios genéticos, empleamos diferentes técnicas de genotipado basadas en la amplificación del DNA:

1. para detectar inserciones o deleciones, tras dicha amplificación solo era necesario observar el tamaño de la región amplificada mediante electroforesis horizontal en gel de agarosa;
2. PCRs específicas de alelo (AS-PCR), en las que para analizar un polimorfismo se llevaban a cabo dos amplificaciones paralelas pero con diferente combinación de primers: se empleaba un primer común en ambos casos pero, la pareja, en la cadena opuesta, era complementaria en 3´, según el caso, a una de las dos variantes posibles del polimorfismo a analizar. Ambos productos de PCR eran posteriormente analizados mediante electroforesis horizontal en gel de agarosa para determinar si uno o las dos amplificaciones había tenido éxito y, determinar, por tanto, el genotipo.

3. RFLPs, o polimorfismos de longitud de fragmento de restricción, es decir, polimorfismos que crean o destruyen sitios de reconocimiento de enzimas de restricción y, por tanto, cuando regiones de DNA que los contienen son amplificadas y digeridas con dichas enzimas de restricción, muestran un patrón de bandas determinado, en geles de agarosa sometidos a electroforesis horizontal, según la variante del polimorfismo presente.
4. finalmente, empleamos la pirosecuenciación en algunos casos concretos, así como electroforesis capilar en polimorfismos de longitud, es decir, en aquellos basados en el diferente número de repeticiones de determinadas secuencias, como dinucleótidos, u homopolímeros.

A pesar de las condiciones específicas en cada técnica de genotipado, el diseño de primers siguió unas normas generales para aumentar la eficiencia de la reacción de amplificación: la longitud de los primers se mantuvo entre 18 y 25 nucleótidos, se intentó que terminaran en 3´en C o en G, así como que no hubiera estructuras secundarias intra- e intermoleculares, que tuvieran una temperatura de fusión inferior a 65ºC y una diferencia inferior a 2ºC en la temperatura de fusión entre los dos miembros de la pareja.

En nuestro **primer estudio**, llevamos a cabo un análisis en población española de factores genéticos de susceptibilidad ampliamente reconocidos, como son el haplotipo H1/H2 de *MAPT* y la presencia de mutaciones en *GBA*, así como de otros que, a día de hoy, y tras múltiples estudios con conclusiones contradictorias, están en entredicho, como son el polimorfismo p.S18Y en *PARK5* y los cambios intrónicos en *NR4A2* (nosotros decidimos analizar uno no analizado previamente en el intrón

3). Además, también se determinó la posible influencia de dos variantes en el gen *TOR1A*, causante de la forma mayoritaria de distonía primaria hereditaria (la distonía es uno de los síntomas motores secundarios presentes en PD), así como el efecto del genotipo en *APOE* (el alelo ε4 es factor de riesgo totalmente reconocido en la enfermedad de Alzheimer) o de la longitud del homopolímero recientemente descrito en *TOMM40*, que se sospecha es el auténtico factor de riesgo para la enfermedad de Alzheimer, y no el alelo ε4.

Al incluir en el estudio individuos vascos y de otros puntos de España (Sevilla, Barcelona y Valencia), decidimos analizarlos por separado, ya que dado el aislamiento que ha mantenido la población vasca y su alta tasa de endogamia, los factores genéticos de susceptibilidad podían ser diferentes entre individuos vascos y no vascos.

A pesar de que en ninguna de las dos poblaciones estudiadas se obtuvieron resultados estadísticamente significativos (y, por tanto, se descartaron como factores genéticos de susceptibilidad para la enfermedad de Parkinson), para el polimorfismo p.S18Y (*PARK5*), el cambio intrónico en *NR4A2*, la longitud del homopolímero en *TOMM40* ni la mutación p.delE302/303 (*TOR1A*) (la cual ni siquiera se encontró en los individuos), observamos que sí había diferencias en el estudio genético entre ambas poblaciones en cuatro casos, para los que los resultados sí tuvieron significatividad estadística, indicando que los factores genéticos de susceptibilidad a la enfermedad podrían ser diferentes entre vascos y no vascos:

- Por lo que respecta al haplotipo H1/H2 (*MAPT*), mientras que en población no vasca, el riesgo de padecer PD de los individuos H1H1 era claramente superior, al de los no portadores, alrededor de dos veces mayor, confirmando la amplia mayoría de estudios

previos, en población vasca, no se obtuvo dicho resultado sino que únicamente se encontró una tendencia del haplotipo H1 (y del genotipo H1H1) a incrementar el riesgo, pero sin llegar a adquirir significatividad estadística.

- También se observó que las mutaciones p.N370S y p.L444P (*GBA*) eran muy frecuentes en población no vasca, confirmando los estudios previos en los que se considera un claro factor de riesgo. De hecho, alcanzaron valores estadísticamente significativos. Sin embargo, en población vasca, únicamente un enfermo presentó la mutación p.N370S, por lo que en dicha población no parece existir una influencia tan clara de estas mutaciones en el desarrollo de la enfermedad (en estos individuos puede ser que haya otras en este gen, que no fueron analizadas, que sean factores de riesgo).
- Únicamente en población no vasca, el polimorfismo p.D216H (*TOR1A*) mostró tendencia a aumentar el riesgo de padecer la enfermedad, concretamente los portadores del genotipo CC eran más frecuentes entre los casos (6 individuos) que entre los controles (ninguno). Dado que el genotipo CC es de muy baja frecuencia, consideramos que, como mucho, representaría un factor de riesgo raro y que probablemente no tenga mayor relevancia en la susceptibilidad a PD.
- Además, mientras que en población vasca, los portadores de un alelo ε4 en *APOE* mostraron tener más riesgo de padecer PD, en población no vasca fueron los portadores de un alelo ε2. Dichos resultados contradictorios son habituales en los estudios que pretenden determinar la relación existente entre *APOE* y PD: tanto el alelo ε4 como el ε2 han sido considerados factores de riesgo, e, incluso, algunos estudios han considerado que ninguno está relacionado con la enfermedad de Parkinson. Además, observamos que los individuos no vascos portadores del

genotipo H1H1 (*MAPT*) y a la vez de un alelo ε2 en *APOE* eran tres veces más susceptibles de padecer la enfermedad de Parkinson que aquellos que no portaban ninguno de los dos factores. Es decir, había un efecto combinado de ambos factores.

En el **segundo estudio** abordamos la determinación de factores genéticos de riesgo para el desarrollo de deterioro cognitivo. Prácticamente desde su comienzo, los enfermos de PD presentan un ligero deterioro cognitivo, el cual, tras décadas de padecer la enfermedad mayoritariamente evoluciona a demencia.

Para ello analizamos casos de PD de Navarra, los cuales fueron clasificados, tras los tests cognitivos llevados a cabo, en cognitivamente normales, con deterioro cognitivo leve o con demencia. Estos dos últimos incluían individuos afectados por procesos diferentes pero conectados de algún modo, como son la demencia y el deterioro cognitivo leve, por lo que, dado que presentaban características demográficas y genéticas similares, los agrupamos en un solo grupo denominado deterioro cognitivo.

También se incluyeron en el estudio individuos sanos. Por ello, se decidió investigar también factores genéticos de susceptibilidad a PD: el haplotipo H1/H2 (*MAPT*), el genotipo de *APOE*, la longitud del homopolímero en *TOMM40*, la presencia de mutaciones en *GBA* (p.N370S y p.L444P) y los polimorfismos Rep1 y rs356219 (*SNCA*). Nuevamente, se obtuvo que la longitud de la poly-T de *TOMM40* no podía considerarse relacionada con la enfermedad de Parkinson y que, por el contrario, los portadores del haplotipo H1 tenían un riesgo elevado de padecer la enfermedad. No se observó que ningún alelo de *APOE* estuviera diferentemente distribuido en casos y en controles (otra

vez los resultados para el genotipo de *APOE* fueron controvertidos), ni tampoco los alelos del microsatélite Rep1. Tampoco las mutaciones de *GBA* se encontraron en una frecuencia suficiente como para ser consideradas factores de riesgo, pero, por el contrario, sí se vio que el alelo G en rs356219 incrementaba el riesgo de padecer PD, como ya se había visto previamente en otros estudios.

En algunos casos, como *GBA*, podemos pensar que los resultados pudieron verse influenciados por únicamente estudiar 2 de las alrededor de 300 mutaciones descritas en el gen, por las diferencias que existían entre los grupos de casos con PD y de controles sanos en cuanto a edad media y porcentaje de hombres y mujeres, así como por su tamaño, aunque ello no fue inconveniente para que dos factores de riesgo ampliamente reconocidos (haplotipo H1 en *MAPT* y alelo G en rs356219, *SNCA*) originaran resultados estadísticamente significativos. Podemos pensar que el resto de factores o no influyen en la enfermedad o lo hacen de un modo menor y que, por tanto, dichas limitaciones alteraron la capacidad de determinar su riesgo.

Por lo que respecta a la influencia de dichas variantes en el proceso de deterioro cognitivo, aunque algunos grupos han observado que el haplotipo H1 y las mutaciones en *GBA* también son más frecuentes en los enfermos de PD que desarrollan demencia que en aquellos que no la padecen, no fue eso lo que observamos. A excepción del genotipo de *APOE*, el resto de polimorfismos y mutaciones no se encontraron distribuidos diferentemente entre los enfermos de PD cognitivamente normales y aquellos con deterioro cognitivo.

Otros estudios han observado que el alelo ε4 de *APOE* aumenta también el riesgo de padecer demencia durante PD; en nuestro caso, observamos que era el alelo ε2 el que aumentaba el riesgo de padecer deterioro

cognitivo. Nuevamente, en *APOE* obtenemos resultados contradictorios. Podemos considerar que las diferencias observadas podrían tener relación con el hecho de que nosotros no analizamos solo demencia sino deterioro cognitivo en general, aunque, como hemos anotado anteriormente, los grupos demencia y deterioro cognitivo leve no presentaban diferencias genéticas en estos factores.

Finalmente, analizamos mutaciones y polimorfismos en genes relacionados con otras enfermedades que cursan con demencia (*PSEN1, PSEN2, APP* y *CALHM1* en la enfermedad de Alzheimer; *HTT* en la enfermedad de Huntington; *PRNP* en la enfermedad de Creutzfeldt-Jakob; *TARDBP* y *GRN* en demencia frontotemporal, *ATP13A2* en el síndrome de Kufor-Rakeb), en aquellos que codifican para factores neurotróficos (*CDNF*, *MANF* -y *DOCK3*-, *BDNF*), en genes implicados en procesos relacionados con la patología de PD como el estrés oxidativo (*NFE2L2* y *KEAP1*), la autofagia mediada por chaperonas (*HSPA8* y *LAMP-2A*), el metabolismo de homocisteína (*MTHFR, MTR* y *CBS*) y *GSK3*β.

Encontramos que en algunos de ellos había diferencias estadísticamente significativas o casi, en las distribuciones alélicas, genotípicas o haplotípicas tanto cuando se compararon pacientes con controles como cuando se comparaban los pacientes de PD entre sí en función de su estado cognitivo, es decir, cognitivamente normales y con deterioro cognitivo. Incluso se vio que polimorfismos en el mismo gen mostraban significatividad en distintas comparaciones.

Pero no se obtuvo una tendencia clara en los resultados. Ello, junto con las limitaciones ya descritas en cuanto a las características de los grupos analizados hace que los resultados observados deban considerarse como tendencias que deberían ser replicadas en otras poblaciones.

RESUMEN.

El **tercer estudio**, pretendía, a pequeña escala, analizar si los genes de la familia *ARMCX* (1 a 6) podrían estar implicados en la enfermedad de Parkinson. Dichos genes se localizan en el cromosoma X, dentro de la zona delimitada como *PARK12*, lo cual podría conllevar un efecto del sexo en la PD. Los 6 genes de la familia están poco descritos, aunque se sabe que comparten ciertas características, como que son cortos, con toda la parte codificante en un intrón y que originan proteínas mitocondriales. Se analizaron 8 polimorfismos distribuidos a lo largo de la región que comprendía el cluster.

Se empleó en este ensayo, una subpoblación no vasca, de la previamente analizada, por lo que al igual que para la población no vasca global, nuevamente se concluyó que ni p.S18Y (*PARK5*), ni p.del302/303 (*TOR1A*), ni la longitud de la poly-T en *TOMM40* ni el cambio intrónico c.864+246C>T (*NR4A2*) podían considerarse factores genéticos de susceptibilidad a PD. Además, los portadores H1H1 (*MAPT*) mostraron tener un riesgo elevado de desarrollar la enfermedad, pero, sin embargo, la presencia de mutaciones en *GBA* no alcanzó significatividad estadística: seguramente ello fue debido al reducido tamaño de muestra ya que sí eran más frecuentes en los casos que en los controles y al hecho de analizar únicamente dos mutaciones. Hubo otras dos diferencias: no se determinó ninguna influencia del polimorfismo p.D216H (*TOR1A*) en el riesgo de padecer la enfermedad, aunque esto era de esperar ya que el resultado obtenido en población no vasca global ya fue considerado un posible artefacto matemático y se observó que los portadores del genotipo ε3ε4 (*APOE*) tenían menor riesgo de PD. Nuevamente, los resultados para el genotipo de APOE fueron controvertidos.

Por lo que respecta a los polimorfismos analizados en los seis miembros del cluster *ARMCX*, únicamente se obtuvieron resultados casi estadísticamente significativos (p≈0.070) para uno de ellos, rs2858162, *ARMCX6*, para el cual el alelo C así como los genotipos que lo contenían eran más frecuentes en pacientes que en controles. Tanto en hombres como en mujeres se observó esa tendencia, aunque no alcanzó significatividad estadística, probablemente por el reducido tamaño de los grupos. Además, se identificó un haplotipo protector (p=0.034) frente a la enfermedad, TACTAC, y otro casi de riesgo (p=0.077), TACCAC, que se diferenciaban justamente en la posición correspondiente a rs2858162, poniendo de manifiesto la importancia de dicho polimorfismo. Al igual que lo observado en los genotipos y alelos, en el haplotipo, la variante C era de riesgo, mientras que la T era protectora.

A pesar de que la familia de genes *ARMCX* eran *a priori* unos buenos candidatos como factores genéticos de susceptibilidad en PD, por su localización celular, su estructura, su patrón de expresión así como su función predicha, no encontramos resultados que apoyaran su implicación en el desarrollo de la enfermedad, más allá de los observados en uno de los marcadores. Nuestros resultados deberían ser replicados para, de ese modo, determinar y confirmar, si procede, la relevancia del polimorfismo rs2858162 en el desarrollo de PD así como la ausencia de la misma por parte de los otros miembros del cluster. Asimismo, dicha réplica determinaría si, al igual que se ha observado en nuestros resultados, el sexo no influye en el papel de rs2858162 ni en el de ninguno de los otros *ARMCX*.

Finalmente, abordamos el objetivo principal desde una perspectiva epigenética y estudiamos si los niveles de metilación de determinados genes podrían verse alterados por la enfermedad y las posibles consecuencias que ello tendría.

La epigenética es el estudio de los cambios heredados mitóticamente o meióticamente que no pueden ser explicados por cambios en la secuencia de DNA. Explica, por ejemplo, porqué, teniendo el mismo genoma, existen diferentes tipos celulares en un individuo. Cada individuo posee un genoma pero cientos o incluso miles de epigenomas que varían con el tiempo y los factores ambientales.

La epigenética incluye las modificaciones de histonas, la metilación del DNA y procesos mediados por RNA.

La metilación del DNA está generalmente asociada con un estado inactivo de la cromatina y, por tanto, con represión transcripcional. Ello podría explicarse bien porque hay proteínas que reconocen el DNA metilado y se unen a él, y, posteriormente, reclutan modificadores de la cromatina para establecer un entorno reprimido, o por la imposibilidad de los factores de transcripción de unirse al DNA metilado lo que conllevaría que la transcripción se inhibiera.

En los mamíferos, la metilación sucede siempre sobre los dinucleótidos CpG. Dichos dinucleótidos, se acumulan en las islas CpG, las cuales se definen por tener una longitud mínima de 200 nucleótidos, un contenido en C+G igual o superior a 0.5 y una relación de dinucleótidos observada/esperada del 0.6 o mayor. Sin embargo, la mayoría de dinucleótidos en las islas no están metilados y sí lo están aquellos distribuidos por el resto del genoma.

RESUMEN.

Según la presencia o no de islas CpG, los genes se clasifican en

- GC rich: todos los genes housekeeping así como la mitad de los genes específicos de tejido tienen una isla CpG en su promotor y bajos niveles de metilación.
- GC poor: la otra mitad de los genes específicos de tejido no tienen isla CpG en el promotor pero los dinucleótidos presentes están muy metilados.

Para el **estudio epigenético**, se llevó a cabo inicialmente un estudio *in silico* de la posición de las islas CpG en los 5 genes a analizar. Dichos genes son aquellos involucrados en las formas familiares de la enfermedad de Parkinson: *DJ-1*, *LRRK2*, *PINK1*, *PRKN*, *SNCA*. Como era de esperar, dada la expresión ubicua de los 5 y ser, por tanto, genes CG rich, únicamente tenían una isla cada uno y localizada en su promotor.

Tras su identificación procedimos a diseñar cebadores que amplificaran dichas islas, teniendo en cuenta los cambios que existirían en la secuencia de DNA tras su tratamiento con bisulfito.

El tratamiento con bisulfito es la técnica más empleada para el estudio de los niveles de metilación de DNA ya que modifica las citosinas no metiladas convirtiéndolas en uracilos, que son reemplazados por timinas, mientras que las citosinas metiladas no se ven modificadas. Dichos cambios en la secuencia pueden ser posteriormente detectados por diferentes métodos, en nuestro caso, tras la amplificación, por pirosecuenciación.

Cabe destacar que el diseño de los primers tuvo en cuenta tanto los requisitos necesarios para amplificar DNA tratado con bisulfito[111] como las limitaciones de la técnica de pirosecuenciación[112]. Todo ello, junto con la gran longitud de las islas CpG predichas, determinó que optáramos por estudiar regiones de la isla en vez de toda entera. Dichas regiones se encontraban antes, en y tras el punto de inicio de transcripción de cada gen.

Antes de comenzar con el estudio de diferentes tejidos cerebrales de personas sanas y enfermas de Parkinson, llevamos a cabo dos pruebas:

1. En primer lugar analizamos los niveles de metilación de DNA en distintos tejidos procedentes de un mismo individuo para conocer aproximadamente el rango en el que se encontrarían nuestros valores así como para probar todo el diseño experimental. Como era de esperar, se observaron bajos niveles de metilación.
2. Posteriormente analizamos sangre de 10 enfermos de PD y los valores estuvieron en el mismo rango.

A continuación estudiamos, en 5 individuos sanos y en 5 pacientes de PD, los niveles de metilación de DNA en *substantia nigra*, córtex parietal y córtex occipital. Se analizó, en cada ensayo tanto los valores individuales

[111] Longitud de los primers entre 18 y 30 nucleótidos; ausencia de estructuras secundarias intra- e intermoleculares; diferencia inferior a 2ºC en la temperatura de fusión entre los dos miembros de la pareja; fragmento amplificado de entre 100 y 500 nucleótidos; inclusión de un número limitado de dinucleótidos CG, no más de 2, localizados lo más lejos posible del extremo 3´en la secuencia del primer; e inclusión de Ts procedentes de la conversión de Cs no metiladas en, o lo más cerca posible, del extremo 3´del primer.

[112] Las reacciones no pueden ser más largas de 50 nucleótidos y no deben incluir homopolímeros de longitud mayor a 4.

por posición como globales y se obtuvieron diferencias estadísticamente significativas entre ambos grupos tras el análisis con tests no paramétricos (test de Mann-Whitney). Aunque la mayoría de los resultados estadísticamente significativos o casi mostraban que los niveles de metilación en enfermos eran menores que en controles, no se observó una tendencia general en los resultados en cuanto a los niveles de metilación.

El hecho de haber analizado tan pocos individuos así como los bajos niveles de metilación obtenidos y la elevada desviación estándar nos hacen considerar estas diferencias como posibles indicios que deberían ser replicados en poblaciones mayores. Además, se debería analizar la relación existente entre los niveles de metilación observados y los niveles de expresión génica.

Además, no solo se encontraron diferencias en la *substantia nigra*, sino también en córtex parietal y córtex occipital, los cuales no se han descrito hasta la fecha como regiones afectadas en la enfermedad de Parkinson. Incluso, en algunos casos, el resultado era compartido entre varios tejidos.

Esto también se ha observado en alguno de los pocos estudios llevados a cabo sobre metilación de DNA en enfermos de PD.

Nuestro estudio se diferencia de los previos principalmente en las regiones analizadas así como en la metodología de cuantificación de la metilación (pirosecuenciación) y en la no observación de niveles de metilación inferiores generalizados en los pacientes respecto a los controles.

Como parte final de este estudio, decidimos hacer una predicción *in silico* para determinar si dichas posiciones (elegimos aquellas en las que se observaron diferencias significativas así como aquellas con valores de $p<0.1$ en *substantia nigra*) eran puntos de unión de factores de transcripción. Las alteraciones en los niveles de metilación podrían conllevar cambios en la unión de determinados factores de transcripción y, consecuentemente, se podrían producir cambios en la expresión de dichos genes, lo que supondría un nuevo mecanismo patogénico en los enfermos de PD.

De entre los candidatos propuestos, únicamente Sp1 parece adecuado ya que los demás, o no se expresaban en cerebro o no incluían a los genes de estudio entre sus dianas. Este factor de transcripción es ubicuo y se une a regiones ricas en GC. Por lo tanto, es habitual en las islas CpG. Además, participa en numerosos procesos celulares como diferenciación celular, apoptosis o remodelación de la cromatina. Interacciona, entre otros, con HDAC1 y DNMT1.

Se deberían llevar a cabo estudios *in vitro* para confirmar si, efectivamente, Sp1 se une en las secuencias predichas y cuál es la influencia de los niveles de metilación en dicha unión.

RESUMEN.

1. Hemos confirmado que, al menos en población no vasca, mutaciones en *GBA* y el genotipo H1H1 en *MAPT* son claros factores genéticos de susceptibilidad que aumentan el riesgo de desarrollar la enfermedad de Parkinson. Su efecto en vascos es menos importante, indicando que puede haber diferencias genéticas entre los dos grupos de población española en cuanto a su riesgo de padecer PD.
   APOE permanece como un factor de riesgo controvertido.

2. A pesar de las limitaciones de nuestro studio, derivadas de las características de los individuos analizados, hemos observado algunos indicios para variantes previamente no relacionadas con PD, o, al menos, no claramente, en relación con su influencia en el riesgo de padecer PD y/o el estado cognitivo durante la enfermedad. Representan importantes aspectos a considerar en futuros estudios.

3. Basándonos en nuestros resultados, hemos concluido que los miembros de la familia génica *ARMCX* no pueden ser considerados factores genéticos de susceptibilidad para la enfermedad de Parkinson. Únicamente el polimorfismo rs2858162 (*ARMCX6*) mostró alguna tendencia que debería ser analizada con mayor profundidad.

4. Y finalmente, en nuestro estudio piloto epigenético, hemos encontrado diferencias entre casos y controles en los niveles de metilación de DNA en algunos dinucleótidos CpG concretos localizados en los promotores (predichos *in silico*) de genes responsables de formas familiares de la enfermedad de Parkinson, aunque no solo en *substantia nigra.* Además, dichos

CpG forman parte de sitios de unión (también predichos *in silico*) del factor de transcripción Sp1 y, consecuentemente, dependiendo de su nivel de metilación, podrían alterar su unión y la expresión génica.

RESUMEN.

Genetics and epigenetics in Parkinson´s disease. (Annexes)

Luis Navarro Sánchez.

Thesis supervisor: Jordi Pérez Tur.
Unitat de Genètica Molecular,
Institut de Biomedicina de València, CSIC.

Doctoral studies: Biotechnology.
Facultat de Ciències Biològiques,
Universitat de València.

INDEX.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms.

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
*APOE** (1)	alleles ε2, ε3 and ε4 are determined by rs429358 and rs7412 (in both cases C>T)	Exon 4 (NM_000041)	F: GGCACGGCTGTCCAAGGA R: CGGGCCCCGGCCTGGTACAC	231	RFLP	2	*Afl*III and *Hae*II
PCR conditions: 95ºC 5´; 5 cycles (95ºC 30", 60ºC 30", 72ºC 45"); 20 cycles (95ºC 30", 60ºC→50ºC[1] 30", 72ºC 45"); 10 cycles (95ºC 30", 50ºC 30", 72ºC 45"); 72ºC 5´; 4ºC ∞							
APP (2)	rs463946 (C>G)	≈3000pb upstream from the 5´ (NM_000484)	F: TACAACAGCATCCCCATCC R: GCAGAGTAGTGGGCAACATC	433	RFLP C: 250+180 G:433	1.5	*Xba*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 62ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							
ARMCX1 (4)	rs6616255 (A>G)	Exon 4 5´UTR (NM_016608)	F: GTGGCAGGGAACTTTAACG R: GCTCCTTTCACAGTCTCAACC	545	RFLP A:545 G:180+365	1.5	*Pvu*II
PCR conditions: 95ºC 10´; 38 cycles (95ºC 30", 52ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							

[1] -0.5ºC/cycle. i.e. TD60-50. The purpose of touchdown PCR is to improve the amplification efficiency in the first few cycles at a relatively high temperature without decreasing the yield of PCR amplification as the annealing temperature drops later.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
ARMCX1 (4)	rs1044275 (C>T) p.G103G	Exon 4 (NM_016608)	F: TCGGGTTAAGAGATTTGTCC R: ATCATTTGTTCGCTCCAGG	687	RFLP C:687 T:340+350	1.2	*Xba*I
PCR conditions: 95°C 10´; 38 cycles (95°C 30", 52°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
ARMCX2 (3)	rs5951282 (C>T)	Intron 5 (NM_177949)	F: CCATGTCTGTACTGCCTCTCG R: GCCCCAACTACTGATTCGG	776	RFLP C:776 T:175+600	1.5	*Tsp590*I
PCR conditions: 95°C 10´; 38 cycles (95°C 30", 52°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
ARMCX3 (3)	rs6995 (A>G)	Exon 5 3´UTR (NM_177948)	F: AATGAATATCACTACTTGTTCTG R: TTAACATACAATTTTATTG***T***TGA A mismatch (***T*** instead of A) was introduced to create the recognition site.	178	RFLP A:178 G:160+20	2.5	*Hinc*II
PCR conditions: 95°C 10´; 42 cycles (95°C 30", 50°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
ARMCX4 ** (3)	rs2179670 (T>C)	Intergenic region	F: GGAGTGAAGGAGTTGGGTGTC R: ATCAGTGAATTAGGAAACAC***G***GT A mismatch (***G*** instead of T) was introduced to create the recognition site.	219	RFLP T:219 C:35+185	2.5	*Hph*I
PCR conditions: 95°C 10´; 42 cycles (95°C 30", 50°C 30", 72°C 1´); 72°C 10´; 4°C ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
ARMCX4 ** (3)	rs6523506 (T>G)	Intergenic region	F: GGTTGGCATACTGGTAA***G***A R: CAGCTTTTTGTTATTCTTTTC A mismatch (***G*** instead of A) was introduced to create the recognition site.	150	RFLP T:150 G:140+10	2.5	*Mnl*I
PCR conditions: 95ºC 10´; 35 cycles (95ºC 30", 52ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
ARMCX5 (4)	rs2235827 (C>T)	Intron 1 (NM_022838)	F: TTGTCTAGGCATTTCGTCCC R: CGTGAGAGCAGGACAAGCAG	582	RFLP C:210+390 T:582	1.2	*Sty*I (*Eco130*I)
PCR conditions: 95ºC 10´; 38 cycles (95ºC 30", 54ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
ARMCX6 (4)	rs2858162 (C>T)	Promoter (NM_019007)	F: TGAATGAAAGGCAGGCTAATC R: TTCTCCTGGCTGAACTCATA***A***T A mismatch (***A*** instead of G) was introduced to create the recognition site.	185	RFLP C:185 T:160+25	2.5	*Ssp*I
PCR conditions: 95ºC 10´; 38 cycles (95ºC 30", 52ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
ATP13A2 (3)	rs2871776 (A>G)	Intron 8 (NM_022089)	F: CTCCTCCTTCTGATGGCTTG R_T: TCTT***T***GCCTCCAATCCCTA**T** R_C: TCTT***T***GCCTCCAATCCCTA**C** A mismatch (***T*** instead of G) was created to disrupt secondary structures. R: GGACTGTTGGGAGCATGAAA	356 F+R for sequencing: 457	AS-PCR	-	-
AS-PCR conditions: 95°C 10´; 40 cycles (95°C 30", 63°C 30", 72°C 45"); 72°C 10´; 4°C ∞ **sequencing conditions:** 95°C 10´; 35 cycles (95°C 30", 63°C 30", 72°C 45"); 72°C 10´; 4°C ∞							
ATP13A2 (3)	rs6684770 (C>T)	Intron 11 (NM_022089)	F: AGTTTCACTATGTTGCCCAGG R: ATTTATGCTGCACCTCTC***G***C A mismatch (***G*** instead of A) was introduced to create the recognition site.	289	RFLP C: 100+170+20 T: 100+190	2.5	*Hha*I
PCR conditions: 95°C 10´; 35 cycles (95°C 30", 57°C 30", 72°C 30"); 72°C 10´; 4°C ∞							
ATP13A2 (3)	rs4920608 (G>A)	Intron 16 (NM_022089)	F: GGTGGTGCATGCCTGTAGTC R: AGGGAAGTTTGGTGTCTGGG	768	RFLP G: 30+130+ 160+240+ 200 A: 30+130+ 160+440	1.5	*Bbv*CI
PCR conditions: 95°C 10´; 40X (95°C 30", 64°C 30", 72°C 1´); 72°C 10´; 4°C ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
BDNF (1)	rs6265 (G>A) p.V66M	Exon 2 (NM_170731)	F: AAAGAAGCAAACATCCGAGGACAAG R: ATTCCTCCAGCAGAAAGAGAAGAGG [1]	274	RFLP G:217+57 A:140+77+57	1.5	*Nla*III
PCR conditions: 94°C 2´; 35 cycles (94°C 1´, 55°C 2´, 72°C 2´); 72°C 4´; 4°C ∞ [1]							
CALHM1 (2)	rs2986017 (C>T) p.P86L	Exon 1 (NM_001001412)	F: GAAGAGTGGAAGCGGCC***A***C R:GACGGCCACCCAGACGACA A mismatch (***A*** instead of G) was introduced to create the recognition site.	114	RFLP C:114 T:94+20	2.5	*Bsr*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 65°C 1´, 72°C 45"); 72°C 10´; 4°C ∞							
CBS[2] (1)	c.844ins68 68bp insertion	Exon 10 (NM_000071)	F:ATAGAATATCGAGGCATGTCCAGGCG R:TGGGGCCCAGGGTCAGCCAGGCTCC	282 (350 with the insertion)	PCR	-	-
PCR conditions: 95°C 10´; 35 cycles (95°C 45", 60°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
CDNF (3)	rs7094179 (G>T)	Intron 1 (NM_001029954)	F: GAAGAAGGGCACATAA***C***TGA R: CTACTTTGCTCAGTGTCAAGG A mismatch (***C*** instead of G) was introduced to create the recognition site.	196	RFLP G:166+20+10 T:186+10	2.5	*Dde*I (*HpyF3*I)
PCR conditions: 95°C 10´; 38 cycles (95°C 30", 51°C 30", 72°C 30"); 72°C 10´; 4°C ∞							

[2] Despite that the insertion is located in an exon, it has no effect on protein sequence.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
CDNF (3)	rs7099185 (T>C)	Intron 3 (NM_001029954)	F_T: CAGTGTATAGCTGACCCTC**T** F_C: AGTGTATAGCTGACCCTC**C** R: TAATCCCAGCATCTCAGG F: CCTTCAGCCTCTTATCTTG	413 F+R for sequencing: 731	AS-PCR	-	-
AS-PCR conditions: 95ºC 10´; 38 cycles (95ºC 30", 60ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞ **sequencing conditions:** 95ºC 10´; 35 cycles (95ºC 30", 54ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
DOCK3 *** (3)	rs4441646 (A>C)	Exon 53 3´UTR region (NM_004947)	F: ATGTGGTCAGGTTCAGCAGG R: ATCCTCAGCAGCCTTCCG	730	RFLP A:387+343 C:730	1.2	*Taq*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 66ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
GBA[3] (1)	p.N370S (A>G)	Exon 9 (NM_000157)	F: GCCTTTGTCCTTACCCT***C***G R:GACAAAGTTACGCACCCAA [2] A mismatch (***C*** instead of A) was introduced to create the recognition site.	105	RFLP A:105 G:89+16	2.5	*Xho*I
PCR conditions: 94ºC 10´; 5 cycles (94ºC 30", 57ºC 30", 72ºC 1´); 11 cycles (94ºC 30", 57ºC→52ºC[4] 30", 72ºC 1´); 22 cycles (94ºC 30", 52ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞ [2]							

[3] Although their correct description is p.N409S and p.L483P, for tradition, they remain named as p.N370S and p.L444P.

[4] -0.5ºC/cycle. i.e. TD57-52.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
GBA (1)	p.L444P (T>C)	Exon 10 (NM_000157)	F:CGTAACTTTGTCGACAGTCC R:TCCCAGACCTCACCATTG [2]	600	RFLP T:600 C:518+82	1.5	*Nci*I
PCR conditions: TD57-52: 94ºC 10´; 5 cycles (94ºC 30", 57ºC 30", 72ºC 1´); 11 cycles (94ºC 30", 57ºC→52ºC[5] 30", 72ºC 1´); 22 cycles (94ºC 30", 52ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞ [2]							
GRN (5)	rs11547442 (T>C) p.L46P	Exon 2 (NM_002087)	R:GCACCCTCCTATCCCCAG F: CTTGGTACTTTGCAGGCAGA	788	RFLP T:788 C:630+160	1.5	*Bsa*WI
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 62ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
GRN (5)	rs34975779 (G>-) p.V452V but the frameshift originates a premature STOP codon downstream	Exon 11 (NM_002087)	F: CCTTATCCCACCCCAGAG R: TTATGTTCCTGTCCCCTCAC	174	RFLP G:170 -:150+20	2.5	*Bgl*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 59ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							

[5] -0.5ºC/cycle. i.e. TD57-52.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
GRN (5)	rs1141754 (C>A) p.Q479K	Exon 12 (NM_002087)	F: GCTAAGCCCAGTGAGGGGAC R: CAGCCGGGCAGCAGT***A***CT A mismatch (***A*** instead of G) was introduced to create the recognition site.	98	RFLP C: 80+18 A:98	2.5	*Bsr*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 62ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							
GSK3β (2)	rs334558 (T>C)	Promoter (NM_002093)	F: TTTATAGACGCCCTCCCTTCGCTT R: TCCTTCCTTCCTTTGTCACTTGGC [3]	721	RFLP T:642+79 C:721	1.2	*Alu*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 63ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
GSK3β (2)	rs6438552 (T>C)	Intron 5 (NM_002093)	F: GCTTTTGGTGCCTTCTTAGGTGAC R: CGAAACATTGGGTTCTCCTCG [3]	287	RFLP T:267+20 C:173+94+20	1.5	*Hpy188*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 50ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							
HSPA8 (3)	rs1461496 (T>C)	Intron 6 (NM_006597)	F: ATTGAAACTGCTGGTGGAG R: GACAGTGCCTCCTTACCC	560	RFLP T: 100+460 C: 100+200+260	1.5	*BspCN*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 58ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
HSPA8 (3)	rs4936770 (G>A)	Intron 8 (NM_006597)	F: AATTCACTTGAGTCCTATGCC R: TCTCCCTGACGCAATCTG	414	RFLP G:200+220 A:414	1.5	*Tsp*RI
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 56ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							
HTT (3)	rs10015979 (A>G)	Intron 6 (NM_002111)	F: TGCTGTGAGTGAGTCTGTG R: ATCTGATTCCACTCTACCC	559	RFLP A:260+300 G:559	1.2	*Mse*I (*Tru1*I)
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 56ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
HTT (3)	rs363066 (T>G)	Intron 18 (NM_002111)	F: TCCAGGTTTTCAGTCAGTTG R: CATCTACCTAAACCACTCGG	848	RFLP T:848 G:400+430	1.2	*BsaX*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 56ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
HTT (3)	rs363096 (T>C)	Intron 34 (NM_002111)	F: CCTGTTAGCTTGATGTGTGC R: ATGATTGCCTCTGATTC***A***CT A mismatch (***A*** instead of C) was introduced to disrupt secondary structures.	136	RFLP T:136 C:115+20	2.5	*Bfa*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 54ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
HTT (3)	rs2298969 (A>G)	Intron 37 (NM_002111)	F: GGGGAGGGAAGGGAGTGAG R: CATGCTGCCAAGGG***G***TGC A mismatch (***G*** instead of A) was introduced to create the recognition site.	211	RFLP A:45+165 G:45+145+20	2	*Nla*IV
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 56ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
HTT (3)	rs110501 (T>C)	Intron 55 (NM_002111)	F: AGAGGTGGTTGTGGGTGTC R: GTATTTAGCACCACAGCCTC	912	RFLP T:400+500 C:912	1.2	*Tsp45*I (*Nmu*CI)
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 57ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
KEAP1 (4)	rs8113472 (G>T)	Intron 2 (NM_203500)	F: AGCACCATACCAACACCAG R: GGCACTCGTCTATGTAATCAG	471	RFLP G:230+240 T:471	1.5	*Kpn*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 57ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
LAMP-2A (3)	rs7057652 (G>T)	Intron 1 (NM_002294)	F: TTCTTCTGTGGTTATTGATTC R: GGCTCACTGCAACTTCCAC	564	RFLP G: 564 T: 460+100	1.5	*Mly*I (*Sch*I)
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 58ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
LAMP-2A (3)	rs42897 (A>G)	Intron 1 (NM_002294)	F: GAGTAAAAGTGGTCCCTGTG R: ATATCTGAGCAATCCACTGTC	802	RFLP A:802 G:430+370	1.2	*BseR*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 54°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
LAMP-2A (3)	rs42895 (C>A)	Intron 3 (NM_002294)	F: TTGATGAATGAGAAAATGC R: ATCCCTTTGTCCAGTATATC	494	RFLP C:494 A:350+150	1.5	*Bsr*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 50°C 30", 72°C 30"); 72°C 10´; 4°C ∞							
LAMP-2A (3)	rs42890 (T>G)	Intron 6 (NM_002294)	F: TGGGTTATGGAGAAGAAAG R: TATACACCTTCCACAGT***C***CT A mismatch (***C*** instead of A) was introduced to create the recognition site.	160	RFLP T:160 G:140+20	2.5	*Mnl*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 54°C 30", 72°C 30"); 72°C 10´; 4°C ∞							
LRRK2 (1)	rs33939927 (C>G) p.R1441G	Exon 31 (NM_198578)	F_C: GTGTCTTTCCCTCCAGGCT**C** F_G: GTGTCTTTCCCTCCAGGCT**G** R: CCCTTGTGATTGAATCACCAC	569	AS-PCR	-	-

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

AS-PCR conditions: 95ºC 10´; 35 cycles (95ºC 30", 62ºC 30", 72ºC 45"); 72ºC 10´; 4ºC ∞ **Sequencing:** F:TCAACAGGAATGTGAGCAGG and R:CCCACAATTTTAAGTGAGTTGC; fragment length: 386bp PCR conditions: 95ºC 5´; 5 cycles (95ºC 30", 60ºC 30", 72ºC 45"); 20 cycles (95ºC 30", 60ºC→50ºC[6] 30", 72ºC 45"); 10 cycles (95ºC 30", 50ºC 30", 72ºC 45"); 72ºC 5´; 4ºC ∞							
LRRK2 (1)	rs33939927 (C>T) p.R1441C	Exon 31 (NM_198578)	F_C: GTCTTTCCCTCCAGGCT**C** F_T: GTCTTTCCCTCCAGGCT**T** R: TGACATTTCTAGGCAGTTGAG	291	AS-PCR	-	-
AS-PCR conditions: 95ºC 10´; 35 cycles (95ºC 30", 62ºC 30", 72ºC 45"); 72ºC 10´; 4ºC ∞ **Sequencing:** F:TCAACAGGAATGTGAGCAGG and R:CCCACAATTTTAAGTGAGTTGC; fragment length: 386bp PCR conditions: 95ºC 5´; 5 cycles (95ºC 30", 60ºC 30", 72ºC 45"); 20 cycles (95ºC 30", 60ºC→50ºC[7] 30", 72ºC 45"); 10 cycles (95ºC 30", 50ºC 30", 72ºC 45"); 72ºC 5´; 4ºC ∞							
LRRK2 (1)	rs34637584 (G>A) p.G2019S	Exon 41 (NM_198578)	F_G: GCAAAGATTGCTGACTAC**G** F_A: TGCAAAGATTGCTGACTAC**A** R: ACAAGTGCCAACAATACCTAG	420	AS-PCR	-	-
AS-PCR conditions: 95ºC 10´; 35 cycles (95ºC 30", 62ºC 30", 72ºC 45"); 72ºC 10´; 4ºC ∞ **Sequencing:** F: TTTTGATGCTTGACATAGTGGAC and R: CACATCTGAGGTCAGTGGTTATC; fragment length: 329bp PCR conditions: 95ºC 5´; 5 cycles (95ºC 30", 60ºC 30", 72ºC 45"); 20 cycles (95ºC 30", 60ºC→50ºC[8] 30", 72ºC 45"); 10 cycles (95ºC 30", 50ºC 30", 72ºC 45"); 72ºC 5´; 4ºC ∞							

[6] -0.5ºC/cycle. i.e. TD60-50.
[7] -0.5ºC/cycle. i.e. TD60-50.
[8] -0.5ºC/cycle. i.e. TD60-50.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
MANF (5)	rs11538558 (A>G) The STOP codon changes to W and thus a larger protein (90aa more) is synthesized	Exon 4 (NM_006010)	F: TGCCTGTAAATGTGTCTGG R: CAGCATCATCAGGAAAGC	472	RFLP A:310+162 G:310+150+12	2.5	*Bsl*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 54ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
MAPT (1)	H1/H2 haplotype (238bp deletion)	Intron 10 (NM_016835)	F: GGAAGACGTTCTCACTGATCTG R: AGGAGTCTGGCTTCAGTCTCTC [4]	H1: 484 H2: 246	PCR	-	-
PCR conditions: 95ºC 10´; 35 cycles (95ºC 30", 66ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
MTHFR (1)	rs1801133 (c.C677T) p.A222V	Exon 5 (NM_005957)	F: TGAAGGAGAAGGTGTCTGCGGGA R:AGGACGGTGCGGTGAGAGTG [5]	198	RFLP C:198 T:174+24	2.5	*Hinf*I
PCR conditions: 95ºC 10´; 35 cycles (95ºC 45", 60ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
MTHFR (1)	rs1801131 (c.A1298C) p.E429A	Exon 8 (NM_005957)	F: GTCTCCCAACTTACCCTTCTCCC R:ATGTGGGGGGAGGAGCTGAC	241	RFLP A:210+30 C:241	2.5	*Mbo*II
PCR conditions: 95ºC 10´; 35 cycles (95ºC 45", 60ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
MTR (1)	rs1805087 (c.A2756G) p.D519G	Exon 26 (NM_000254)	F: TGTTCCCAGCTGTTAGATGAAAATC R: GATCCAAAGCCTTTTACACTCCTC	211	RFLP A:211 G:131+80	1.5	*Bsu*RI
PCR conditions: 95°C 10´; 35 cycles (95°C 45", 60°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
NFE2L2 (3)	rs1806649 (G>A)	Intron 1 (NM_006164)	F: GTATTGTTTATTTGAAGGGG R: AGTCACATTTTCTCATCTGC	602	RFLP G:400+200 A: 602	1.5	*Bpu10*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 53°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
NFE2L2 (3)	rs10183914 (G>A)	Intron 3 (NM_006164)	F: AATAGAAAAACTCCATCATAAC R: CCGAAGAAACCTAAAATTG	653	RFLP G: 300+353 A: 653	1.2	*Hph*I
PCR conditions: 95°C 10´; 40 cycles (95°C 30", 52°C 30", 72°C 1´); 72°C 10´; 4°C ∞							
NR4A2 ****	(C>T)	Intron 3 (NM_006186)	F: GCTGTGTGTGGGGACAACG R:AGTGGAACGTGATGCTGGAG	1073	RFLP C: 120+290+300 +400 T:120+290+30 0+360+30	1.5	*Mva*I
PCR conditions: 95°C 10´; 35 cycles (95°C 30", 60°C 30", 72°C 1´30"); 72°C 10´; 4°C ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
PRNP (3)	rs4815729 (G>A)	Intron 1 (NM_000311)	F: CTAAGTGAATAGTGTGTGCAGC R: CTTTTCACTGAAGGATTAC***G***TC A mismatch (***G*** instead of C) was introduced to create the recognition site.	150	RFLP G: 150 A: 130+20	2.5	*Bsm*AI (*Alw26*I)
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 54ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							
PRNP (3)	rs1799990 (A>G) p.M129V	Exon 2 (NM_000311)	F: CAGCTGATACCATTGCTATGC R: AGGGTATTGATTAGCCTATCCG	1006	RFLP A: 70+440+490 G: 440+560	1.5	*Nsp*I (*Xce*I)
PCR conditions: 95ºC 10´; 42 cycles (95ºC 30", 58ºC 30", 72ºC 1´); 72ºC 10´; 4ºC ∞							
PSEN1 (3)	rs165932 (G>T)	Intron 8 (NM_000021)	F: ATTTAGTGGCTGTTTTGTG R: CACTGATTACTAATTCAA***G***ATC A mismatch (***G*** instead of T) was introduced to create the recognition site.	140	RFLP G:115+25 T:140	2.5	*Bst*YI
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 56ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							
PSEN2 (3)	rs6426554 (A>G)	Intron 12 (NM_000447)	F: TCCTGCCGTGACTTCATCTC R: GGATGGTAAGATGCCCTCAG	452	RFLP A: 452 G: 70+380	1.5	*Bsa*HI
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 65ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
SNCA (1)	Rep1 microsatellite	10kb upstream (NM_000345)	F: [9]CCTGGCATATTTGATTGCAA R: GACTGGCCCAAGATTAACCA [6]	267/269/271 are the most frequent	Fragment analysis	-	-
PCR conditions: 95ºC 10´; 10 cycles (94ºC 15", 55ºC 15", 72ºC 30"); 30 cycles (89ºC 15", 55ºC 15", 72ºC 30"); 72ºC 10´; 4ºC ∞							
SNCA (1)	rs356219 (A>G)	9kb downstream (NM_000345)	F: AATGTGAGGGCTCAAAAAC R: AAAACAAACACAAAATTC***G***A A mismatch (***G*** instead of C) was introduced to create the recognition site.	140	RFLP A: 140. G: 120 + 20.	2.5	*BsaH*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 50ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							
TARDBP or TDP-43 (5)	rs11547736 (C>A) p.S92X	Exon 3 (NM_007375)	F: GGATGAGACAGATGCTT***G***AT R: ACCTGCACCATAAGAACTTC A mismatch (***G*** instead of C) was introduced to create the recognition site.	150	RFLP C:130+20 A:150	2.5	*Mbo*I
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 61ºC 1´, 72ºC 45"); 72ºC 10´; 4ºC ∞							
TOMM40 ***** (2)	rs10524523 Poly-T	Intron 6 (NM_006114)	F: CTCCAACTGCTGACCTCAAG R: [10]GCTGAGAAGGGAGGATTGC	from 165 (12T) to 188 (35T)	Fragment analysis	-	-
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 65ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							

[9] This primer was modified in 5´ with 6FAM.
[10] This primer was modified in 5´ with 6FAM.

Annex I. Primers genetics.

Table A. Description of the parameters used for the genetic analysis of mutations/polymorphisms. (Continued)

Gene	Mutation or polymorphism	Genic location (reference sequence)	Primer sequences (5´→ 3´)	Size of fragment (bp)	Genotyping method	% agarose	Restriction enzyme
TOR1A (1)	rs1801968 (G>C) p.D216H	Exon 4 (NM_000113)	F:[11]AACCCTGTCCTTACCCACTG R: TCTGCTTTCCACTCCTCCAG Pyroseq (R): GAAATCCAAAGCCACA [7]	97	Pyrosequencing	-	-
PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 57ºC 30", 72ºC 30"); 72ºC 10´; 4ºC ∞							
TOR1A (1)	p.delE302/303 (GAG/-)	Exon 5 (NM_000113)	F: CCTGGAATACAAACACCTA R: [12]GGCTGCCAATCATGACTGTC [8] Pyroseq (F): ATTGTAAGCAGAGTGGC	225	Pyrosequencing	-	-
PCR conditions: 94ºC 2´; 40 cycles (94ºC 1´, 55ºC 1´, 72ºC 1´30"); 72ºC 10´; 4ºC ∞ [8]							
UCHL1 (1)	rs5030732 (C>A) p.S18Y	Exon 3 (NM_004181)	F_A: CTCTCCGCAGGTGCTGT**A** F_C: CTCTCCGCAGGTGCTGT**C** R: AGCCCAGGGAGTAGGTACC F: CTTTGTGCTGTGTCATTGC	303 F+R for sequencing: 465	AS-PCR	-	-
AS-PCR conditions: 95ºC 10´; 40 cycles (95ºC 30", 60ºC 30", 72ºC 45"); 72ºC 10´; 4ºC ∞ **sequencing conditions:** 95ºC 10´; 5 cycles (95ºC 30", 55ºC 30", 72ºC 45"); 25 cycles (95ºC 30", 55ºC→50ºC[13] 30", 72ºC 45"); 5 cycles (95ºC 30", 50ºC 30", 72ºC 45"); 72ºC 10´; 4ºC ∞							

[11] This primer was biotinylated in 5´.
[12] This primer was biotinylated in 5´.
[13] -0.2ºC/cycle. i.e. TD55-50.

Annex I. Primers genetics.

For each gene, we decided to study (1) mutations or polymorphisms that were relevant and had been extensively analyzed in genetics of Parkinson´s disease. However, in some cases, that was not possible and we focused on (2) variants reported in some article about genetics in PD or AD. When even that was not possible, we considered that, in order of priority, (3) tag SNPs, (4) common polymorphisms or even (5) mutations, could be part of our study.

Tag SNPs were selected by using the data released by the HapMap (Haplotype Map) project[14] which was launched in 2002 to describe the common patterns of genetic variation in humans. By the global analysis of SNPs[15] and of patterns of linkage disequilibrium[16] in worldwide population, this study highlighted haplotypes[17] and, therefore, the possibility to select tSNPs[18], thus reducing the need for genotyping, with little loss of information, in association studies [9]. With this information it was possible to analyze larger genomic regions with the minimum number of SNPs (Figure I) [10][19].

[14] The information is freely available in **http://hapmap.ncbi.nlm.nih.gov/**.
[15] SNP or single nucleotide polymorphism is a site in the genome where individuals differ by a single base.
[16] Linkage disequilibrium (LD) refers to the phenomenon that alleles that are close together in the genome tend to be inherited together.
[17] Haplotype is a combination of alleles at multiple linked sites on a single chromosome, all of which are transmitted together.
[18] tSNP or tag SNP is a representative SNP in a region of the genome with high LD to other variants.
[19] It is noteworthy that there is no clinical information about any of the individuals that enrolled in the HapMap project. The only available data about people are their sex and their origin.

Annex I. Primers genetics.

HapMap project has developed during the time we have been conducting our study and thus new data have been released. For this reason, nowadays, some of our selected tag SNPs[20] could have lost their category and just be SNPs. If any of the tSNPs gives significant results, we will take this fact into account in order to give the appropriate interpretation to results.

Common polymorphisms and mutations were chosen among those described by the HapMap project or the Single Nucleotide Polymorphism database, dbSNP (**http://www.ncbi.nlm.nih.gov/projects/SNP/**).

Taken from [10]

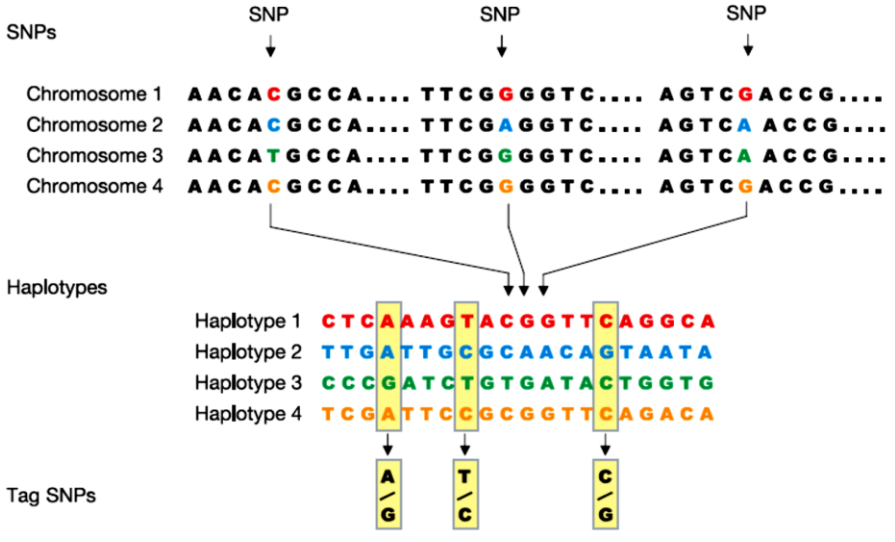


Fig. I. SNPs, haplotypes and tag SNPs.
Most of the DNA sequence is identical but there are bases that differ. Here, there are three SNPs in the four chromosomes. These and other SNPs define haplotypes: combinations of alleles. However, due to the strong association among the SNPs in most chromosomal regions, only a few carefully chosen SNPs (known as tag SNPs) need to be typed to predict the likely variants at the rest of the SNPs in each region. In this example, the four haplotypes can be differentiated just by genotyping the three marked tSNPs.

[20] Cut off values defined for tSNP selection: $r^2 \geq 0.8$; MAF≥ 0.05.

Annex I. Primers genetics.

*APOE**: two restrictions (one with *Afl*III and other with *Hae*II) were conducted and the genotype was determined by the combined restriction fragments pattern obtained visualized in an electrophoresis (Table B).

Table B. *APOE* restriction fragments

	***Afl*III**		***Hae*II**	
Genotype	**171**	**231**	**205**	**231**
ε2/ ε2	+	-	-	+
ε2/ ε3	+	-	+	+
ε2/ ε4	+	+	+	+
ε3/ ε3	+	-	+	-
ε3/ ε4	+	+	+	-
ε4/ ε4	-	+	+	-

The assay with *Afl*III determined the nucleotide sequence in rs429358 (T>C) and thus the codon 112: **C**GC (R)→**T**GC (C) [C: 231; T: 171+62].
Whereas, the assay with *Hae*II determined the nucleotide sequence in rs7412 (C>T) and thus the codon 158: **C**GC (R)→**T**GC (C) [C: 205+26; T: 231]

+: presence; -: absence.

*ARMCX4***: when we started our genetic analysis, ARMCX4 had been considered a big pseudogene (≈115 kb) by *in silico* predictions and comparison with murine genome, so we decided to analyze the three tag SNPSs that covered the region. However, it has been recently described and considered a functional gene that only covers ≈8kb. Under these new circumstances, tag SNP rs5951332 should have been analyzed instead of our tag SNPSs which are located far from the gene: rs2179670 is ≈60kb upstream, whereas rs6523506 is ≈10kb downstream.

*DOCK3****: this gene is upstream *MANF* and has a tag SNP (rs4441646) in high linkage disequilibrium with it. Moreover, *DOCK3* presents some promising features: it is expressed in brain and *in vitro* experiments have concluded that it binds PSEN1 [11], it decreases the activity of GSK3β on tau

but increases its phosphorylation by other kinases [12] and, moreover, DOCK3 modulates the neurite growth due to its important role downstream of BDNF signaling (it works as a guanine nucleotide exchange factor -GEF- in the cytoskeleton network) by stimulating dual pathways: actin polymerization and microtubule assembly [13, 14].

*NR4A2*****: its intron 3 is highly conserved in mammals (data obtained from the UCSC genome browser) and, for that reason we thought it could be interesting to analyze it rather than to focus on previously described and studied variants. Initially, we sequenced few healthy controls without finding any change. Nevertheless, when we analyzed 48 randomly chosen PD patients (16 from Donosti, 16 from Barcelona and 16 from Sevilla) one of them presented a new mutation that we decided to study in more detail. Its correct description is c.864+246C>T (or IVS3+246C>T).

*TOMM40******: we classified the homopolymers in three groups as [15] did, i.e. short (s; T≤19), long (l; 20≤T≤29) and very long (vl; T≥30) alleles. Mostly of ε4 carriers (*APOE*) presented long alleles, whereas ε3 (and ε2) carriers presented short and very long alleles.

Annex I. Primers genetics.

Bibliography:

1. Sen, S., et al., *A BDNF coding variant is associated with the NEO personality inventory domain neuroticism, a risk factor for depression.* Neuropsychopharmacology, 2003. **28**(2): p. 397-401.
2. Toft, M., et al., *Glucocerebrosidase gene mutations and Parkinson disease in the Norwegian population.* Neurology, 2006. **66**(3): p. 415-7.
3. Kwok, J.B., et al., *GSK3B polymorphisms alter transcription and splicing in Parkinson's disease.* Ann Neurol, 2005. **58**(6): p. 829-39.
4. Baker, M., et al., *Association of an extended haplotype in the tau gene with progressive supranuclear palsy.* Hum Mol Genet, 1999. **8**(4): p. 711-5.
5. Frosst, P., et al., *A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase.* Nat Genet, 1995. **10**(1): p. 111-3.
6. Xia, Y., et al., *Genetic studies in Alzheimer's disease with an NACP/alpha-synuclein polymorphism.* Ann Neurol, 1996. **40**(2): p. 207-15.
7. Risch, N.J., et al., *Intragenic Cis and Trans modification of genetic susceptibility in DYT1 torsion dystonia.* Am J Hum Genet, 2007. **80**(6): p. 1188-93.
8. Leung, J.C., et al., *Novel mutation in the TOR1A (DYT1) gene in atypical early onset dystonia and polymorphisms in dystonia and early onset parkinsonism.* Neurogenetics, 2001. **3**(3): p. 133-43.
9. Manolio, T.A., L.D. Brooks, and F.S. Collins, *A HapMap harvest of insights into the genetics of common disease.* J Clin Invest, 2008. **118**(5): p. 1590-605.
10. HapMap Consortium, T.I., *The International HapMap Project.* Nature, 2003. **426**(6968): p. 789-96.

Annex I. Primers genetics.

11. Kashiwa, A., et al., *Isolation and characterization of novel presenilin binding protein.* J Neurochem, 2000. **75**(1): p. 109-16.
12. Chen, Q., et al., *Presenilin binding protein is associated with neurofibrillary alterations in Alzheimer's disease and stimulates tau phosphorylation.* Am J Pathol, 2001. **159**(5): p. 1597-602.
13. Chen, Q., et al., *Loss of modifier of cell adhesion reveals a pathway leading to axonal degeneration.* J Neurosci, 2009. **29**(1): p. 118-30.
14. Namekata, K., et al., *Dock3 stimulates axonal outgrowth via GSK-3beta-mediated microtubule assembly.* J Neurosci, 2012. **32**(1): p. 264-74.
15. Linnertz, C., et al., *Characterization of the poly-T variant in the TOMM40 gene in diverse populations.* PLoS One, 2012. **7**(2): p. e30994.

Annex I. Primers genetics.

Annex II. Primers epigenetics.

Table C. Description of the parameters used for the epigenetic analysis of DNA methylation levels by pyrosequencing.

Gene	Assay	Primers sequence (5´→ 3´)	Amplicon (bp)	CpGs analyzed	Number of PCR cycles (X)	Annealing temperature (Y; ^{o}C)	Target
DJ-1	1	F: GGGAGG**TT**TGGA**TT**AGAGT**TT** R: *ACCCCCCAC***C**AAT**AA**CACA**A**TCC Pyroseq (F): GG**TT**TGGA**TT**AGAGT**TT**TAA**T**AG	229	6	38	61	"Promoter"
DJ-1	2	F: GGTGG***A***GGTAGAGA**T**TGTTAAG**TTT** R: *CACCCCACACCA**AA**CT**A**A Pyroseq (F): TGTGGGG**T**TGAGGGA	273	8	45	60	Exon 1 •
DJ-1	3; 3.1	F: GTGTGGGGTGAGTGGTA**T** R: ***A**TCAACCCAACTACATCT**A**TCTCT Pyroseq (F; 3): GGTGT**TT*****A***G**T**TGG**TTT**AG Pyroseq (F; 3.1): **T**TGG***A***GTTGGATTTGA**T**TGA	465	5; 4	45	65	Intron 1 (• only for 3, not for 3.1)
LRRK2	1	F: GGGG**T**TT***A***GGG**TT**TGTGGA**T** R: *TCCCTCTCC***CA**AA**A**CCCTCCT**A**C Pyroseq (F): AG**TT**AGG**TT*****A***GG**T**T**TT**AG**T**AG**T**	307	9	45	65	"Promoter"
LRRK2	2	F: **TT**TGAGTGGGGGAGGAGGAA R: ***A**CCACTA**A**CCAT**AA**T**AA**CACCT**A**CTTC Pyroseq (F): AG**T**TG**TT**T**TT**TT**TT**T**T**ATAAA**T**AGG	254	9	45	63	Exon 1
PINK1	1	F: GTG**T**AAAGGGAAAGT**T**A**T**TG**T**TAGAG R: ***A**TCCT**AC*****CA**CCCA**A**CCT**AA**AC Pyroseq (F): GGG**T*****A***G***A***GG**TT**TG**T**AG**T**TGG	358	7	45	61	Exon 1

Annex II. Primers epigenetics.

Table C. Description of the parameters used for the epigenetic analysis of DNA methylation levels by pyrosequencing. (Continued)

Gene	Assay	Primers sequence (5´→ 3´)	Amplicon (bp)	CpGs analyzed	Number of PCR cycles (X)	Annealing temperature (Y; ^{o}C)	Target
PINK1	2	F: TGGTG***A***GGG**TT**TGGGG**T**TG R: *A**C**CCC***C***CTCACCT**AA**ATCTCCT**A**AC Pyroseq (F): **T**TGGG**TT**T**T**AT***A***GAGGAAAAA**T**AG	142	5	38	61	Exon 1 •
PRKN	1	F: AGAG**T**TG**T**AA**T**AAG**T**TT**TT**AAAGGTAAG**T** R: *CTCCCACCA**AC*****CA**CTCTCCT**AAA**TTA Pyroseq (F): GGGGGG**T**TGGGGG**T**A	284	4	45	60	"Promoter"
PRKN	2	F: GATAGGTA***A***GTGGGTA**TT**TG**TT**AGGTA**T**AG R: *ACTTT**AA**CCCC***C***TCATT**A**ACA**A**TT**AA**CACC Pyroseq (F): A**TT**TG**TT**AGGTA**T**AG**TT**T**T**T**T**TG	124	9	38	58	Intron 1 (• partially overlaps)
SNCA	1	F: GGGGAAAGAGGAAGAGG**T** R: *CCCTCTCTT**AAA**CCCCTTCT**A** Pyroseq (F): GGAG**T**A***A***G**T**TG**T**AGGGAAAG**T**A	340	6	45	63	"Promoter"
SNCA	2	F: AGG**T**AGG***A***GG**T**TGGAGTTGAT R: *T**AA**CCACTCCCA**A**TTCTCC Pyroseq (F): GGG**TTT**AAGAGAGGGGG	380	8	38	61	Exon 1[1] (• partially overlaps)

[1] In isoform NM_000345.

Annex II. Primers epigenetics.

Table C. Description of the parameters used for the epigenetic analysis of DNA methylation levels by pyrosequencing. (Continued)

Gene	Assay	Primers sequence (5´→ 3´)	Amplicon (bp)	CpGs analyzed	Number of PCR cycles (X)	Annealing temperature (Y; ^{o}C)	Target
SNCA	3	F: GGAGAA**T**TGGGAGTGG**TT**AT R: *CACA**AA**TACTTACCT**A**A**A**TCCCTCT**A**C Pyroseq (F): GGGTTTG**TT**T**TTT**A**TT**T**TTTT**AG	262	5	45	60	Exon 1[2]
SNCA	4	F: GAGAAG**T**AGAGGGA**TTT**AGGTAAGTA**T** R: *ACCT**AA**CCCA**AC*C*AA**TCCTCAT**AA**CA Pyroseq (F): GGG***A***GT**T**TTTGGAAAT**TT**TGGAG	139	5	38	58	Intron 1

*All the reverse primers were biotinylated in 5´.

T (in forward primers) and **A** (in reverse primers) denote the converted unmethylated cytosines whereas ***A*** (in forward primers) and ***C*** (in reverse primers) correspond to cytosines in CpG dinucleotides and are thus introduced as mismatches to overpass those variable positions.

PCR conditions: 95^{o}C 15´; X cycles (94^{o}C 30”, Y^{o}C 30”, 72^{o}C 30”); 72^{o}C 10´; 4^{o}C ∞

PCR mix[3] per one reaction (1X) for a final volume of 25µL: 17.25µL Milli-Q water + 2.5µL 10X buffer + 1µL dNTPs 5mM each + 2.5µL $MgCl_2$ 25mM + 0.5µL primerF 10 µM + 0.5µL primerR 10 µM + 0.25µL Maxima Hot Start *Taq* DNA polymerase (Thermo Scientific) 5U/µL + 0.5µL bisulfite treated DNA 50ng/µL.

[2] In isoforms NM_007308 and NM_001146054.

Annex II. Primers epigenetics.

"Promoter" indicates that our trial predictions located it in this area. However, in some assays, the predicted promoter overlapped with exons or introns (marked with a •).

Assays DJ-1 3, DJ-1 3.1, PINK1 1 and SNCA 4 were finally discarded because it was not possible to obtain methylation values that passed the quality controls established by the program. PCR amplifications were successful but the pyrosequencing reactions did not work. We do not know the reason because their features were similar to the other assays.

[3] When trying to amplify bisulfite treated DNA it is essential to use Hot Start polymerase and $MgCl_2$. In addition, because biotinylated template strands as well as unincorporated biotinylated primers will be captured on streptavidin-coated beads, only a small amount of primer is used in the PCR amplification and a large number of amplification cycles is performed to exhaust the primers. These amplification conditions are necessary; otherwise, the unincorporated biotinylated primer will compite with the amplicons for the streptavidin sites, reducing the amount of template bound to the beads and, therefore, the number of molecules being sequenced, lessening the strength of the results.

Basque population:

Table D. Genotypic frequency of H1/H2 haplotype (*MAPT*).

Gene and polymorphism	Genotypes		
MAPT, H1/H2 haplotype	H1H1	H1H2	H2H2
Controls	36	41	9
PD cases	81	58	11
	X^2 = 3.315; 2df; p=0.191		

With regard to the mutation c.864+246C>T in *NR4A2*, any of the individuals, neither cases nor controls, carried it.

Table E. Genotypic and allelic frequency of p.S18Y (*UCHL1*) and p.D216H (*TOR1A*).

Gene and polymorphism	Genotypes			Alleles	
UCHL1, rs5030732 (p.S18Y)	SS	SY	YY	S	Y
Controls	61	21	4	143	29
PD cases	103	43	4	249	51
	X^2 = 1.039; 2df; p=0.595			X^2 = 0.002; 1df; p=0.969	
TOR1A, rs1801968 (p.D216H)	GG	GC	CC	G	C
Controls	74	12	0	160	12
PD cases	129	21	0	279	21
	X^2 = 0; 1df; p=0.992			X^2 = 0; 1df; p=0.992	

Annex III. Analysis of genetic susceptibility factors in Spanish population.

Table F. Genotypic and allelic frequency of *APOE* genotype.

Gene and polymorphism	Genotypes						Alleles		
APOE, genotype	ε2ε2	ε2ε3	ε3ε3	ε3ε4	ε4ε4	ε2ε4	ε2	ε3	ε4
Controls	3	17	54	8	0	4	27	133	12
PD cases	5	24	83	24	0	14	48	214	38
	X^2 = 4.354; 4df; p=0.360						X^2 = 3.881; 2df; p=0.144		

Table G. Frequency of *APOE* genotype when considering the number of ε2 alleles.

APOE, genotype	ε2 +/+	ε2 +/-	ε2 -/-	ε2 +	ε2 -
Controls	3	21	62	24	62
PD cases	5	38	107	43	107
	X^2 = 0.027; 2df; p=0.987			X^2 = 0.016; 1df; p=0.901	

Table H. Genotypic and allelic frequency of poly-T homopolymer (*TOMM40*).

Gene and polymorphism	Genotypes						Alleles		
TOMM40, rs10524523	ss	sl	ll	vll	vlvl	svl	s	l	vl
Controls	19	3	5	2	17	40	81	15	76
PD cases	45	14	3	3	22	63	167	23	110
	X^2 = 7.341; 5df; p=0.196						X^2 = 3.249; 2df; p=0.197		

Non-Basque population:

With regard to the mutation c.864+246C>T in *NR4A2*, there were 3 heterozygous controls and 2 heterozygous cases (Fisher´s exact test, p=0.673). Therefore, there was no statistical significance.

Table I. Genotypic and allelic frequency of p.S18Y (*UCHL1*).

Gene and polymorphism	Genotypes			Alleles	
UCHL1, rs5030732 (p.S18Y)	SS	SY	YY	S	Y
Controls	201	86	14	488	114
PD cases	234	86	12	554	110
	X^2 = 1.142; 2df; p=0.565			X^2 = 1.218; 1df; p=0.270	

Table J. Allelic frequency of p.D216H (*TOR1A*).

Gene and polymorphism	Alleles	
TOR1A, rs1801968 (p.D216H)	G	C
Controls	550	52
PD cases	595	69
	X^2 = 1.123; 1df; p=0.289	

Table K. Frequency of *APOE* genotype when considering the number of ε4 alleles.

APOE, genotype	ε4 +/+	ε4 +/-	ε4 -/-	ε4 +	ε4 -
Controls	2	73	226	75	226
PD cases	4	79	249	83	249
	X^2 = 0.500; 2df; p=0.779			X^2 = 0.001; 1df; p=0.981	

Table L. Genotypic and allelic frequency of poly-T homopolymer (*TOMM40*).

Gene and polymorphism	**Genotypes**						**Alleles**		
TOMM40, rs10524523	ss	sl	ll	vll	vlvl	svl	s	l	vl
Controls	64	25	11	25	59	117	270	72	260
PD cases	61	30	15	19	73	134	286	79	299
	X^2 = 3.086; 5df; p=0.687						X^2 = 0.471; 2df; p=0.790		

Annex IVa. Analysis of genetic determinants in cognitive status in Parkinson´s disease. PD risk.

Navarrese healthy controls and PD patients.

i. Possible genetic susceptibility factors

a. <u>in PD:</u>

Table M. Genotypic and allelic frequency of Rep1 microsatellite (*SNCA*).

Gene and polymorphism	Genotypes						Alleles		
SNCA, Rep1 microsatellite	267-267	267-269	269-269	269-271	271-271	267-271	267	269	271
Controls	5	26	25	2	0	0	36	78	2
PD cases	13	42	49	5	0	2	70	145	7
	X^2 = 1.961; 4df; p=0.743						X^2 = 0.633; 2df; p=0.729		

b. <u>in AD:</u>

Table N. Genotypic and allelic frequency of *APOE* genotype.

Gene and polymorphism	Genotypes						Alleles		
APOE, genotype	ε2ε2	ε2ε3	ε3ε3	ε3ε4	ε4ε4	ε2ε4	ε2	ε3	ε4
Controls	0	6	43	8	0	1	7	100	9
PD cases	1	14	75	19	1	1	17	183	22
	X^2 = 1.928; 5df; p=0.859						X^2 = 0.797; 2df; p=0.671		

Annex IVa. Analysis of genetic determinants in cognitive status in Parkinson´s disease. PD risk.

Table O. Frequency of *APOE* genotype when considering the number of ε2 or ε4 alleles.

	ε2 +/+	ε2 +/-	ε2 -/-	ε2 +	ε2 -
Controls	0	7	51	7	51
PD cases	1	15	95	16	95
	X^2 = 0.608; 2df; p=0.738			X^2 = 0.178; 1df; p=0.673	
	ε4 +/+	ε4 +/-	ε4 -/-	ε4 +	ε4 -
Controls	0	9	49	9	49
PD cases	1	20	90	21	90
	X^2 = 0.715; 2df; p=0.699			X^2 = 0.302; 1df; p=0.583	

Table P. Genotypic and allelic frequency of poly-T homopolymer (*TOMM40*).

Gene and polymorphism	**Genotypes**						**Alleles**		
TOMM40, rs10524523	ss	sl	ll	vll	vlvl	svl	s	l	vl
Controls	15	5	2	4	9	23	58	13	45
PD cases	24	15	3	4	22	43	106	25	91
	X^2 = 2.404; 5df; p=0.791						X^2 = 0.171; 2df; p=0.918		

ii. Genes related to other diseases where dementia is consubstantial such as

a. AD:

Only one healthy control and one PD case were heterozygous for the SNP rs463946 (*APP*). The rest of individuals were CC. Therefore, there was no statistical significance (Fisher´s exact test, p=1).

Table Q. Genotypic and allelic frequencies of *PSEN1*, *PSEN2* and *CALHM1* variants.

Gene and polymorphism	Genotypes			Alleles	
PSEN1, rs165932	TT	TG	GG	T	G
Controls	21	22	15	64	52
PD cases	34	54	23	122	100
	X^2 = 1.785; 2df; p=0.410			X^2 = 0.001; 1df; p=0.970	
PSEN2, rs6426554	AA	AG	GG	A	G
Controls	40	17	1	97	19
PD cases	68	40	3	176	46
	X^2 = 1.019; 2df; p=0.601			X^2 = 0.925; 1df; p=0.336	
CALHM1, rs2986017 (p.P86L)	TT	TC	CC	T	C
Controls	3	17	38	23	93
PD cases	9	21	81	39	183
	X^2 = 2.593; 2df; p=0.274			X^2 = 0.260; 1df; p=0.610	

b. HD:

Table R. Genotypic and allelic frequencies of *HTT* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
HTT, rs10015979	AA	AG	GG	A	G
Controls	27	25	6	79	37
PD cases	44	47	20	135	87
	X^2 = 1.896; 2df; p=0.387			X^2 = 1.744; 1df; p=0.187	
HTT, rs363066	TT	TG	GG	T	G
Controls	39	19	0	97	19
PD cases	72	37	2	181	41
	X^2 = 1.082; 2df; p=0.582			X^2 = 0.228; 1df; p=0.633	
HTT, rs2298969	AA	AG	GG	A	G
Controls	30	21	7	81	35
PD cases	39	57	15	135	87
	X^2 = 4.522; 2df; p=0.104			X^2 = 2.685; 1df; p=0.101	
HTT, rs110501	TT	TC	CC	T	C
Controls	29	23	6	81	35
PD cases	46	51	14	143	79
	X^2 = 1.139; 2df; p=0.566			X^2 = 0.999; 1df; p=0.318	

c. CJD:

Table S. Genotypic and allelic frequencies of *PRNP* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
PRNP, rs4815729	AA	AG	GG	A	G
Controls	6	23	29	35	81
PD cases	10	48	53	68	154
	X^2 = 0.228; 2df; p=0.892			X^2 = 0.008; 1df; p=0.931	
PRNP, rs1799990 (p.M129V)	AA	AG	GG	A	G
Controls	19	29	10	67	49
PD cases	33	62	16	128	94
	X^2 = 0.554; 2df; p=0.758			X^2 =0.0003; 1df; p=0.986	

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table S). The result was the same when considering haplotypes (Table T):

Table T. Haplotypes in *PRNP*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value
GA	0.570	0.567, 0.577	0.031	0.859
AG	0.298	0.296, 0.301	0.007	0.932
GG	0.125	0.127, 0.122	0.021	0.885

The order of the markers at the haplotype is: rs4815729 – rs1799990.

d. Kufor-Rakeb syndrome, that is also characterized by parkinsonism:

Table U. Genotypic and allelic frequencies of *ATP13A2* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
ATP13A2, rs2871776	AA	AG	GG	A	G
Controls	10	21	27	41	75
PD cases	29	31	51	89	133
	X^2 = 2.155; 2df; p=0.340			X^2 = 0.725; 1df; p=0.395	
ATP13A2, rs4920608	AA	AG	GG	A	G
Controls	21	28	9	70	46
PD cases	44	57	10	145	77
	X^2 = 1.624; 2df; p=0.444			X^2 = 0.813; 1df; p=0.367	

iii. Genes that encode neurotrophic factors:

Table V. Genotypic and allelic frequencies of *CDNF* and *BDNF* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
CDNF, rs7094179	GG	GT	TT	G	T
Controls	28	28	2	84	32
PD cases	46	52	13	144	78
	X^2 = 3.354; 2df; p=0.187			X^2 = 1.978; 1df; p=0.160	
BDNF, rs6265 (p.V66M)	AA	AG	GG	A	G
Controls	4	17	37	25	91
PD cases	6	47	58	59	163
	X^2 = 2.754; 2df; p=0.252			X^2 = 1.030; 1df; p=0.310	

Annex IVa. Analysis of genetic determinants in cognitive status in Parkinson´s disease. PD risk.

iv. Genes involved in processes that are related to PD such as

a. homocysteine metabolism:

Table W. Genotypic and allelic frequencies of *CBS* and *MTR* variants.

Gene and polymorphism	Genotypes			Alleles	
CBS, c.844ins68	ins/ins	ins/-	-/-	ins	-
Controls	0	9	49	9	107
PD cases	1	17	93	19	203
	X^2 = 0.526; 2df; p=0.769			X^2 = 0.064; 1df; p=0.800	
MTR, rs1805087 (c.A2756G)	AA	AG	GG	A	G
Controls	38	18	2	94	22
PD cases	79	28	4	186	36
	X^2 = 0.651; 2df; p=0.722			X^2 = 0.405; 1df; p=0.524	

b. protection against oxidative stress:

Table X. Genotypic and allelic frequencies of *NFE2L2* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
NFE2L2, rs1806649	AA	AG	GG	A	G
Controls	5	22	31	32	84
PD cases	6	53	52	65	157
	X^2 = 1.770; 2df; p=0.413			X^2 = 0.107; 1df; p=0.744	
NFE2L2, rs10183914	AA	AG	GG	A	G
Controls	8	20	30	36	80
PD cases	16	55	40	87	135
	X^2 = 4.223; 2df; p=0.121			X^2 = 2.189; 1df; p=0.139	

Annex IVa. Analysis of genetic determinants in cognitive status in Parkinson´s disease. PD risk.

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table X). The result was the same even when considering haplotypes (Table Y):

Table Y. Haplotypes in *NFE2L2*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value
GG	0.592	0.566, 0.642	1.828	0.176
AA	0.243	0.251, 0.228	0.207	0.649
GA	0.121	0.141, 0.082	2.513	0.113
AG	0.044	0.042, 0.047	0.054	0.817

The order of the markers at the haplotype is: rs1806649 – rs10183914.

Table Z. Genotypic and allelic frequency of *KEAP1* polymorphism.

Gene and polymorphism	Genotypes			Alleles	
KEAP1, rs8113472	GG	GT	TT	G	T
Controls	49	9	0	107	9
PD cases	96	14	1	206	16
	X^2 = 0.777; 2df; p=0.678			X^2 = 0.034; 1df; p=0.854	

c. <u>chaperone-mediated autophagy:</u>

Table AA. Genotypic and allelic frequency of *HSPA8* polymorphism.

Gene and polymorphism	Genotypes			Alleles	
HSPA8, rs4936770	AA	AG	GG	A	G
Controls	3	18	37	24	92
PD cases	7	35	69	49	173
	X^2 = 0.102; 2df; p=0.950			X^2 = 0.086; 1df; p=0.769	

Annex IVa. Analysis of genetic determinants in cognitive status in Parkinson´s disease. PD risk.

Table AB. Genotypic and allelic frequencies of *LAMP-2A* polymorphisms.

Gene and polymorphism	**Genotypes**			**Alleles**	
LAMP-2A, rs42897	AA	AG	GG	A	G
Controls, overall	52	2	4	81	9
PD cases, overall	97	7	7	138	14
	X^2 = 0.627; 2df; p=0.731 Z_{mfG}^2 = 1.384; 1df; p=0.239 Z_C^2 = 1.400; 2df; p=0.497			$X^2 = Z_A^2$ = 0.041; 1df; p=0.840 Z_{mfA}^2 = 1.485; 1df; p=0.223	
Controls, women	27	2	3	56	8
PD cases, women	34	7	0	75	7
	X^2 = 5.556; 2df; **p=0.062**[1] Z_{fG}^2 = 0.460; 1df; p=0.498			$X^2 = Z_{fA}^2$ = 0.613; 1df; p=0.433	
Controls, men	-	-	-	25	1
PD cases, men	-	-	-	63	7
	-			$X^2 = Z_m^2$ = 0.940; 1df; p=0.332	

[1] As the result obtained with the parameter Z_{fG}^2, which is more appropriate to analyze markers at chromosome X, was not even almost statistically significant, we did not consider that there were relevant results to our study in this polymorphism and thus, we did not include it in the subsequent analysis. The p-value observed in the X^2 test may be originated by the absence of homozygous GG female patients.

Table AC. Genotypic and allelic frequencies of *LAMP-2A* polymorphisms. (Continued)

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs42895	AA	AC	CC	A	C
Controls, overall	29	14	15	55	35
PD cases, overall	62	14	35	93	59
	X^2 = 3.711; 2df; p=0.156 Z_{mfG}^2 = 0.019; 1df; p=0.891 Z_C^2 = 0.019; 2df; p=0.991			$X^2 = Z_A^2$ = 0.0001; 1df; p=0.991 Z_{mfA}^2 = 0.020; 1df; p=0.887	
Controls, women	12	14	6	38	26
PD cases, women	17	14	10	48	34
	X^2 = 0.764; 2df; p=0.682 Z_{fG}^2 = 0.009; 1df; p=0.926			$X^2 = Z_{fA}^2$ = 0.010; 1df; p=0.919	
Controls, men	-	-	-	17	9
PD cases, men	-	-	-	45	25
	-			$X^2 = Z_m^2$ = 0.010; 1df; p=0.920	

It is noteworthy that the gene *LAMP-2A* is located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [1][2].

[2] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

d. and *GSK3β*, which is mainly associated with tauopathies, but also with PD.

Table AD. Genotypic and allelic frequencies of *GSK3β* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
GSK3β, rs334558	TT	TC	CC	T	C
Controls	21	30	7	72	44
PD cases	48	54	9	150	72
	X^2 = 1.166; 2df; p=0.558			X^2 = 1.022; 1df; p=0.312	
GSK3β, rs6438552	TT	TC	CC	T	C
Controls	22	24	12	68	48
PD cases	50	49	12	149	73
	X^2 = 3.138; 2df; p=0.208			X^2 = 2.393; 1df; p=0.121	

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table AD). The result was the same when considering haplotypes (Table AE):

Table AE. Haplotypes in *GSK3β*.

Haplotype	Overall frequency	Frequencies: Case, Control	Chi square (X^2)	p-value
TT	0.584	0.604, 0.547	1.017	0.313
CC	0.285	0.257, 0.340	2.574	0.109
TC	0.073	0.072, 0.074	0.004	0.948
CT	0.058	0.067, 0.039	1.100	0.294

The order of the markers at the haplotype is: rs334558 – rs6438552.

Bibliography:

1. Zheng, G., et al., *Testing association for markers on the X chromosome.* Genet Epidemiol, 2007. **31**(8): p. 834-43.

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

Parkinson´s disease patients:

normal cognition (N.C.) – cognitive impairment (C.I.).

i. Possible genetic susceptibility factors

a. <u>in PD:</u>

Table AF. Genotypic and allelic frequencies of H1/H2 haplotype (*MAPT*) and rs356219 (*SNCA*).

Gene and polymorphism	Genotypes			Alleles	
MAPT, H1/H2 haplotype	H1H1	H1H2	H2H2	H1	H2
N.C.	27	15	8	69	31
C.I.	42	11	8	95	27
	X^2 = 2.814; 2df; p=0.245			X^2 = 2.240; 1df; p=0.135	
SNCA, rs356219	AA	AG	GG	A	G
N.C.	17	26	7	60	40
C.I.	21	32	8	74	48
	X^2 = 0.019; 2df; p=0.991			X^2 = 0.010; 1df; p=0.921	

Table AG. Genotypic and allelic frequency of Rep1 microsatellite (*SNCA*).

Gene and polymorphism	Genotypes						Alleles		
SNCA, Rep1 microsatellite	267-267	267-269	269-269	269-271	271-271	267-271	267	269	271
N.C.	4	21	21	3	0	1	30	66	4
C.I.	9	21	28	2	0	1	40	79	3
	X^2 = 2.053; 4df; p=0.726						X^2 = 0.562; 2df; p=0.755		

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

b. <u>in AD:</u>

Table AH. Genotypic and allelic frequency of *APOE* genotype.

Gene and polymorphism	Genotypes						Alleles		
APOE, genotype	ε2ε2	ε2ε3	ε3ε3	ε3ε4	ε4ε4	ε2ε4	ε2	ε3	ε4
N.C.	1	3	34	11	1	0	5	82	13
C.I.	0	11	41	8	0	1	12	101	9
	X^2 = 7.684; 5df; p=0.175						X^2 = 3.436; 2df; p=0.179		

Table AI. Frequency of *APOE* genotype when considering the number of ε4 alleles.

	ε4 +/+	ε4 +/-	ε4 -/-	ε4 +	ε4 -
N.C.	1	11	38	12	38
C.I.	0	9	52	9	52
	X^2 = 2.310; 2df; p=0.315			X^2 = 1.531; 1df; p=0.216	

Table AJ. Genotypic and allelic frequency of poly-T homopolymer (*TOMM40*).

Gene and polymorphism	Genotypes						Alleles		
TOMM40, rs10524523	ss	sl	ll	vll	vlvl	svl	s	l	vl
N.C.	11	5	2	2	11	19	46	11	43
C.I.	13	10	1	2	11	24	60	14	48
	X^2 = 1.674; 5df; p=0.892						X^2 = 0.307; 2df; p=0.858		

ii. Genes related to other diseases where dementia is consubstantial such as

a. AD:

Table AK. Genotypic and allelic frequencies of *PSEN1*, *PSEN2* and *CALHM1* variants.

Gene and polymorphism	Genotypes			Alleles	
PSEN1, rs165932	TT	TG	GG	T	G
N.C.	15	22	13	52	48
C.I.	19	32	10	70	52
	X^2 = 1.640; 2df; p=0.440			X^2 = 0.642; 1df; p=0.423	
PSEN2, rs6426554	AA	AG	GG	A	G
N.C.	34	15	1	83	17
C.I.	34	25	2	93	29
	X^2 = 1.761; 2df; p=0.415			X^2 = 1.533; 1df; p=0.216	
CALHM1, rs2986017 (p.P86L)	TT	TC	CC	T	C
N.C.	4	10	36	18	82
C.I.	5	11	45	21	101
	X^2 = 0.069; 2df; p=0.966			X^2 = 0.023; 1df; p=0.878	

Only one patient with cognitive impairment was heterozygous for the SNP rs463946 (*APP*). The rest of individuals were CC. Therefore, there was no statistical significance (Fisher´s exact test, p=1).

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

b. HD:

Table AL. Genotypic and allelic frequencies of *HTT* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
HTT, rs10015979	AA	AG	GG	A	G
N.C.	23	18	9	64	36
C.I.	21	29	11	71	51
	X^2 = 1.793; 2df; p=0.408			X^2 = 0.777; 1df; p=0.378	
HTT, rs363066	TT	TG	GG	T	G
N.C.	30	19	1	79	21
C.I.	42	18	1	102	20
	X^2 = 0.946; 2df; p=0.623			X^2 = 0.774; 1df; p=0.379	
HTT, rs363096	TT	TC	CC	T	C
N.C.	10	24	16	44	56
C.I.	12	30	19	54	68
	X^2 = 0.016; 2df; p=0.992			X^2 = 0.002; 1df; p=0.969	
HTT, rs2298969	AA	AG	GG	A	G
N.C.	18	26	6	62	38
C.I.	21	31	9	73	49
	X^2 = 0.181; 2df; p=0.913			X^2 = 0.108; 1df; p=0.742	
HTT, rs110501	TT	TC	CC	T	C
N.C.	21	22	7	64	36
C.I.	25	29	7	79	43
	X^2 = 0.221; 2df; p=0.896			X^2 = 0.014; 1df; p=0.907	

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

There were not statistically or almost statistically significant results either for genotypes or for alleles or for haplotypes (Tables AL and AM and Figure II, which shows the high linkage disequilibrium observed across the gene).

Table AM. Haplotypes in *HTT*.

Haplotype	Frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value
GTTAT	0.346	0.367, 0.319	0.573	0.449
ATCAT	0.208	0.200, 0.218	0.107	0.744
AGCGC	0.168	0.141, 0.201	1.424	0.233
ATCGC	0.153	0.188, 0.110	2.611	0.106
ATTGT	0.037	0.026, 0.050	0.837	0.360
ATTAT	0.027	0.018, 0.039	0.948	0.330
GTTGT	0.017	0.023, 0.010	0.504	0.478
GGCGC	0.011	0.014, 0.006	0.326	0.568

The order of the markers at the haplotype is: rs10015979 – rs363066 – rs363096 – rs2298969 – rs110501.

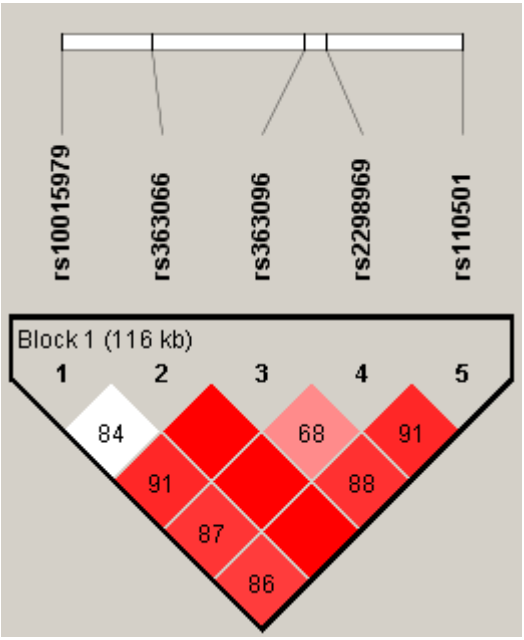


Fig. II. Linkage disequilibrium map in *HTT*.

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

c. CJD:

Table AN. Genotypic and allelic frequencies of *PRNP* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
PRNP, rs4815729	AA	AG	GG	A	G
N.C.	4	23	23	31	69
C.I.	6	25	30	37	85
	X^2 = 0.321; 2df; p=0.852			X^2 = 0.012; 1df; p=0.914	
PRNP, rs1799990 (p.M129V)	AA	AG	GG	A	G
N.C.	15	28	7	58	42
C.I.	18	34	9	70	52
	X^2 = 0.013; 2df; p=0.993			X^2 = 0.009; 1df; p=0.926	

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table AN). There was not any effect for the haplotypes either (Table AO).

Table AO. Haplotypes in *PRNP*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value
GA	0.566	0.564, 0.569	0.004	0.949
AG	0.296	0.294, 0.299	0.006	0.938
GG	0.128	0.133, 0.121	0.060	0.807
AA	0.010	0.010, 0.011	0.020	0.887

The order of the markers at the haplotype is: rs4815729 – rs1799990.

d. Kufor-Rakeb syndrome, that is also characterized by parkinsonism:

Table AP. Genotypic and allelic frequencies of *ATP13A2* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
ATP13A2, rs2871776	AA	AG	GG	A	G
N.C.	13	17	20	43	57
C.I.	16	14	31	46	76
	X^2 = 1.902; 2df; p=0.386			X^2 = 0.642; 1df; p=0.423	
ATP13A2, rs6684770	TT	TC	CC	T	C
N.C.	8	22	20	38	62
C.I.	3	32	26	38	84
	X^2 = 3.855; 2df; p=0.146			X^2 = 1.146; 1df; p=0.284	
ATP13A2, rs4920608	AA	AG	GG	A	G
N.C.	24	22	4	70	30
C.I.	20	35	6	75	47
	X^2 = 2.665; 2df; p=0.264			X^2 = 1.763; 1df; p=0.184	

Table AQ. Haplotypes in *ATP13A2*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value
ACA	0.300	0.302, 0.299	0.002	0.961
ATG	0.191	0.168, 0.217	0.858	0.354
GCG	0.189	0.226, 0.144	2.423	0.120
ACG	0.122	0.122, 0.123	0.001	0.978
GTG	0.098	0.107, 0.086	0.277	0.599
GCA	0.046	0.040, 0.055	0.296	0.586
ATA	0.040	0.023, 0.061	2.060	0.151
GTA	0.014	0.013, 0.015	0.025	0.874

The order of the markers at the haplotype is: rs4920608 – rs6684770 – rs2871776.

There were not statistically or almost statistically significant results either for genotypes or for alleles (Table AP). There was not any effect for the haplotypes either (Table AQ, Figure III).

The linkage disequilibrium was low across the region.

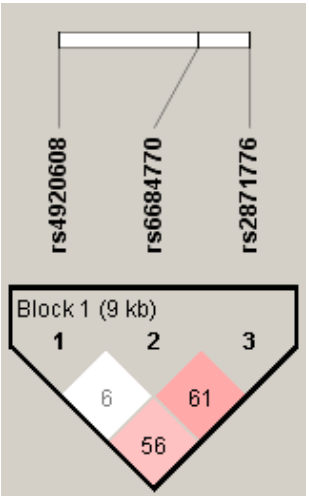


Fig. III. Linkage disequilibrium map in *ATP13A2*.

iii. Genes that encode neurotrophic factors:

Table AR. Genotypic and allelic frequencies of *CDNF*, *DOCK3* and *BDNF* variants.

Gene and polymorphism	Genotypes			Alleles	
CDNF, rs7099185	TT	TC	CC	T	C
N.C.	35	15	0	85	15
C.I.	38	23	0	99	23
	X^2 = 0.725; 1df; p=0.395			X^2 = 0.575; 1df; p=0.448	
DOCK3, rs4441646	AA	AC	CC	A	C
N.C.	33	11	6	77	23
C.I.	42	16	3	100	22
	X^2 = 1.935; 2df; p=0.380			X^2 = 0.839; 1df; p=0.360	
BDNF, rs6265 (p.V66M)	AA	AG	GG	A	G
N.C.	3	22	25	28	72
C.I.	3	25	33	31	91
	X^2 = 0.207; 2df; p=0.902			X^2 = 0.189; 1df; p=0.664	

iv. Genes involved in processes that are related to PD such as

a. homocysteine metabolism:

Table AS. Genotypic and allelic frequencies of *MTHFR* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
MTHFR, rs1801133 (c.C677T)	CC	CT	TT	C	T
N.C.	23	21	6	67	33
C.I.	23	34	4	80	42
	X^2 = 2.406; 2df; p=0.300			X^2 = 0.050; 1df; p=0.823	
MTHFR, rs1801131 (c.A1298C)	AA	AC	CC	A	C
N.C.	22	24	4	68	32
C.I.	31	22	8	84	38
	X^2 = 1.877; 2df; p=0.391			X^2 = 0.018; 1df; p=0.892	

Any of the polymorphisms analyzed showed statistically or almost statistically significant results neither for genotypes nor for alleles (Table AS). The study of haplotypes reported again no statistical significance (Table AT).

Table AT. Haplotypes in *MTHFR*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value
CA	0.354	0.347, 0.363	0.060	0.806
TA	0.331	0.342, 0.317	0.147	0.701
CC	0.308	0.309, 0.307	0.001	0.980

The order of the markers at the haplotype is: rs1801133 (c.C677T) – rs1801131 (c.A1298C).

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

b. protection against oxidative stress:

Table AU. Genotypic and allelic frequencies of *NFE2L2* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
NFE2L2, rs1806649	AA	AG	GG	A	G
N.C.	3	28	19	34	66
C.I.	3	25	33	31	91
	X^2 = 2.877; 2df; p=0.237			X^2 = 1.958; 1df; p=0.162	
NFE2L2, rs10183914	AA	AG	GG	A	G
N.C.	10	25	15	45	55
C.I.	6	30	25	42	80
	X^2 = 2.893; 2df; p=0.235			X^2 = 2.578; 1df; p=0.108	

c. chaperone-mediated autophagy:

Table AV. Genotypic and allelic frequency of *HSPA8* polymorphism.

Gene and polymorphism	Genotypes			Alleles	
HSPA8, rs1461496	TT	TC	CC	T	C
N.C.	9	16	25	34	66
C.I.	11	14	36	36	86
	X^2 = 1.239; 2df; p=0.538			X^2 = 0.514; 1df; p=0.474	

Annex IVb. Analysis of genetic determinants in cognitive status in Parkinson´s disease. Cognitive impairment.

Table AW. Genotypic and allelic frequencies of *LAMP-2A* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs7057652	TT	TG	GG	T	G
N.C., overall	21	8	21	33	33
C.I., overall	27	9	25	44	42
	X^2 = 0.067; 2df; p=0.967 Z_{mfG}^2 = 0.034; 1df; p=0.854 Z_C^2 = 0.054; 2df; p=0.973			$X^2 = Z_A^2$ = 0.020; 1df; p=0.887 Z_{mfA}^2 = 0.025; 1df; p=0.875	
N.C., women	4	8	4	16	16
C.I., women	8	9	8	25	25
	X^2 = 0.788; 2df; p=0.674 Z_{fG}^2 = 0; 1df; p=1.000			$X^2 = Z_{fA}^2$ = 0; 1df; p=1.000	
N.C., men	-	-	-	17	17
C.I., men	-	-	-	19	17
	-			$X^2 = Z_m^2$ = 0.054; 1df; p=0.816	

Table AX. Genotypic and allelic frequencies of *LAMP-2A* polymorphisms. (Continued)

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs42895[1]	AA	AC	CC	A	C
N.C., overall	27	8	15	42	24
C.I., overall	35	6	20	51	35
	X^2 = 0.952; 2df; p=0.621 Z_{mfG}^2 = 2.364; 1df; p=0.124 Z_C^2 = 2.599; 2df; p=0.273			$X^2 = Z_A^2$ = 0.295; 1df; p=0.587 Z_{mfA}^2 = 3.000; 1df; **p=0.083** OR: 0.83, $CI_{95\%}$=[0.43-1.61]	
N.C., women	7	8	1	22	10
C.I., women	10	6	9	26	24
	X^2 = 5.505; 2df; **p=0.064** Z_{fG}^2 = 1.740; 1df; p=0.187			$X^2 = Z_{fA}^2$ = 2.256; 1df; p=0.133	
N.C., men	-	-	-	20	14
C.I., men	-	-	-	25	11
	-			$X^2 = Z_m^2$ = 0.859; 1df; p=0.354	

[1] As the result obtained with the parameter Z_{fG}^2, which is more appropriate to analyze markers at chromosome X, was not even almost statistically significant, and there were opposite results for the Z_A^2 and Z_{fmA}^2, we did not consider that there were relevant results to our study in this polymorphism and thus, we did not include it in the subsequent analysis.

Table AY. Genotypic and allelic frequencies of *LAMP-2A* polymorphisms. (Continued)

Gene and polymorphism	Genotypes			Alleles	
LAMP-2A, rs42890[2]	GG	GT	TT	G	T
N.C., overall	27	7	16	42	24
C.I., overall	34	9	18	52	34
	X^2 = 0.082; 2df; p=0.960 Z_{mfG}^2 = 3.377; 1df; **p=0.066** Z_C^2 = 3.598; 2df; p=0.165			$X^2 = Z_A^2$ = 0.159; 1df; p=0.690 Z_{mfA}^2 = 3.942; 1df; **p=0.047** OR: 0.87, $CI_{95\%}$=[0.45-1.69]	
N.C., women	8	7	1	23	9
C.I., men	9	9	7	27	23
	X^2 = 2.977; 2df; p=0.226 Z_{fG}^2 = 2.220; 1df; p=0.136			$X^2 = Z_{fA}^2$ = 2.620; 1df; p=0.106	
N.C., men	-	-	-	19	15
C.I., men	-	-	-	25	11
	-			$X^2 = Z_m^2$ = 1.378; 1df; p=0.241	

It is noteworthy that *LAMP-2A* is located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [1][3].

[2] The opposite results observed for the different parameters calculated in the overall genotype and allele distribution made us consider that there were not relevant results to our study in this polymorphism and thus, we did not include it in the subsequent analysis.

[3] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

d. and *GSK3β*, which is mainly associated with tauopathies, but also with PD.

Table AZ. Genotypic and allelic frequencies of *GSK3β* polymorphisms.

Gene and polymorphism	Genotypes			Alleles	
GSK3β, rs334558	TT	TC	CC	T	C
N.C.	21	24	5	66	34
C.I.	27	30	4	84	38
	X^2 = 0.442; 2df; p=0.802			X^2 = 0.204; 1df; p=0.651	
GSK3β, rs6438552	TT	TC	CC	T	C
N.C.	24	18	8	66	34
C.I.	26	31	4	83	39
	X^2 = 3.810; 2df; p=0.149			X^2 = 0.103; 1df; p=0.748	

There were not relevant results either for genotypes or for alleles (Table AZ).

There was not any effect for the haplotypes either (Table BA).

Table BA. Haplotypes in *GSK3β*.

Haplotype	Overall frequency	Frequencies: C.I., N.C.	Chi square (X^2)	p-value
TT	0.603	0.617, 0.585	0.227	0.634
CC	0.256	0.248, 0.265	0.087	0.768
TC	0.073	0.072, 0.075	0.007	0.933
CT	0.069	0.064, 0.075	0.107	0.744

The order of the markers at the haplotype is: rs334558 – rs6438552.

Bibliography:

1. Zheng, G., et al., *Testing association for markers on the X chromosome.* Genet Epidemiol, 2007. **31**(8): p. 834-43.

We considered the genotypes, neither the alleles nor the haplotypes, that reported relevant results, i.e. p-value < 0.05 (X^2 test) to find a mathematical model to quantify the contribution of the variants on the risk to develop dementia[1]:

3. **Risk of developing dementia during PD (61 PD patients: 36 mild cognitive impairment (M.C.I.) and 25 dementia (PDD))[2]:**
   A. rs6426554 (*PSEN2*): AA vs AG+GG, p=0.039.
   B. rs2986017 (p.P86L, *CALHM1*): CC vs TT+CT, p=0.042.

Again, we employed a binary logistic regression model to quantify the influence of these two variables on the risk to develop dementia during PD.

The following equation explains what a binary logistic regression is:

$$\frac{P}{1-P} = e^{a+b_1X_1+b_2X_2+b_3X_3+b_4X_4+\cdots}$$

$$= e^{a} \cdot e^{b_1X_1} \cdot e^{b_2X_2} \cdot e^{b_3X_3} \cdot e^{b_4X_4} \cdot \ldots$$

In our case,

- P is the probability that an individual presents dementia (to belong to the PDD group), whereas 1-P is the probability that an individual presents mild cognitive impairment (to be part of the M.C.I. group).

[1] Due to the small number of individuals analyzed, the relevance of this model is compromised.
[2] The risk variants are underlined.

- The term P / (1-P) is the odds ratio, i.e. the increase or decrease in the risk to evolve from M.C.I. to dementia during PD.
- Each X_i represents a variable (genotype) that influences on the risk (value 1 for the underlined risk variants and 0 for the others) and a, b_1, b_2… are numerical coefficients.

Table BB shows how many subjects were correctly assigned by the model.

Table BB. Classification of the subjects analyzed under the binary logistic regression model.

Observed		**Predicted**		
		cognitive status		Percentage Correct
		M.C.I.	PDD	
cognitive status	M.C.I.	30	6	83.3 (**specificity**)[3]
	PDD	14	9	39.1 (**sensitivity**)
Overall Percentage				66.1

The results we obtained were[4] (Table BC):

[3] Sensitivity can be defined as the proportion of PDD cases that are correctly assigned by the model, that is, capacity to identify positives. Specificity is defined as the proportion of M.C.I. patients that are correctly assigned by the model, i.e. capacity to identify negatives.
The values obtained were low.
As there were two subjects from the PDD group with unknown age at onset, they were not included in the model.

[4] Sex distribution and mean age at onset were similar between both groups and were, therefore, considered in the model but the model did not include them. *PSEN2* was excluded too because was not useful either.

Annex IVc. Analysis of genetic determinants in cognitive status in Parkinson´s disease. M.C.I. vs PDD.

Table BC. Results of the binary logistic regression model to determine the risk of developing dementia.

	b	df	p-value	OR (e^b)	$CI_{95\%}$	
					Lower	Upper
rs2986017 (p.P86L, *CALHM1*) $\mathbf{X_1}$	1.168 $\mathbf{b_1}$	1	**0.059**	3.214 $\mathbf{e^{b1}}$	0.956	10.803
constant	-0.762 **a**	1	**0.019**	0.467		

What would happen if a person carries the risk variant in variable B (X_1)?

$$\text{odds ratio} = e^{a+b_1 X_1} = e^{-0.762+1.168\cdot 1} = e^{0.406} = 1.50$$

$$\frac{P}{1-P} = 1.50;\ \mathrm{P} = 0.60$$

This person would have 1.50 times more possibilities to develop dementia or would develop it with a probability of 60%.

Finally, we analyzed the distribution of the two variables in patients with mild cognitive impairment (M.C.I.) and with dementia (PDD)[5] (Table BD and Figure IV):

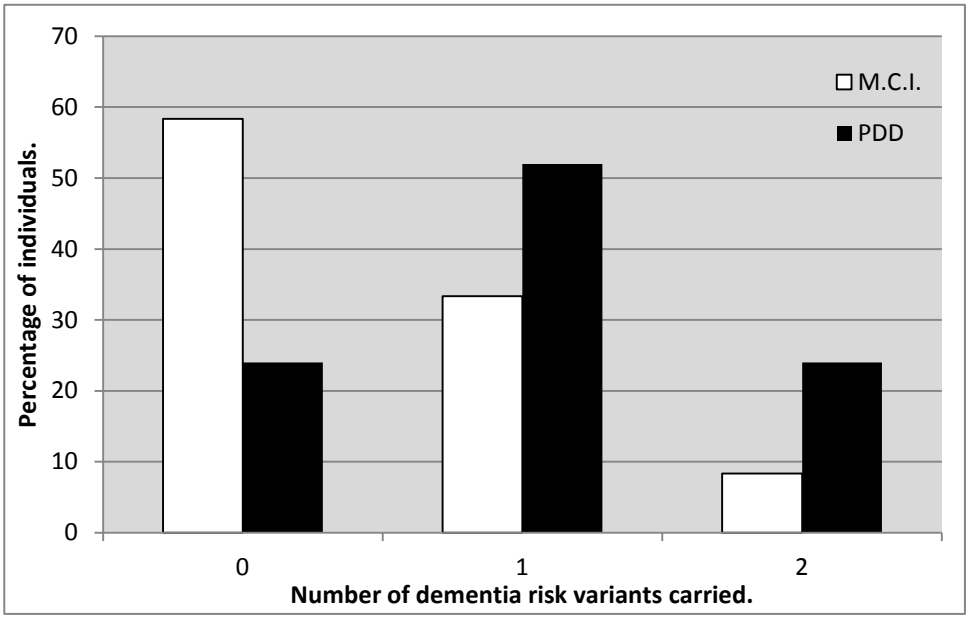
[5] We also tried to determine if there was some correlation between them, but we observed that there was not any. There was not any correlation between these two with the nine PD risk variants or with the three cognitive impairment risk variants either.

Annex IVc. Analysis of genetic determinants in cognitive status in Parkinson´s disease. M.C.I. vs PDD.

Table BD. Percentage of M.C.I. and PDD subjects that carry each number of dementia risk variants. Number of individuals are in parenthesis.

Number of dementia risk variants carried	0	1	2	mean	median
M.C.I.	58.33 (21)	33.33 (12)	8.33 (3)	0.50	1
PDD	24.00 (6)	52.00 (13)	24.00 (6)	1.00	1

Fig. IV. Distribution of the number of dementia risk variants carried by M.C.I. and PDD groups.



Furthermore, we studied how many variants affecting PD (Figure V and Table BE) or cognitive impairment (Figure VI and Table BF) presented individuals of both groups.

Annex IVc. Analysis of genetic determinants in cognitive status in Parkinson´s disease. M.C.I. vs PDD.

Fig. V. Distribution of the number of PD risk variants carried by M.C.I. and PDD groups.

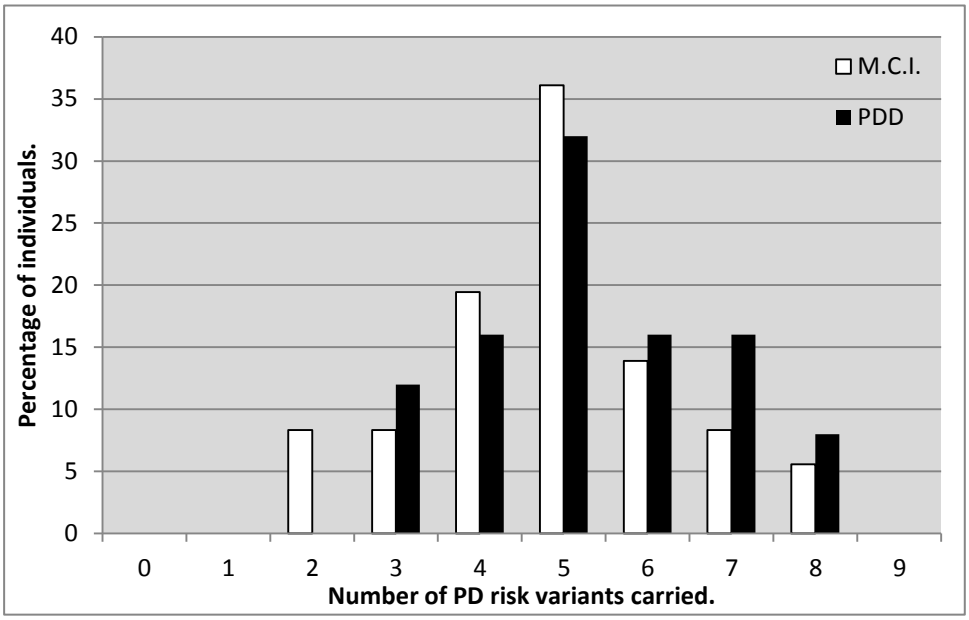


Table BE. Percentage of M.C.I. and PDD cases that carry each number of PD risk variants.
Number of individuals are in parenthesis.

Number of PD risk variants carried	0	1	2	3	4	5	6	7	8	9
M.C.I.	0 (0)	0 (0)	8.33 (3)	8.33 (3)	19.44 (7)	36.11 (13)	13.89 (5)	8.33 (3)	5.56 (2)	0 (0)
PDD	0 (0)	0 (0)	0 (0)	12.00 (3)	16.00 (4)	32.00 (8)	16.00 (4)	16.00 (4)	8.00 (2)	0 (0)

The mean number of PD risk variants carried by the subjects from the M.C.I. group was 4.86, whereas for PDD was 5.32. Medians were 5 and 5, respectively.

Annex IVc. Analysis of genetic determinants in cognitive status in Parkinson´s disease. M.C.I. vs PDD.

Fig. VI. Distribution of the number of cognitive impairment risk variants carried by M.C.I. and PDD groups.

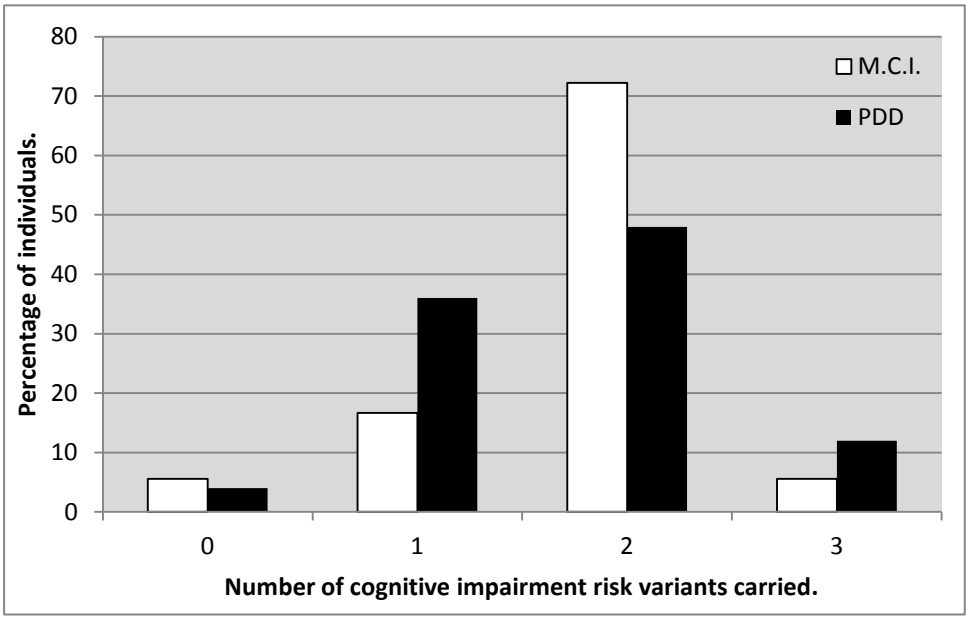


Table BF. Percentage of M.C.I. and PDD cases that carry each number of cognitive impairment risk variants. Number of individuals are in parenthesis.

Number of cognitive impairment risk variants carried	0	1	2	3
M.C.I.	5.56 (2)	16.67 (6)	72.22 (26)	5.56 (2)
PDD	4.00 (1)	36.00 (9)	48.00 (12)	12.00 (3)

The mean number of cognitive impairment risk variants carried by the cases with mild cognitive impairment was 1.78, whereas for those with dementia was 1.68. Medians were 2 and 2, respectively.

Table BG. Frequency of *APOE* genotype when considering the number of ε2 alleles.

	ε2 +/+	ε2 +/-	ε2 -/-	ε2 +	ε2 -
Controls	0	8	87	8	87
PD cases	3	10	79	13	79
	X^2 = 3.561; 2df; p=0.169			X^2 = 1.528; 1df; p=0.216	

Table BH. Genotypic and allelic frequency of poly-T homopolymer (*TOMM40*).

Gene and polymorphism	**Genotypes**						**Alleles**		
TOMM40, rs10524523	ss	sl	ll	vll	vlvl	svl	s	l	vl
Controls	20	7	2	10	21	35	82	21	87
PD cases	21	6	1	6	16	42	90	14	80
	X^2 = 2.699; 5df; p=0.746						X^2 = 1.970; 2df; p=0.373		

Table BI. Genotypic and allelic frequencies of p.S18Y (*UCHL1*) and p.D216H (*TOR1A*).

Gene and polymorphism	**Genotypes**			**Alleles**	
UCHL1, rs5030732 (p.S18Y)	SS	SY	YY	S	Y
Controls	62	29	4	153	37
PD cases	66	23	3	155	29
	X^2 = 0.912; 2df; p=0.634			X^2 = 0.887; 1df; p=0.346	
TOR1A, rs1801968 (p.D216H)	GG	GC	CC	G	C
Controls	78	17	0	173	17
PD cases	69	21	2	159	25
	X^2 = 2.925; 2df; p=0.232			X^2 = 2.018; 1df; p=0.155	

With regard to the mutation c.864+246C>T in *NR4A2*, 2 controls were heterozygous: there was no statistical significance (Fisher´s exact test, p=0.497).

Table BJ. Genotypic and allelic frequency of rs6616255 (*ARMCX1*).

Gene and polymorphism	**Genotypes**			**Alleles**	
ARMCX1, rs6616255	AA	AG	GG	A	G
Controls, overall	72	13	10	110	23
PD cases, overall	75	9	8	109	19
	X^2 = 0.963; 2df; p=0.618 Z_{mfG}^2 = 0.821; 1df; p=0.365 Z_C^2 = 1.108; 2df; p=0.575			$X^2 = Z_A^2$ = 0.289; 1df; p=0.590 Z_{mfA}^2 = 0.639; 1df; p=0.424	
Controls, women	25	13	0	63	13
PD cases, women	25	9	2	59	13
	X^2 = 2.675; 2df; p=0.262 Z_{fG}^2 = 0.024; 1df; p=0.878			$X^2 = Z_{fA}^2$ = 0.023; 1df; p=0.879	
Controls, men	-	-	-	47	10
PD cases, men	-	-	-	50	6
	-			$X^2 = Z_m^2$ =1.084; 1df; p=0.298	

Table BK. Genotypic and allelic frequency of rs1044275 (*ARMCX1*).

Gene and polymorphism	**Genotypes**			**Alleles**	
ARMCX1, rs1044275	CC	CT	TT	C	T
Controls, overall	83	10	2	121	12
PD cases, overall	78	8	6	114	14
	X^2 = 2.330; 2df; p=0.312 Z_{mfG}^2 = 2.013; 1df; p=0.156 Z_C^2 = 2.398; 2df; p=0.302			$X^2 = Z_A^2$ = 0.267; 1df; p=0.606 Z_{mfA}^2 = 1.611; 1df; p=0.204	
Controls, women	28	10	0	66	10
PD cases, women	28	8	0	64	8
	X^2 = 0.168; 1df; p=0.682 Z_{fG}^2 = 0.168; 1df; p=0.681			$X^2 = Z_{fA}^2$ = 0.145; 1df; p=0.703	
Controls, men	-	-	-	55	2
PD cases, men	-	-	-	50	6
				$X^2 = Z_m^2$ = 2.229; 1df; p=0.135	

Table BL. Genotypic and allelic frequency of rs5951282 (*ARMCX2*).

Gene and polymorphism	**Genotypes**			**Alleles**	
ARMCX2, rs5951282	CC	CT	TT	C	T
Controls, overall	49	22	24	85	48
PD cases, overall	55	20	17	90	38
	X^2 = 1.589; 2df; p=0.452 Z_{mfG}^2 = 1.519; 1df; p=0.218 Z_C^2 = 1.592; 2df; p=0.451			$X^2 = Z_A^2$ = 1.210; 1df; p=0.271 Z_{mfA}^2 = 1.209; 1df; p=0.272	
Controls, women	14	22	2	50	26
PD cases, women	15	20	1	50	22
	X^2 = 0.409; 2df; p=0.815 Z_{fG}^2 = 0.320; 1df; p=0.572			$X^2 = Z_{fA}^2$ = 0.225; 1df; p=0.635	
Controls, men	-	-	-	35	22
PD cases, men	-	-	-	40	16
	-			$X^2 = Z_m^2$ = 1.272; 1df; p=0.259	

Table BM. Genotypic and allelic frequency of rs6995 (*ARMCX3*).

Gene and polymorphism	Genotypes			Alleles	
ARMCX3, rs6995	AA	AG	GG	A	G
Controls, overall	75	13	7	112	21
PD cases, overall	68	10	14	102	26
	X^2 = 3.020; 2df; p=0.221 Z_{mfG}^2 = 1.585; 1df; p=0.208 Z_C^2 = 2.509; 2df; p=0.285			$X^2 = Z_A^2$ = 0.904; 1df; p=0.341 Z_{mfA}^2 = 1.157; 1df; p=0.282	
Controls, women	24	13	1	61	15
PD cases, women	24	10	2	58	14
	X^2 = 0.671; 2df; p=0.715 Z_{fG}^2 = 0.002; 1df; p=0.965			$X^2 = Z_{fA}^2$ = 0.002; 1df; p=0.964	
Controls, men	-	-	-	51	6
PD cases, men	-	-	-	44	12
	-			$X^2 = Z_m^2$ = 2.507; 1df; p=0.113	

Table BN. Genotypic and allelic frequency of rs2179670 (*ARMCX4*).

Gene and polymorphism	**Genotypes**			**Alleles**	
ARMCX4, rs2179670	CC	CT	TT	C	T
Controls, overall	65	14	16	99	34
PD cases, overall	59	15	18	89	39
	X^2 = 0.394; 2df; p=0.821 Z_{mfG}^2 = 0.485; 1df; p=0.486 Z_C^2 = 1.097; 2df; p=0.578			$X^2 = Z_A^2$ = 0.779; 1df; p=0.377 Z_{mfA}^2 = 0.747; 1df; p=0.387	
Controls, women	20	14	4	54	22
PD cases, women	15	15	6	45	27
	X^2 = 1.096; 2df; p=0.578 Z_{fG}^2 = 1.095; 1df; p=0.295			$X^2 = Z_{fA}^2$ = 1.221; 1df; p=0.269	
Controls, men	-	-	-	45	12
PD cases, men	-	-	-	44	12
	-			$X^2 = Z_m^2$ = 0.002; 1df; p=0.961	

Table BO. Genotypic and allelic frequency of rs6523506 (*ARMCX4*).

Gene and polymorphism	Genotypes			Alleles	
ARMCX4, rs6523506	GG	TG	TT	G	T
Controls, overall	72	14	9	110	23
PD cases, overall	65	14	13	99	29
	X^2 = 1.037; 2df; p=0.595 Z_{mfG}^2 = 1.139; 1df; p=0.286 Z_C^2 = 1.357; 2df; p=0.507			$X^2 = Z_A^2$ = 1.176; 1df; p=0.278 Z_{mfA}^2 = 1.177; 1df; p=0.278	
Controls, women	24	14	0	62	14
PD cases, women	20	14	2	54	18
	X^2 = 2.311; 2df; p=0.315 Z_{fG}^2 = 1.069; 1df; p=0.301			$X^2 = Z_{fA}^2$ = 0.944; 1df; p=0.331	
Controls, men	-	-	-	48	9
PD cases, men	-	-	-	45	11
	-			$X^2 = Z_m^2$ = 0.288; 1df; p=0.592	

It is noteworthy that the *ARMCX* genes are located at chromosome X and, therefore, males present only one allele whereas females present two. Pearson´s X^2 may not be appropriate to test for association of X chromosome markers but there are not standardized association tests. We decided to calculate the statistical tests proposed by [1][1].

[1] Z_{mfA}^2 and Z_{mfG}^2 are the weighted sum of Z_m^2 plus Z_{fA}^2 or Z_{fG}^2, respectively. Both are modified tests to allow the differential allele effects in males and females. Z_C^2 is a genotype-based test where allele frequencies are estimated separately for males and females.

Table BP. Genotypic and allelic frequency of rs2235827 (*ARMCX5*).

Gene and polymorphism	Genotypes			Alleles	
ARMCX5, rs2235827	CC	CT	TT	C	T
Controls, overall	85	6	4	123	10
PD cases, overall	76	9	7	112	16
	X^2 = 1.874; 2df; p=0.392 Z_{mfG}^2 = 1.915; 1df; p=0.166 Z_C^2 = 1.937; 2df; p=0.380			$X^2 = Z_A^2$ = 1.804; 1df; p=0.179 Z_{mfA}^2 = 1.810; 1df; p=0.178	
Controls, women	32	6	0	70	6
PD cases, women	27	9	0	63	9
	X^2 = 0.970; 1df; p=0.325 Z_{fG}^2 = 0.970; 1df; p=0.325			$X^2 = Z_{fA}^2$ = 0.861; 1df; p=0.353	
Controls, men	-	-	-	53	4
PD cases, men	-	-	-	49	7
	-			$X^2 = Z_m^2$ = 0.966; 1df; p=0.326	

Bibliography:

1. Zheng, G., et al., *Testing association for markers on the X chromosome.* Genet Epidemiol, 2007. **31**(8): p. 834-43.

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

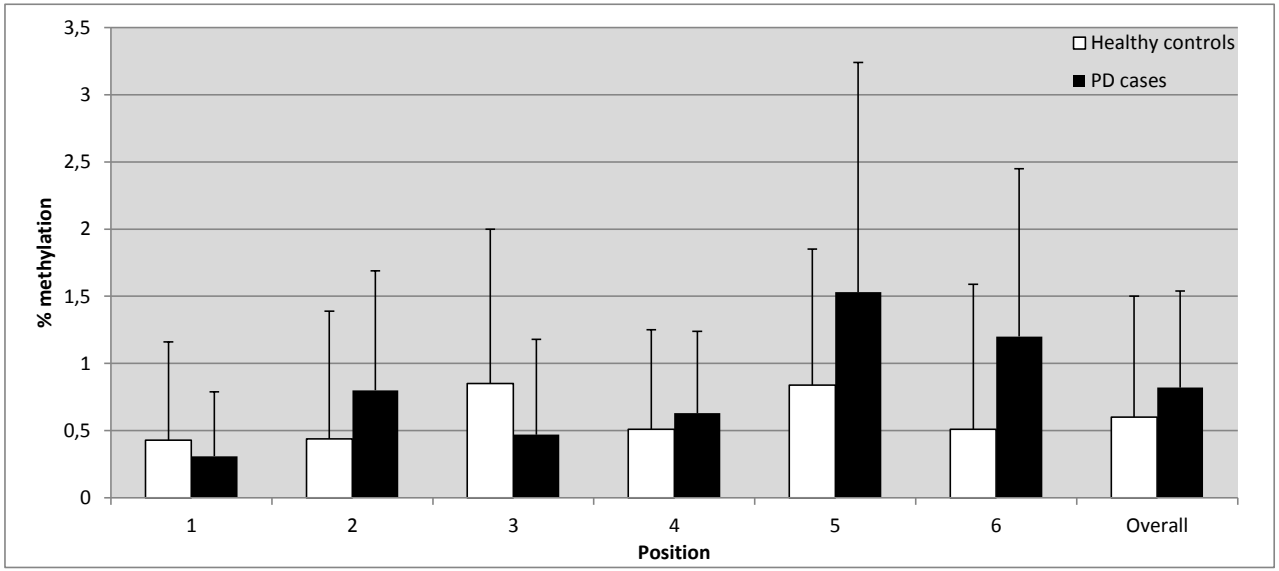
***DJ-1* 1**

Parietal cortex

Table BQ. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	Overall
Healthy controls	0,43 (0,73)	0,44 (0,95)	0,85 (1,15)	0,51 (0,74)	0,84 (1,01)	0,51 (1,08)	0,60 (0,90)
PD cases	0,31 (0,48)	0,80 (0,89)	0,47 (0,71)	0,63 (0,61)	1,53 (1,71)	1,20 (1,25)	0,82 (0,72)
p-value	0,841	0,841	0,841	0,841	0,421	0,548	0,548

Figure VII. Graphical representation of the results detailed in Table BQ.



Blood

Table BR. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	Overall
Other PD cases	1,14 (1,34)	0,58 (0,76)	1,10 (0,97)	0,54 (0,75)	1,21 (1,55)	0,70 (0,99)	0,88 (1,00)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

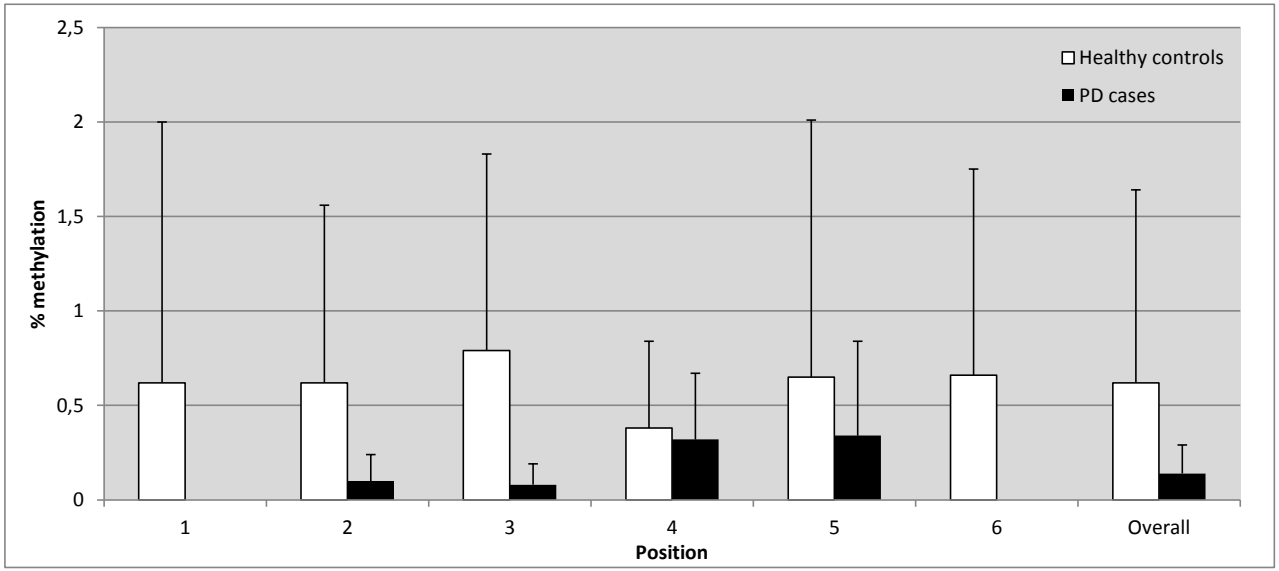
***DJ-1* 1**

Occipital cortex

Table BS. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Healthy controls	0,62 (1,38)	0,62 (0,94)	0,79 (1,04)	0,38 (0,46)	0,65 (1,36)	0,66 (1,09)	0,62 (1,02)
PD cases	0	0,10 (0,14)	0,08 (0,11)	0,32 (0,35)	0,34 (0,50)	0	0,14 (0,15)
p-value	0,690	0,548	0,310	0,841	0,841	0,310	0,690

Figure VIII. Graphical representation of the results detailed in Table BS.



Blood

Table BT. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Other PD cases	1,14 (1,34)	0,58 (0,76)	1,10 (0,97)	0,54 (0,75)	1,21 (1,55)	0,70 (0,99)	0,88 (1,00)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

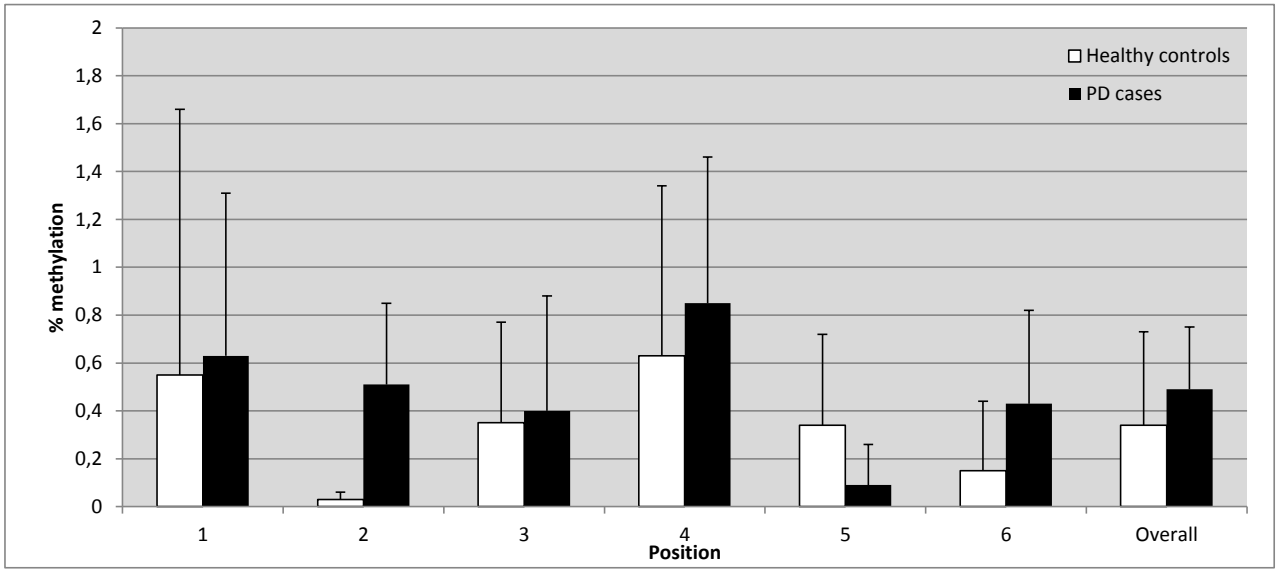
***DJ-1* 1**

Substantia nigra

Table BU. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Healthy controls	0,55 (1,11)	0,03 (0,03)	0,35 (0,42)	0,63 (0,71)	0,34 (0,38)	0,15 (0,29)	0,34 (0,39)
PD cases	0,63 (0,68)	0,51 (0,34)	0,40 (0,48)	0,85 (0,61)	0,09 (0,17)	0,43 (0,39)	0,49 (0,26)
p-value	0,730	0,111	1,000	0,556	0,190	0,111	0,556

Figure IX. Graphical representation of the results detailed in Table BU.



Blood

Table BV. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Other PD cases	1,14 (1,34)	0,58 (0,76)	1,10 (0,97)	0,54 (0,75)	1,21 (1,55)	0,70 (0,99)	0,88 (1,00)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

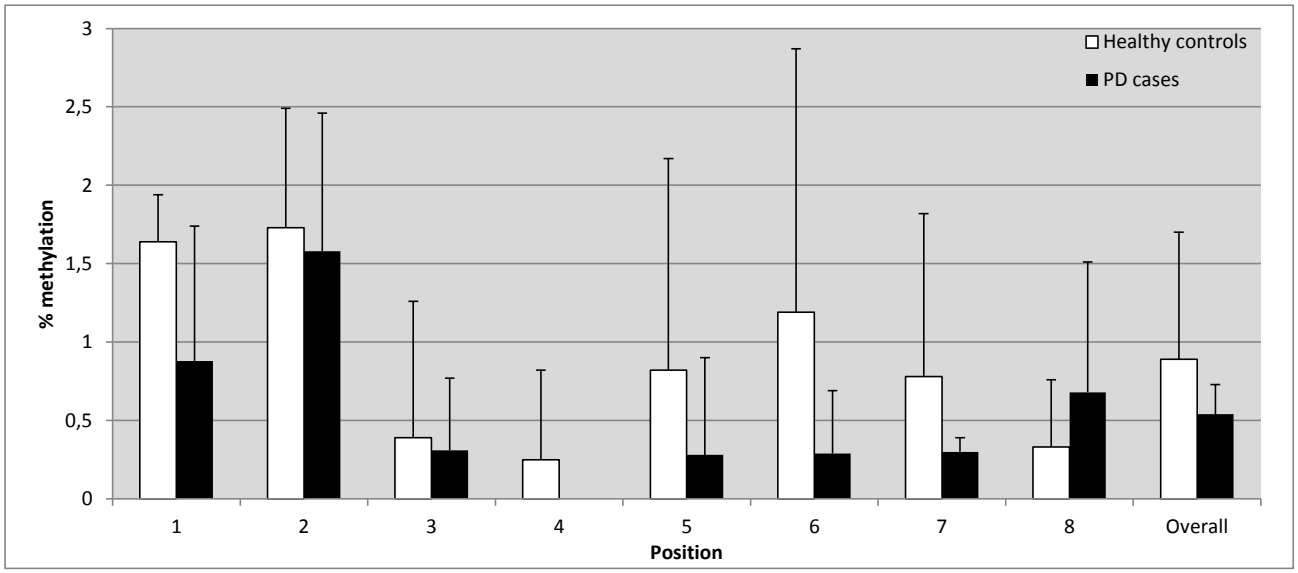
***DJ-1* 2**

Parietal cortex

Table BW. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	Overall
Healthy controls	1,64 (0,30)	1,73 (0,76)	0,39 (0,87)	0,25 (0,57)	0,82 (1,35)	1,19 (1,68)	0,78 (1,04)	0,33 (0,43)	0,89 (0,81)
PD cases	0,88 (0,86)	1,58 (0,88)	0,31 (0,46)	0	0,28 (0,62)	0,29 (0,40)	0,30 (0,09)	0,68 (0,83)	0,54 (0,19)
p-value	0,151	0,841	0,841	0,690	0,421	0,548	0,690	0,690	0,690

Figure X. Graphical representation of the results detailed in Table BW.



Blood

Table BX. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	Overall
Other PD cases	0,63 (1,28)	2,16 (1,71)	1,20 (1,15)	0,63 (1,35)	0,71 (0,72)	0,09 (0,24)	0,73 (0,99)	1,32 (1,17)	0,93 (0,52)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

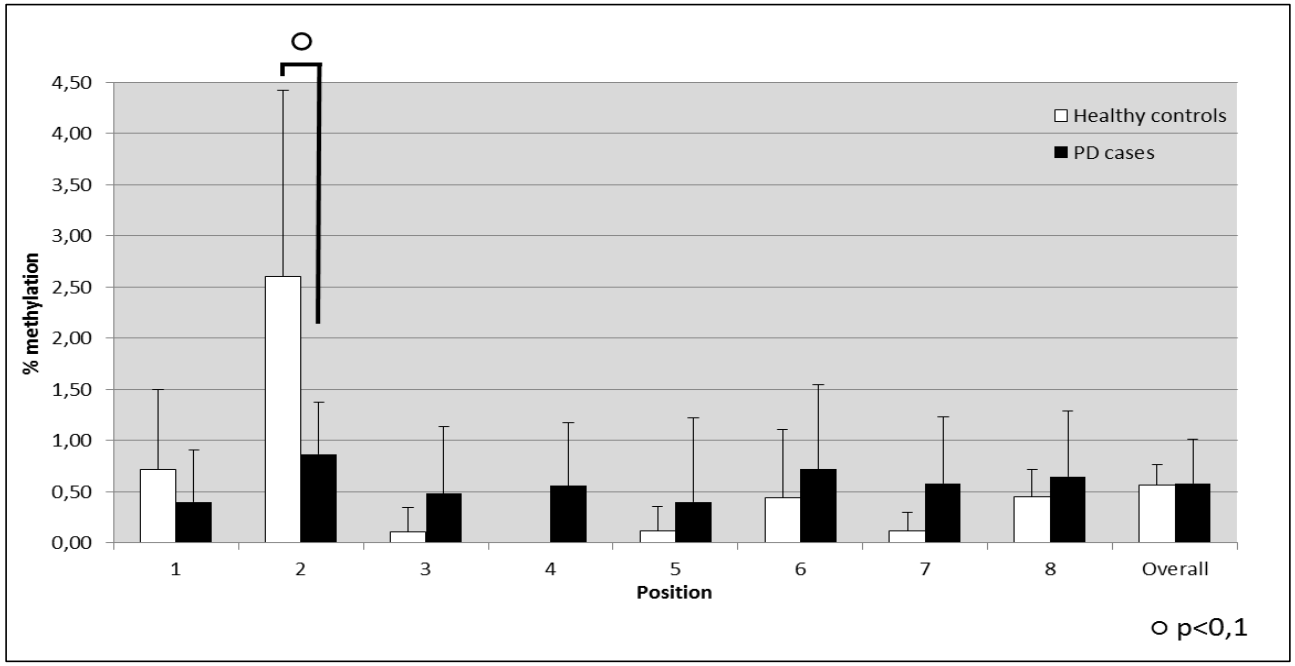
***DJ-1* 2**

Occipital cortex

Table BY. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	Overall
Healthy controls	0,72 (0,78)	2,60 (1,82)	0,11 (0,24)	0	0,12 (0,24)	0,44 (0,67)	0,12 (0,18)	0,45 (0,27)	0,57 (0,20)
PD cases	0,40 (0,51)	0,87 (0,50)	0,48 (0,66)	0,56 (0,61)	0,40 (0,82)	0,72 (0,83)	0,58 (0,65)	0,65 (0,64)	0,58 (0,43)
p-value	0,548	**0,095**	0,548	0,151	0,841	0,548	0,222	0,841	0,690

Figure XI. Graphical representation of the results detailed in Table BY.



Blood

Table BZ. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	Overall
Other PD cases	0,63 (1,28)	2,16 (1,71)	1,20 (1,15)	0,63 (1,35)	0,71 (0,72)	0,09 (0,24)	0,73 (0,99)	1,32 (1,17)	0,93 (0,52)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

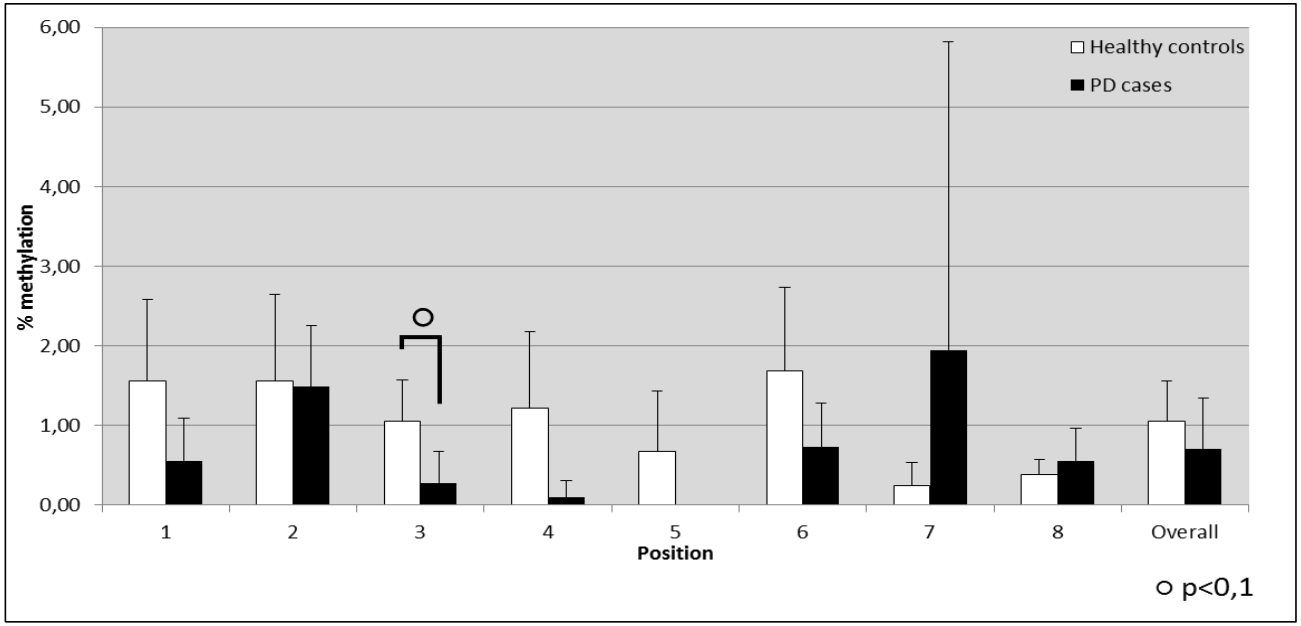
***DJ-1* 2**

Substantia nigra

Table CA. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Healthy controls	1,56 (1,02)	1,56 (1,09)	1,05 (0,52)	1,22 (0,96)	0,67 (0,77)	1,69 (1,04)	0,24 (0,30)	0,39 (0,18)	1,05 (0,51)
PD cases	0,56 (0,53)	1,49 (0,77)	0,28 (0,39)	0,10 (0,21)	0	0,73 (0,55)	1,95 (3,87)	0,56 (0,41)	0,71 (0,63)
p-value	0,190	1,000	**0,063**	0,111	0,286	0,286	1,000	0,556	0,413

Figure XII. Graphical representation of the results detailed in Table CA.



Blood

Table CB. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Other PD cases	0,63 (1,28)	2,16 (1,71)	1,20 (1,15)	0,63 (1,35)	0,71 (0,72)	0,09 (0,24)	0,73 (0,99)	1,32 (1,17)	0,93 (0,52)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

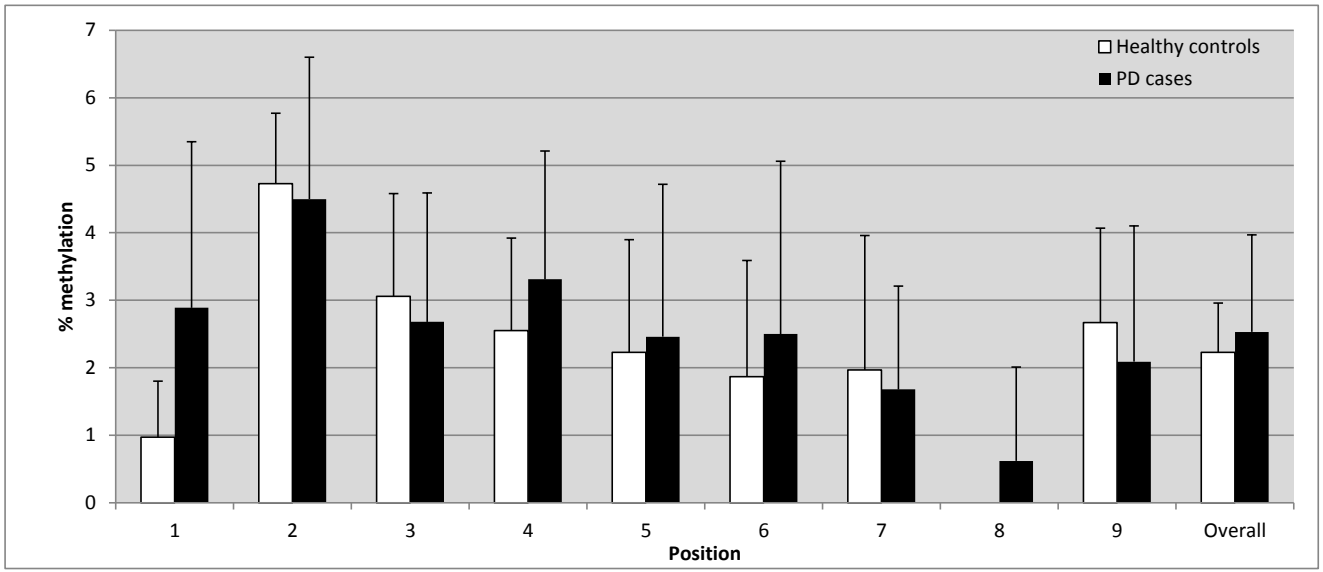
***LRRK2* 1**

Parietal cortex

Table CC. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,97 (0,83)	4,73 (1,04)	3,06 (1,52)	2,55 (1,37)	2,23 (1,67)	1,87 (1,72)	1,97 (1,99)	0	2,67 (1,40)	2,23 (0,73)
PD cases	2,89 (2,46)	4,50 (2,10)	2,68 (1,91)	3,31 (1,90)	2,46 (2,26)	2,50 (2,56)	1,68 (1,53)	0,62 (1,39)	2,09 (2,01)	2,53 (1,44)
p-value	0,222	1,000	0,690	0,690	0,841	0,841	0,690	0,690	0,690	0,841

Figure XIII. Graphical representation of the results detailed in Table CC.



Blood

Table CD. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	3,76 (2,82)	2,71 (1,72)	3,57 (2,27)	3,74 (3,09)	2,15 (1,83)	3,63 (2,66)	2,87 (1,74)	2,65 (1,85)	3,30 (3,14)	3,15 (1,54)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

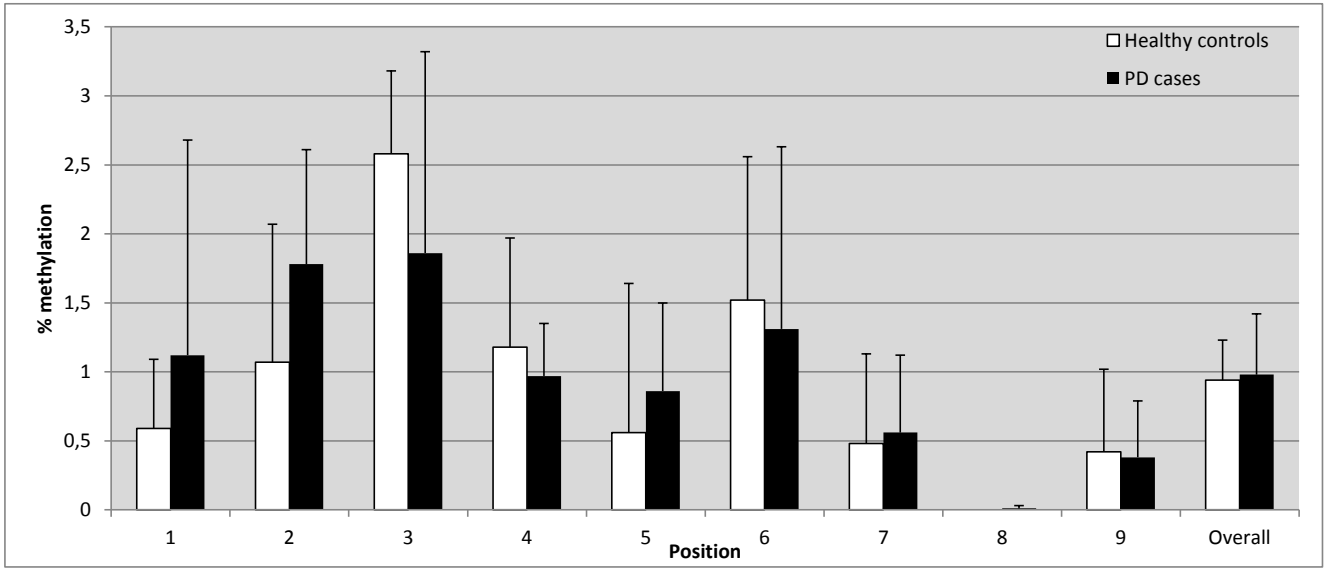
***LRRK2* 1**

Occipital cortex

Table CE. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,59 (0,50)	1,07 (1,00)	2,58 (0,60)	1,18 (0,79)	0,56 (1,08)	1,52 (1,04)	0,48 (0,65)	0	0,42 (0,60)	0,94 (0,29)
PD cases	1,12 (1,56)	1,78 (0,83)	1,86 (1,46)	0,97 (0,38)	0,86 (0,64)	1,31 (1,32)	0,56 (0,56)	0,01 (0,02)	0,38 (0,41)	0,98 (0,44)
p-value	0,841	0,310	0,310	0,548	0,310	1,000	1,000	0,690	1,000	1,000

Figure XIV. Graphical representation of the results detailed in Table CE.



Blood

Table CF. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	3,76 (2,82)	2,71 (1,72)	3,57 (2,27)	3,74 (3,09)	2,15 (1,83)	3,63 (2,66)	2,87 (1,74)	2,65 (1,85)	3,30 (3,14)	3,15 (1,54)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

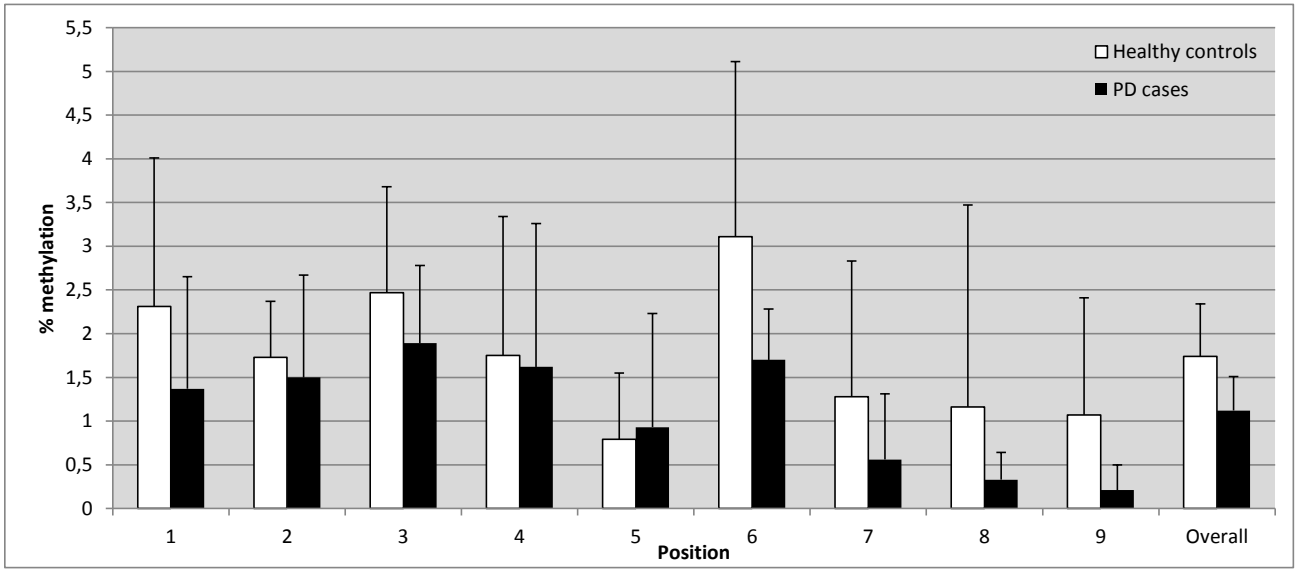
***LRRK2* 1**

Substantia nigra

Table CG. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	2,31 (1,70)	1,73 (0,64)	2,47 (1,21)	1,75 (1,59)	0,79 (0,76)	3,11 (2,00)	1,28 (1,55)	1,16 (2,31)	1,07 (1,34)	1,74 (0,60)
PD cases	1,37 (1,28)	1,50 (1,17)	1,89 (0,89)	1,62 (1,64)	0,93 (1,30)	1,70 (0,58)	056 (0,75)	0,33 (0,31)	0,21 (0,59)	1,12 (0,39)
p-value	0,413	0,905	0,556	0,905	0,905	0,286	0,730	0,413	0,556	0,111

Figure XV. Graphical representation of the results detailed in Table CG.



Blood

Table CH. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	3,76 (2,82)	2,71 (1,72)	3,57 (2,27)	3,74 (3,09)	2,15 (1,83)	3,63 (2,66)	2,87 (1,74)	2,65 (1,85)	3,30 (3,14)	3,15 (1,54)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

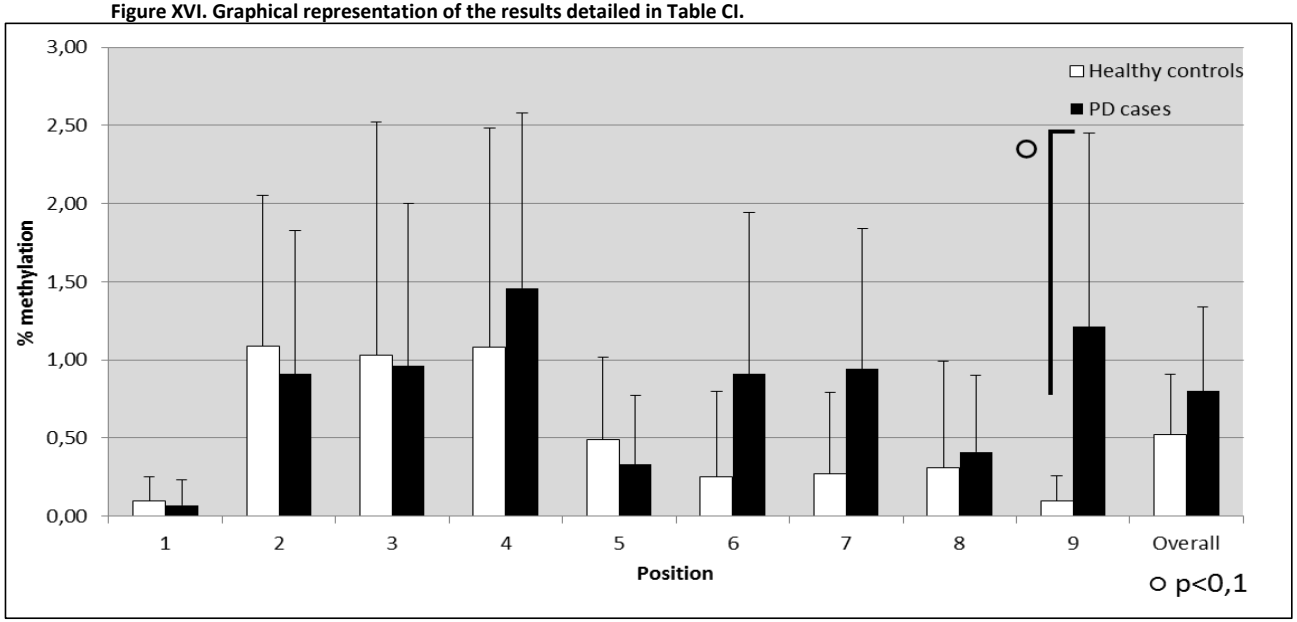
***LRRK2* 2**

Parietal cortex

Table CI. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Overall**
Healthy controls	0,10 (0,15)	1,09 (0,96)	1,03 (1,49)	1,08 (1,40)	0,49 (0,53)	0,25 (0,55)	0,27 (0,52)	0,31 (0,68)	0,10 (0,16)	0,52 (0,39)
PD cases	0,07 (0,16)	0,91 (0,92)	0,96 (1,04)	1,46 (1,12)	0,33 (0,44)	0,91 (1,03)	0,94 (0,90)	0,41 (0,49)	1,21 (1,24)	0,80 (0,54)
p-value	0,841	0,841	0,841	0,421	0,841	0,310	0,151	0,548	**0,095**	0,421

Figure XVI. Graphical representation of the results detailed in Table CI.



Blood

Table CJ. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Overall**
Other PD cases	1,57 (0,82)	1,31 (1,00)	0,71 (0,71)	0,57 (0,54)	1,34 (0,52)	1,05 (0,62)	0,77 (0,49)	3,31 (1,74)	1,13 (0,79)	1,31 (0,45)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

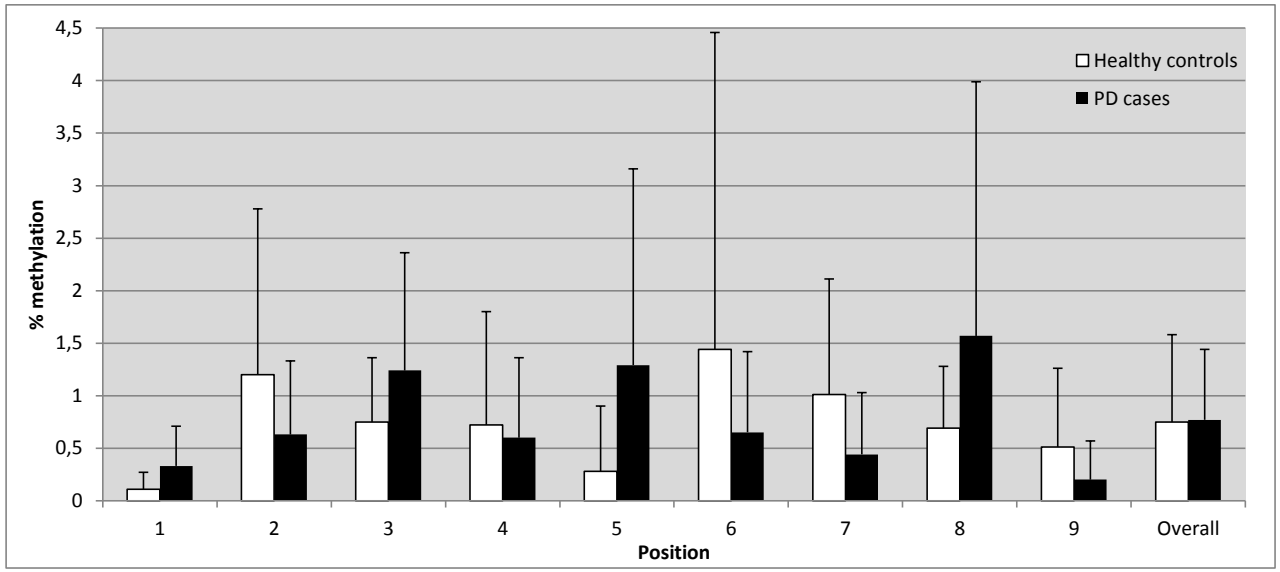
LRRK2 2

Occipital cortex

Table CK. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,11 (0,16)	1,20 (1,58)	0,75 (0,61)	0,72 (1,08)	0,28 (0,62)	1,44 (3,02)	1,01 (1,10)	0,69 (0,59)	0,51 (0,75)	0,75 (0,83)
PD cases	0,33 (0,38)	0,63 (0,70)	1,24 (1,12)	0,60 (0,76)	1,29 (1,87)	0,65 (0,77)	0,44 (0,59)	1,57 (2,42)	0,20 (0,37)	0,77 (0,67)
p-value	0,421	0,690	0,421	0,690	0,222	0,421	0,841	1,000	0,690	0,690

Figure XVII. Graphical representation of the results detailed in Table CK.



Blood

Table CL. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	1,57 (0,82)	1,31 (1,00)	0,71 (0,71)	0,57 (0,54)	1,34 (0,52)	1,05 (0,62)	0,77 (0,49)	3,31 (1,74)	1,13 (0,79)	1,31 (0,45)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

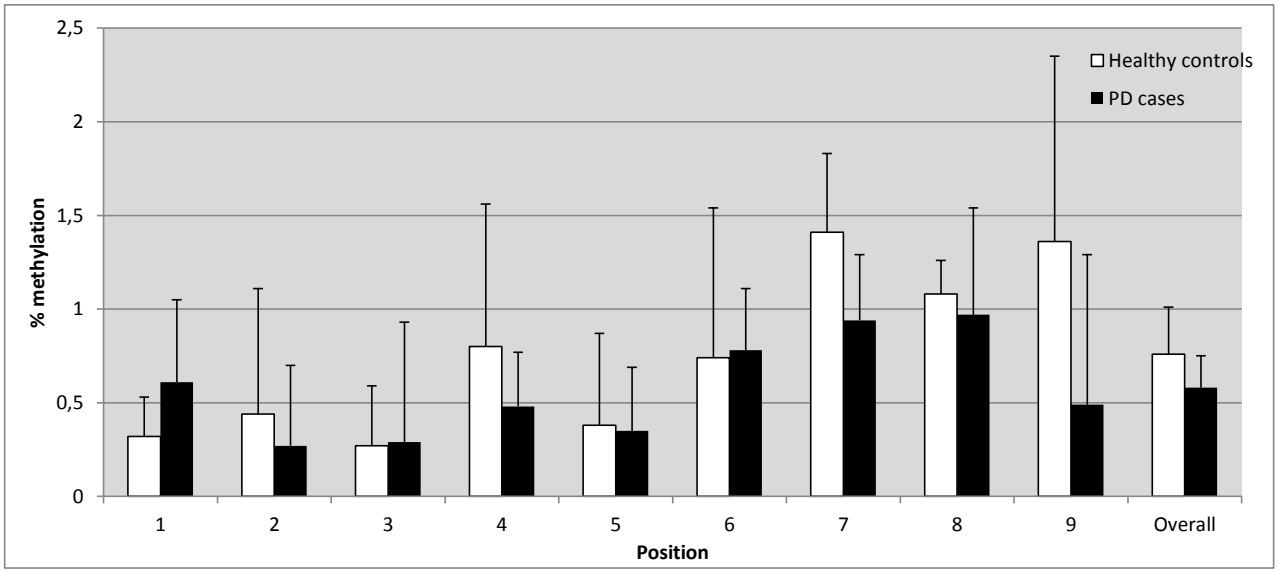
LRRK2 2

Substantia nigra

Table CM. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,32 (0,21)	0,44 (0,67)	0,27 (0,32)	0,80 (0,76)	0,38 (0,49)	0,74 (0,80)	1,41 (0,42)	1,08 (0,18)	1,36 (0,99)	0,76 (0,25)
PD cases	0,61 (0,44)	0,27 (0,43)	0,29 (0,64)	0,48 (0,29)	0,35 (0,34)	0,78 (0,33)	0,94 (0,35)	0,97 (0,57)	0,49 (0,80)	0,58 (0,17)
p-value	0,413	0,905	0,905	0,556	1,000	0,556	0,190	0,730	0,286	0,190

Figure XVIII. Graphical representation of the results detailed in Table CM.



Blood

Table CN. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	1,57 (0,82)	1,31 (1,00)	0,71 (0,71)	0,57 (0,54)	1,34 (0,52)	1,05 (0,62)	0,77 (0,49)	3,31 (1,74)	1,13 (0,79)	1,31 (0,45)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

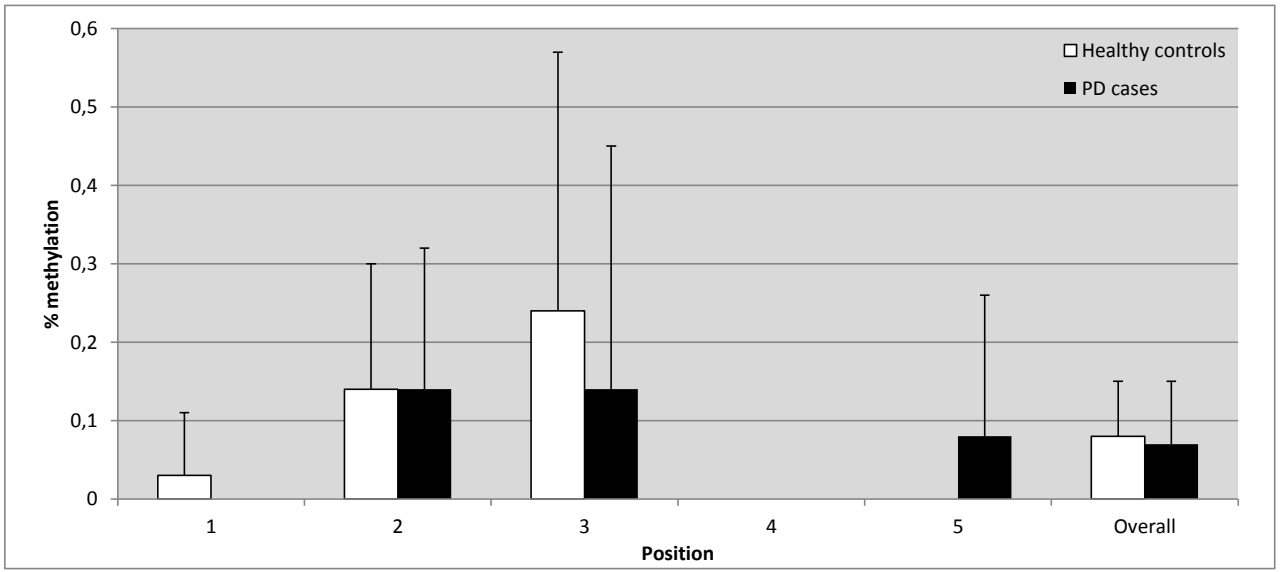
***PINK1* 2**

Parietal cortex

Table CO. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	Overall
Healthy controls	0,03 (0,08)	0,14 (0,16)	0,24 (0,33)	0	0	0,08 (0,07)
PD cases	0	0,14 (0,18)	0,14 (0,31)	0	0,08 (0,18)	0,07 (0,08)
p-value	0,690	1,000	0,690	1,000	0,690	0,841

Figure XIX. Graphical representation of the results detailed in Table CO.



Blood

Table CP. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	Overall
Other PD cases	0,72 (1,07)	0,09 (0,24)	0,04 (0,10)	0,14 (0,25)	0,01 (0,05)	0,20 (0,30)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

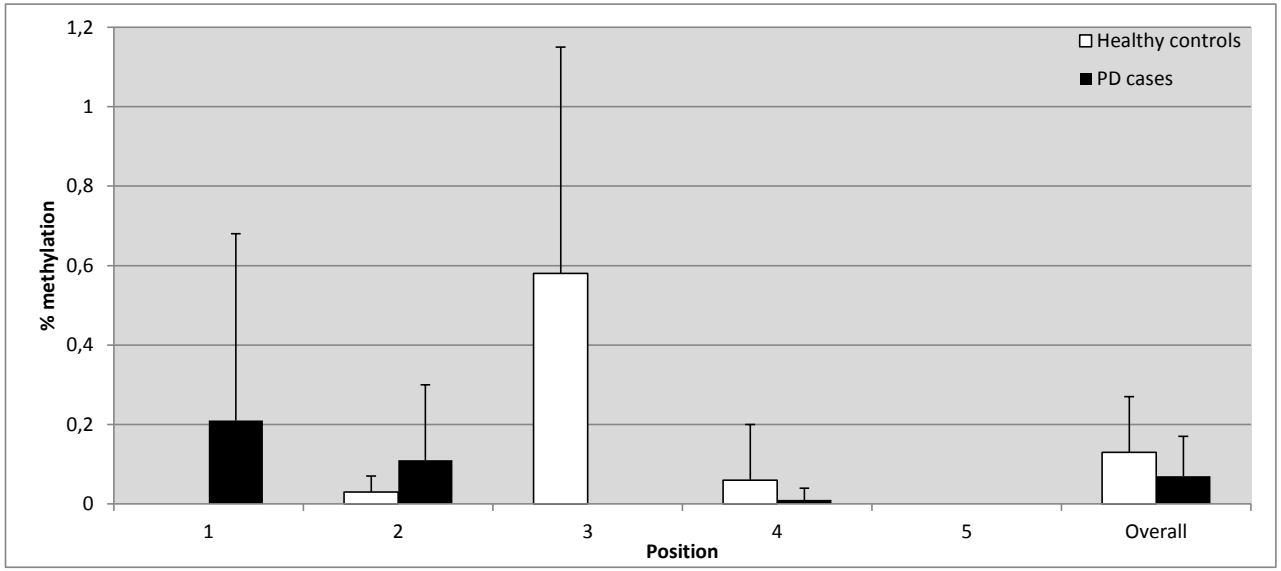
***PINK1* 2**

Occipital cortex

Table CQ. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Healthy controls	0	0,03 (0,04)	0,58 (0,57)	0,06 (0,14)	0	0,13 (0,14)
PD cases	0,21 (0,47)	0,11 (0,19)	0	0,01 (0,03)	0	0,07 (0,10)
p-value	0,690	0,690	0,222	1,000	1,000	0,421

Figure XX. Graphical representation of the results detailed in Table CQ.



Blood

Table CR. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Other PD cases	0,72 (1,07)	0,09 (0,24)	0,04 (0,10)	0,14 (0,25)	0,01 (0,05)	0,20 (0,30)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

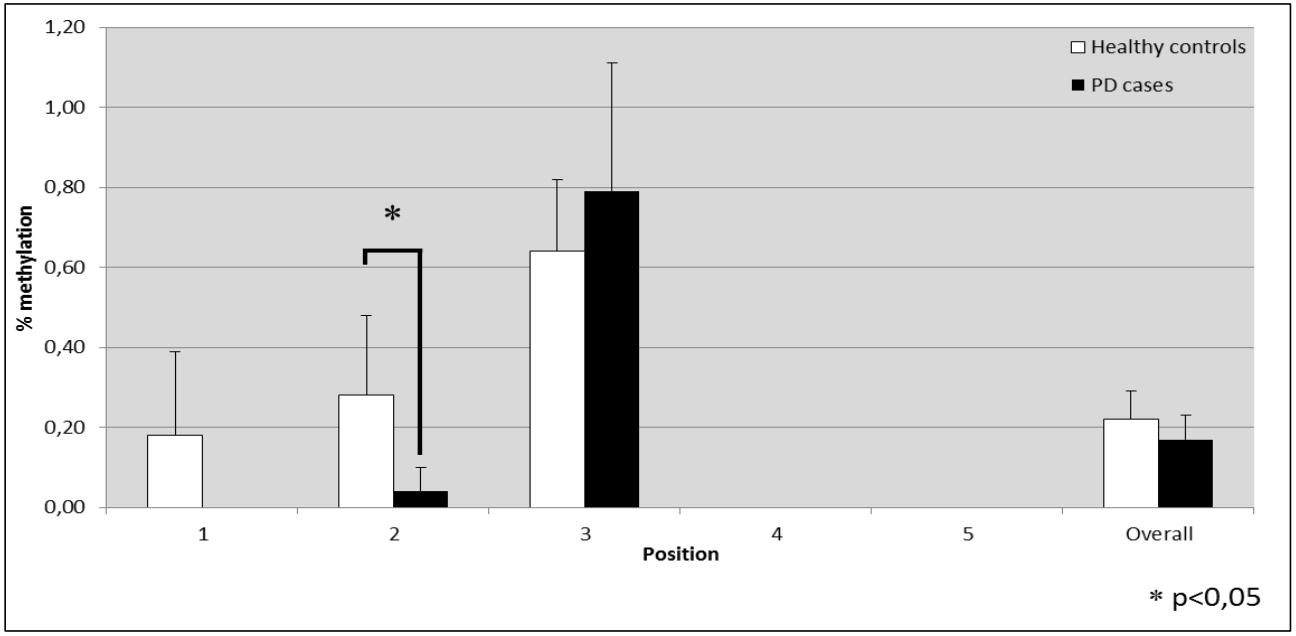
***PINK1* 2**

Substantia nigra

Table CS. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	Overall
Healthy controls	0,18 (0,21)	0,28 (0,20)	0,64 (0,18)	0	0	0,22 (0,07)
PD cases	0	0,04 (0,06)	0,79 (0,32)	0	0	0,17 (0,06)
p-value	0,286	**0,016**	0,556	1,000	1,000	0,286

Figure XXI. Graphical representation of the results detailed in Table CS.



Blood

Table CT. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	Overall
Other PD cases	0,72 (1,07)	0,09 (0,24)	0,04 (0,10)	0,14 (0,25)	0,01 (0,05)	0,20 (0,30)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

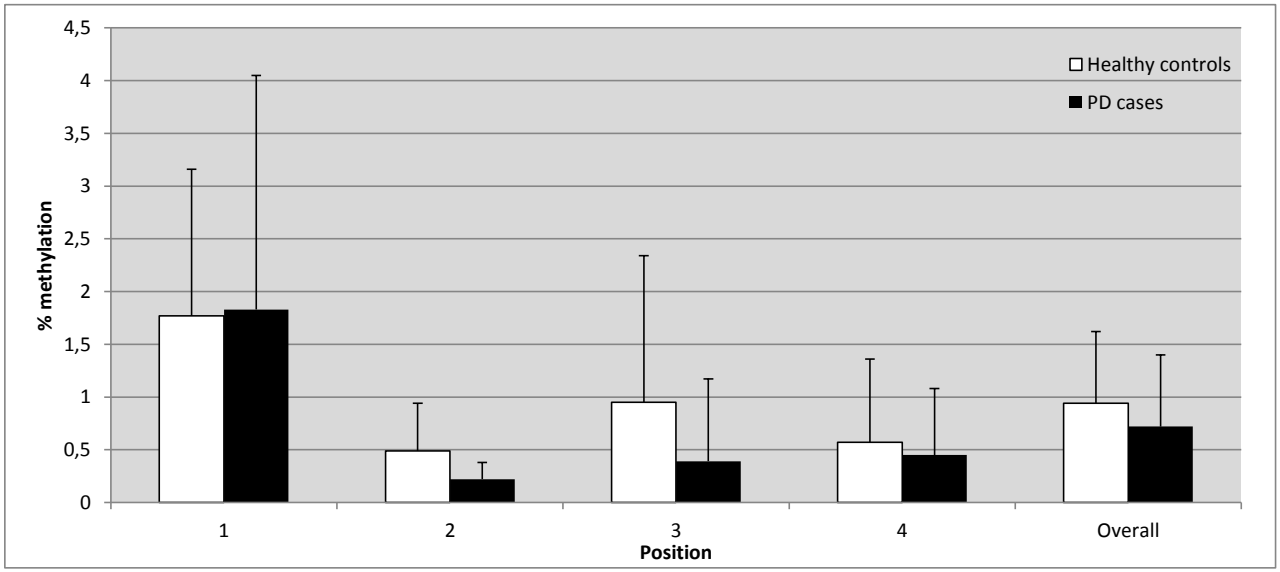
***PRKN* 1**

Parietal cortex

Table CU. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**Overall**
Healthy controls	1,77 (1,39)	0,49 (0,45)	0,95 (1,39)	0,57 (0,79)	0,94 (0,68)
PD cases	1,83 (2,22)	0,22 (0,16)	0,39 (0,78)	0,45 (0,63)	0,72 (0,68)
p-value	0,841	0,421	0,841	0,841	0,548

Figure XXII. Graphical representation of the results detailed in Table CU.



Blood

Table CV. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**Overall**
Other PD cases	0,65 (0,41)	0,26 (0,26)	1,15 (2,09)	0,46 (0,39)	0,63 (0,59)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

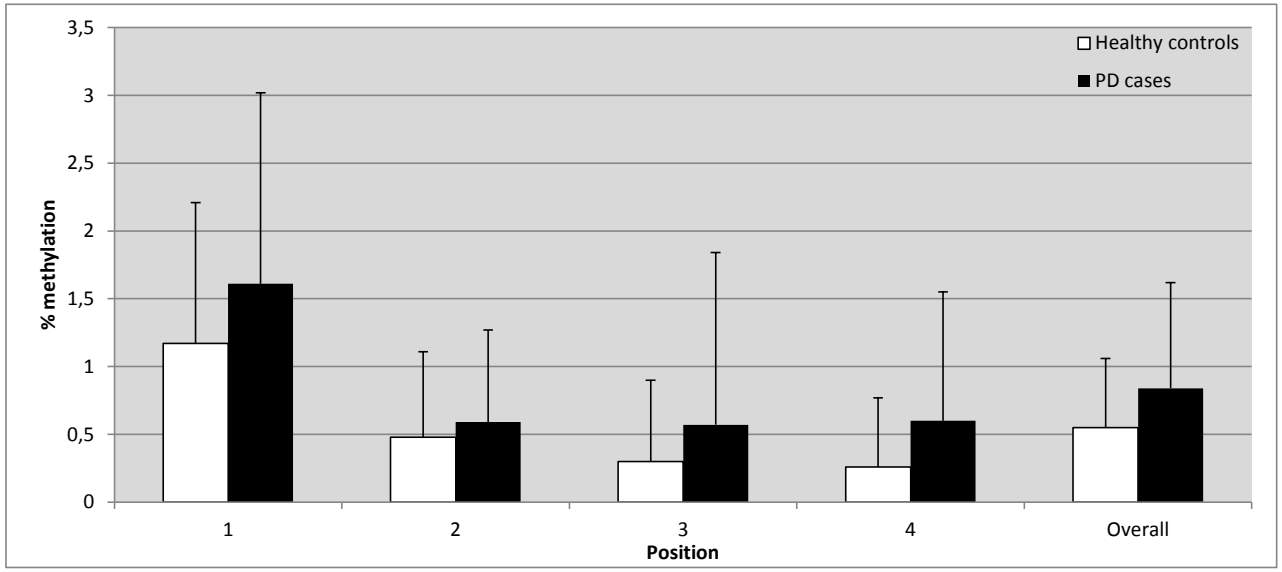
***PRKN* 1**

Occipital cortex

Table CW. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**Overall**
Healthy controls	1,17 (1,04)	0,48 (0,63)	0,30 (0,60)	0,26 (0,51)	0,55 (0,51)
PD cases	1,61 (1,41)	0,59 (0,68)	0,57 (1,27)	0,60 (0,95)	0,84 (0,78)
p-value	0,690	0,548	0,841	0,690	0,548

Figure XXIII. Graphical representation of the results detailed in Table CW.



Blood

Table CX. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**Overall**
Other PD cases	0,65 (0,41)	0,26 (0,26)	1,15 (2,09)	0,46 (0,39)	0,63 (0,59)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

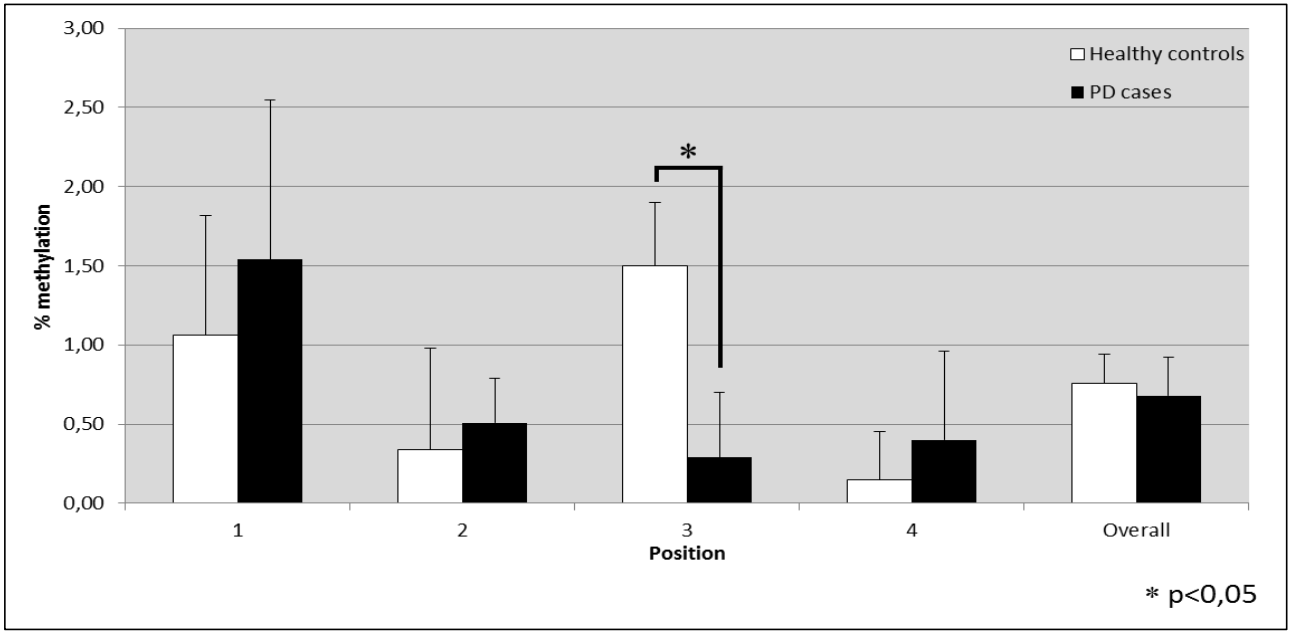
***PRKN* 1**

Substantia nigra

Table CY. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	Overall
Healthy controls	1,06 (0,76)	0,34 (0,64)	1,50 (0,40)	0,15 (0,30)	0,76 (0,18)
PD cases	1,54 (1,01)	0,51 (0,28)	0,29 (0,41)	0,40 (0,56)	0,68 (0,24)
p-value	0,556	0,286	**0,016**	0,556	0,730

Figure XXIV. Graphical representation of the results detailed in Table CY.



Blood

Table CZ. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	Overall
Other PD cases	0,65 (0,41)	0,26 (0,26)	1,15 (2,09)	0,46 (0,39)	0,63 (0,59)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

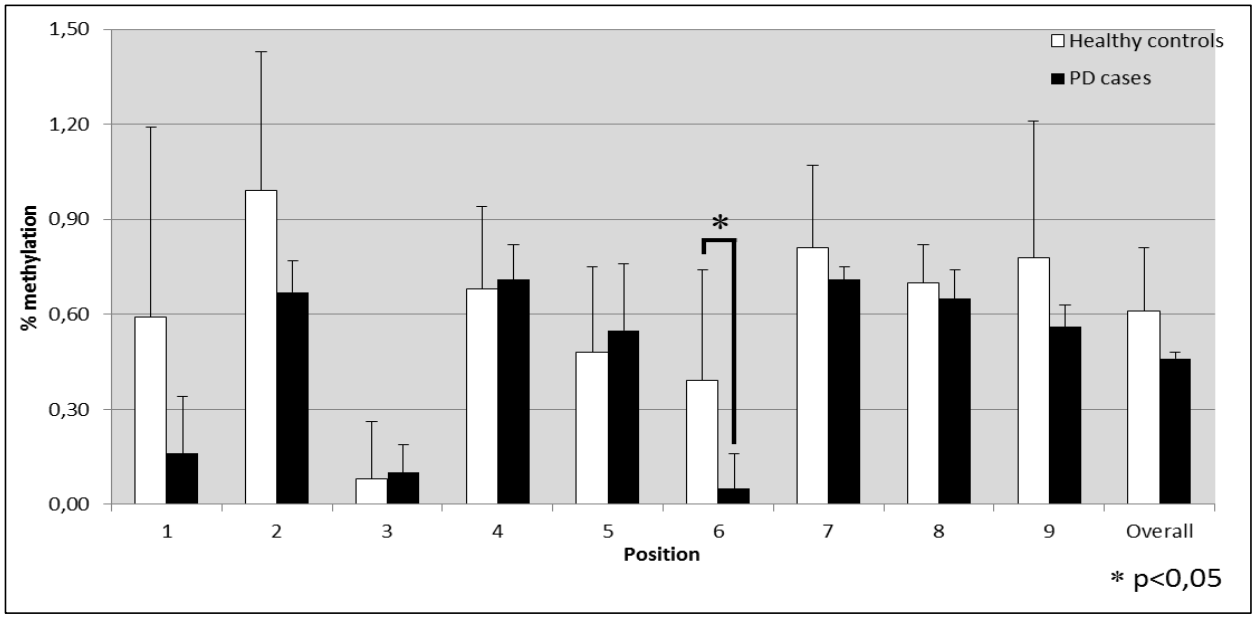
PRKN 2

Parietal cortex

Table DA. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,59 (0,60)	0,99 (0,44)	0,08 (0,18)	0,68 (0,26)	0,48 (0,27)	0,39 (0,35)	0,81 (0,26)	0,70 (0,12)	0,78 (0,43)	0,61 (0,20)
PD cases	0,16 (0,18)	0,67 (0,10)	0,10 (0,09)	0,71 (0,11)	0,55 (0,21)	0,05 (0,11)	0,71 (0,04)	0,65 (0,09)	0,56 (0,07)	0,46 (0,02)
p-value	0,222	0,222	0,151	0,841	0,690	**0,032**	0,421	0,310	0,548	0,151

Figure XXV. Graphical representation of the results detailed in Table DA.



Blood

Table DB. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	0,09 (0,16)	0,71 (0,30)	0	0	0,40 (0,25)	0,01 (0,04)	0,69 (0,30)	0,59 (0,25)	0,00	0,28 (0,12)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

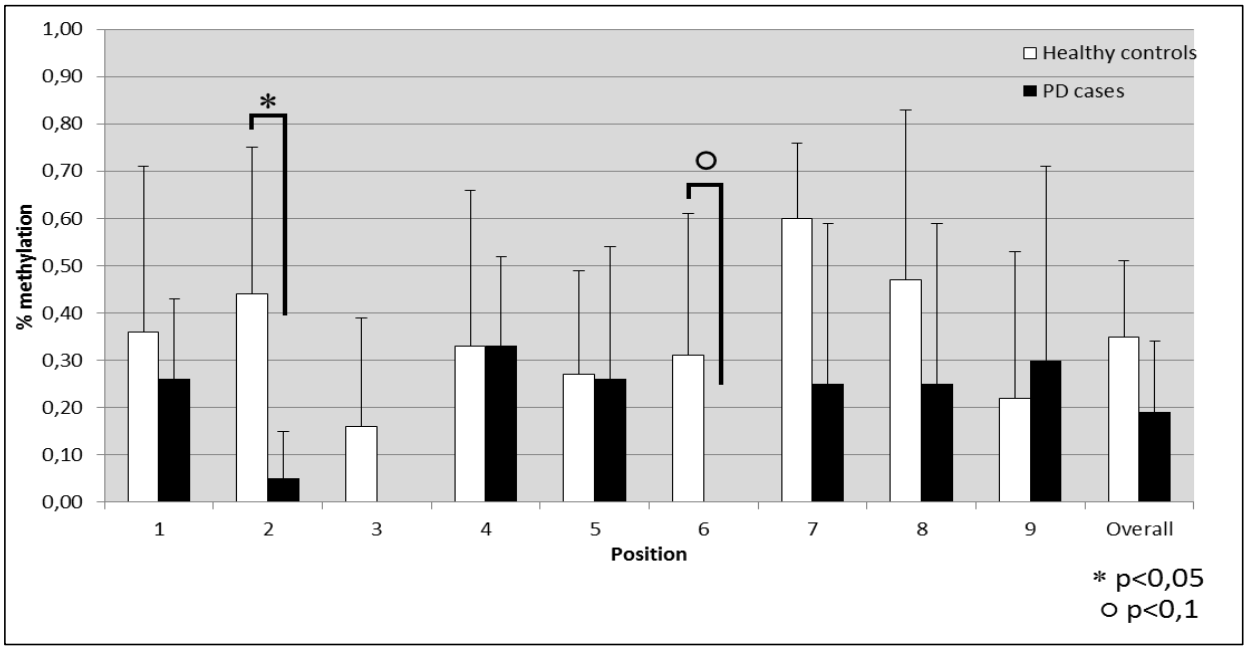
***PRKN* 2**

Occipital cortex

Table DC. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,36 (0,35)	0,44 (0,31)	0,16 (0,23)	0,33 (0,33)	0,27 (0,22)	0,31 (0,30)	0,60 (0,16)	0,47 (0,36)	0,22 (0,31)	0,35 (0,16)
PD cases	0,26 (0,17)	0,05 (0,10)	0	0,33 (0,19)	0,26 (0,28)	0	0,25 (0,34)	0,25 (0,34)	0,30 (0,41)	0,19 (0,15)
p-value	0,690	**0,032**	0,310	1,000	1,000	**0,056**	0,151	0,310	0,841	0,222

Figure XXVI. Graphical representation of the results detailed in Table DC.



Blood

Table DD. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	0,09 (0,16)	0,71 (0,30)	0	0	0,40 (0,25)	0,01 (0,04)	0,69 (0,30)	0,59 (0,25)	0,00	0,28 (0,12)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

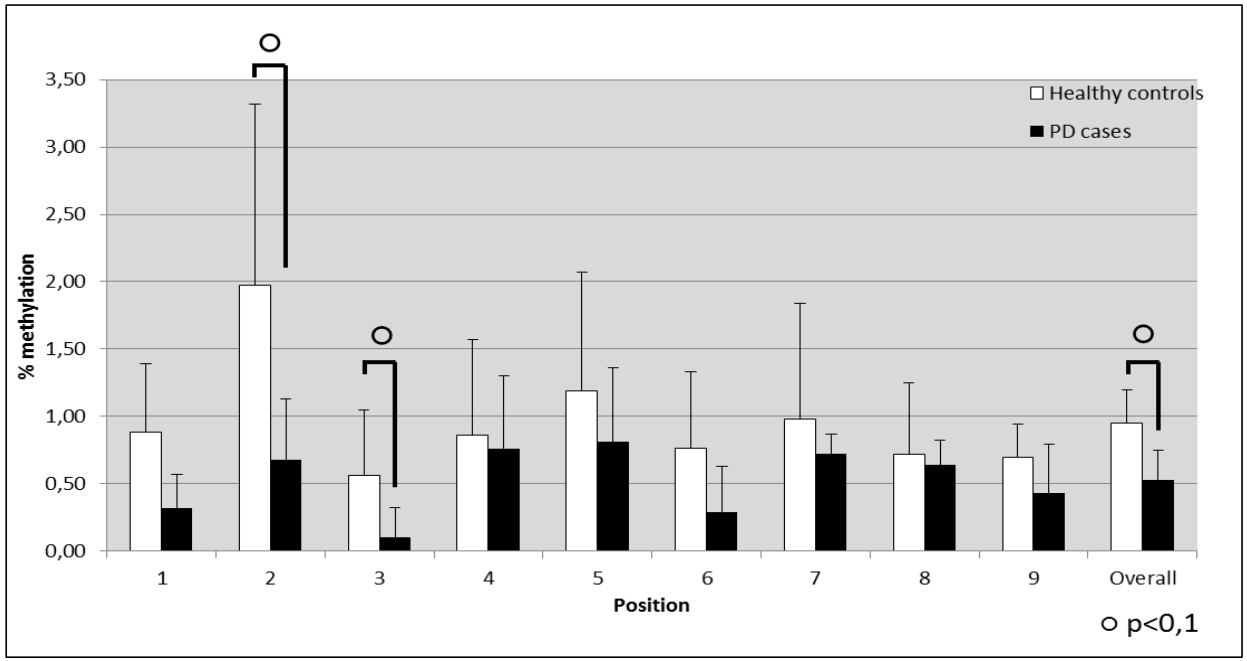
***PRKN* 2**

Substantia nigra

Table DE. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	9	Overall
Healthy controls	0,88 (0,51)	1,97 (1,35)	0,56 (0,49)	0,86 (0,71)	1,19 (0,88)	0,76 (0,57)	0,98 (0,86)	0,72 (0,53)	0,70 (0,24)	0,95 (0,25)
PD cases	0,32 (0,25)	0,68 (0,45)	0,10 (0,22)	0,76 (0,54)	0,81 (0,55)	0,29 (0,34)	0,72 (0,15)	0,64 (0,18)	0,43 (0,36)	0,53 (0,22)
p-value	0,111	**0,063**	**0,063**	0,730	0,905	0,286	0,905	0,730	0,286	**0,063**

Figure XXVII. Graphical representation of the results detailed in Table DE.



Blood

Table DF. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	9	Overall
Other PD cases	0,09 (0,16)	0,71 (0,30)	0	0	0,40 (0,25)	0,01 (0,04)	0,69 (0,30)	0,59 (0,25)	0,00	0,28 (0,12)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

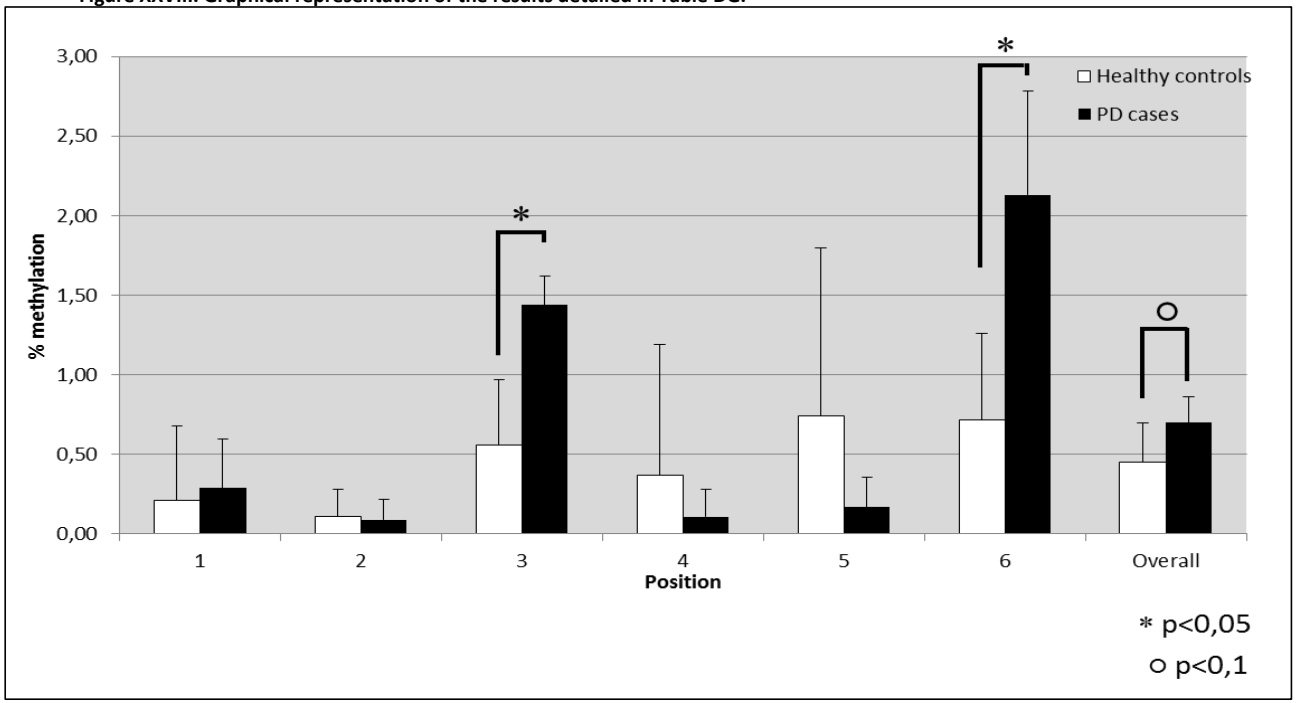
***SNCA* 1**

Parietal cortex

Table DG. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Healthy controls	0,21 (0,47)	0,11 (0,17)	0,56 (0,41)	0,37 (0,82)	0,74 (1,06)	0,72 (0,54)	0,45 (0,25)
PD cases	0,29 (0,31)	0,09 (0,13)	1,44 (0,18)	0,11 (0,17)	0,17 (0,19)	2,13 (0,65)	0,70 (0,16)
p-value	0,310	1,000	**0,008**	0,548	0,310	**0,016**	**0,095**

Figure XXVIII. Graphical representation of the results detailed in Table DG.



Blood

Table DH. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Other PD cases	3,62 (3,01)	2,49 (2,31)	2,20 (2,53)	1,77 (1,31)	1,44 (1,08)	1,56 (1,13)	2,18 (1,76)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

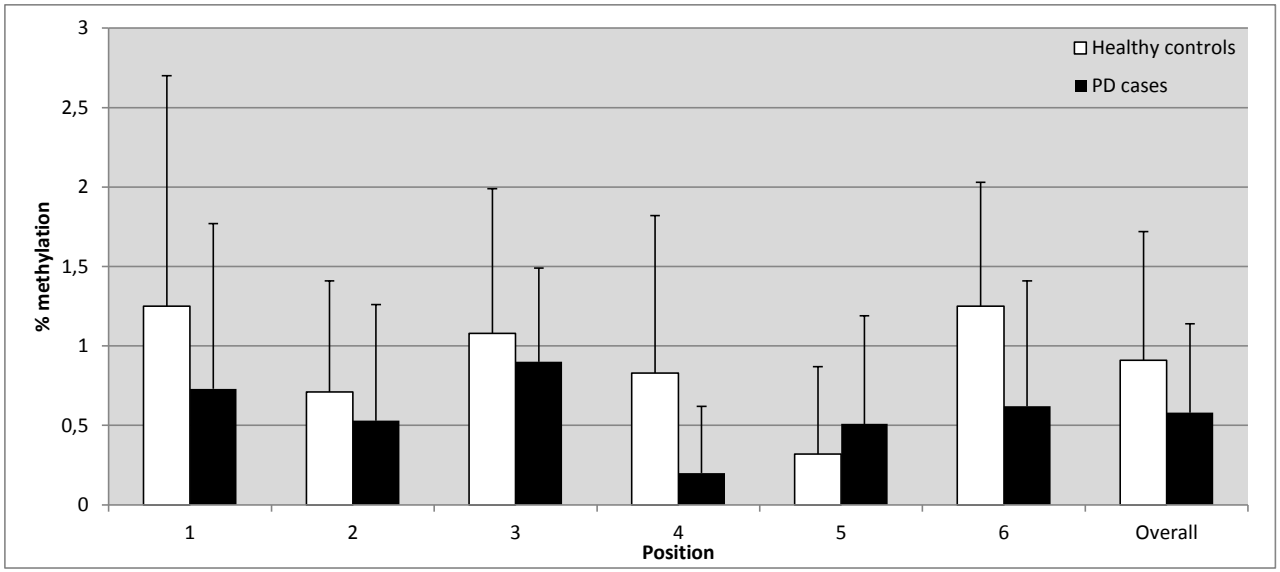
***SNCA* 1**

Occipital cortex

Table DI. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	Overall
Healthy controls	1,25 (1,45)	0,71 (0,70)	1,08 (0,91)	0,83 (0,99)	0,32 (0,55)	1,25 (0,78)	0,91 (0,81)
PD cases	0,73 (1,04)	0,53 (0,73)	0,90 (0,59)	0,20 (0,42)	0,51 (0,68)	0,62 (0,79)	0,58 (0,56)
p-value	0,841	1,000	1,000	0,151	0,690	0,421	0,548

Figure XXIX. Graphical representation of the results detailed in Table DI.



Blood

Table DJ. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	Overall
Other PD cases	3,62 (3,01)	2,49 (2,31)	2,20 (2,53)	1,77 (1,31)	1,44 (1,08)	1,56 (1,13)	2,18 (1,76)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

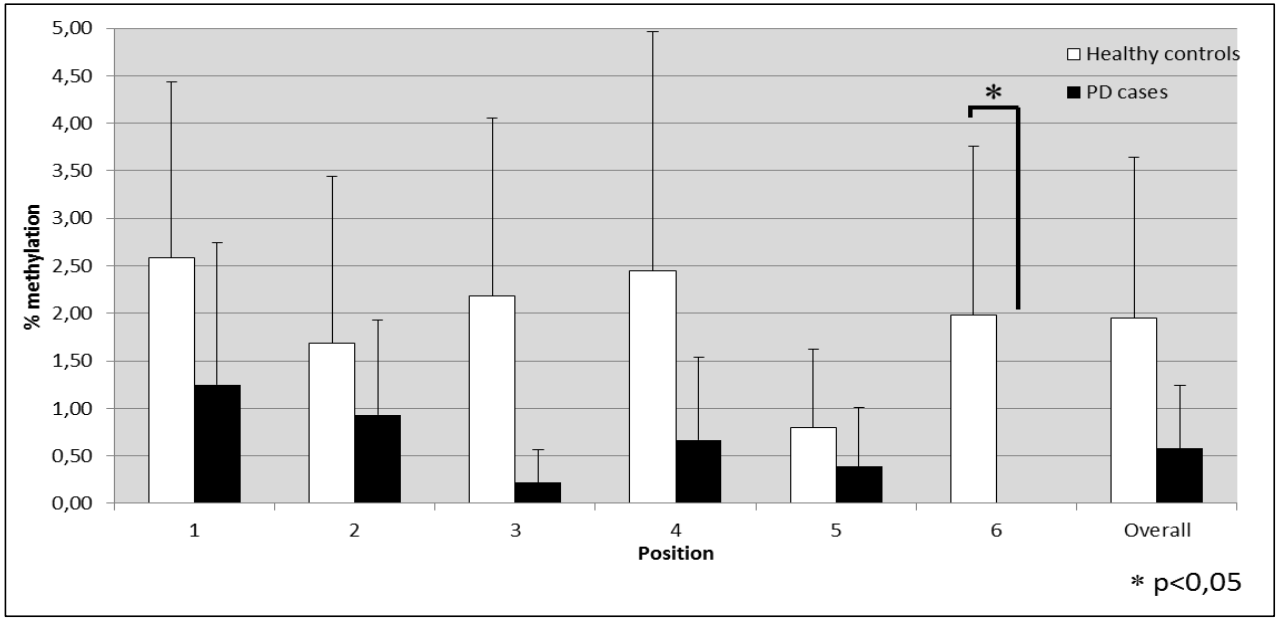
***SNCA* 1**

Substantia nigra

Table DK. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Healthy controls	2,59 (1,85)	1,69 (1,75)	2,18 (1,87)	2,45 (2,51)	0,80 (0,82)	1,98 (1,78)	1,95 (1,69)
PD cases	1,25 (1,49)	0,93 (1,00)	0,22 (0,34)	0,67 (0,87)	0,39 (0,62)	0	0,58 (0,66)
p-value	0,286	0,730	0,190	0,413	0,556	**0,016**	0,190

Figure XXX. Graphical representation of the results detailed in Table DK.



Blood

Table DL. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**Overall**
Other PD cases	3,62 (3,01)	2,49 (2,31)	2,20 (2,53)	1,77 (1,31)	1,44 (1,08)	1,56 (1,13)	2,18 (1,76)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

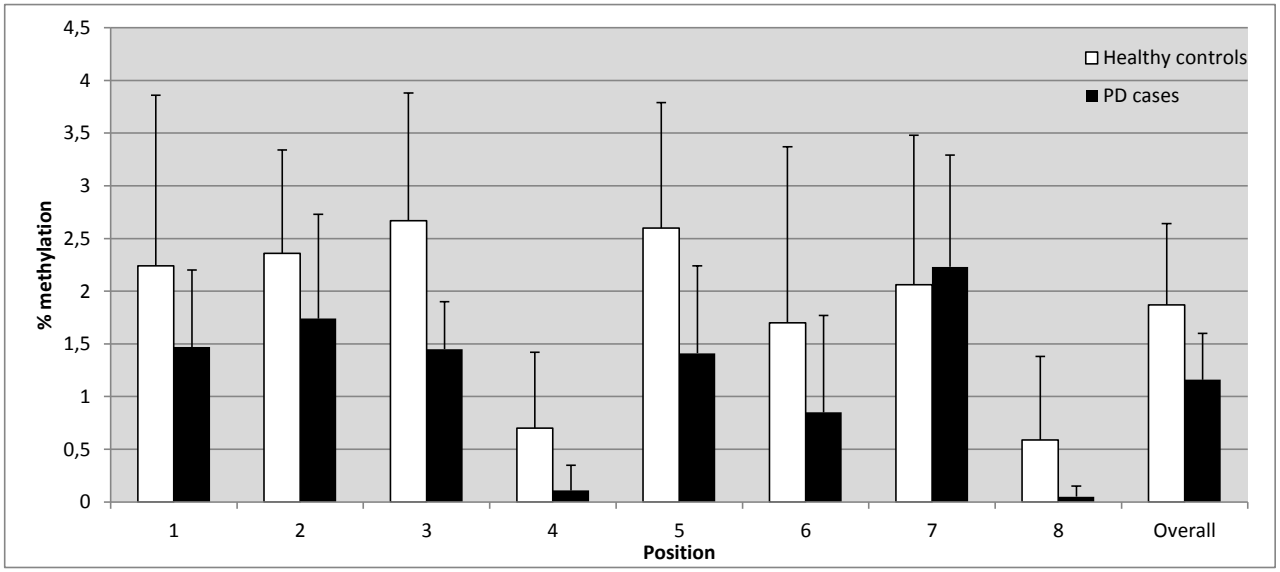
***SNCA* 2**

Parietal cortex

Table DM. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Healthy controls	2,24 (1,62)	2,36 (0,98)	2,67 (1,21)	0,70 (0,72)	2,60 (1,19)	1,70 (1,67)	2,06 (1,42)	0,59 (0,79)	1,87 (0,77)
PD cases	1,47 (0,73)	1,74 (0,99)	1,45 (0,45)	0,11 (0,24)	1,41 (0,83)	0,85 (0,92)	2,23 (1,06)	0,05 (0,10)	1,16 (0,44)
p-value	0,548	0,310	0,151	0,222	0,421	0,548	1,000	0,310	0,151

Figure XXXI. Graphical representation of the results detailed in Table DM.



Blood

Table DN. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Other PD cases	3,21 (2,03)	5,77 (3,29)	5,48 (3,43)	1,39 (1,83)	2,86 (2,16)	3,11 (2,58)	3,24 (2,10)	0,34 (0,36)	3,17 (1,65)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

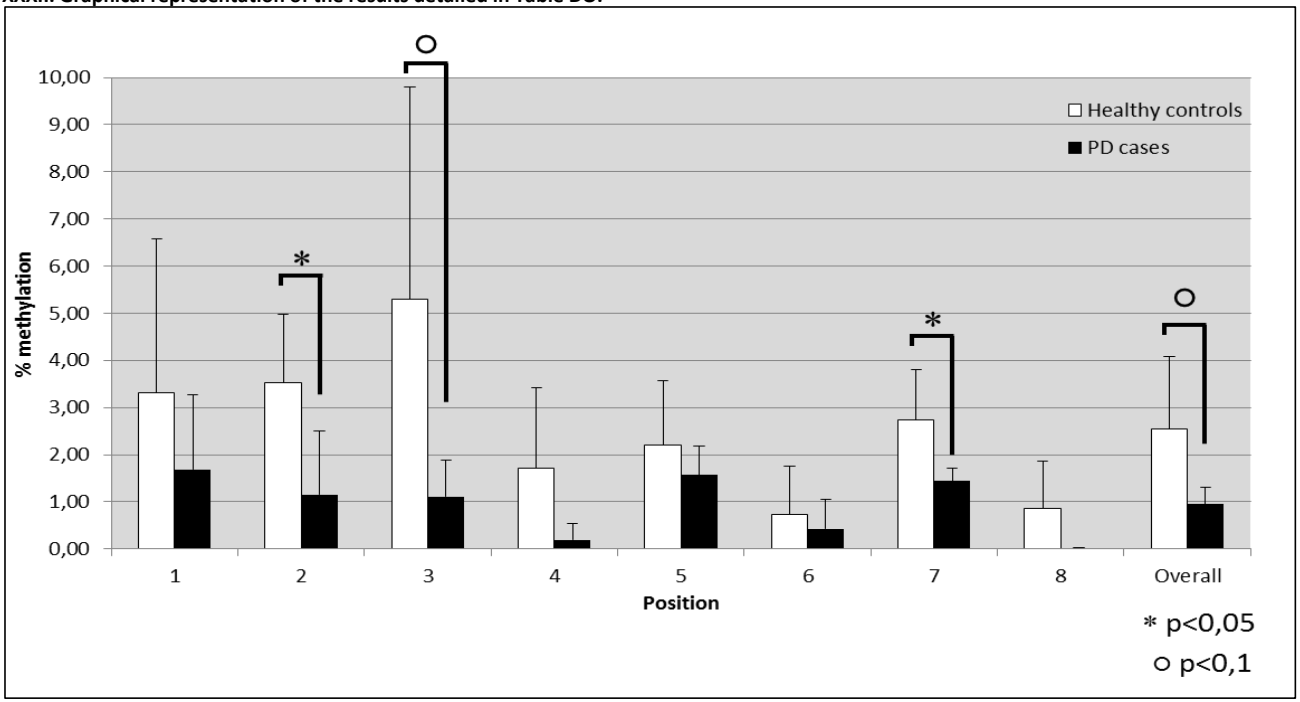
***SNCA* 2**

Occipital cortex

Table DO. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	1	2	3	4	5	6	7	8	Overall
Healthy controls	3,32 (3,26)	3,53 (1,45)	5,29 (4,52)	1,72 (1,69)	2,20 (1,36)	0,74 (1,02)	2,74 (1,07)	0,85 (1,01)	2,55 (1,53)
PD cases	1,69 (1,58)	1,15 (1,35)	1,11 (0,77)	0,19 (0,35)	1,57 (0,61)	0,43 (0,63)	1,44 (0,27)	0,01 (0,02)	0,95 (0,35)
p-value	0,548	**0,032**	**0,095**	0,310	0,421	0,841	**0,008**	0,222	**0,095**

Figure XXXII. Graphical representation of the results detailed in Table DO.



Blood

Table DP. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	1	2	3	4	5	6	7	8	Overall
Other PD cases	3,21 (2,03)	5,77 (3,29)	5,48 (3,43)	1,39 (1,83)	2,86 (2,16)	3,11 (2,58)	3,24 (2,10)	0,34 (0,36)	3,17 (1,65)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

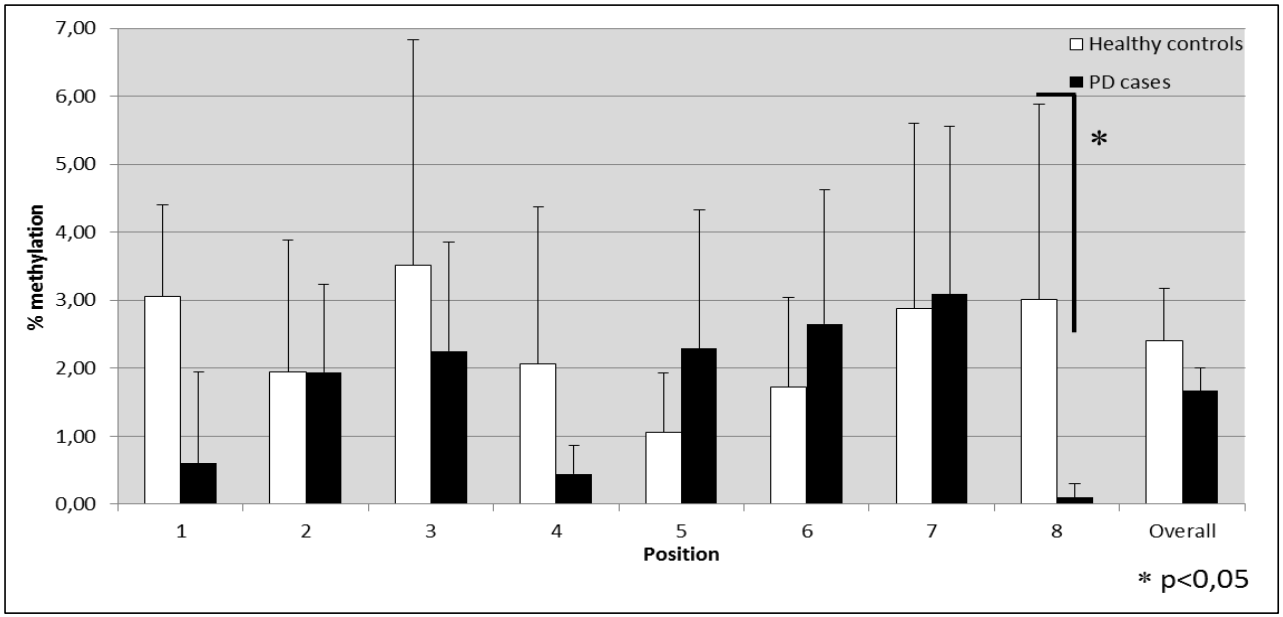
***SNCA* 2**

Substantia nigra

Table DQ. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Healthy controls	3,05 (1,36)	1,95 (1,94)	3,51 (3,32)	2,07 (2,30)	1,06 (0,87)	1,72 (1,32)	2,88 (2,73)	3,01 (2,88)	2,40 (0,77)
PD cases	0,60 (1,35)	1,94 (1,30)	2,25 (1,60)	0,45 (0,42)	2,29 (2,04)	2,65 (1,97)	3,09 (2,47)	0,10 (0,20)	1,67 (0,34)
p-value	0,111	1,000	0,556	0,190	0,413	0,556	1,000	**0,016**	0,111

Figure XXXIII. Graphical representation of the results detailed in Table DQ.



Blood

Table DR. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Overall**
Other PD cases	3,21 (2,03)	5,77 (3,29)	5,48 (3,43)	1,39 (1,83)	2,86 (2,16)	3,11 (2,58)	3,24 (2,10)	0,34 (0,36)	3,17 (1,65)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

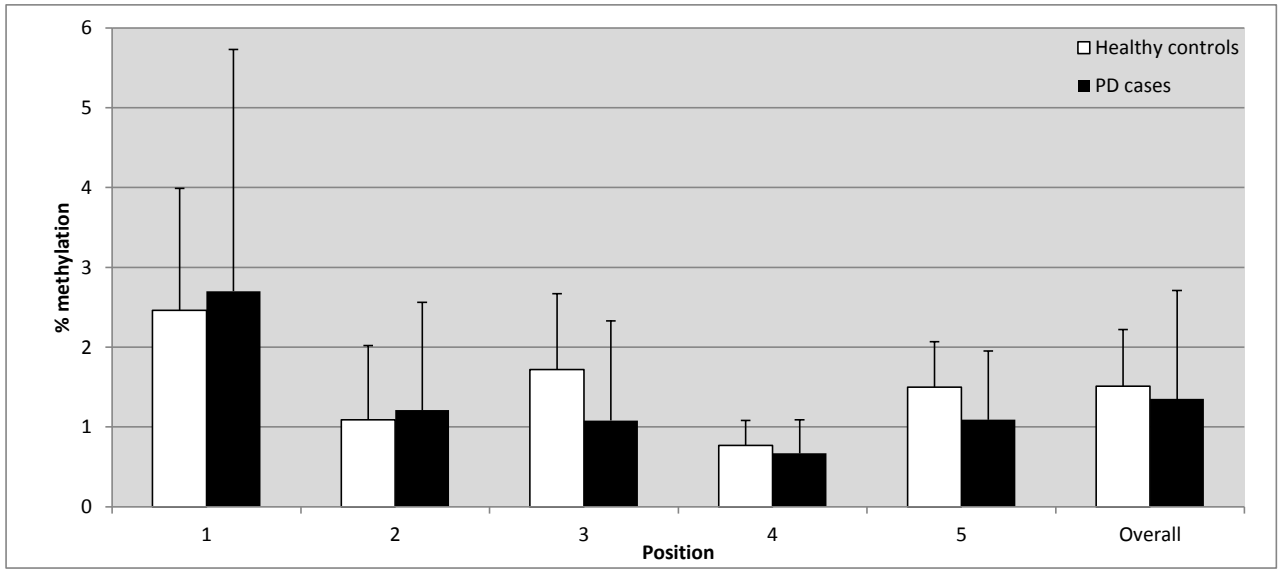
***SNCA* 3**

Parietal cortex

Table DS. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Healthy controls	2,46 (1,53)	1,09 (0,93)	1,72 (0,95)	0,77 (0,31)	1,50 (0,57)	1,51 (0,71)
PD cases	2,70 (3,03)	1,21 (1,35)	1,08 (1,25)	0,67 (0,42)	1,09 (0,86)	1,35 (1,36)
p-value	0,841	0,841	0,310	1,000	0,421	0,841

Figure XXXIV. Graphical representation of the results detailed in Table DS.



Blood

Table DT. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Other PD cases	2,03 (1,12)	2,36 (1,48)	2,20 (0,97)	2,46 (1,20)	2,45 (1,24)	2,30 (0,74)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

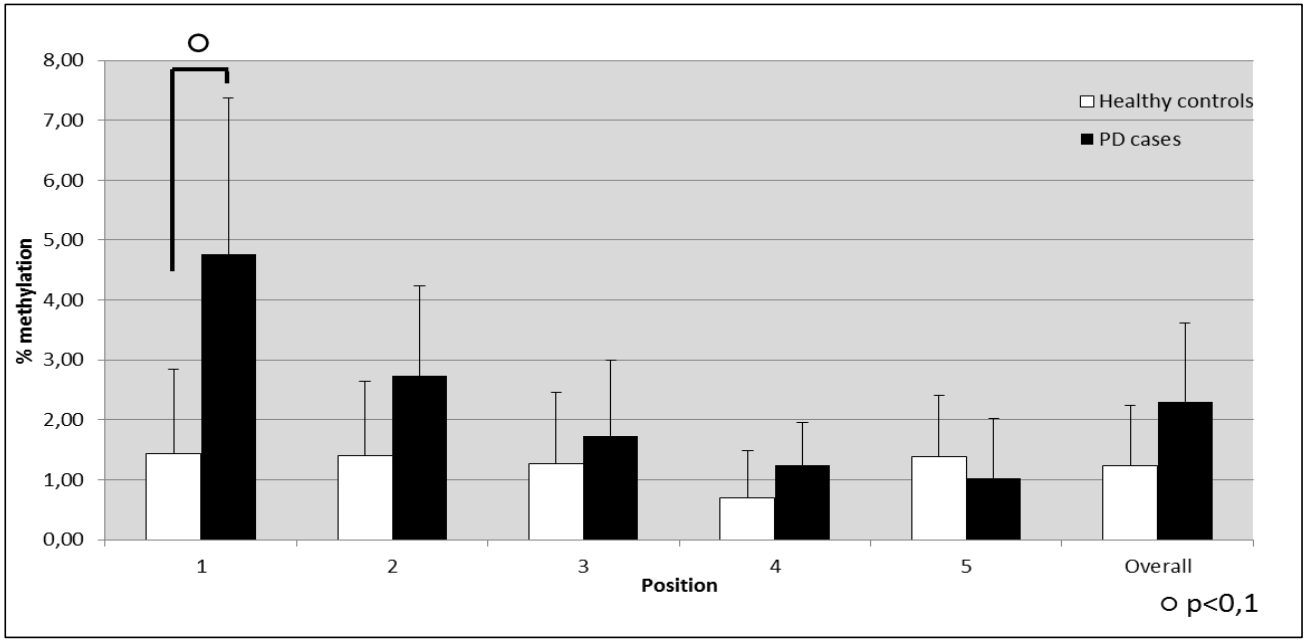
***SNCA* 3**

Occipital cortex

Table DU. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Healthy controls	1,43 (1,42)	1,41 (1,23)	1,27 (1,19)	0,69 (0,80)	1,38 (1,02)	1,24 (1,00)
PD cases	4,76 (2,62)	2,73 (1,50)	1,73 (1,27)	1,24 (0,71)	1,03 (0,99)	2,30 (1,31)
p-value	**0,095**	0,222	0,548	0,548	0,421	0,310

Figure XXXV. Graphical representation of the results detailed in Table DU.



Blood

Table DV. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Other PD cases	2,03 (1,12)	2,36 (1,48)	2,20 (0,97)	2,46 (1,20)	2,45 (1,24)	2,30 (0,74)

Annex VI. Quantitation of DNA methylation levels in promoters of genes responsible for familial Parkinson´s disease. Pilot analysis.

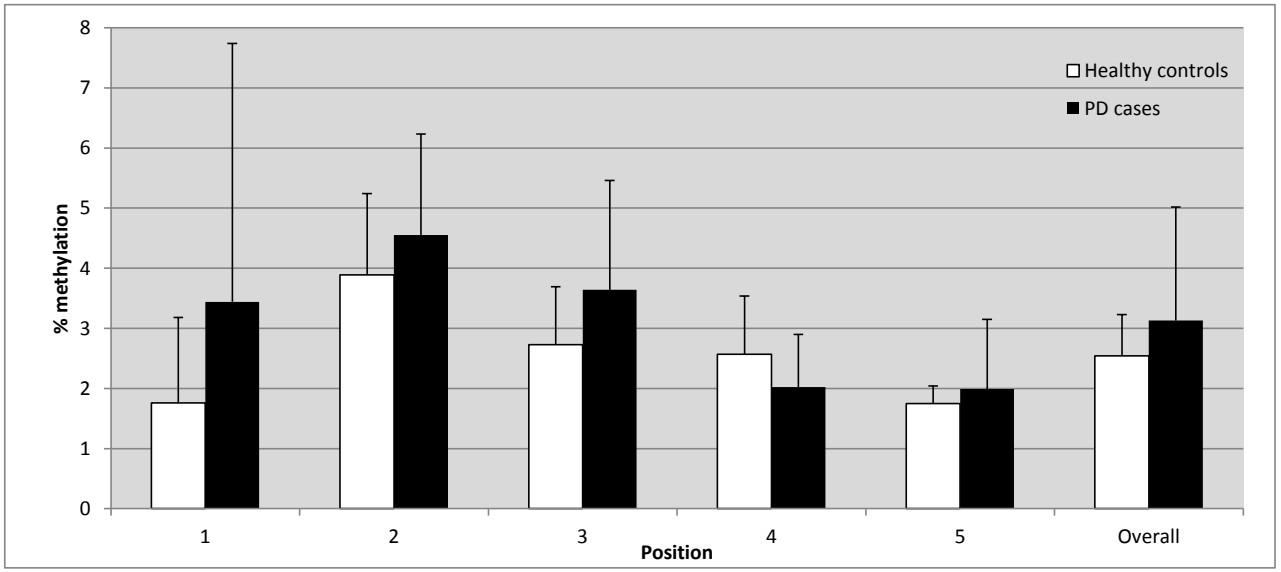
***SNCA* 3**

Substantia nigra

Table DW. Mean percentages of DNA methylation (and the standard deviation) per position and group. The statistical results are also described.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Healthy controls	1,76 (1,42)	3,89 (1,35)	2,73 (0,96)	2,57 (0,97)	1,75 (0,29)	2,54 (0,69)
PD cases	3,44 (4,30)	4,55 (1,68)	3,64 (1,82)	2,02 (0,88)	1,99 (1,16)	3,13 (1,89)
p-value	1,000	0,730	0,413	0,556	0,730	1,000

Figure XXXVI. Graphical representation of the results detailed in Table DW.



Blood

Table DX. Mean percentages of DNA methylation (and the standard deviation) per position.

Position	**1**	**2**	**3**	**4**	**5**	**Overall**
Other PD cases	2,03 (1,12)	2,36 (1,48)	2,20 (0,97)	2,46 (1,20)	2,45 (1,24)	2,30 (0,74)