



ESTUDIO DE LAS BASES GENÉTICAS IMPLICADAS EN LA ENFERMEDAD DE PARKINSON

DEPARTAMENTO DE MEDICINA

CRISTINA TEJERA PARRADO

Tesis doctoral

Sevilla, 2019

Memoria presentada por Cristina Tejera Parrado para optar al grado de Doctor con mención internacional por la Universidad de Sevilla.

El presente trabajo de investigación ha sido realizado en el Instituto de Biomedicina de Sevilla (IBiS), bajo la supervisión del Doctor Pablo Mir Rivera y la Doctora M^a del Pilar Gómez Garre.

Cristina Tejera Parrado ha recibido financiación de los contratos predoctorales o de Personal Investigador en Formación (PIF) para el desarrollo del programa propio de I+D+i de la Universidad de Sevilla.

D. Pablo Mir Rivera, Profesor Asociado del Departamento de Medicina de la Facultad de Medicina de la Universidad de Sevilla y Facultativo Especialista de Área en Neurología en el Hospital Universitario Virgen del Rocío, y Dña. M^a del Pilar Gómez Garre, Doctora en Biología e investigadora de la UGC de Neurología del Hospital Virgen del Rocío,

CERTIFICAN:

Que la memoria del trabajo titulada “Estudio de las bases genéticas implicadas en la Enfermedad de Parkinson” elaborada por Cristina Tejera Parrado, graduada en Biotecnología, ha sido realizada bajo su dirección y cumple las condiciones exigidas para optar al grado de Doctor con Mención Internacional por la Universidad de Sevilla.

Para que así conste, firman la presente en Sevilla Abril 2019.

Fdo: D. Pablo Mir Rivera

M^a del Pilar Gómez Garre

RESUMEN

La enfermedad de Parkinson (EP) es la segunda enfermedad neurodegenerativa más común caracterizada por la degeneración de neuronas dopaminérgicas en la sustancia negra. Se trata de una enfermedad multifactorial causada por la interacción de factores genéticos y tóxico-metabólicos. Se han identificado hasta 11 genes responsables de formas monogénicas de la enfermedad, tanto dominantes (*SNCA*, *LRRK2* y *VPS35*) como recesivos (*PARK2*, *PINK1*, *DJ-1*, *ATP13A2*, *PLA2G6*, *FBXO7* y *DNAJC6*). Sin embargo, solo un pequeño porcentaje de los casos de EP son debido a una única mutación causativa. La mayoría se deben a la interacción de variaciones genéticas que confieren un riesgo bajo-moderado. Hasta la fecha, se han identificado varios *loci* y genes de riesgo (*GBA* el mejor conocido) asociados a la EP. A pesar de todos estos estudios, la mayor parte de la heredabilidad de la enfermedad es todavía desconocida. Se espera que la tecnología de última generación de secuenciación masiva y los estudios de asociación de genoma completo, junto a la aplicación de las técnicas genéticas tradicionales, contribuyan a ampliar el conocimiento sobre la etiología de la EP. Además, es necesario la replicación de los hallazgos encontrados en cohortes independientes para comprobar su robustez.

Debido al aumento de la prevalencia de la enfermedad, en una población cada vez más envejecida, al gasto económico y social que conlleva, y a la falta de terapias que frenen el proceso degenerativo, resulta fundamental conocer en profundidad las bases moleculares involucradas en el desarrollo de la enfermedad. La hipótesis de esta tesis es que a través del estudio de las bases genéticas de la EP se contribuirá a ampliar el conocimiento sobre su etiopatogenia. Para ello, nuestro primer objetivo ha sido estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas

de la enfermedad. Con este objetivo se han analizado mediante secuenciación masiva 17 genes relacionados con la EP, causativos y de riesgo, confirmados y no confirmados (*SNCA*, *LRRK2*, *PINK1*, *PARK2*, *ATP13A2*, *FBXO7*, *VPS35*, *DJI*, *PLA2G6*, *EIF4G1*, *DNAJC6*, *HTRA2*, *SMPD1*, *SNYJ1*, *UCHL1*, *GIGYF2* y *GBA*). Para ello se ha reclutado una cohorte de 117 pacientes de EP de inicio temprano de la Comunidad de Madrid. Se han identificado variaciones con potencial efecto deletéreo en 29,06% de los pacientes estudiados, concentrado los genes *PARK2*, *LRRK2*, *PINK1* y *GBA* el 23,08% de los casos. El gen en el que se localizaron más variaciones fue *PARK2*, siendo la variación p.N52fs la más frecuente. *LRRK2* también parecía tener un papel más relevante que el que se pensaba en la población estudiada, debido a su alta frecuencia (presente en el 6,84% de los pacientes). Además, se ha evidenciado la co-ocurrencia de variaciones en los genes *ATP13A2* y *PARK2* y entre *LRRK2* y *GIGYF2*, sugiriendo efectos aditivos en ambos casos.

Con este objetivo también se estudió la implicación en la fisiopatogenia de la EP de los genes *CHCHD2* y *TMEM230*, dos nuevos genes propuestos causativos de EP, en una cohorte de pacientes de EP del Hospital Universitario Virgen del Rocío y controles sanos. Se han identificado 4 variaciones no exónicas en el gen *CHCHD2*, pero ninguna parece ser causativa o factor de riesgo para la EP. En el gen *TMEM230* se han identificado 13 variaciones genéticas, pero tampoco parecen ser causativas o factor de riesgo para la enfermedad. Estos resultados sugieren que los genes *CHCHD2* y *TMEM230* no están implicados en el desarrollo de la EP, al menos en población del sur de España.

El segundo objetivo fue la replicación de 5 *loci* de riesgo para la EP en una cohorte de pacientes del Hospital Universitario Virgen del Rocío, que previamente se habían asociado a una cohorte de EP de Granada. En este estudio se ha replicado la asociación

de los *loci* *ACMSD/TMEM16* y *MAPT* con la EP, de manera que polimorfismos en *MAPT* parecen jugar un papel protector frente a la EP, mientras que variaciones en *ACMSD/TMEM16* pueden ser un factor de riesgo para la enfermedad.

Los resultados de este estudio ponen de manifiesto la complejidad genética que hay detrás de la EP. A pesar de todos los esfuerzos realizados en esta materia, en éste y otros trabajos, todavía queda mucho por conocer sobre su etiopatonegia. Aumentar el número de pacientes en los estudios de genoma completo, así como la secuenciación cada vez más extendida del genoma o exoma humano, junto con los análisis adecuados, contribuirán a ampliar nuestro conocimiento sobre la EP.

SUMMARY

Parkinson's disease is the second most common neurodegenerative disorder characterized by the loss of dopaminergic neurons in the substantia nigra. It is a multifactorial etiology disorder in which both environmental and genetic factors are involved. Eleven genes have been identified as a genetic cause of PD. *SNCA*, *LRRK2* and *VPS35* have been linked to autosomal dominant forms of PD. Otherwise, *PARK2*, *PINK1*, *DJ-1*, *ATP13A2*, *PLA2G6*, *FBXO7* and *DNAJC6* have been linked to autosomal recessive forms of PD. However, monogenic PD cases are rare. Indeed, risk variants with a moderate effect size are more usual. To date, several risk *loci* and genes (like a well-known *GBA* gene) have been identified related to PD. However, a large proportion of genetic heritability remains unknown. Current technologies like next generation sequencing of whole genome or just exome, genome wide association studies and also traditional technologies, are expected to contribute to elucidating the missing heritability of the disease. Replication studies in independent cohorts are also necessary to prove the robustness of new identified variants.

Because the global prevalence of the disease and the population aging are increasing, also the economic and social costs of the disease are getting bigger and bigger, and because the lack of curative therapies, is important to understand the molecular mechanisms involved in the development of PD. Our hypothesis was that through the study of genetic bases of PD we will be able to better understand the underlined mechanisms involved in the disease.

The first aim of this work was to assess the contribution of known PD related genes in monogenic forms of the disease. To this end, we employed target sequencing to analyze 17 previously related to PD genes, and unconfirmed genes in a cohort of 117 early onset PD patients from the Comunidad Autónoma de Madrid (Spain). These genes were:

SNCA, LRRK2, PINK1, PARK2, ATP13A2, FBXO7, VPS35, DJI, PLA2G6, EIF4G1, DNAJC6, HTRA2, SMPD1, SNYJ1, UCHL1, GIGYF2 and *GBA*. We identified 29,06% of the studied patients carrying potential pathogenic variants. From them, 23,08% carried putative pathogenic variants in *PARK2, LRRK2, PINK1* and *GBA*. The most frequently mutated gene was *PARK2*, and p.N52Mfs in *PARK2* was the most frequent common variation. 6,84% of patients carried some potential pathogenic variants in *LRRK2*, which suggests an important role for *LRRK2* in early onset PD in Spanish population. Also, a co-occurrence of mutations involving two genes suggests additive effects of *ATP13A2* and *PARK2* genes, as well as of *LRRK2* and *GIGYF2* genes.

To this aim, we also studied the contribution to PD of two new identified related to PD genes, *CHCHD2* and *TMEM230*, in a cohort of PD patients from the Hospital Universitario Virgen del Rocío and in control subjects. We have reported 4 non-exonic variants in *CHCHD2*, with no significant association in PD cases compared with control subjects. Also, we have reported 13 variants in *TMEM230* but neither of them seems to be related with the development of PD. These results suggest that *CHCHD2* and *TMEM230* may not play a mayor role in the development of PD, at least in population from southern Spain.

Our second aim was to investigate whether there was a link between previously risk variants related to PD in a cohort from the South of Spain, in another larger independent PD cohort, also from the South of Spain. We got replicated the association with PD of two *loci* *ACMSD/TMEM16* and *MAPT*, and suggested that *MAPT* gene polymorphisms might be a protective factor against PD while *ACMSD/TMEM16* gene polymorphisms could be a risk factor for PD.

Those results show the genetic complexity of this pathology. Despite of the contribution of this work and others in the area, a big proportion of the etiopathogenesis of PD is

still unknown. Larger cohorts in genome wide association studies, more sequencing data of exome or genome of PD patients, and also an adequate bioinformatic analysis will contribute to provide a more comprehensive picture of the complexity of this pathology.

ABREVIATURAS

ACMG: *American College of Medical and Genomics*

ACMSD: *Aminocarboxymuconate Semialdehyde Decarboxylase*

ADORA1: *Adenosine A1 Receptor*

ATP13A2: *ATPase Cation Transporting 13A2*

CAV1: *Caveolin 1*

CAV2: *Caveolin 2*

CHCHD2: *Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 2*

CHR: Cromosoma (del inglés: *Chromosome*)

CL: Cuerpos de Lewy

DJ-1: *Deglycase DJ-1*

DNAJC13: *DnaJ Heat Shock Protein Family (Hsp40) Member C13*

DNAJC6: *DnaJ Heat Shock Protein Family (Hsp40) Member C6*

EIF4G1: *Eukaryotic Translation Initiation Factor 4 Gamma 1*

EP: Enfermedad de Parkinson

EPIJ: Enfermedad de Parkinson de inicio juvenil

EPIT: Enfermedad de Parkinson de inicio temprano

eVAI: *Expert Variant Interpreter*

FBXO7: *F-Box Only Protein 7*

GBA: *Beta-Glucocerebrosidase*

GCH1: *GTP Cyclohydrolase 1*

GIGYF2: *GRB10 Interacting GYF Protein 2*

GWAS: Estudio de asociación de genoma completo (del inglés: *genome-wide association study*)

GWA: Asociación de genoma completo (del inglés: *genome-wide association*)

HRM: *High Resolution Melting*

HTRA2: *HTRA Serine Peptidase 2*

LPA: *Lipoprotein (A)*

LRRK2: *Leucine-Rich Repeat Kinase 2*

MAPT: *Microtubule Associated Protein Tau*

MIR4697: *MicroRNA 4697*

MLPA: *Multiplex Ligation-dependent Probe Amplification*

NL: *Neuritas de Lewy*

OR: *Odds ratio*

PARK11: *Parkinson Disease 11*

PB: *Par de bases (del inglés: Pair base)*

PINK1: *PTEN Induced Kinase 1*

PLA2G6: *Phospholipase A2*

PODXL: *Podocalyxin Like*

PARK2: *Parkin*

PTRHD1: *Peptidyl-TRNA Hydrolase Domain Containing 1*

RAB39B: *Ras-Related Protein Rab-39B*

RAI1: *Retinoic Acid Induced 1*

REM: *Rapid Eye Movement*

RIC3: *RIC3 Acetylcholine Receptor Chaperone*

SMPD1: *Sphingomyelin Phosphodiesterase 1*

SNCA: *α -Synuclein*

SNP: *Polimorfismo de un solo nucleótido (del inglés: single nucleotide polymorphism)*

SNPs: *Polimorfismos de un solo nucleótido (del inglés: single nucleotide polymorphisms)*

SREBF: Sterol Regulatory Element Binding Transcription Factor

STK39: Serine/Threonine Kinase 39

SYNJ1: Synaptojanin 1

TMEM163: Transmembrane Protein 163

TMEM230: Transmembrane Protein 230

TNFRSF9: Tumor Necrosis Factor Receptor Superfamily Member 9

UCHL1: Ubiquitin C-Terminal Hydrolase L1

VOUS: Variaciones genéticas de significado incierto (del inglés: variants of unknown significance)

VPS13C: Vacuolar Protein Sorting 13 Homolog C

VPS35: Vacuolar Protein Sorting 35

ÍNDICE DE CONTENIDOS

I.	INTRODUCCIÓN.....	1
1.	ENFERMEDAD DE PARKINSON: CARACTERÍSTICAS GENERALES.....	1
1.1.	CARACTERÍSTICAS CLÍNICAS.....	2
1.2.	CARACTERÍSTICAS ANATOMOPATOLÓGICAS	2
1.3.	EPIDEMIOLOGÍA.....	4
2.	ETIOPATOGENIA DE LA EP.....	5
2.1.	FACTORES GENÉTICOS EN LA EP.....	5
2.1.1.	GENES CAUSATIVOS DE FORMAS MONOGÉNICAS DE EP.....	10
2.1.1.1.	GENES CON HERENCIA AUTOSÓMICA DOMINANTE.....	11
2.1.1.1.1.	<i>SNCA</i>	11
2.1.1.1.2.	<i>LRRK2</i>	12
2.1.1.1.3.	<i>VPS35</i>	12
2.1.1.2.	GENES CON HERENCIA AUTOSÓMICA RECESIVA.....	13
2.1.1.2.1.	<i>PARK2</i>	13
2.1.1.2.2.	<i>PINK1</i>	13
2.1.1.2.3.	<i>DJ-1</i>	14
2.1.1.2.4.	<i>ATP13A2</i>	14
2.1.1.2.5.	<i>PLA2G6</i>	15
2.1.1.2.6.	<i>FBXO7</i>	15
2.1.1.2.7.	<i>DNAJC6</i>	15
2.1.1.2.8.	<i>SYNJ1</i>	15
2.1.2.	GENES PROPUESTOS CAUSATIVOS DE FORMAS MONOGÉNICAS DE EP PERO NO CONFIRMADOS.....	16
2.1.2.1.	GENES CON HERENCIA AUTOSÓMICA DOMINANTE.....	16

2.1.2.1.1.	<i>DNAJC13</i>	16
2.1.2.1.2.	<i>CHCHD2</i>	17
2.1.2.1.3.	<i>TMEM230</i>	17
2.1.2.1.4.	<i>UCHL1</i>	18
2.1.2.1.5.	<i>GIGYF2</i>	18
2.1.2.1.6.	<i>HTRA2</i>	18
2.1.2.1.7.	<i>EIF4G1</i>	18
2.1.2.1.8.	<i>RIC3</i>	19
2.1.2.2.	GENES CON HERENCIA AUTOSÓMICA RECESIVA O LIGADOS AL CROMOSOMA X.....	19
2.1.2.2.1.	<i>PODXL</i>	19
2.1.2.2.2.	<i>PTRHDI</i>	20
2.1.2.2.3.	<i>RAB39B</i>	20
2.1.2.2.4.	<i>VPS13C</i>	22
2.1.3.	<i>LOCI</i> DE RIESGO EN LA EP.....	22
2.1.4.	GENES DE RIESGO EN LA EP.....	25
2.1.4.1.	<i>SNCA</i>	25
2.1.4.2.	<i>LRRK2</i>	26
2.1.4.3.	<i>GBA</i>	26
2.1.4.4.	<i>MAPT</i>	27
2.1.5.	GENES NO CONFIRMADOS DE RIESGO EN LA EP.....	27
2.1.5.1.	<i>SMPDI</i>	27
2.1.5.2.	<i>HTRA2</i>	28
2.2.	RUTAS METABÓLICAS IMPLICADAS EN LA EP.....	28
2.3.	FACTORES AMBIENTALES EN LA EP.....	30

II. HIPÓTESIS, OBJETIVOS Y DISEÑO EXPERIMENTAL	33
1. HIPÓTESIS DE TRABAJO.....	33
2. OBJETIVOS.....	34
3. DISEÑO EXPERIMENTAL.....	35
III. RESULTADOS	39
1. Objetivo 1: Estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas de EP.....	39
1.1. Sub-objetivo 1: Estudiar exhaustivamente 17 genes relacionados con la EP y otros parkinsonismos en pacientes de inicio temprano de la Comunidad de Madrid.....	39
1.1.1. <i>Manuscrito en vías de publicación: Mutational landscape of early onset Parkinson’s disease in a Spanish population by targeted exome sequencing. Tejera-Parrado C, Mir P, Periñán MT, Vela-Desojo L, Abreu-Rodríguez I, Alonso-Cánovas A, Bernal-Bernal I, Bonilla-Toribio M, Buiza-Rueda D, Catalán MJ, García-Ramos R, García-Ruiz PJ, Huertas-Fernández I, Jesús S, López-Manzanares L, Martínez-Castrillo JC, Posada IJ, Rojo-Sebastián A, Ruiz-Huete C, Del Val J, Gómez-Garre P</i>	39
1.1.2. <i>Resumen</i>	77
1.2. Sub-objetivo 2: Estudiar la implicación en la fisiopatogenia de la EP de <i>CHCHD2</i> , nuevo gen propuesto causativo de EP, en población del sur de España.....	81

1.2.1. <i>Publicación: Genetic analysis of CHCHD2 in a southern Spanish population. Tejera-Parrado C, Jesús S, Huertas-Fernández I, Bernal-Bernal I, Bonilla-Toribio M, Córdoba-Tevar I, Abreu-Rodríguez I, Carrillo F, Bernal-Escudero M, Vargas-González L, Carballo M, Gómez-Garre P, Mir P.....</i>	81
1.2.2. <i>Resumen</i>	87
1.3. Sub-objetivo 3: Estudiar la implicación en la fisiopatogenia de la EP de <i>TMEM230</i> , nuevo gen propuesto causativo de EP, en población del sur de España.....	89
1.3.1. <i>Publicación: TMEM230 in Parkinson's disease in a southern Spanish population. Tejera-Parrado C, Jesús S, López-Ruiz A, Buiza-Rueda D, Bonilla-Toribio M, Bernal.Bernal I, Periñán MT, Vargas-González L, Gómez-Garre P, Mir P.....</i>	89
1.3.2. <i>Resumen</i>	101
2. Objetivo 2: Replicación de 5 <i>loci</i> de riesgo para la EP, previamente asociados a una cohorte de EP de Granada, en otra cohorte independiente del sur de España.....	103
2.1. <i>Manuscrito en vías de publicación: A replication study of GWAS-genetic risk variants associated with Parkinson's disease in a Spanish population: a case-control study. Tejera-Parrado C, Oliva-Ariza G, Jesús S, Adarmes-Gómez AD, Bernal-Bernal I, Bonilla-Torbio M, Buiza-Rueda D, Carrillo F, Vargas-González L, Carballo M, Gómez-Garre P, Mir P.....</i>	103

2.2. Resumen	123
IV. DISCUSIÓN GENERAL.....	125
1. Objetivo 1: Estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas de EP.	126
1.1.Sub-objetivo 1: Estudiar exhaustivamente 17 genes relacionados con la EP y otros parkinsonismos en pacientes de inicio temprano de la Comunidad de Madrid.....	126
1.2.Sub-objetivo 2: Estudiar la implicación en la fisiopatogenia de la EP de <i>CHCHD2</i> , nuevo gen propuesto causativo de EP, en población del sur de España.....	129
1.3.Sub-objetivo 3: Estudiar la implicación en la fisiopatogenia de la EP de <i>TMEM230</i> , nuevo gen propuesto causativo de EP, en población del sur de España.....	131
2. Objetivo 2: Replicación de 5 <i>loci</i> de riesgo para la EP, previamente asociados a una cohorte de EP de Granada, en otra cohorte independiente del sur de España.....	133
V. CONCLUSIONES.....	139
VI. CONCLUSIONS.....	141
VII. BIBLIOGRAFÍA.....	143

ÍNDICE DE FIGURAS

Figura 1. Principales características anatomopatológicas de la EP.....	3
Figura 2. Representación esquemática del tipo de variaciones genéticas involucradas en la etiopatogenia de la EP.....	7
Figura 3. Esquema de las distintas estrategias empleadas para el estudio de los factores genéticos implicados en las enfermedades humanas complejas como la EP.....	9
Figura 4. Representación esquemática de la arquitectura genética de la EP.....	21
Figura 5. Esquema del diseño experimental seguido en la Tesis Doctoral.....	35

ÍNDICE DE TABLAS

Tabla 1. Genes causativos de formas monogénicas de EP.....	10
Tabla 2. 28 señales independientes en 24 <i>loci</i> asociados con la EP en el estudio de meta-análisis de Nalls y col., 2014.....	23
Tabla 3. 17 <i>loci</i> asociados con la EP en el estudio de meta-análisis de Chang y col., 2017.....	24

I. INTRODUCCIÓN

1. ENFERMEDAD DE PARKINSON: CARACTERÍSTICAS GENERALES

La enfermedad de Parkinson (EP) es la segunda enfermedad neurodegenerativa más común después de la enfermedad de Alzheimer, con una prevalencia estimada del 0,3% en la población general y del 1% en mayores de 60 años. Se trata de una enfermedad multifactorial, en la que la interacción entre los factores genéticos y tóxico-metabólicos lleva a la degeneración de neuronas en regiones concretas del cerebro. Sin embargo, su fisiopatología no está aún clara. Parece que disfunciones mitocondriales, el estrés oxidativo y un mal funcionamiento de determinadas proteínas están involucrados en el desarrollo de la enfermedad (de Lau y Breteler., 2006).

Las primeras menciones de la EP se remontan al Antiguo Egipto, e incluso ya en la Biblia se hace referencia al temblor. Pero no fue hasta 1817 cuando James Parkinson definió por primera vez esta enfermedad como “parálisis agitante”. En su ensayo *The Shaking Palsy* describió sus principales características clínicas, que más tarde pasarían a formar parte de los criterios diagnósticos del *United Kingdom Parkinson's Disease Society Brain Bank*. Estas características incluían el temblor en reposo, la bradicinesia, la rigidez muscular y la pérdida de reflejos posturales (Parkinson, 1817; Gibb y col., 1988). Más tarde, William Sanders propuso el término enfermedad de Parkinson para esta patología. Ya en 1960, Arvid Carlsson evidenció el déficit dopaminérgico asociado a la enfermedad, dando paso a los primeros ensayos con levodopa (Carlsson y col., 1957). Este fármaco, precursor de la dopamina, se convertirá en el eje del tratamiento de la EP hasta nuestros días.

1.1. CARACTERÍSTICAS CLÍNICAS

La EP tradicionalmente se ha clasificado como un trastorno del movimiento, sin embargo, esta enfermedad neurodegenerativa causa un amplio espectro de síntomas tanto motores como no motores. Los síntomas motores más frecuentes incluyen la bradicinesia, el temblor en reposo, la rigidez o la inestabilidad postural. Estos síntomas van normalmente precedidos por manifestaciones no motoras que pueden desarrollarse incluso 20 años antes de ser diagnosticado (Pont-Sunyer y col., 2015; Schrag y col., 2015). Algunos de ellos son disfunción olfatoria, trastornos del sueño REM, estreñimiento, sudor excesivo, fatiga o comportamiento ansioso/depresivo. A medida que la enfermedad avanza, pueden aparecer otros síntomas relacionados con el tratamiento con levodopa, tanto motores: problemas en el habla y la deglución o congelación de la marcha; como no motores: trastornos psiquiátricos o demencia (Obeso y col., 2017). La respuesta a los tratamientos y la evolución de la enfermedad es muy variable, con casos muy benignos, de buena respuesta a levodopa y poca sintomatología derivada del tratamiento, a casos de poca respuesta y de evolución muy severa (Lang y col., 2004). La heterogenicidad de la enfermedad en cuanto a sintomatología clínica, progresión, edad de inicio y respuesta a los tratamientos médicos, sugiere la existencia de distintos subtipos dentro de la patología. Estos subtipos vienen determinados genéticamente principalmente (Obeso y col., 2017).

1.2. CARACTERÍSTICAS ANATOMOPATOLÓGICAS

Anatomopatológicamente, la EP se caracteriza principalmente por la pérdida de neuronas dopaminérgicas en la sustancia negra, provocando la degeneración de sus

proyecciones, que se dirigen fundamentalmente al glóbulo pálido y al putamen. Esta pérdida dopaminérgica ocurre predominantemente en el componente ventral lateral de la *pars compacta*. Aproximadamente se han perdido entre un 60-80% de las neuronas dopaminérgicas cuando aparecen los primeros síntomas motores. En este punto la sustancia negra aparece pálida, perdiendo su pigmentación original (Figura 1). Por otro lado, la presencia de cuerpos de Lewy (CL) y neuritas de Lewy (NL) en el citoplasma de las neuronas supervivientes es otro de los principales signos neuropatológicos en la EP. Los CL son inclusiones redondas, intracitoplasmáticas, eosinófilas e inmunoreactivas para la α -sinucleína (Mackenzie y col., 2001; Yarnall y col., 2012; Figura 1).

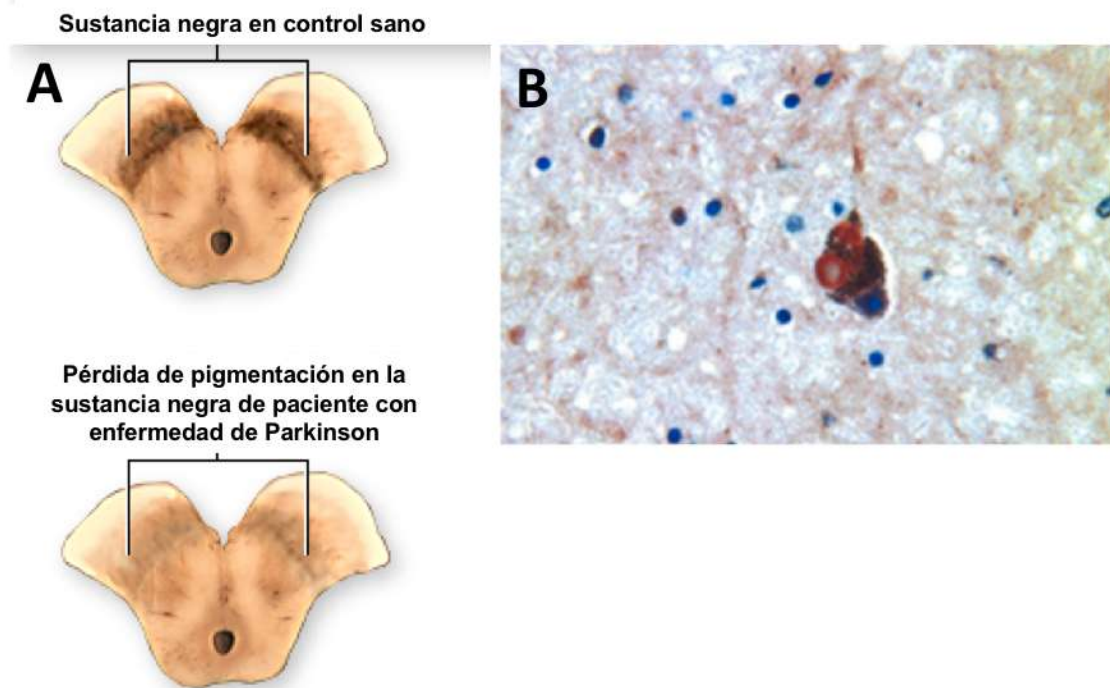


Figura 1. Principales características anatomopatológicas de la EP. A) Hemisección transversal del cerebro de un control sano (arriba), y de un paciente con enfermedad de Parkinson (abajo). B) Cuerpos de Lewy en una neurona degenerada. Tinción de α -sinucleína. (Modificado de Sveinbjornsdttir, 2016; y de Enciclopedia Médica A.D.A.M., 2005).

Además, otras neuronas se ven afectadas en esta patología, como las neuronas acetilcolinérgicas, serotoninérgicas y noradrenérgicas. La neurodegeneración de estas neuronas probablemente esté implicada en sus síntomas no motores, como la disfunción autonómica, problemas de sueño y complicaciones psiquiátricas (Yarnall y col., 2012).

1.3. EPIDEMIOLOGÍA

La EP es la segunda enfermedad neurodegenerativa más frecuente después de la enfermedad de Alzheimer. La prevalencia de la EP se estima entre 100-300/100000 habitantes y su incidencia entre 10-50/100000 habitantes (Elbaz y col., 2016; Pringsheim y col., 2014). Tanto la prevalencia como la incidencia aumentan a partir de 60 años, y así se refleja en un meta-análisis donde la prevalencia en personas de entre 50-59 años resulta en 107/100000 habitantes, pasando a 1087/100000 habitantes en el rango de 70-79 años (Pringsheim y col., 2014). Por otra parte se ha descrito que la EP es más frecuente en hombres que en mujeres, con un ratio de 1,5 (Wooten y col., 2004; Elbaz y col., 2002; Taylor y col., 2007).

La prevalencia de la enfermedad parece no ser constante en diferentes regiones geográficas. Se ha observado en varios estudios que es más frecuente en población caucásica que en africana o asiática, sin embargo, estos resultados son contradictorios (Okubadejo y col., 2006; Schoenberg, 1988; Rijk y col., 1997; Pringsheim y col., 2014). Las diferencias observadas pueden deberse a diferencias en la tasa de supervivencia según la región, en diferencias en la detección de casos, o en diversas metodologías empleadas en el diseño de los estudios.

En España, se ha observado una prevalencia de entre 161-1500/100000 habitantes en todas las edades, aumentando hasta 973-10400/100000 habitantes en mayores de 80

años. En cuanto a la incidencia, se estimaron valores de 8,2/100000 habitantes en mayores de 50 años y de hasta 186,8/100000 habitantes en población mayor de 65 años (Claveria y col., 2002; Martínez-Suárez y col., 2000M; Vines, 1999; Benito-León y col., 2003). Debido al aumento en la edad de la población, se espera que el número de casos se doble en los próximos años (Dorsey y col., 2007).

2. ETIOPATOGENIA DE LA EP

La EP es una enfermedad multifactorial en la que intervienen factores tóxico-metabólicos y genéticos. Aunque en las últimas décadas se han descrito numerosos genes involucrados en la fisiopatología de la enfermedad, tan sólo en una pequeña minoría de casos la EP es debida a una única mutación causativa. La mayoría de las veces la causa es multifactorial. Los factores genéticos de bajo-moderado riesgo junto con factores ambientales y distintos estilos de vida son la causa de la degeneración neuronal.

2.1. FACTORES GENÉTICOS

La EP se ha considerado durante mucho tiempo una enfermedad esporádica causada por estrés o condiciones ambientales, como virus o neurotoxinas. Sin embargo, en 1997 se descubrió su carácter genético a través de la identificación de la primera mutación asociada a la EP en el gen de la α -sinucleína (*SNCA*; Polymeropoulos y col., 1997). Este hallazgo marcó un antes y un después en el conocimiento de la enfermedad, y le siguieron numerosos estudios enfocados en dilucidar los mecanismos genéticos implicados en el desarrollo de la EP (Kitada y col., 1998; Valente y col., 2001; Bonifati y col., 2003).

Hasta la fecha, se han identificado 11 genes responsables de formas monogénicas de la enfermedad, tanto dominantes como recesivos. Sin embargo, sólo un pequeño porcentaje de los casos de EP son formas monogénicas causadas por una única mutación, con penetrancia variable. Se estima que tan sólo el 30% de los casos monogénicos familiares y sobre el 3-5% de los casos complejos esporádicos pueden explicarse por este tipo de mutaciones (Kumar y col., 2011). La gran mayoría de casos son genéticamente complejos, donde hay una interacción entre variaciones genéticas que confieren cierto riesgo y condiciones ambientales, con múltiples presentaciones clínicas. Esta naturaleza de las enfermedades complejas ha llevado a elaborar dos teorías (Figura2):

- *Common disease rare variant (CDRV)*: mutaciones de alta penetrancia con baja frecuencia en la población (<1%) son la causa de la enfermedad.
- *Common disease common variant (CDCV)*: variaciones comunes presentes en la población (>1%) conllevan un aumento del riesgo en el desarrollo de enfermedades complejas como la EP.

Ambas teorías no son excluyentes, sino que las dos conjuntamente pueden explicar la complejidad etiológica que hay detrás de las enfermedades complejas como la EP (Bras y col., 2009).

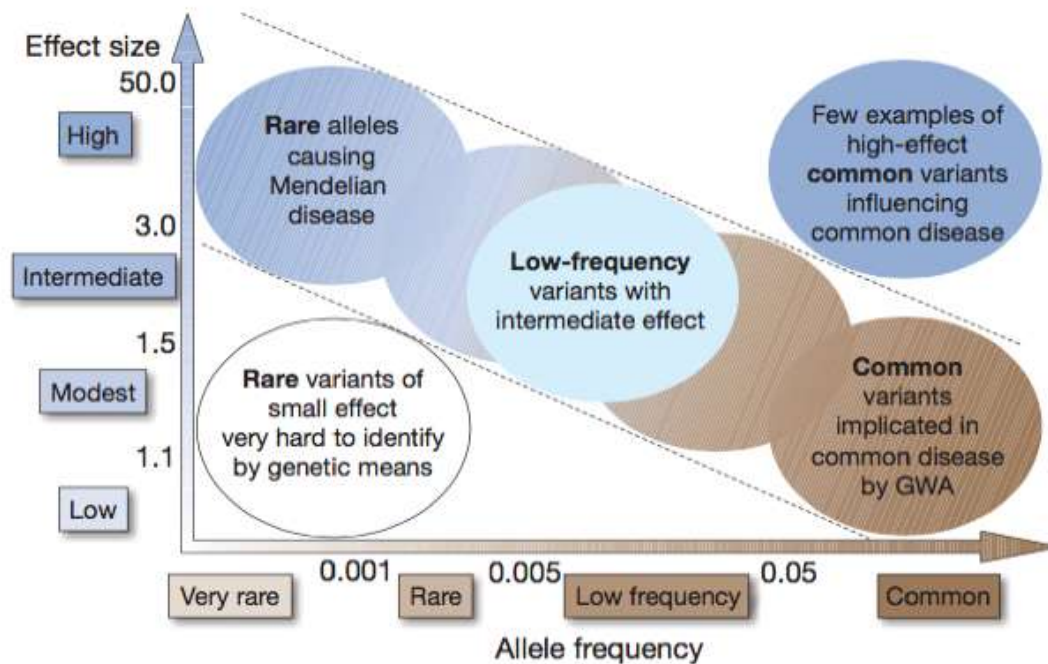


Figura 2. Representación esquemática del tipo de variaciones genéticas involucradas en la etiopatogenia de la EP. Estas variaciones comprenden desde mutaciones de herencia monogénica de alta penetrancia raras en la población, a variaciones comunes de herencia compleja asociadas a la EP, cuyo papel se ve interrelacionado con otras variaciones genéticas, así como con factores ambientales (Modificado de Manolio y col., 2009).

Tradicionalmente se han empleado los estudios de ligamiento en familias con varios miembros afectados para identificar mutaciones causativas de alta penetrancia. Sin embargo, en el caso de la EP, los casos de herencia mendeliana son raros. Por otra parte, la accesibilidad a muestras de familias con varios miembros afectados es compleja. Además, otros problemas como la penetrancia incompleta o la variabilidad clínica de los enfermos entran en juego en este tipo de abordaje. A pesar de todo, estos estudios son útiles para la detección de los casos de mutaciones causativas de Parkinson, y no se descarta la identificación de nuevas formas monogénicas de la EP en futuros estudios de ligamiento en familias (Bonifati y col., 2014). Sin embargo, para el estudio de variaciones comunes de bajo-moderado riesgo para la EP, el abordaje por excelencia es

el estudio de asociación de genoma completo (GWAS). Esta tecnología representa una mejora frente a los estudios tradicionales de genes candidatos, en los que los tamaños poblacionales eran menores, y en los que las variaciones genéticas estudiadas se limitaban a aquellas zonas genómicas que *a priori* podían estar involucradas en la etiopatogenia de la EP. Para realizar los estudios de asociación de genoma completo (GWA) se emplean chips de miles de polimorfismos de un solo nucleótido (SNPs) comunes en la población, localizados a lo largo de todo el genoma para genotiparlos en grandes cohortes poblacionales. Así se puede comprobar si su frecuencia alélica está asociada con la patología. Esta técnica ha permitido identificar nuevos *loci* de riesgo para la EP, lo que ha llevado a un mejor entendimiento de las bases moleculares que están detrás de la enfermedad (Manolio y col., 2009; Bonifati y col., 2014).

La aparición de la tecnología de secuenciación masiva de nueva generación (Figura 3) está contribuyendo aún más a descifrar la compleja red genética implicada en la EP. Por un lado, se están secuenciando regiones genómicas identificadas *a priori* por medio de estudios de ligamiento o mediante estudios GWA. Por otro lado, a través de la secuenciación del exoma o genoma completo de pacientes de EP y mediante el filtrado de variaciones genéticas nuevas y potencialmente patogénicas, se posibilita la identificación de nuevas mutaciones causativas para la EP. Con el coste cada vez menor de esta tecnología su uso está cada vez más extendido (Bonifati y col., 2014).

La identificación de variaciones genéticas mediante las diferentes estrategias descritas, hace necesarios los estudios de replicación en nuevas cohortes que validen los resultados (Bonifati y col., 2014). Además, los estudios funcionales son fundamentales para verificar la implicación de los nuevos genes identificados en la EP, sin embargo, estos estudios son difíciles de llevar a cabo en esta patología debido a la gran

complejidad del sistema nervioso y a la falta de modelos *in vitro* que repliquen la complejidad del cerebro humano (Billingsley y col., 2018).

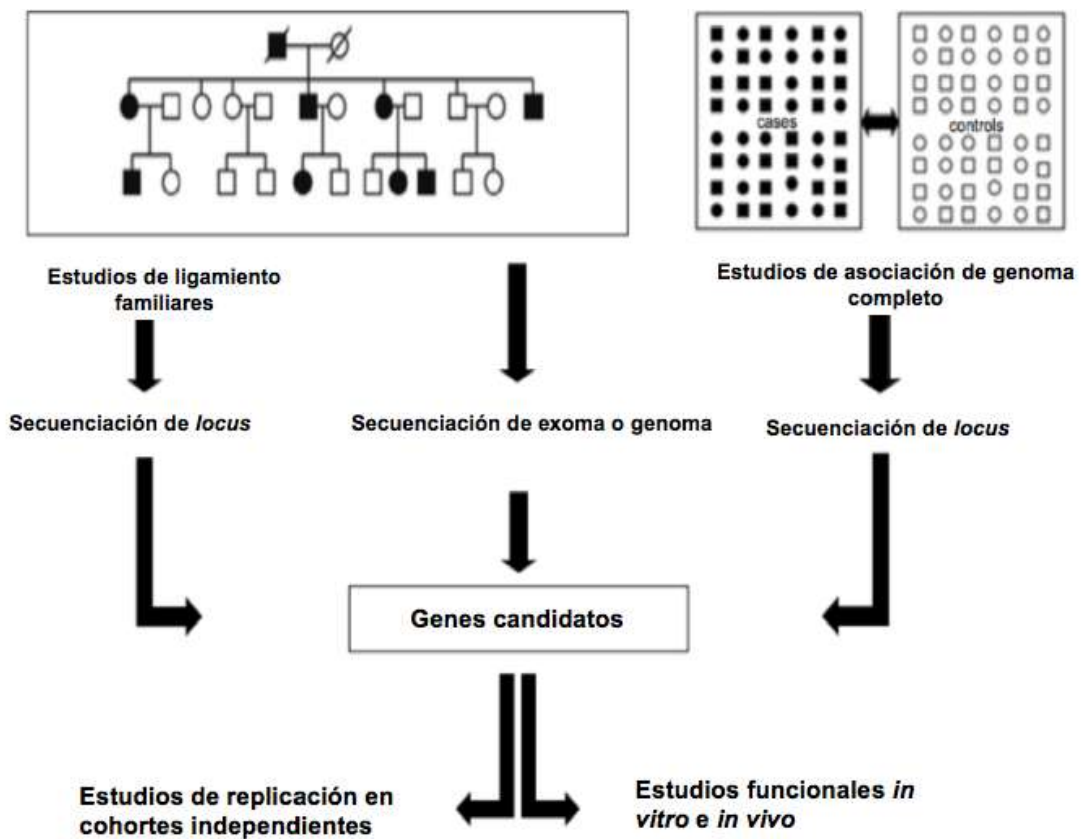


Figura 3. Esquema de las distintas estrategias empleadas para el estudio de los factores genéticos implicados en las enfermedades humanas complejas como la EP. Los estudios de ligamiento familiares permiten identificar mutaciones causativas de alta penetrancia en regiones a priori relacionadas con la enfermedad. Mediante los estudios de asociación de genoma completo se pueden identificar loci de riesgo asociados a la enfermedad. La tecnología de secuenciación masiva permite el estudio exhaustivo de regiones previamente identificadas susceptibles de estar involucradas en el desarrollo de la enfermedad, o la secuenciación de todo el exoma o genoma humano, para identificar nuevas variaciones candidatas a ser patogénicas. Los resultados de todas estas técnicas son necesarios validarlos mediante estudios en cohortes independientes y mediante estudios funcionales *in vitro* e *in vivo* (Modificado de Bonifati y col., 2014).

2.1.1. GENES CAUSATIVOS DE FORMAS MONOGÉNICAS DE EP

Hasta ahora, se ha demostrado que mutaciones en los genes, *SNCA*, *LRRK2* y *VPS35* causan formas dominantes de EP, mientras que mutaciones en *PINK1*, *DJ-1*, *PARK2*, *ATP13A2*, *FBXO7*, *PLA2G6*, *DNAJC6* y *SYNJ1* son responsables de formas recesivas de la enfermedad. Existen otros genes descritos cuya implicación en la etiopatogenia de la EP aún no se ha demostrado (Tabla 1).

Tabla 1. Genes causativos de formas monogénicas de la EP

FORMAS MONOGÉNICAS EN LA ENFERMEDAD DE PARKINSON		
GENES CONFIRMADOS		
GEN	PROTEÍNA	MODELO
<i>SNCA</i>	<i>α-Synuclein</i>	AD
<i>LRRK2</i>	<i>Leucine-Rich Repeat Kinase 2</i>	AD
<i>VPS35</i>	<i>Vacuolar Protein Sorting 35</i>	AD
<i>PARK2</i>	<i>Parkin</i>	AR
<i>PINK1</i>	<i>PTEN Induced Kinase 1</i>	AR
<i>DJ-1</i>	<i>Deglycase DJ-1</i>	AR
<i>ATP13A2</i>	<i>ATPase Cation Transporting 13A2</i>	AR
<i>PLA2G6</i>	<i>Phospholipase A2</i>	AR
<i>FBXO7</i>	<i>F-Box Only Protein 7</i>	AR
<i>DNAJC-6</i>	<i>DnaJ Heat Shock Protein Family (Hsp40) Member C6</i>	AR
<i>SYNJ1</i>	<i>Synaptojanin 1</i>	AR
GENES SIN CONFIRMAR		
GEN	PROTEÍNA	MODELO
<i>DNAJC13</i>	<i>DnaJ Heat Shock Protein Family (Hsp40) Member C13</i>	AD
<i>CHCHD2</i>	<i>Coiled-Coil-Helix-Coiled-Coil-Helix Domain Containing 2</i>	AD
<i>TMEM230</i>	<i>Transmembrane Protein 230</i>	AD
<i>UCHL1</i>	<i>Ubiquitin C-Terminal Hydrolase L1</i>	AD
<i>GIGYF2</i>	<i>GRB10 Interacting GYF Protein 2</i>	AD
<i>HTRA2</i>	<i>HTRA Serine Peptidase 2</i>	AD
<i>EIF4G1</i>	<i>Eukaryotic Translation Initiation Factor 4 Gamma 1</i>	AD
<i>RIC3</i>	<i>RIC3 Acetylcholine Receptor Chaperone</i>	AD
<i>PODXL</i>	<i>Podocalyxin Like</i>	AR

GEN	PROTEÍNA	MODELO
<i>PTRHD1</i>	<i>Peptidyl-TRNA Hydrolase Domain Containing 1</i>	AR
<i>RAB39B</i>	<i>Ras-Related Protein Rab-39B</i>	ligado al X
<i>VPS13C</i>	<i>Vacuolar Protein Sorting 13 Homolog C</i>	AR

AD: Autosómico dominante; AR: Autosómico recesivo.

2.1.1.1. GENES CON HERENCIA AUTOSÓMICA DOMINANTE

2.1.1.1.1. SNCA: Las mutaciones en el gen *SNCA* son una causa rara de EP. La primera mutación identificada en este gen que sienta las bases del carácter genético de la EP, es también la más frecuente (p.A53T). Se identificó por primera vez en una familia de origen italiano (Polymeropoulos y col., 1997). Su implicación ha sido demostrada en estudios posteriores en familias de diferente procedencia (Athanassiadou y col. 1999; Spira y col., 2001; Choi y col., 2008; Ki y col., 2007; Puschmann y col., 2009). Además de ésta, otras variaciones puntuales en este gen se han descrito asociadas a la EP (Kruger y col., 1998; Zarranz y col., 2004). También son frecuentes las variaciones en el número de copia del gen, tanto duplicaciones (Chartier-Harlin y col., 2004; Fuchs y col., 2007; Ibanez y col., 2004; Ikeuchi y col., 2008; Nishioka y col., 2006; Nuytemans y col., 2009; Troiano y col., 2008), como triplicaciones (Singleton y col., 2003; Farrer y col., 2004; Ibanez y col., 2004). Parece que la gravedad de la EP desarrollada por este tipo de variaciones genéticas es dosis dependiente, siendo la EP debida a triplicaciones en el gen más severa que la debida a duplicaciones (Hernández y col., 2016). Poco después del primer descubrimiento del

gen *SNCA* en relación con la EP, se descubrió que la α -sinucleína era el componente principal de los CL, la característica anatomopatológica más representativa de la enfermedad, aunque su papel en la EP no está aún claro (Spillantini y col., 1997).

2.1.1.1.2. *LRRK2*: Las mutaciones en este gen son la causa genética más frecuente de la EP de inicio tardío. Son responsables del 10% de los casos de EP familiar de herencia autosómica dominante, y del 3,6% de los casos esporádicos (Hernández y col., 2016). Se han descrito más de 100 mutaciones en este gen, pero sólo 7 han demostrado tener significancia clínica: p.R1441C/G/H, p.Y1699C, p.S1761R, p.I2012T, p.G2019S, y p.I2020T. En concreto, la mutación p.G2019S es la más frecuente en la población, siendo responsable del 0,5-4% de los casos en población caucásica, del 28% en judíos Askenazí y llega hasta el 39% en árabes del norte de África (Bandrés-Ciga y col., 2016; Lill, 2016). La penetrancia de la mutación es también muy variable según la edad (Hernandez y col., 2016; Healy y col., 2008).

2.1.1.1.3. *VPS35*: La primera mutación en este gen, p.D620N, fue identificada en 2011, en una familia de Austria, y posteriormente este hallazgo fue confirmado por otro estudio en una familia suiza (Haugarvoll y Wszolek, 2009; Vilariño-Güell y col., 2011). Esta mutación es responsable de alrededor del 1% de los casos de EP familiar y sólo del 0,2% de los casos de EP esporádica (Hernández y col., 2016; Kumar y col., 2012). La penetrancia de la mutación no es completa, pues se ha

encontrado en personas no afectas de edad avanzada (Vilariño-Güell y col., 2011). Se han identificado otras mutaciones en este gen, pero su patogenicidad no está clara todavía (Sheerin y col., 2012).

2.1.1.2. GENES CON HERENCIA AUTOSÓMICA RECESIVA

Las mutaciones en estos genes causan principalmente enfermedad de Parkinson de inicio temprano (EPIT), siendo responsables del 20% de estos casos y de un 3% de los casos de EP de inicio tardío (Hernandez y col., 2016).

2.1.1.2.1. *PARK2*: Las mutaciones en *PARK2* representan la primera causa de EP de herencia autosómica recesiva. Fue el segundo gen identificado causativo de EP y se identificó por primera vez en una familia japonesa con EP de inicio juvenil (EPIJ, edad de inicio menor de 35 años). Kitada y col. (1998) identificaron una delección homocigota de varios exones del gen *PARK2*. Posteriormente se han identificado alrededor de 150 mutaciones en este gen, de las cuales el 33% son mutaciones puntuales, el 13% pequeñas delecciones y hasta el 54% grandes delecciones y duplicaciones de varios exones (Hernández y col., 2016). Estas mutaciones se encuentran presentes en aproximadamente el 50% de los pacientes con EPIT de herencia autosómica recesiva y hasta en un 77% de los casos esporádicos de EPIJ (Lucking y col., 2000).

2.1.1.2.2. *PINK1*: Valente y col. (2004) describen por primera vez dos mutaciones homocigotas en el gen *PINK1*, p.G309D y p.W437X, identificadas en

una familia española y otra italiana respectivamente. Actualmente hay unas 10 mutaciones homocigotas y heterocigotas compuesta identificadas en este gen, siendo la segunda causa más frecuente de EP de herencia autosómica recesiva (Bonifati y col., 2005; Hernández y col., 2016). La presencia de mutaciones en este gen en heterocigosis en casos de EP esporádicos (2-4%) hace pensar que pueda suponer también un factor de riesgo para el desarrollo de la enfermedad (Nuytemans y col., 2010; Kumazawa y col., 2008).

2.1.1.2.3. *DJ-1*: En 2003, *DJ-1* se convirtió en el tercer gen asociado a la EP de herencia autosómica recesiva (Bonifati, 2003). Se han identificado tanto mutaciones puntuales como deleciones exónicas en homocigosis o en heterocigosis compuesta en varios estudios. A pesar de esto, sigue siendo una causa rara de EP, representando menos del 1% de los casos de EP de herencia autosómica recesiva (Bonifati y col., 2003; Annesi y col., 2005; Abou-Sleiman y col., 2003).

2.1.1.2.4. *ATP13A2*: Las mutaciones en *ATP13A2* se han descrito causativas del síndrome Kufor-Rakeb, una forma rara hereditaria de EPIJ. Tanto mutaciones en homocigosis como en heterocigosis compuesta han sido descritas en un total de 30 casos en 11 familias. Las mutaciones en este gen son muy poco frecuentes y el fenotipo es muy variable dependiendo de la mutación heredada (Eiberg y col., 2012; Park y col., 2015).

2.1.1.2.5. *PLA2G6*: En 2006 se asoció este gen por primera vez con enfermedades neurodegenerativas pediátricas que conllevan acumulación de hierro, aparentemente no relacionadas con la EP. Sin embargo, más tarde se identificaron mutaciones en este gen asociadas a un tipo de parkinsonismo con distonia, que responde a levodopa (Morgan y col., 2006; Paisán-Ruiz y col., 2009; Sina y col., 2009). Estas mutaciones son poco frecuentes y se han descrito casos tanto en homocigosis como en heterocigosis compuesta.

2.1.1.2.6. *FBXO7*: Se han identificado mutaciones en este gen en homocigosis y en heterocigosis compuesta en casos de EPIJ familiar en Irán y en Italia, ausentes en controles. Los casos se presentaban con distintos grados de severidad respecto a los síntomas asociados (Shojaee y col., 2008; Di Fonzo y col., 2009).

2.1.1.2.7. *DNAJC6*: En 2012 y 2013 se identificaron por primera vez casi simultáneamente mutaciones en el gen *DNAJC6* en dos familias con consanguinidad procedentes de Palestina y Turquía de EPIJ (Edvardson y col., 2012; Köroglu y col., 2013). Más tarde Olgiati identificó mutaciones en este gen en dos pacientes con EPIT familiar y de forma heterocigota en dos casos esporádicos de EPIT, confirmando este gen como causa de EP (Olgiati y col., 2015).

2.1.1.2.8. *SYNJI*: Una mutación idéntica, p.R258G, fue identificada en dos estudios publicados al mismo tiempo en dos familias, una de Sicilia y

otra de Irán, con dos hermanos afectados de padres con consanguinidad (Quadri y col., 2013; Krebs y col., 2013), y más tarde en un tercer estudio del sur de Italia, y similares características clínicas (Olgati y col., 2014). Su implicación en la EP ha sido corroborada en estudios posteriores (Kirola y col., 2016; Rauschendorf y col., 2017; Taghavi y col., 2018).

2.1.2. GENES PROPUESTOS CAUSATIVOS DE FORMAS MONOGÉNICAS DE EP PERO NO CONFIRMADOS

Diferentes estudios han identificado genes con mutaciones candidatas a ser responsables de formas monogénicas de EP, pero la ausencia de evidencias en posteriores estudios de replicación que confirmen estos resultados hace que no se tenga certeza de la relación de algunos de los genes propuestos con la EP.

2.1.2.1. GENES CON HERENCIA AUTOSÓMICA DOMINANTE

2.1.2.1.1. *DNAJC13*: En 2013 se identificó por primera vez la mutación p.N855S en el gen *DNAJC13* en una familia con EP de herencia autosómica dominante (Vilariño-Güell y col., 2014). Esta mutación se identificó también en otras dos familias y en tres casos esporádicos en Canadá, pero no en otras poblaciones (Gustavsson y col.2015; Foo y col., 2014; Lorenzo-Betancor y col., 2015; Ross y col., 2016).

2.1.2.1.2. *CHCHD2*: El gen *CHCHD2* se encuentra bajo estudio desde que en 2015 Funayama y col. proponen la mutación T61I como causa de EP familiar de herencia autosómica dominante en una familia japonesa (Funayama y col., 2015). Además, en este estudio se describen otras dos mutaciones posiblemente causativas de EP (p.R145Q y c.300+5G>A) y dos polimorfismos como posibles factores de riesgo. Para confirmar estos resultados se han realizado posteriormente numerosos estudios del gen *CHCHD2* en diferentes poblaciones. En población de Taiwan, China, Canadá o Europa no se han podido confirmar estos resultados (Fan y col, 2016; Jansen y col., 2015; Lu y col., 2016; Zhang y col., 2016). Otros autores si confirman la presencia de variaciones en el gen *CHCHD2* como factores de riesgo para la EP en población japonesa y otras cohortes de China (Foo y col., 2015; Funayama y col., 2015; Shi y col., 2016).

2.1.2.1.3. *TMEM230*: Recientemente, se ha descrito una mutación en el gen *TMEM230* en una familia norte-americana con EP de herencia autosómica dominante (Deng y col., 2016). En el mismo estudio se describe una mutación localizada en dos pacientes no relacionados con EPIT; y otra en población china en 7 casos de EP familiar. Sin embargo, posteriores estudios han fallado en la replicación de estos hallazgos (Giri y col., 2017; Yan y col., 2017; Quadri y col., 2017; Shi y col., 2017; Yang y col., 2017; Wu y col., 2017; Baumann y col., 2017; He y col., 2017; Ibanez y col., 2017; Buongarzone y col., 2017).

- 2.1.2.1.4. *UCHL1*:** Fue descrito en 1998 en una familia alemana con EP de herencia autosómica dominante en la que se detectó una mutación en homocigosis (Leroy y col., 1998). Hasta la fecha, no se ha publicado ningún otro trabajo que confirme este resultado.
- 2.1.2.1.5. *GIGYF2*:** Este gen se encuentra en el locus *PARK11*, que fue asociado a la EP en 2002. Mutaciones en este gen en pacientes de Italia y Francia se han asociado al desarrollo de la EP (Lautier y col., 2008). Sin embargo, no se han podido replicar estos resultados en otros estudios en diferentes cohortes (Bras y col., 2009; Bonetti y col., 2009; Guella y col., 2011; Nichols y col., 2009).
- 2.1.2.1.6. *HTRA2*:** En un estudio alemán de 2005 realizado en pacientes de EP se detectó una mutación en el gen *HTRA2* ausente en controles, y un polimorfismo asociado con la EP (Strauss y col., 2005). Posteriormente se estudió la presencia de mutaciones en este gen en pacientes de EP de herencia autosómica dominante sin hallarse ningún resultado, y sin ninguna correlación de variaciones en el gen con una mayor susceptibilidad a la EP (Ross y col., 2008).
- 2.1.2.1.7. *EIF4G1*:** En 2011 se describen por primera vez 5 mutaciones en el gen *EIF4G1*, pero la cosegregación con la enfermedad se demuestra sólo para una de las mutaciones, la p.R1205H. Esta mutación se identificó también en pacientes de Estados Unidos, Canadá, Irlanda, Italia y Túnez (Chartier-Harlin y col., 2011). Sin embargo, que *EIF4G1* sea un gen

implicado en el desarrollo de la EP está aún por confirmar. Se han publicado estudios en otros grupos étnicos que localizan ésta y otras mutaciones del gen *EIF4G1* en controles (Tucci y col., 2012; Lesage y col., 2012; Siitonen y col., 2013).

2.1.2.1.8. *RIC3*: En 2016 se describe una mutación en el gen *RIC3*, p.P57T, presente en 9 miembros de una familia de 3 generaciones (Sudhaman y col., 2016). Sólo los miembros de la familia con sintomatología son portadores de la mutación descrita. Sin embargo, la temprana edad a la que se realizó el examen de los miembros de la tercera generación no permite discernir realmente si los no portadores no desarrollarán EP o algún síntoma relacionado. Tampoco se ha confirmado la asociación de ninguna variación en *RIC3* en estudios poblacionales (Ross y col., 2017).

2.1.2.2. GENES CON HERENCIA AUTOSÓMICA RECESIVA O LIGADOS AL CROMOSOMA X

2.1.2.2.1. *PODXL*: Recientemente se ha identificado una mutación que conlleva la completa pérdida de la proteína *PODXL* en homocigosis en 3 hermanos con EP (Sudhaman y col., 2016). El resto de hermanos, así como los padres no afectados, eran portadores en heterocigosis de la mutación o carecían de ella, pero no hay otros estudios que corroboren estos hallazgos.

2.1.2.2.2. *PTRHDI*: Dos estudios independientes sitúan a este gen como posible causa de EP de herencia autosómica recesiva. Por una parte, Khodadadi y col., describen una mutación en homocigosis en *PTRHDI* presente en dos hermanos con discapacidad intelectual y trastornos motores desde la niñez, desarrollando parkinsonismo a los 20-30 años (Khodadadi y col., 2017). Jabery y col., por otra parte, describen el caso de dos hermanos con características clínicas similares a los del estudio de Khodadadi con otra mutación en el gen *PTRHDI*, aunque también presentan otra mutación en el gen *ADORA1*, que podría también ser la causa de este fenotipo (Jabery y col., 2016). *PTRHDI* por su parte es un gen cuya implicación en síndromes de discapacidad intelectual es conocida, pero su papel en el desarrollo de la EP debe ser mejor explorada.

2.1.2.2.3. *RAB39B*: En el gen *RAB39B* se han identificado mutaciones asociadas a discapacidad intelectual ligada al cromosoma X y a EPIT. En 2014 se identificaron las dos primeras mutaciones, en Australia y en una familia de Estados Unidos (Wilson y col., 2014). Más tarde, en 2015 se identificó una nueva mutación en este gen en un caso aislado de EP (Lesage y col., 2015).

2.1.2.2.4. *VPS13C*: Hasta ahora, tres mutaciones se han identificado en este gen, en tres casos independientes de EP, uno en Turquía y otros dos en Francia (Lesage y col., 2016). Parece que la función de esta proteína podría interferir en la función de *PARK2*, pero los datos publicados todavía no sustentan su implicación en la EP.

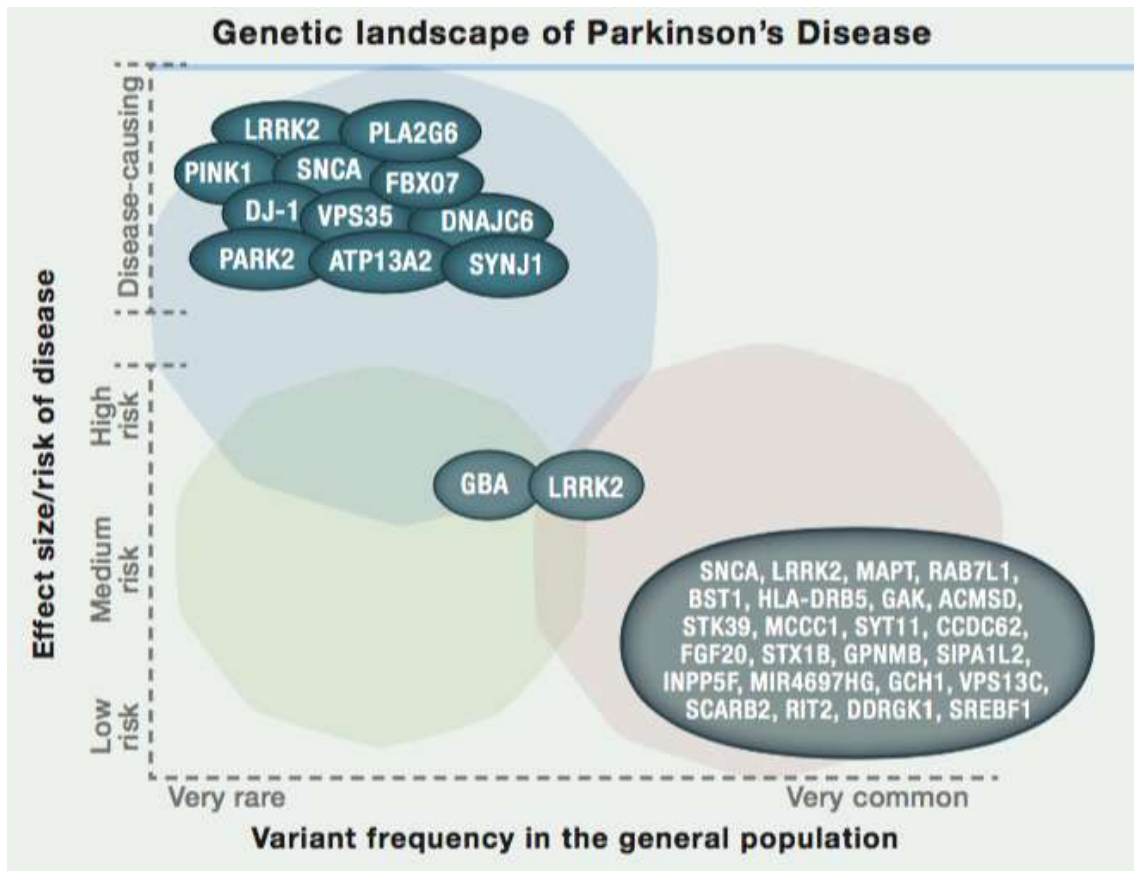


Figura 4. Representación esquemática de la arquitectura genética de la EP. El eje Y muestra el efecto de la mutación y el X indica la frecuencia en la población. El color azul indica genes involucrados en la EP descubiertos mediante estudios de ligamiento en familias, en verde genes identificados mediante secuenciación de exoma, y en rosa genes identificados mediante estudios GWA. Un total de 11 genes están implicados en formas monogénicas de la EP (zona superior izquierda). De éstos, 8 siguen un patrón de herencia autosómica recesivo, causando EP de inicio temprano (PARK2, PINK1 y DJ-1) o de inicio juvenil (ATP13A2, PLA2G6, FBX07, DNAJC6 y SYNJ1). Los 3 restantes siguen un patrón de herencia autosómica dominante (SNCA, LRRK2 y VPS35). En la zona central del gráfico se localizan genes en los que se han identificado mutaciones de riesgo alto para el desarrollo de la EP (GBA y LRRK2). Por último, la presencia de variaciones de bajo riesgo para la EP se han identificado mediante estudios GWA en grandes poblaciones (zona inferior derecha). (Modificado de Bras y col., 2015).

2.1.3. *LOCI DE RIESGO EN LA EP*

Los estudios GWA se han aplicado en el campo del Parkinson desde que en 2005 se identificara mediante esta técnica un *locus* de riesgo para la degeneración macular (Klein y col., 2005). Los primeros estudios GWA en la EP fallaron debido a que el número de muestras no era suficiente para detectar variaciones comunes con un *odds ratio* (OR) bajo o moderado. Por ello, los trabajos colaborativos entre diferentes grupos y laboratorios resultan fundamentales para contar con cohortes suficientes que permitan identificar este tipo de variaciones. En 2009, dos trabajos colaborativos identificaron los primeros *loci* de riesgo para la EP, como *SNCA* o *MAPT* (Satake y col., 2009; Simón-Sánchez y col., 2009). Y desde entonces se han publicado numerosos trabajos identificando nuevos *loci* de riesgo (Edwards y col., 2010; Hernández y col., 2012; Pihlstrom y col.2013). Pero no fue hasta 2014 cuando Nalls y col. realizan el primer meta-análisis de estudios publicados GWA en población de origen europeo, contando por primera vez con una cohorte de varios miles de muestras: 19061 pacientes de EP y 100833 controles (Nalls y col., 2014). En este estudio se identificaron en total 28 señales independientes situadas en 24 *loci* de riesgo para la EP (Tabla 2).

A este trabajo le sigue un posterior meta-análisis realizado por Chang y col. con 26035 pacientes de EP y 403190 controles (Chang y col., 2017), en el que se identifican 17 nuevos *loci* de riesgo para la EP (Tabla 3).

En España en 2016, Bandrés-Ciga y col. estudian la arquitectura genética de esta enfermedad en población del sur de España (Bandrés-Ciga y col., 2016). De los 24 *loci* de riesgo identificados previamente, sólo 5 se replicaron en este estudio (Tabla 2). Sin embargo, este análisis en población española tiene la limitación de contar con una cohorte pequeña (240 casos de EP y 192 controles). Más estudios en ésta y otras

poblaciones son necesarios para confirmar estos *loci* de riesgo para la EP, así como estudios funcionales que logren identificar los genes implicados.

Tabla 2. 28 señales independientes en 24 *loci* asociados con la EP en el estudio de meta-análisis de Nalls y col., 2014.

CHR:BP	SNP	Gen candidato	OR
1:155135036	rs35749011	<i>GBA/SYT11</i>	1,82
1:154898185	rs114138760	<i>GBA/SYT11</i>	1,59
1:205723572	rs823118	<i>RAB7L1/NUCKS1</i>	1,12
1:232664611	rs10797576	<i>SIPAIL2</i>	1,13
2:135539967	rs6430538	<i>ACMSD/TMEM163*</i>	0,88
2:169110394	rs1474055	<i>STK39 *</i>	1,21
3:182762437	rs12637471	<i>MCCC1</i>	0,84
4:951947	rs34311866	<i>TMEM175/GAK/DGKQ</i>	0,79
4:944210	rs34884217	<i>TMEM175/GAK/DGKQ</i>	1,11
4:15737101	rs11724635	<i>BST1</i>	1,13
4:77198986	rs6812193	<i>FAM47E/SCARB2</i>	0,91
4:90626111	rs356182	<i>SNCA</i>	0,76
4:90763703	rs7681154	<i>SNCA</i>	0,93
6:32666660	rs9275326	<i>HLA-DQB1</i>	0,83
6:32343604	rs13201101	<i>HLA-DQB1</i>	1,22
7:23293746	rs199347	<i>GPNMB</i>	1,11
8:16697091	rs591323	<i>FGF20</i>	0,92
10:121536327	rs117896735	<i>INPP5F</i>	1,62
11:133765367	rs329648	<i>MIR4697*</i>	1,11
12:40614434	rs76904798	<i>LRRK2</i>	1,16
12:123303586	rs11060180	<i>CCDC62</i>	1,11
14:55348869	rs11158026	<i>GCHI</i>	0,90
15:61994134	rs2414739	<i>VPSI3C</i>	1,11
16:31121793	rs14235	<i>BCKDK/STX1B</i>	1,10
17:17715101	rs11868035	<i>SREBF/RAI1*</i>	0,94
17:43994648	rs17649553	<i>MAPT*</i>	0,77
18:40673380	rs12456492	<i>RIT2</i>	0,90
20:3168166	rs8118008	<i>DDRGI1</i>	1,11

**loci* replicados en población de Granada (sur de España) en el estudio de Bandrés-Ciga y col., 2016

CHR: *chromosome*; BP: *pair base*; SNP: *single nucleotide polymorphism*; OR: *odds-ratio*

Hasta la fecha, se sabe que entre un 6-7% de la heredabilidad de la EP puede explicarse a *loci* ya identificados, pero se estima que hasta un 27% de la heredabilidad del Parkinson podría deberse a este tipo de variaciones comunes por lo que queda todavía mucho trabajo que realizar en este sentido. Serán necesarios esfuerzos comunes entre laboratorios para contar con grandes cohortes que permitan identificarlas (Keller y col., 2012; Bras y Singleton, 2009). Además, es probable que una parte de esa heredabilidad perdida se deba a variaciones raras con bajo o alto riesgo para la EP, junto a variaciones estructurales que se escapan en los estudios GWA, por lo que son necesarias otras técnicas genéticas con el fin de desentrañar toda la etiología involucrada en la EP.

Tabla 3. 17 *loci* asociados con la EP en el estudio de meta-análisis de Chang y col., 2017

CHR:PB	SNP	Gen Candidato	OR
1:226916078	rs4653767	<i>ITPKB</i>	0,92
2:102413116	rs34043159	<i>IL1R2</i>	1,08
2:166133632	rs353116	<i>SCN3A</i>	0,94
3:18277488	rs4073221	<i>SATB1</i>	1,1
3:48748989	rs12497850	<i>NCKIPSD/CDC71</i>	0,93
3:52816840	rs143918452	<i>ALAS1/TLR9/DNAH1/BAP1/PHF7/NISCH/STAB1/ITIH3/ITIH4</i>	0,68
4:114360372	rs78738012	<i>ANK2/CAMK2D</i>	1,13
5:60273923	rs2694528	<i>ELOVL7</i>	1,15
6:27681215	rs9468199	<i>ZNF184</i>	1,11
8:11707174	rs2740594	<i>CTSB</i>	1,09
8:22525980	rs2280104	<i>ORBS3/PDLIM2/C8orf58/BIN3</i>	1,07
9:17579690	rs13294100	<i>SH3GL2</i>	0,92
10:15569598	rs10906923	<i>FAM171A1</i>	0,93
14:88472612	rs8005172	<i>GALC</i>	1,08
16:19279464	rs11343	<i>COQ7</i>	1,07
16:52599188	rs4784227	<i>TOX3</i>	1,09
17:40698158	rs601999	<i>ATP6V0A1/PSMC3IP/TUBG2</i>	0,93

CHR: *chromosome*; PB: *pair base*; SNP: *single nucleotide polymorphism*; OR: *odds-ratio*

2.1.4. GENES DE RIESGO EN LA EP

Mientras que los estudios GWA señalan los *loci* asociados a la EP, el gen responsable de la patogenicidad es desconocido. En numerosos estudios, se ha tratado de dilucidar como estos factores de riesgo pueden estar involucrados en la patología de la enfermedad a través de las rutas metabólicas en las que participan. El hecho es que muchos de estos factores de riesgo identificados en grandes estudios GWA aparecen también en formas monogénicas de la enfermedad por lo que se piensa que los mismos mecanismos moleculares están involucrados en las dos vertientes de la enfermedad (Parikshak y col., 2015). Variaciones en los genes *SNCA*, *LRRK2* y *MAPT*, así como variaciones menos frecuentes en *GBA*, son las señales más asociadas estadísticamente en estos estudios, así como en estudios de genes candidatos.

2.1.4.1. SNCA: A través de numerosos estudios se ha confirmado que este gen, además de mutaciones deletéreas causativas de EP, contiene variaciones en regiones no codificantes que contribuyen al riesgo de desarrollo de la enfermedad. El primer estudio publicado en este sentido identificó la asociación de un polimorfismo, REP1, ubicado en el promotor de este gen (Maraganore y col., 2006). Posteriores estudios GWA han señalado repetidamente variaciones en regiones no codificantes de *SNCA* asociadas con la EP (Simón-Sánchez y col., 2009; Lill y col.; 2012, Nalls y col.; 2014, Chang y col., 2017), e incluso se han propuesto varias señales independientes dentro del mismo *locus*, aunque todavía no se conoce con certeza cuántas hay.

2.1.4.2. LRRK2: Dos líneas de investigación sugieren que este gen está implicado en la EP como un factor de riesgo, además de contener mutaciones causativas. Por un lado en diversos estudios se han detectado dos polimorfismos, p.G2385R y p.R1628P, cuya presencia implica el doble de riesgo de padecer EP, sobre todo en población asiática (Di Fonzo y col., 2006, Farrer y col., 2007, Tan y Schapira 2008, Tan y col., 2007). Por otra parte, numerosos estudios GWA señalan varios polimorfismos situados en regiones no codificantes cerca de este gen que incrementan el riesgo de padecer EP en 1,2, sobre todo en población caucásica (Chang y col., 2017). Parece que alteraciones en la expresión del gen o en el *splicing* podrían estar implicados en este caso incrementando el riesgo de desarrollo de EP, aunque para conocer las variaciones exactas implicadas se necesita un estudio más profundo.

2.1.4.3. GBA: Debido a que las variaciones en *GBA* tienen una frecuencia menor del 5% en la población, este gen no fue detectado en los primeros estudios GWA realizados en la EP, sino que fue a través de estudios de genes candidatos. En primer lugar se identificaron mutaciones en homocigosis causativas de la enfermedad de Gaucher. Sin embargo, con el tiempo observaron que familiares de primer y segundo grado de estos enfermos tenían una incidencia alta de EP (Neudorfer y col., 1996; Aharon-Peretz y col., 2004; Halperin y col., 2006). Más tarde, Sindransky y col., mediante un estudio de meta-análisis de los datos publicados hasta la fecha identificaron dos mutaciones en el gen *GBA* en heterocigosis con una alta frecuencia en casos de EP frente a controles, que conferían un riesgo de

hasta cinco veces superior de desarrollo de EP (Sidransky y col., 2009). Se estima que en torno al 5-10% de pacientes de EP son portadores de mutaciones en *GBA*, siendo particularmente frecuente en judíos Askenazí (Mitsui y col., 2009). La penetrancia varía dependiendo de la mutación, y con ella el fenotipo asociado (Ferreira y col., 2017).

2.1.4.4. *MAPT*: La asociación entre la EP y *MAPT* resulta intrigante puesto que *MAPT* codifica para la proteína Tau que es el mayor componente de los depósitos de neurofibrillas características de la enfermedad de Alzheimer, pero no de la EP. Se ha observado, sin embargo, la asociación entre variaciones en el haplotipo H1 de *MAPT* con el riesgo de desarrollo de EP en diversos grupos poblacionales (Fung y col., 2006; Kwok y col., 2004; Skipper y col., 2004; Zabetian y col., 2007), aumentando hasta un 1,3 el riesgo de la enfermedad. De hecho, *MAPT* es una de las señales más consistentes en los estudios GWA, aunque más en población caucásica que en asiática.

2.1.5. GENES NO CONFIRMADOS DE RIESGO EN LA EP

2.1.5.1. *SMPD1*: Se sabe que mutaciones en este gen, p.R496L, p.L302P y p.fs330P, son causativas de la enfermedad de Niemann-Pick en judíos Askenazí (Levran y col., 1991; Levran y col., 1991; Levran y col., 1993), aunque algunas de ellas han sido asociadas posteriormente con la EP en algunos estudios (Gan-Or y col., 2013; Dagan y col., 2015), pero no en otros (Foo y col., 2013; Gan-Or y col., 2013).

2.1.5.2. *HTRA2*: El papel de *HTRA2* como factor de riesgo en la EP sigue siendo controvertido, ya que se ha descrito que actúa como factor de riesgo para la EP e incluso se han corroborado estos resultados con estudios *in vitro* (Strauss y col., 2005). Sin embargo, en 2008, Simón-Sánchez y Singleton, no encuentran ninguna de las mutaciones previamente descritas relacionadas con la EP (Simón-Sánchez y Singleton, 2008).

2.2. RUTAS METABÓLICAS IMPLICADAS EN LA EP

Más allá de identificar las causas genéticas implicadas en la enfermedad, esclarecer los mecanismos por los cuales estos factores están implicados en el desarrollo de la EP es el objetivo último, incluyendo si actúan independientemente o conjuntamente a través de determinadas rutas metabólicas. Tras los últimos meta-análisis, los genes propuestos asociados a la EP parecen pertenecer en su mayoría a unas pocas rutas ya propuestas relacionadas con la patología como son la autofagia, la biología mitocondrial, el estrés oxidativo, el tráfico de vesículas sinápticas o la neuroinflamación. Estos mecanismos no son excluyentes y normalmente hay interacción entre ellos. Por ello, mucho de los genes están involucrados en más de una de estas rutas.

La α -sinucleína es una proteína presináptica que en condiciones normales se encuentra sin plegar. Mutaciones en el gen *SNCA* provocan un mal plegamiento de la proteína, así como un aumento en su concentración si se trata de duplicaciones o triplicaciones en el gen (Cacabelos, 2017). En estas dos condiciones, la proteína tiende a plegarse y formar agregados que formarán los CL característicos de la EP. Parece que estos acúmulos fomentan la salida de la dopamina de las vesículas presinápticas de manera que se acumula en el citoplasma. La dopamina en el citoplasma tiende a formar especies oxidativas, produciendo daño oxidativo y muerte celular (Fujita y col., 2014).

Por otra parte, la vía de reciclaje de proteínas mal plegadas o dañadas del sistema ubiquitina proteosoma está involucrada en el desarrollo de la EP. La proteína codificada por *PARK2* es una E3 ubiquitina ligasa encargada de ubiquitinar la α -sinucleína para su degradación en el proteosoma. La proteína codificada por *UCHL1* también está involucrada en este proceso, a través del mantenimiento de ubiquitinas disponibles (Fujita y col., 2014; Karimi-Moghadam y col., 2018). La reducción en el funcionamiento del proteosoma por mutaciones en estos genes afecta la degradación de proteínas, y en concreto a la inadecuada eliminación de fibrillas disfuncionales de α -sinucleína, aumentando su concentración y rompiendo la homeostasis de la dopamina, lo que produce neurotoxicidad por el incremento del estrés oxidativo.

Los defectos bioenergéticos, la disrupción de la cadena de transporte de electrones y el aumento de las especies reactivas de oxígeno debido a un mal funcionamiento de la mitocondria juegan también un papel central en la patogénesis tanto de EP familiar como esporádico. Parece ser que *PARK2* junto a *PINK1* están involucrado en el sistema de control de la mitocondria a través de la vía de degradación dependiente del lisosoma (autofagia o mitofagia) de las mitocondrias dañadas. *DJ-1* se ha sugerido también que juega un papel similar a *PINK1-PARK2* en el mantenimiento de la función mitocondrial en presencia de estrés oxidativo (Rocha y col., 2018). Es más, la causa más común de EP monosómico es la mutación G2019S en *LRRK2* y se ha visto implicada en el mal funcionamiento de la mitocondria en fibroblastos humanos que portaban esa mutación (Mortiboys y col., 2010). Otros genes más recientemente relacionados con la EP también precen estar implicados en este proceso como *PLA2G6*, *CHCHD2*, *HTRA2*, *EIF4G1* y *VPS13C*. El estrés oxidativo, como consecuencia de la disfunción mitocondrial, lleva a la degeneración de neuronas dopaminérgicas, especialmente vulnerables a este daño (Poewe y col., 2017).

Por otra parte, la neuroinflamación y la activación del sistema inmune también son procesos asociados a enfermedades neurodegenerativas como la enfermedad de Alzheimer o la EP (Fujita y col., 2014). La microglía, presente en abundancia en zonas del cerebro como la *pars compacta*, detecta la α -sinucleína mal plegada y produce neurotoxicidad mediante la producción de estrés oxidativo y citoquinas proinflamatorias (Zhang y col., 2005; Alvarez-Erviti y col., 2011). Se ha visto como la expresión de genes relacionados con la EP, como *LRRK2*, *PARK2* o *DJI*, están involucrados en esta respuesta (Gillardon y col., 2012; Frank-Cannon y col., 2008; Waak y col., 2009).

2.3. FACTORES AMBIENTALES EN LA EP

Vivir en ambientes rurales, dedicarse a la agricultura o consumir agua de pozo, son estilos de vida que aumentan la exposición a determinados tóxicos asociados a un mayor riesgo de padecer EP, como pesticidas, herbicidas o metales pesados (Obeso y col., 2017; Fleming y col., 2017). Dos pesticidas en concreto son ampliamente conocidos factores de riesgo para la EP, el Paraquat y la Rotenona. Ambos interfieren en la cadena de transporte de electrones creando estrés oxidativo, lo que lleva a la degeneración de neuronas dopaminérgicas (Ramachandiran y col., 2007, Liou y col., 1997, Semchuk y col., 1992). Los dos metales pesados cuya relación con la EP se encuentra mejor descrita son el hierro y el manganeso. El manganeso es una metal que se encuentra de forma libre en la naturaleza y una excesiva exposición a él, como en la minería o en la soldadura, se ha visto que es un factor de riesgo para el desarrollo de una patología neurodegenerativa relacionada con la EP (Kwakye y col., 2015). La relación entre el hierro y la EP es conocida desde hace décadas. Parece que actúa a nivel de la

neuromelamina de las neuronas dopaminérgicas, pero su papel en la EP aún no está claro (Dexter y col., 1987; Sofic y col., 1988; Gorell y col., 1995; Double y col., 2003). Otros factores, sin embargo, han demostrado tener un efecto protector a la hora de desarrollar EP. En múltiples estudios se demuestra que los fumadores tienen menos de la mitad de riesgo de padecer Parkinson en comparación con los no fumadores, siendo el efecto además dosis dependiente (Hernán y col., 2002; Fleming y col., 2017; Ritz y col., 2007). Por otra parte, consumir grandes cantidades de cafeína ha sido asociada a una disminución de hasta el 60% de riesgo de padecer Parkinson en hombres, aunque no en mujeres (Palacios y col., 2012; Yamada-Fowler y col., 2015). Los niveles altos de ácido úrico en sangre han demostrado también tener un efecto protector en el desarrollo de la EP (Weisskopf y col., 2007), y así se ha confirmado en un estudio en población española (Jesús S y col., 2013). Los mecanismos implicados en los efectos protectores de estas sustancias continúan bajo estudio.

II. HIPÓTESIS, OBJETIVOS Y DISEÑO EXPERIMENTAL

1. HIPÓTESIS DE TRABAJO

La EP es una enfermedad de etiopatogenia multifactorial, en la que intervienen factores tóxico-metabólicos y genéticos, aunque las bases moleculares subyacentes a la enfermedad no se conocen con exactitud. Debido al aumento de su prevalencia en una población cada vez más envejecida, a la falta de terapias que frenen el proceso neurodegenerativo así como al creciente consumo de recursos sanitarios y sociales que conlleva, es fundamental hacer un diagnóstico precoz de la enfermedad e identificar nuevas dianas terapéuticas contra las que diseñar fármacos que frenen o reviertan esta degeneración asociada. Por otra parte, la heterogenicidad existente en la enfermedad en cuanto a sintomatología clínica, progresión, edad de inicio y respuesta a los tratamientos médicos, sugiere la existencia de distintos subtipos dentro de la patología. Poder distinguir a los diferentes pacientes de EP llevará también a un mejor tratamiento, pudiendo ofrecer una medicina personalizada adaptada a cada caso. Para dar respuesta a todas estas cuestiones resulta fundamental conocer en profundidad la fisiopatogenia que hay detrás de la EP. Nuestra hipótesis es que a través del estudio de las claves genéticas implicadas en esta enfermedad se contribuirá a completar el conocimiento sobre la intrincada red molecular que contribuye al desarrollo de la EP con el fin de proporcionar nuevos datos y contribuir al conocimiento de las bases moleculares que están implicadas en la patología.

2. OBJETIVOS

El objetivo general de este trabajo es profundizar en las bases moleculares que subyacen en la EP. Para ello se han reclutado dos cohortes de origen nacional, en la que se han evaluado el papel de genes conocidos asociados a la EP, el de otros genes cuya implicación no está confirmada, y de *loci* de riesgo asociados a la enfermedad. El estudio se dividió en los siguientes objetivos concretos:

Objetivo 1: Estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas de EP. Para la consecución del objetivo 1 se realizaron los siguientes sub-objetivos:

- **Sub-objetivo 1:** Estudiar exhaustivamente 17 genes relacionados con la EP y otros parkinsonismos en pacientes de inicio temprano de la Comunidad de Madrid.
- **Sub-objetivo 2:** Estudiar la implicación en la fisiopatogenia de la EP de *CHCHD2*, nuevo gen propuesto causativo de EP, en población del sur de España.
- **Sub-objetivo 3:** Estudiar la implicación en la fisiopatogenia de la EP de *TMEM230*, nuevo gen propuesto causativo de EP, en población del sur de España.

Objetivo 2: Replicación de 5 *loci* de riesgo para la EP, previamente asociados a una cohorte de EP de Granada, en otra cohorte independiente del sur de España.

3. DISEÑO EXPERIMENTAL

Con el propósito de alcanzar los objetivos propuestos, se planteó el siguiente diseño experimental (Figura 5).

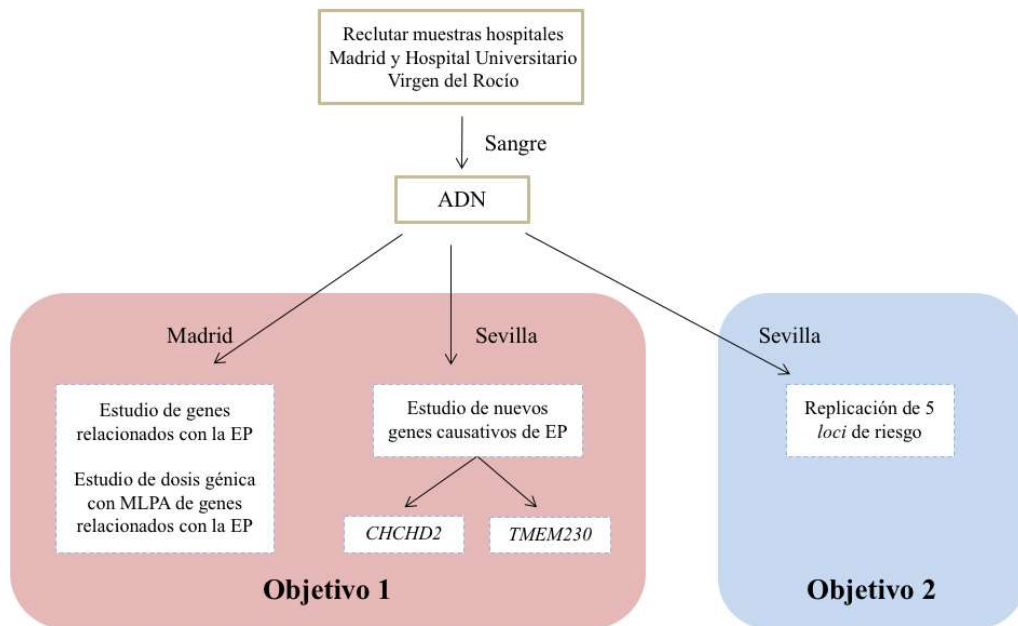


Figura 5. Esquema del diseño experimental seguido en la Tesis Doctoral. En el primer objetivo se ha realizado el estudio de genes relacionados con la EP. La primera parte ha consistido en un análisis de secuenciación masiva y estudio de dosis génica mediante Multiplex Ligation-dependent Probe Amplification (MLPA) en pacientes de EP de inicio temprano de la Comunidad de Madrid. La segunda parte ha consistido en estudiar dos genes recientemente propuestos causativos de EP en pacientes procedentes del sur de España. Para el objetivo 2 se ha realizado el genotipado de 5 loci previamente asociados a pacientes de EP del sur de España, en una cohorte distinta de similar procedencia.

Para alcanzar el **objetivo 1**, el primer paso fue reclutar a los pacientes de EPIT procedentes de la Comunidad Autónoma de Madrid, que procedían de los siguientes hospitales: Hospital Príncipe de Asturias, Fundación Jiménez Díaz, Hospital Clínico San Carlos, Hospital 12 de Octubre, Hospital Ramón y Cajal, Hospital Fundación

Alcorcón, Hospital de la Princesa y Hospital Nuestra Señora del Rosario. Posteriormente, se llevó a cabo la secuenciación masiva de genes descritos relacionados con la EP y otros parkinsonismos, tanto causativos como de riesgo, confirmados y no confirmados. Para el diseño del panel customizado y la librería se empleó *Haloplex Target Enrichment System*, así como la plataforma NEXTSeq de Illumina para la secuenciación de las muestras. Se realizó un análisis de priorización del efecto deletéreo de las variaciones identificadas mediante el software *Expert Variant Interpreter (eVAI)*, de acuerdo con las guías internacionales del *American College of Medical and Genomics (ACMG)*. Para la predicción del significado clínico de las variaciones identificadas se siguieron varios criterios según el tipo de mutación, la predicción de herramientas bioinformáticas sobre su efecto o la co-ocurrencia de variaciones en el mismo paciente. Las frecuencias alélicas de las variaciones en las principales bases de datos también se tuvieron en cuenta en la selección de variaciones. Las variaciones que fueron clasificadas como patogénicas se validaron mediante secuenciación de Sanger. Además se completó el análisis de variaciones genéticas con un estudio de dosis génica mediante *Multiplex Ligation-dependent Probe Amplification (MLPA)* en los genes: *SNCA, PARK2, PINK1, DJ-1, LRRK2, ATP13A2, GCHI, UCHL1, CAV1, CAV2, LPA* y *TNFRSF9*. Se validaron estos resultados con el análisis de variaciones en el número de copias de los datos procedentes de la secuenciación masiva.

Por otra parte, se estudió de manera aislada un gen recientemente propuesto causativo de EP de herencia autosómica dominante, el gen *CHCHD2*, en una cohorte de pacientes de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla) y controles sanos. Para ello se analizaron todos los exones codificantes y las uniones exón-intrón del gen *CHCHD2* mediante *High Resolution Melting (HRM)*, seguido de la secuenciación directa de aquellas muestras que mostraban curvas de

melting diferentes. Las variaciones identificadas fueron genotipadas en todos los pacientes y controles mediante tecnología TaqMan y SimpleProbe. Se empleó el test χ^2 en el estudio de asociación de las variaciones genéticas con la EP.

Por último, se procedió al análisis del gen *TMEM230*, recientemente propuesto como cuasativo de EP de herencia autómica dominante. Para ello se analizaron todos los exones codificantes y las uniones exón-intrón del gen *TMEM230* mediante HRM, seguido de la secuenciación directa de aquellas muestras que mostraban curvas de *melting* diferentes en una cohorte de pacientes de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla, España) y controles sanos.

Para alcanzar el **objetivo 2**, se genotiparon en una cohorte procedente de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla, España), y en controles sanos, 5 SNPs situados en 5 *loci* de riesgo asociados a la EP: rs6430538 (en el *locus ACMSD/TMEM163*), rs329648 (en el *locus MIR4697*), rs11868035 (en el *locus SREBF/RAII*), rs17649553 (en el *locus MAPT*) y rs1955337 (en el *locus STK39*). Para los primeros 4 se usó la tecnología Taqman y para el último HRM seguido de la secuenciación directa de aquellas muestras que mostraban curvas de *melting* diferentes. Se empleó el test χ^2 en el estudio de asociación de las variaciones genéticas con la EP. Además se realizó un meta-análisis con los resultados conjuntos de la cohorte de Bandrés-Ciga y col. y los de nuestro estudio.

III. RESULTADOS

1. **Objetivo 1: Estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas de EP.**

1.1. Sub-objetivo 1: Estudiar exhaustivamente 17 genes relacionados con la EP y otros parkinsonismos en pacientes de inicio temprano de la Comunidad de Madrid.

1.1.1. Manuscrito en vías de publicación: Mutational landscape of early-onset Parkinson's disease in a Spanish population by targeted exome sequencing. Tejera-Parrado C, Mir P, Periñán MT, Vela-Desojo L, Abreu-Rodríguez I, Alonso-Cánovas A, Bernal-Bernal I, Bonilla-Toribio M, Buiza-Rueda D, Catalán MJ, García-Ramos R, García-Ruiz PJ, Huertas-Fernández I, Jesús S, López-Manzanares L, Martínez-Castrillo JC, Posada IJ, Rojo-Sebastián A, Ruiz-Huete C, Del Val J, Gómez-Garre P.

Mutational landscape of early onset Parkinson's disease in a Spanish population by targeted exome sequencing

Tejera-Parrado Cristina¹, Mir Pablo^{1,2*}, Periñán María Teresa^{1,2}, Vela-Desojo Lydia³, Abreu-Rodríguez Irene¹, Alonso-Cánovas Araceli⁴, Bernal-Bernal Inmaculada¹, Bonilla-Toribio Marta¹, Buiza-Rueda Dolores^{1,2}, Catalán María José⁵, García-Ramos Rocío⁵, García-Ruiz Pedro José⁶, Huertas-Fernández Ismael¹, Jesús Silvia^{1,2}, López-Manzanares Lydia⁷, Martínez-Castrillo Juan Carlos⁴, Posada Ignacio Javier⁸, Rojo-Sebastián Ana⁹, Ruiz-Huete Cristina¹⁰, Del Val Javier⁶, Gómez-Garre Pilar^{1,2*}

¹Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología. Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Seville, Spain

²Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain

³Servicio de Neurología, Hospital Fundación Alcorcón, Madrid, Spain

⁴Servicio de Neurología, Hospital Ramón y Cajal, Madrid, Spain

⁵Servicio de Neurología. Hospital Clínico San Carlos, Madrid, Spain

⁶Servicio de Neurología, Fundación Jiménez Díaz, Madrid, Spain

⁷Servicio de Neurología, Hospital La Princesa, Madrid, Spain

⁸Servicio de Neurología, Hospital Universitario 12 de Octubre, Madrid, Spain

⁹ Servicio de Neurología, Hospital Príncipe de Asturias, Madrid, Spain

¹⁰ Servicio de Neurología, Clínica Nuestra Señora del Rosario, Madrid, Spain

*** Corresponding authors:**

Dr. Pilar Gómez-Garre and Dr. Pablo Mir

Unidad de Trastornos del Movimiento

Instituto de Biomedicina de Sevilla (IBiS)

Hospital Universitario Virgen del Rocío

Avda. Manuel Siurot s/n. 41013 Sevilla, Spain

Fax: +34-955923101

Telephone: +34-955923039

email: mgomez-ibis@us.es / pmir@us.es/

Short Title: Mutational landscape of early onset Parkinson's disease in a Spanish population

Keywords: Parkinson's disease, targeted sequencing, MLPA, Clinical Genetics, Spanish population, Genetic variants.

ABSTRACT

Introduction: Several genes have been described as responsible for both recessive and dominant forms of Parkinson's disease (PD). The aim of this study was to assess the contribution of several genes to the pathophysiology of early onset PD in a cohort from central Spain.

Methods: We analyzed a cohort of 117 patients with early onset PD (age at onset < 50 years old) using a pipeline, based on a combination of a next-generation sequencing panel of 17 genes previously related with PD and other Parkinsonisms (*PARK7*, *ATP13A2*, *PINK1*, *HTRA2*, *DNAJC6*, *GIGYF2*, *EIF4G1*, *UCHL1*, *SNCA*, *PARK2*, *SMPD1*, *LRRK2*, *VPS35*, *SNYJ1*, *FBXO7*, *PLA2G6* and *GBA*), and MLPA.

Results: Twenty-seven patients (23.08%) carried putative pathogenic variants in *PARK2*, *LRRK2*, *PINK1*, or *GBA*. The gene most frequently mutated was *PARK2*, and p.Asn52Metfs was the most common variation in this gene. Pathogenic variants were not observed in genes *SNCA*, *FBXO7*, *PARK7*, *HTRA2*, *DNAJC6*, *PLA2G6*, and *UCHL1*. No significant differences relating to age at onset were observed between patients with pathogenic variants in *LRRK2* in comparison with those with pathogenic variants in *PARK2*. Co-occurrence of mutations involving two genes suggests additive effects of *ATP13A2* and *PARK2* genes, as well as of *LRRK2* and *GIGYF2* genes.

Conclusions: Our results contribute to the understanding of the genetic architecture associated with early onset PD, showing different associated backgrounds. However, a large proportion of genetic components remains unknown. All in all, this study might contribute to genetic diagnosis and counseling for families with early onset PD.

INTRODUCTION

PD is the second most common neurodegenerative disorder characterized by the loss of dopaminergic neurons in the *substantia nigra*. It is a chronic and progressive disorder of multifactorial etiology, in which causative and susceptibility genetic factors are involved. Emerging evidence has provided support for the hypothesis that PD is the result of complex interactions among genetic abnormalities, environmental toxins, mitochondrial dysfunction, and other cellular processes.

Since Polymeropoulos et al. identified a causative variant in gene encoding Synuclein alpha (*SNCA*) [1], many efforts have been made to identify genes involved in the development of PD. Mutations in *SNCA*, *LRRK2*, and Vacuolar Protein Sorting 35 (*VPS35*) genes have been linked to autosomal dominant forms of PD (ADPD) [2, 3]. In fact, the mutations in *LRRK2* are the most common cause of ADPD. In addition, the Eukaryotic Translation Initiation Factor 4 Gamma 1 (*EIF4G1*), was initially related to ADPD, but subsequent studies have failed to replicate this association [4]. Furthermore, genes related to autosomal recessive forms of PD (ARPD) have also been discovered. Thus, the genes coding for Parkin (*PARK2*), PTEN-induced kinase 1 (*PINK1*), and the Protein deglycase *DJ1* (*PARK7*) have been related to typical early-onset PD (EOPD; age at onset <50 years old). Other genes, such as those coding for the ATPase Cation Transporting 13A2 (*ATP13A2*), the Phospholipase A2 Group VI (*PLA2G6*), the F-Box Protein 7 (*FBXO7*), HSP40 Auxilin (*DNAJC6*) and the Synaptojanin-1 (*SNYJ1*), have been related to atypical PD with juvenile onset (age at onset <35 years old) [4]. On the other hand, the gene *GIGYF2* (encoding Grb10-Interacting GYF Protein 2) has been described as responsible for typical autosomal dominant PD [5]. Its pathogenic

contribution to PD is still not clear but it has been suggested that some variations are risk factors for PD in Caucasians [6].

Nevertheless, monogenic mutations are not a very common cause of PD. Indeed, risk variants with a moderate effect size are more usual.

Variations in genes such as *GBA* (encoding glucocerebrosidase) or *SMPD1* (encoding sphingomyelin phosphodiesterase 1) constitute risk factors for PD. In fact, mutations in *GBA* are the most important risk factor for idiopathic PD [7].

To date, genome wide association studies (GWAS) have allowed identification of several PD risk loci [8]. However, a large proportion of genetic heritability remains unknown. Nowadays, next generation sequencing (NGS) of whole genome (WGS) or just the exome (WES) are expected to contribute to elucidating the missing heritability through the identification of new risk variants related to PD.

Since genetic background has a higher impact on PD with early disease onset, the aim of this study was to assess the contribution of the main seven genes previously related to PD, and unconfirmed genes, to the pathophysiology of the EOPD in a cohort of patients from central Spain. For that purpose a combination of NGS-based targeted sequencing and Multiplex Ligation-Dependent Probe Amplification (MLPA) were applied.

PATIENTS AND METHODS

Subjects and Clinical Assessments

We included 117 EOPD (age at onset younger than 50 years old) patients. Among them, thirty-three reported a family history of PD. The demographic characteristics of the participants are summarized in Table 1. Patients were recruited from the

Neurology outpatient clinic at different hospitals in the Comunidad Autónoma de Madrid (Spain). PD was diagnosed by Movement Disorders neurologists according to the United Kingdom Parkinson's Disease Society Brain Bank criteria [9]. Patients were clinically evaluated and an extensive set of clinical features was obtained. After clinical diagnosis, peripheral blood samples were collected from each subject.

Ethics Statements

The study was approved by the ethics committees from all participating centers and it was conducted according to the principles expressed in the Helsinki Declaration. Each individual who participated in the study signed a written informed consent form prior to blood withdrawal.

Genetic analysis

DNA isolation

Genomic DNA was isolated from peripheral blood samples from each subject according to established protocols, by manual and automated commercial procedures (Roche Applied Science, Indianapolis, IN, USA). The quantity and purity of the DNA were determined by Qubit 3.0 fluorometer (Invitrogen) and NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Florida, USA), respectively.

Targeted gene enrichment and next-generation sequencing (NGS)

Targeted re-sequencing was performed using a customized Haloplex Target Enrichment Panel, which was designed using Agilent's online Sure Design tool, following the manufacturer's protocol (Agilent Technologies, Inc. Santa Clara, CA). This customized panel covered all coding exons, exon-intron boundary regions, and 3' and 5' untranslated regions of 17 selected genes. These genes were those related to dominant and recessive forms of PD, as well as risk genes, and unconfirmed genes:

SNCA, LRRK2, PINK1, PARK2, ATP13A2, FBXO7, VPS35, DJ1, PLA2G6, EIF4G1, DNAJC6, HTRA2, SMPD1, SNYJ1, UCHL1, GIGYF2, and GBA.

Sample preparation was carried out according to the manufacturer's protocol. The concentration of the enriched and amplified samples was determined using a Bioanalyzer High Sensitivity chip (Agilent Technologies). Then, samples were pooled in equimolar amounts and sequenced at the Illumina NextSeq platform (Illumina Inc., San Diego, CA, USA).

Bioinformatics analysis for NGS results

Data analysis and prioritization of variants were performed using the workflow summarized in Suppl. Fig. 1.

Briefly, sequenced reads quality were examined by FastQC (Babraham Bioinformatics, Cambridge, UK) and poor quality bases were trimmed by Cutadapt v1.9.1. The reference genome was downloaded from NCBI, version GRCh37/hg19. The Burrows-Wheeler Aligner v0.7.12 was used to align the remaining clean reads. All variants, single nucleotide polymorphisms (SNP) and small indels, were detected using SAMtools v1.2. Variants were filtered by Bcftools v1.2 and GATK v4.0, and only those with depth read over 30 and quality over 20 remained. Subsequently, variants were annotated using SeattleSeq Annotation 138 (<http://snp.gs.washington.edu/SeattleSeqAnnotation138/>) and wANNOVAR (<http://wannovar.wglab.org/index.php>). The variants where the fraction of non-reference reads was <0.3 were rejected. In addition, a scoring and prioritization analysis of deleterious alleles were performed using the software Expert Variant Interpreter (eVAI), from enGenome (www.engenome.com/product) and variants were categorized

according to the international guidelines of American College of Medical Genetics and Genomics (ACMG) [10].

For the selection of candidate variants we followed the combination of several criteria (Suppl. Methods). Missense variants were analyzed with a total of 13 bioinformatic tools. Furthermore, all variants were analyzed using the Batch version of the annotation software package Alamut (Interactive Biosoftware, Rouen, France) for splicing predictions. A threshold was set for each tool above, in which the variant was classified as damaging (Suppl. Table 1). The impact was interpreted in relation to all known isoforms of each gene. To restrict the analysis, allele frequencies on population databases were also taken into account.

Sanger resequencing

Filtered variants predicted as pathogenic were validated by Sanger sequencing. Exons containing each selected variant were amplified using standard PCR protocol. Amplicons were sequenced on both strands using the BigDye terminator cycle sequencing kit (Applied Biosystem, Foster City, CA, USA) and, subsequently, they were resolved on an ABI3500 genetic analyzer and analyzed by software Variant Reporter v1.1 (Applied Biosystem).

Copy Number Variation (CVN) Analysis

Multiplex ligation-Dependent Probe Amplification (MLPA) was carried out using two commercially available SALSA MLPA kits (P051 and P052; MRC-Holland, Amsterdam, The Netherlands), following the manufacturer's recommendations. Data were analyzed with the Coffalyser.Net software (MRC-Holland).

In addition, a reanalysis of NGS data for CNV identification was performed with VisCap software. Losses were defined by a maximum log₂ ratio of -0.55 and a minimum log₂ ratio of 0.4 for gains [11].

RESULTS

In this study, a total of 224 variants (categorized as VOUS, likely pathogenic or pathogenic according to the ACMG) were detected by bioinformatics analysis of NGS data, with 82 coding variants and 142 non-coding variants (Fig. 1). Subsequently, all of them were further characterized using several algorithms, following the aforementioned analysis, to analyze their putative functional impact (Suppl. Fig.1).

Identification of the most interesting variants

Firstly, 32 variants, that included the nonsense variants and those missense variants, were kept. These variants were predicted as likely pathogenic at 54% (7 out of 13) of the bioinformatics tools (Suppl. Table 2). However, due to prediction tools, in general, have a tendency of over-predicting deleterious impacts, only those variants with >69% *in silico* predictions in favor of pathogenicity, and/or described as pathogenic in ClinVar were considered as very interesting variants (Suppl. Table 3).

Three variants, all of them in *PARK2*, were directly assigned as loss-of-function because they disrupt the protein: two of them were nonsense variants, and one was a deletion of one nucleotide resulting in a frameshift (Suppl. Table 3).

In addition, four different deletions affecting to entire exons (all of the gene *PARK2*) were identified: three single-exon deletions and one affecting four consecutive exons (Suppl. Table 3).

Insertions or deletions of a few bases were found, but these together with those variations located in the untranslated regions were considered for further analysis.

Thirteen variants were predicted bioinformatically to have a putative effect on splicing (Suppl. Table 4). Of them, four variants (rs560897844, rs112019125, rs148944108, and rs41286476) were predicted as likely pathogenic by all tools with high scores, and they presented allele frequencies higher than those in databases (Suppl. Table 4).

Two intronic variants had a score of 1f (those that are known eQTLs for genes and have TF binding or a DNase peak) in the RegulomeDB scoring system: rs6437074 (in *GIGYF2*) and rs2298298 (in *PINK1*). However, their allele frequencies in our population were very high and similar to those previously described in databases, thus they were excluded.

Lastly, we kept only those variants that were in a heterozygous state in dominant genes, and in homozygous or compound heterozygous states for recessive ones.

In this way, we identified 34 of the 117 studied patients (29.06%) to be carrying potential pathogenic variants across 8 genes (Fig. 2). Three patients carried potential pathogenic variations in *ATP13A2*, 2 patients in *EIF4G1*, 1 in *VPS35*, 8 in *LRRK2*, 16 in *PARK2*, and 1 in *PINK1*. In addition, 6 patients carried putatively pathogenic variations in the risk factors: 3 in *SMPD1*, 2 in *GBA*, and 1 in *GIGYF2* (Table 2).

Mutational findings in PD-associated genes

Twenty-seven patients (23.08%) carried putative pathogenic variants in *PARK2*, *LRRK2*, *PINK1*, or *GBA*.

PARK2 was the gene most frequently mutated (Fig. 2; Table 2). Therefore, 16 patients (13.68%) presented pathogenic variants in *PARK2*, 8 of them carried deletions of entire exons of *PARK2*. The most often mutation in *PARK2* was a previously described

frameshift variation (p.Asn52Metfs). To the best of our knowledge, all carriers for this variation are unrelated patients.

Eight patients (6.84%) carried previously described pathogenic variations in *LRRK2*. Of them, 5 patients (4.27%) carried the mutation p.G2019S, 1 patient (0.86%) carried the mutation p.R1441C, 1 patient (0.86%) carried the mutation p.T1410M. Lastly, 1 patient (0.86%) carried the intronic variation c.4536+3A>G.

No significant difference ($P=0.205$) was shown between age at onset of patients with pathogenic variants in *LRRK2* (age at onset 40.1 ± 5.6 years old) and those with pathogenic variants in *PARK2* (age at onset 36.7 ± 5.7 years old).

We found one patient carrying a homozygous mutation in *PINK1* (p.Leu347Pro) and another patient with a heterozygous mutation in *VPS35* (p.Ala654Ser). Therefore, mutations in these genes are a rare cause of disease in our population.

Finally, two patients carried two non-previously described putative pathogenic variants in *GBA* (p.Ile347Thr and p.Ser403Asn).

No pathogenic variant was observed in *SNCA*, and *PARK7*, illustrating their rarity in our population.

Mutational findings in other genes

A total of 8 variants in *ATP13A2* (p.Gly217Ser, p.Arg745His; and c.2762+21T>C), *EIF4G1* (p.Gly693Cys and p.Met1362Thr), *GIGYF2* (p.Arg1225Cys), and *SMPD1* (p.Val147Val and p.Arg378Cys), were found as putatively contributing to PD (Table 2).

Pathogenic variants were not observed in *FBXO7*, *DNAJC6*, *PLA2G6*, *HTRA2*, and *UCHL1*.

DISCUSSION

In this study, a pipeline, based on a combination of a next-generation sequencing panel of 17 genes and MLPA, has been applied to detect PD-related variations in a population with EOPD from Central Spain. Twenty-seven patients out of 117 (23.08%) showed putative pathogenic variants in known PD-associated genes.

Pathogenic variants in the *PARK2* gene were the most common genetic cause of early-onset ARPD, with a variable prevalence in previous studies in EOPD [12-14]. Therefore, *PARK2* variants were the most frequently identified explanation for PD in our cohort also (13.68%; 16 patients out of 117); with the most frequent mutation (6.4%; 15/234) being a previously described frameshift variation (p.Asn52Metfs) [15]. This result is different in comparison to a study with PD patients from northern Spain, which described this variation, in homozygous, in one patient from a population of 72 EOPD patients (1.4%; 2/144) [16]. However, in another previous study, that variation was found in 2 patients (in a homozygous and heterozygous state) from a 26 EOPD patient cohort (5.8%; 3/52) from southern Spain [17], similar to what it was obtained in our population. Those results suggest that, although p.Asn52Metfs is geographically widespread, it is less prevalent in the Basque Country and it could be associated with a founder effect and subsequent dispersion of the mutation.

Interestingly, one patient (EP-17) to be carrying the pathogenic variation p.Asn52Metfs in *PARK2* in a homozygous state, also carried a variation in *ATP13A2* (c.2762+21T>C), which was predicted by the bioinformatic tools as affecting splicing due the apparition of a stronger cryptic 5' donor site. This patient had an age at onset that was younger than the other carriers of the same variation p.Asn52Metfs, suggesting a cumulative effect between both genes on age at onset. The *ATP13A2* gene has been found mutated in some types of early-onset Parkinsonism [18]. Although the pathogenicity of

single heterozygous mutations is a matter of debate, the reduction of functional ATP13A2 has been related to the development of PD as a risk factor. Moreover, it has been suggested that dopaminergic neurons expressing higher ATP13A2 protein levels are less susceptible to cell death [19]. Given the rarity of *ATP13A2* mutations, confirmation of this point will require sequencing of much larger sample sizes.

On the other hand, to date, mutations in *LRRK2* (a multi-domain protein with both GTPase and kinase functionality) are the most common genetic determinant of ADPD, with at least 100 genetic variants linked to PD. The most frequent mutation described is p.G2019S, although its prevalence varies considerably [17, 20, 21]. We found this mutation in 5 patients, which represents the 4.27% of the 117 studied samples. This frequency is consistent with the prevalence in Caucasians. In total, in this study 6.84% of the studied patients carried pathogenic variants in *LRRK2*, without differences relating to age at onset with those presenting mutations in *PARK2*. Therefore, our results provide a very important role for *LRRK2* in typical EOPD in our Spanish population.

Mutations in the *GIGYF2* gene have been related to familial PD [5], but some studies in diverse populations showed that there is no correlation between the presence of variants in this gene with the disease [22-24]. In this analysis, we found a patient (EP-33) with co-occurrence of c.4536+3A>G (predicted affecting splicing) in *LRRK2* and p.Arg1225Cys in *GIGYF2* pathogenic variants. This patient presented cognitive impairment, as previously described in a Spanish family featuring late-onset PD and cognitive impairment and with a genetic variant, in *GIGYF2*, identified as potential disease-causing mutation [25]. It is interesting to note that both genes (*LRRK2* and *GIGYF2*) have been described as playing some roles in the autophagy process [26]. As

such, this kind of co-occurrence involving both genes has been previously described and associated with an earlier age at onset of PD, suggesting additive effects [5]. Further genetic studies and functional analysis are necessary to conclude the implications in PD of variants located in *GIGYF2*.

Among risk factors in PD, *GBA* is well known [27]. Mutations in this gene are present in multiple ethnicities [28]. Heterozygous mutations are dominantly inherited, and the penetrance is high, related to EOPD [29]. Importantly, our *GBA* mutational spectrum did not include the most frequent pathogenic variants L444P and N370S [30]. It might be due to difficulties in *GBA* sequencing using targeted next generation sequencing because the existence of the pseudo-*GBA*, with very similar sequence. In addition, it has been shown that *GBA* variants have a different impact on the PD phenotype according to their pathogenicity (both deleterious and benign). Because of preselecting for pathogenic, likely pathogenic or VOUS variants, it may very well be an underestimate of the true yield of variants in *GBA* affecting Parkinson's disease in our population.

This study focused on exonic and intronic variants; however, most of the variants found were located in the untranslated regions (mainly in 3'UTR). Interestingly, heterozygous pathogenic mutations in recessive genes were detected in some patients that also presented variants in untranslated regions. Therefore, it is important to perform further analysis in order to study their contribution to disease pathogenesis.

In conclusion, this study, in terms of cohort size, number of included genes and applied methods is the first systematic study of genetic variability in PD-related genes in EOPD patients from Spain. Our results substantially expand on previous investigations, and provide a comprehensive genomics profile of EOPD. In addition, further evidence for

gene interaction involving *ATP13A2* and *PARK2* are provided in an individual with co-occurrence of mutations in both genes, who had a much earlier onset of PD than patients with single-gene mutations. Perhaps, most importantly, our study suggests a greater role for *LRRK2* in typical, idiopathic EOPD than previously believed in the Spanish population. Thus, our results could weight utility in personalized genetic counselling.

Contributorship statement

- Cristina Tejera-Parrado: Analysis and interpretation of data; writing of the first draft, review and critique of the manuscript.
- Pablo Mir: Conception and design of the study; review and critique of the manuscript.
- María Teresa Perriñán: Analysis and interpretation of data; review and critique of the manuscript.
- Lydia Vela-Desojo: Acquisition of data; review and critique of the manuscript.
- Irene Abreu-Rodríguez: Analysis and interpretation of data; review and critique of the manuscript.
- Araceli Alonso-Cánovas: Acquisition of data; review and critique of the manuscript.
- Inmaculada Bernal-Bernal: Acquisition of data; review and critique of the manuscript.
- Marta Bonilla-Toribio: Acquisition of data; review and critique of the manuscript.
- Dolores Buiza-Rueda: Acquisition of data; review and critique of the manuscript.
- José Catalán María: Acquisition of data; review and critique of the manuscript.
- Rocío García-Ramos: Acquisition of data; review and critique of the manuscript.
- Pedro José García-Ruiz: Acquisition of data; review and critique of the manuscript.
- Ismael Huertas-Fernández: Analysis of data; review and critique of the manuscript.

- Silvia Jesús: Acquisition of data; review and critique of the manuscript.
- Lydia López-Manzanares: Acquisition of data; review and critique of the manuscript.
- Juan Carlos Martínez-Castrillo: Acquisition of data; review and critique of the manuscript.
- Javier Posada Ignacio: Acquisition of data; review and critique of the manuscript.
- Ana Rojo-Sebastián: Acquisition of data; review and critique of the manuscript.
- Cristina Ruiz-Huete: Acquisition of data; review and critique of the manuscript.
- Javier Del Val: Acquisition of data; review and critique of the manuscript.
- Pilar Gómez-Garre: Conception and design of the study; analysis and interpretation of data; review and critique of the manuscript.

All the listed authors gave their final approval of the final version of the manuscript.

Acknowledgments:

We would like to thank patients who participated in this study. We also thank Raquel Gómez (Multiple Uses Laboratory), and Juan Antonio Cordero (Bioinformatic and Computational Biology Service), both of them from the Instituto de Biomedicina de Sevilla (IBiS), for their technical support, and the *HUVR-IBiS Biobank (Andalusian Public Health System Biobank and ISCIII-Red de Biobancos PT13/0010/0056)* for the human specimens used in this study.

Financial Disclosure/Conflict of Interest concerning the research related to the manuscript:

This study was supported by grants from the Spanish Ministry of Economy and Competitiveness [PI14/01823, PI16/01575, PI18/01898] co-founded by ISCIII (Subdirección General de Evaluación y Fomento de la Investigación) and by Fondo Europeo de Desarrollo Regional (FEDER), the Consejería de Economía, Innovación, Ciencia y Empleo de la Junta de Andalucía [CVI-02526, CTS-7685], the Consejería de Salud y Bienestar Social de la Junta de Andalucía [PI-0437-2012, PI-0471-2013], the Sociedad Andaluza de Neurología, the Fundación Alicia Koplowitz, the Fundación Mutua Madrileña. Pilar Gómez-Garre was supported by the "Miguel Servet" (from ISCIII-FEDER) and "Nicolás Monardes" (from Andalusian Ministry of Health) programs. Silvia Jesús Maestre was supported by the "Juan Rodés" program (from ISCIII-FEDER). Cristina Tejera was supported by VPPI-US from the Universidad de Sevilla.

Conflict of interests:

The authors declare no conflict of interest.

REFERENCES

- [1] M.H. Polymeropoulos, C. Lavedan, E. Leroy, S.E. Ide, A. Dehejia, A. Dutra, B. Pike, H. Root, J. Rubenstein, R. Boyer, E.S. Stenroos, Mutation in the alpha-Synuclein Gene Identified in Families with Parkinson's Disease, *Science* 276(5321) (1997) 2045-7.
- [2] C. Paisan-Ruiz, S. Jain, E.W. Evans, W.P. Gilks, J. Simon, M. van der Brug, A. Lopez de Munain, S. Aparicio, A.M. Gil, N. Khan, J. Johnson, J.R. Martinez, D. Nicholl, I.M. Carrera, A.S. Pena, R. de Silva, A. Lees, J.F. Marti-Masso, J. Perez-Tur, N.W. Wood, A.B. Singleton, Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease, *Neuron* 44(4) (2004) 595-600.
- [3] A. Zimprich, A. Benet-Pages, W. Struhal, E. Graf, S.H. Eck, M.N. Offman, D. Haubenberger, S. Spielberger, E.C. Schulte, P. Lichtner, S.C. Rossle, N. Klopp, E. Wolf, K. Seppi, W. Pirker, S. Presslauer, B. Mollenhauer, R. Katzenschlager, T. Foki, C. Hotzy, E. Reinthaler, A. Harutyunyan, R. Kralovics, A. Peters, F. Zimprich, T. Brucke, W. Poewe, E. Auff, C. Trenkwalder, B. Rost, G. Ransmayr, J. Winkelmann, T. Meitinger, T.M. Strom, A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease, *American journal of human genetics* 89(1) (2011) 168-75.
- [4] V. Bonifati, Genetics of Parkinson's disease – state of the art, 2013, *Parkinsonism & Related Disorders* 20 (2014) S23-S28.
- [5] C. Lautier, S. Goldwurm, A. Durr, B. Giovannone, W.G. Tsiaras, G. Pezzoli, A. Brice, R.J. Smith, Mutations in the GIGYF2 (TNRC15) gene at the PARK11 locus in familial Parkinson disease, *Am J Hum Genet* 82(4) (2008) 822-33.
- [6] Y. Zhang, Q.Y. Sun, R.H. Yu, J.F. Guo, B.S. Tang, X.X. Yan, The contribution of GIGYF2 to Parkinson's disease: a meta-analysis, *Neurol Sci* 36(11) (2015) 2073-9.
- [7] J. Trinh, M. Farrer, Advances in the genetics of Parkinson disease, *Nat Rev Neurol* 9(8) (2013) 445-54.

- [8] D. Chang, M.A. Nalls, I.B. Hallgrimsdottir, J. Hunkapiller, M. van der Brug, F. Cai, C. International Parkinson's Disease Genomics, T. andMe Research, G.A. Kerchner, G. Ayalon, B. Bingol, M. Sheng, D. Hinds, T.W. Behrens, A.B. Singleton, T.R. Bhangale, R.R. Graham, A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci, *Nat Genet* 49(10) (2017) 1511-1516.
- [9] W. Gibb, A. Lee, The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease, *Journal of Neurology, Neurosurgery, and Psychiatry* 51(6) (1988) 745-752.
- [10] S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W.W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H.L. Rehm, A.L.Q.A. Committee, Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology, *Genetics in medicine : official journal of the American College of Medical Genetics* 17(5) (2015) 405-24.
- [11] T.J. Pugh, S.S. Amr, M.J. Bowser, S. Gowrisankar, E. Hynes, L.M. Mahanta, H.L. Rehm, B. Funke, M.S. Lebo, VisCap: inference and visualization of germ-line copy-number variants from targeted clinical sequencing data, *Genet Med* 18(7) (2016) 712-9.
- [12] M.G. Macedo, D. Verbaan, Y. Fang, S.M. van Rooden, M. Visser, B. Anar, A. Uras, J.L. Groen, P. Rizzu, J.J. van Hilten, P. Heutink, Genotypic and phenotypic characteristics of Dutch patients with early onset Parkinson's disease, *Movement Disorders* 24(2) (2009) 196-203.
- [13] D. Kozirowski, D. Hoffman-Zacharska, J. Sławek, Z. Jamrozik, P. Janik, A. Potulska-Chromik, A. Roszmann, R. Tataj, J. Bal, A. Friedman, Incidence of mutations in the PARK2, PINK1, PARK7 genes in Polish early-onset Parkinson disease patients, *Neurologia i Neurochirurgia Polska* 47(4) (2013) 319-324.
- [14] R.N. Alcalay, E. Caccappolo, H. Mejia-Santana, M.X. Tang, L. Rosado, B.M. Ross, M. Verbitsky, S. Kisselev, E.D. Louis, C. Comella, A. Colcher, D. Jennings, M.A. Nance, S.B. Bressman, W.K. Scott, C. Tanner, S. Mickel, H. Andrews, C. Waters, S. Fahn, L. Cote, S. Frucht,

- B. Ford, M. Rezak, K. Novak, J.H. Friedman, R. Pfeiffer, L. Marsh, B. Hiner, A. Siderowf, R. Ottman, K. Marder, L.N. Clark, Frequency of known mutations in early-onset Parkinson disease: implication for genetic counseling: the consortium on risk for early onset Parkinson disease study, *Archives of neurology* 67(9) (2010) 1116-22.
- [15] N. Spataro, A. Roca-Umbert, L. Cervera-Carles, M. Valles, R. Anglada, J. Pagonabarraga, B. Pascual-Sedano, A. Campolongo, J. Kulisevsky, F. Casals, J. Clarimon, E. Bosch, Detection of genomic rearrangements from targeted resequencing data in Parkinson's disease patients, *Mov Disord* 32(1) (2017) 165-169.
- [16] A. Gorostidi, J.F. Marti-Masso, A. Bergareche, M.C. Rodriguez-Oroz, A. Lopez de Munain, J. Ruiz-Martinez, Genetic Mutation Analysis of Parkinson's Disease Patients Using Multigene Next-Generation Sequencing Panels, *Mol Diagn Ther* 20(5) (2016) 481-91.
- [17] S. Bandres-Ciga, N.E. Mencacci, R. Duran, F.J. Barrero, F. Escamilla-Sevilla, S. Morgan, J. Hehir, F. Vives, J. Hardy, A.M. Pittman, Analysis of the genetic variability in Parkinson's disease from Southern Spain, *Neurobiology of aging* 37 (2016) 210 e1-5.
- [18] A. Ramirez, A. Heimbach, J. Grundemann, B. Stiller, D. Hampshire, L.P. Cid, I. Goebel, A.F. Mubaidin, A.L. Wriekat, J. Roeper, A. Al-Din, A.M. Hillmer, M. Karsak, B. Liss, C.G. Woods, M.I. Behrens, C. Kubisch, Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase, *Nat Genet* 38(10) (2006) 1184-91.
- [19] D. Ysselstein, J.M. Shulman, D. Krainc, Emerging links between pediatric lysosomal storage diseases and adult parkinsonism, *Mov Disord* (2019).
- [20] C.M. Lill, Genetics of Parkinson's disease, *Molecular and cellular probes* 30(6) (2016) 386-396.
- [21] D.G. Hernandez, X. Reed, A.B. Singleton, Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance, *J Neurochem* 139 Suppl 1 (2016) 59-74.

- [22] J. Bras, J. Simon-Sanchez, M. Federoff, A. Morgado, C. Januario, M. Ribeiro, L. Cunha, C. Oliveira, A.B. Singleton, Lack of replication of association between GIGYF2 variants and Parkinson disease, *Human molecular genetics* 18(2) (2009) 341-6.
- [23] I. Guella, A. Pistocchi, R. Asselta, V. Rimoldi, A. Ghilardi, F. Sironi, L. Trotta, P. Primignani, M. Zini, A. Zecchinelli, D. Coviello, G. Pezzoli, L. Del Giacco, S. Duga, S. Goldwurm, Mutational screening and zebrafish functional analysis of GIGYF2 as a Parkinson-disease gene, *Neurobiology of aging* 32(11) (2011) 1994-2005.
- [24] M. Bonetti, A. Ferraris, M. Petracca, A.R. Bentivoglio, B. Dallapiccola, E.M. Valente, GIGYF2 variants are not associated with Parkinson's disease in Italy, *Movement Disorders* 24(12) (2009) 1867-1868.
- [25] J. Ruiz-Martinez, C.E. Krebs, V. Makarov, A. Gorostidi, J.F. Marti-Masso, C. Paisan-Ruiz, GIGYF2 mutation in late-onset Parkinson's disease with cognitive impairment, *J Hum Genet* 60(10) (2015) 637-40.
- [26] M. Kim, I. Semple, B. Kim, A. Kiers, S. Nam, H.W. Park, H. Park, S.H. Ro, J.S. Kim, G. Juhasz, J.H. Lee, *Drosophila* Gyf/GRB10 interacting GYF protein is an autophagy regulator that controls neuron and muscle homeostasis, *Autophagy* 11(8) (2015) 1358-72.
- [27] S. Jesus, I. Huertas, I. Bernal-Bernal, M. Bonilla-Toribio, M.T. Caceres-Redondo, L. Vargas-Gonzalez, M. Gomez-Llamas, F. Carrillo, E. Calderon, M. Carballo, P. Gomez-Garre, P. Mir, GBA Variants Influence Motor and Non-Motor Features of Parkinson's Disease, *PloS one* 11(12) (2016) e0167749.
- [28] E. Sidransky, M.A. Nalls, J.O. Aasly, J. Aharon-Peretz, G. Annesi, E.R. Barbosa, A. Bar-Shira, D. Berg, J. Bras, A. Brice, C.M. Chen, L.N. Clark, C. Condroyer, E.V. De Marco, A. Durr, M.J. Eblan, S. Fahn, M.J. Farrer, H.C. Fung, Z. Gan-Or, T. Gasser, R. Gershoni-Baruch, N. Giladi, A. Griffith, T. Gurevich, C. Januario, P. Kropp, A.E. Lang, G.J. Lee-Chen, S. Lesage, K. Marder, I.F. Mata, A. Mirelman, J. Mitsui, I. Mizuta, G. Nicoletti, C. Oliveira, R. Ottman, A. Orr-Urtreger, L.V. Pereira, A. Quattrone, E. Rogaeva, A. Rolfs, H. Rosenbaum, R. Rozenberg, A. Samii, T.

Samaddar, C. Schulte, M. Sharma, A. Singleton, M. Spitz, E.K. Tan, N. Tayebi, T. Toda, A.R.
Troiano, S. Tsuji, M. Wittstock, T.G. Wolfsberg, Y.R. Wu, C.P. Zabetian, Y. Zhao, S.G. Ziegler,
Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease, *The New England
journal of medicine* 361(17) (2009) 1651-61.

[29] M. Ferreira, J. Massano, An updated review of Parkinson's disease genetics and
clinicopathological correlations, *Acta neurologica Scandinavica* (2016).

[30] F. Zhao, L. Bi, W. Wang, X. Wu, Y. Li, F. Gong, S. Lu, F. Feng, Z. Qian, C. Hu, Y. Wu, Y. Sun,
Mutations of glucocerebrosidase gene and susceptibility to Parkinson's disease: An updated
meta-analysis in a European population, *Neuroscience* 320 (2016) 239-46.

Figure legends

Fig. 1. Number and distribution of detected variants in our population using a 17-gene sequencing panel. Variants in all 17 genes were evaluated and only those described *a priori* as a variant of unknown significance (VOUS), likely pathogenic or pathogenic, according to ACMG, are shown. Genes with variants and the variant distribution are displayed (A), and the proportions of coding and non-coding variants are shown for all genes (B).

Fig. 2. Pie plot showing the distribution and frequencies of putatively pathogenic variants found. In our population, we identified 34 patients carrying putative pathogenic variants (A), of which more than a half presented variations in PARK2 (41.02%) or LRRK2 (20.51%) (B). RF: Risk factor. SF: Susceptibility factor. N: number of subjects. n: number of carriers subjects with pathogenic variations in each gene.

Suppl. Fig. 1. Flowchart of sequencing data analysis. Horizontal boxes represent steps in the workflow. *In silico* analysis were carried out in 224 SNVs. SIFT: Sorting Intolerant From Tolerant. Polyphen-2: Polymorphism Phenotyping 2. GERP: Genomic Evolutionary Rate Profiling. CADD: Combined Annotation Dependent Depletion. RadialSVM: Radial Support Vector Machine. LR: Logistic Regression. LRT: Likelihood Ratio Test. FATHMM: Functional Analysis Through Hidden Markov Models. GWAVA: Genome Wide Annotation of Variants. ReMM: Regulatory Mendelian Mutation. IW-score: Integrative Weighted score (with an associated p-value<0.5).

FIGURE 1

A)

Gene	in 5' UTR	coding variants				Intronic (+/- 30 bp)	in 3' UTR	Total
		Synonymous	Nonsense	Missense	Ins / Del			
<i>ATP13A2</i>	1	3	0	4	0	3	3	14
<i>DNAJC6</i>	0	2	0	3	0	2	5	12
<i>EIF4G1</i>	0	7	0	3	0	8	2	20
<i>FBXO7</i>	1	2	0	0	0	5	0	8
<i>GBA</i>	0	1	0	0	0	4	1	6
<i>GIGYF2</i>	0	3	0	5	3	8	8	27
<i>HTRA2</i>	0	1	0	0	0	1	0	2
<i>LRRK2</i>	0	2	0	4	0	5	4	15
<i>PARK2</i>	0	2	2	4	1	5	18	32
<i>PARK7 (DJ1)</i>	0	0	0	0	0	1	1	2
<i>PINK1</i>	0	1	0	1	0	1	4	7
<i>PLA2G6</i>	2	2	0	4	0	2	3	13
<i>SMPD1</i>	1	0	0	3	3	0	0	7
<i>SNCA</i>	3	0	0	0	0	1	9	13
<i>SYNJ1</i>	0	1	0	6	0	6	3	16
<i>UCHL1</i>	0	1	0	6	0	6	3	16
<i>VPS35</i>	0	1	0	1	0	2	10	14
Total	8	29	2	44	7	60	74	224

B)

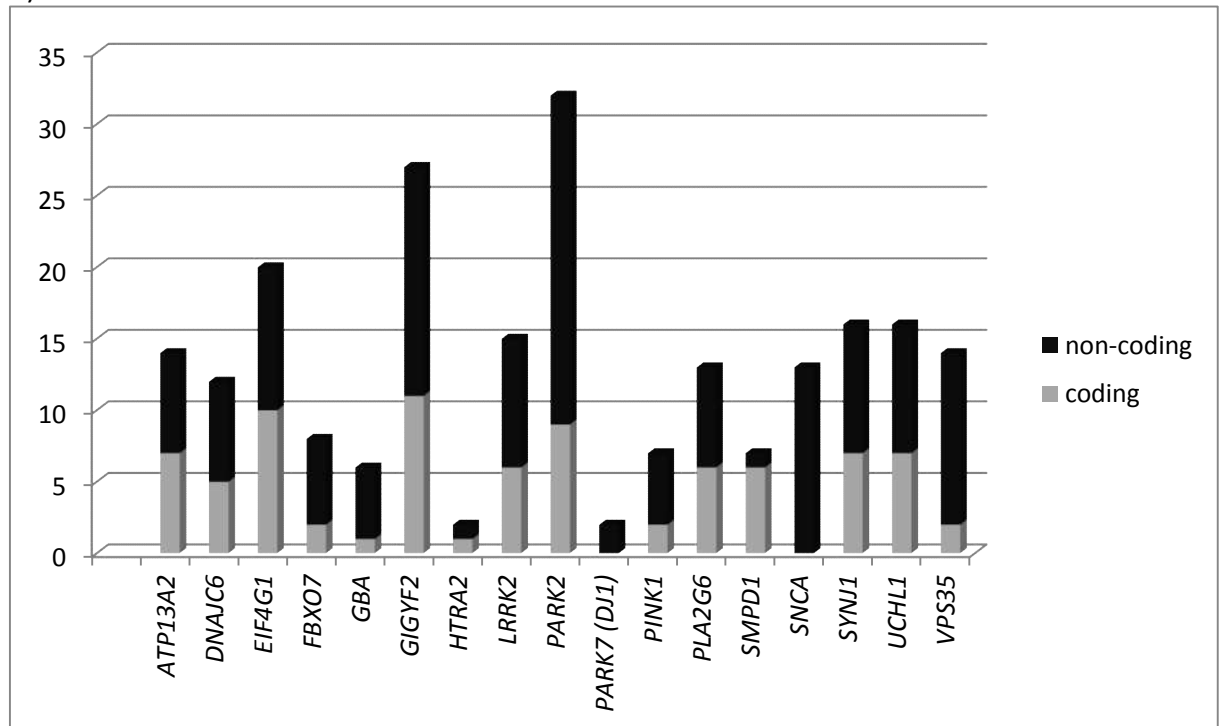


FIGURE 2

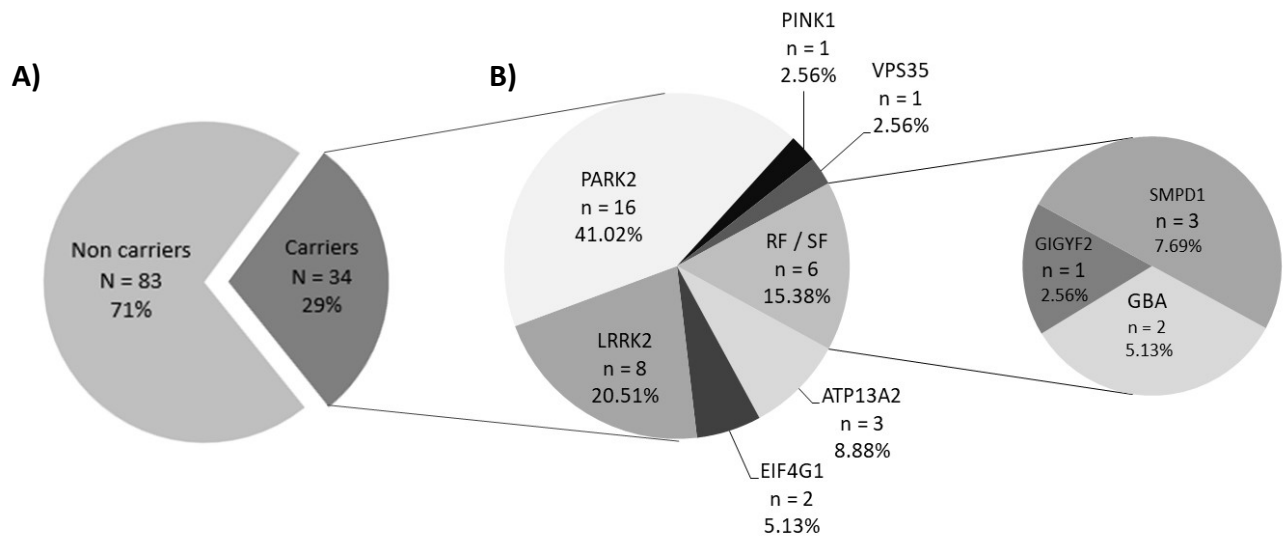


Table 1. Demographic characteristics of PD patients

		Sex (M/F)	Age (y)	AO (y)	N
EOPD	Total	64/53	51.7 ± 9	40.2 ± 5.3	117
	Familial PD	17/16	52.2 ± 9.2	39.2 ± 5.5	33

EOPD: early-onset Parkinson's disease (age at onset <50 years old); PD: Parkinson's disease; M: males; F: females; y: years; AO: age at onset; N: number of samples.

Table 2. Patients carrying putative pathogenic variants. Genotype-phenotype correlations.

Patient	Gene	Variation (HGVS)		Zigosity	VAF	Coverage (Alt/Ref)	Sex	Age at onset	Family history	Dopaminergic complications (AAOS)			Non-motor symptoms		
		CDS	AA							MF	DK	N-MF	Depression	CI	VH
EP-1	<i>ATP13A2</i>	c.649G>A	p.Gly217Ser	Het	0.467	353 (165/188)	F	48	No	No	No	No	No	No	No
EP-2	<i>ATP13A2</i>	c.2234G>A	p.Arg745His	Het	0.466	562 (262/300)	F	45	Yes	Yes (8)	No	No	No	No	No
EP-3	<i>EIF4G1</i>	c.2077G>T	p.Gly693Cys	Het	0.425	858 (365/493)	M	43	No	Yes (3)	Yes (3)	-	No	No	No
EP-4	<i>EIF4G1</i>	c.4085T>C	p.Met1362Thr	Het	0.457	1216 (556/660)	M	45	No	Yes (6)	Yes (8)	Yes (6)	Yes (10)	Yes (24)	Yes (20)
EP-5	<i>SMPD1</i>	c.441G>A	p.Val147Val	Het	0.469	1453 (681/772)	M	45	No	Yes (4)	Yes (5)	No	Yes	No	No
EP-6	<i>SMPD1</i>	c.441G>A	p.Val147Val	Het	0.497	1264 (628/636)	M	49	No	Yes (4)	Yes (1)	Yes	No	No	No
EP-7	<i>VPS35</i>	c.1960G>T	p.Ala654Ser	Het	0.323	563 (182/381)	F	45	-	Yes	Yes	Yes	yes	No	Yes
EP-8	<i>GBA</i>	c.1040T>C	p.Ile347Thr	Het	0.490	1117 (547/570)	M	36	Yes	No	No	No	No	No	No
EP-9	<i>GBA</i>	c.1208G>A	p.Ser403Asn	Het	0.462	764 (353/411)	M	40	Yes	Yes (3)	Yes (4)	Yes (2)	No	No	No
EP-10	<i>PINK1</i>	c.1040T>C	p.Leu347Pro	Hom	0.981	1776 (1743/33)	M	25	No	No	Yes	No	No	No	No
EP-11	<i>PARK2</i>	c.79A>T	p.Lys27Stop	Hom	0.997	567 (565/2)	F	41	No	No	No	No	No	No	No
EP-12	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	0.985	1092 (1076/16)	M	35	Yes	Yes	No	No	No	No	No
EP-13	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	1	450 (450/0)	F	32	No	Yes	Yes	No	No	No	No
EP-14	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	0.985	391 (385/6)	F	37	-	Yes	Yes	Yes	No	Yes	No
EP-15	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	0.983	591 (581/10)	F	31	No	No	No	No	Yes	Yes	No
EP-16	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	0.998	875 (873/2)	F	42	No	Yes	No	No	No	No	No
EP-17	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Hom	0.997	387 (386/1)	F	24	Yes	No	No	No	Yes	No	No
	<i>ATP13A2</i>	c.2762+21T>C	-	Het	0.531	443 (235/208)									
EP-18	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Het	0.336	935 (314/621)	F	37	No	Yes	Yes	Yes	Yes	Yes	Yes
	<i>PARK2</i>	Del ex 7	-	Het	-	-									
EP-19	<i>PARK2</i>	c.155delA	p.Asn52Metfs	Het	0.318	1407 (447/960)	F	37	No	Yes	Yes	Yes	Yes	No	No
	<i>PARK2</i>	Del ex 7	-	Het	-	-									

Patient	Gene	Variation (HGVS)		Zigosity	VAF	Coverage (Alt/Ref)	Sex	Age at onset	Family history	Dopaminergic complications (AAOS)			Non-motor symptoms		
		CDS	AA							MF	DK	N-MF	Depression	CI	VH
EP-20	PARK2	c.155delA	p.Asn52Metfs	Het	0.348	811 (282/529)	F	28	No	Yes	Yes	Yes	No	No	No
	PARK2	c.719C>T	p.Thr240Met	Het	0.413	1532 (632/900)									
EP-21	PARK2	c.766C>T	p.Arg256Cys	Het	0.525	1959 (1029/930)	M	44	No	No	No	No	No	No	No
	PARK2	Del ex 7	-	Het	-	-									
EP-22	PARK2	c.1205G>A	p.Arg402His	Het	0.512	1256 (643/613)	M	40	No	Yes (12)	Yes (13)	Yes (10)	No	No	No
	PARK2	Del ex 7	-	Het	-	-									
EP-23	PARK2	Del ex 3	-	Hom	-	-	F	33	Yes	Yes	Yes	Yes	No	No	
EP-24	PARK2	Del ex 3-4-5-6	-	Hom	-	-	M	34	Yes	Yes	No	No	No	No	
EP-25	PARK2	Del ex 3-4-5-6	-	Hom	-	-	F	47	Yes	Yes (9)	Yes (7)	-	-	-	
EP-26	PARK2	c.1334G>A	p.Trp445Stop	Het	0.520	321 (167/154)	F	34	Yes	No	No	No	No	Yes	No
	PARK2	Del ex 3-4-5-6	-	Het	-	-									
EP-27	LRRK2	c.6055G>A	p.Gly2019Ser	Het	0.442	981 (434/547)	M	29	No	Yes (3)	Yes (4)	No	No	No	No
EP-28	LRRK2	c.6055G>A	p.Gly2019Ser	Het	0.393	1011 (397/614)	M	40	Yes	Yes (12)	No	No	No	No	No
EP-29	LRRK2	c.6055G>A	p.Gly2019Ser	Het	0.442	1025 (453/572)	F	39	No	Yes (4)	Yes (5)	-	No	No	No
EP-30	LRRK2	c.6055G>A	p.Gly2019Ser	Het	0.369	1154 (426/728)	F	42	No	Yes (1)	Yes (1)	-	No	No	No
EP-31	LRRK2	c.6055G>A	p.Gly2019Ser	Het	0.419	1046 (438/608)	F	40	No	Yes (2)	Yes (2)	Yes (2)	No	No	No
EP-32	LRRK2	c.4321C>T	p.Arg1441Cys	Het	0.524	708 (371/337)	M	45	No	Yes (4)	Yes (5)	Yes (4)	No	No	No
EP-33	LRRK2	c.4536+3A>G	-	Het	0.460803	523 (241/282)	M	45	Yes	Yes	Yes	-	No	Yes	No
	GIGYF2	c.3673C>T	p.Arg1225Cys	Het	0.500606	825 (413/412)									
EP-34	LRRK2	c.4229C>T	p.Thr1410Met	Het	0.477394	752 (359/393)	F	46	No	No	No	No	Yes	Yes	No
	SMPD1	c.1132C>T	p.Arg378Cys	Het	0.500808	1857 (930/927)									

HGVS: Human Genome Variation Society (Variants have been described using HGVS nomenclature). CDS: Coding sequence. AA: Amino acid. Het: Heterozygous. Hom: Homozygous. VAF: Variant allele frequency. Alt: Alternate allele. Ref: Reference allele. M: Male. F: Female. MF: Motor fluctuations. NMF: Non-motor fluctuations. DK: Dyskinesia. AAOS: Age at onset of symptom. CI: Cognitive impairment. VH: Visual hallucinations.

SUPPLEMENTARY METHODS

Prediction of the significance of variants

For the prediction of the significance of variants we followed the combination of several criteria:

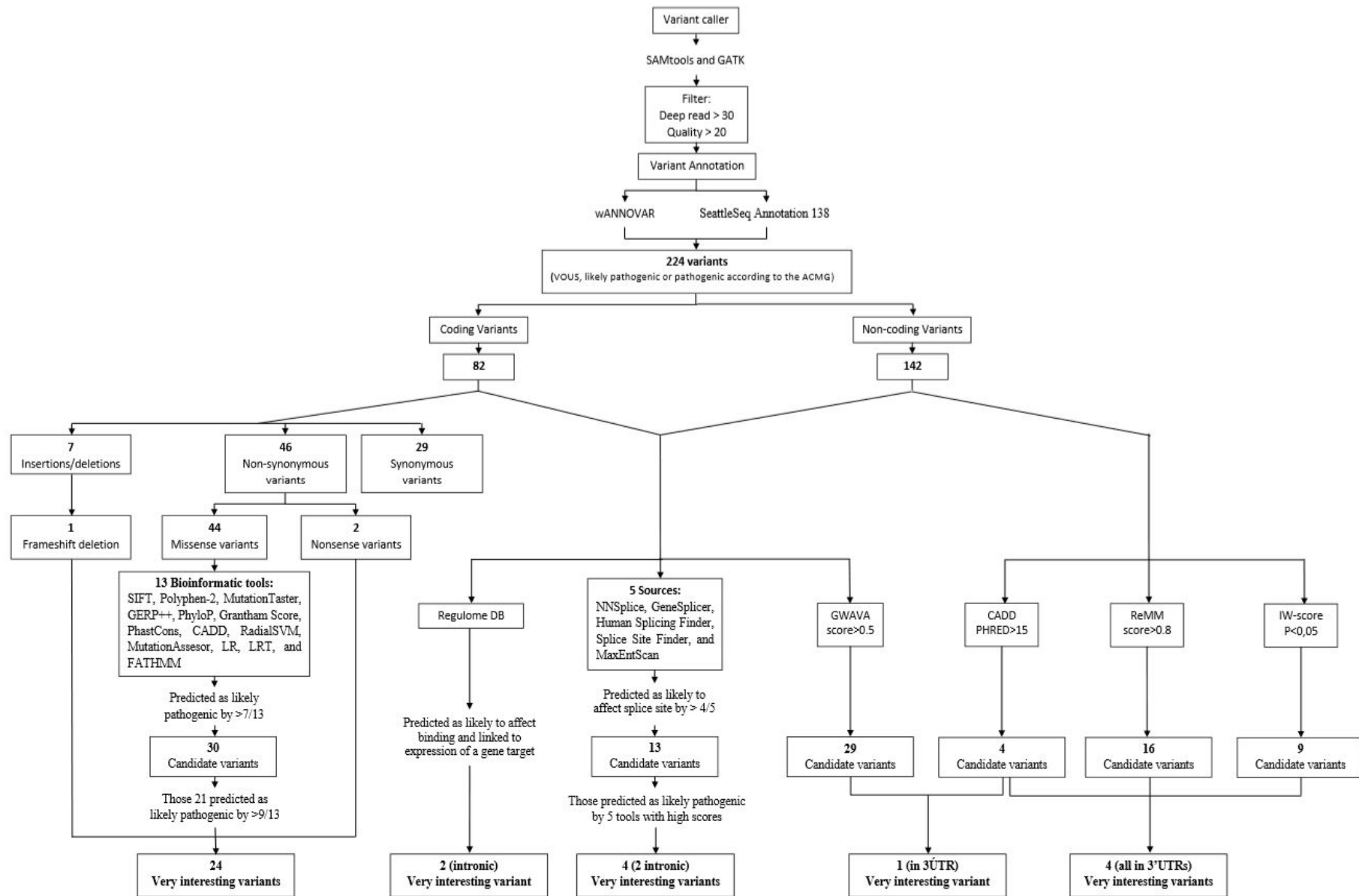
1) Type of the variation: whether variation results in a null variant: that is nonsense, frameshift or affecting to canonical splice sites that result in frameshift of the coding sequence, mutations affecting the initiation codon, and those leading to exon(s) deletion.

2) In-silico prediction: We used multiple bioinformatics tools and the impact was interpreted with respect to all known isoforms of each gene. For missense variants, a total of thirteen tools were examined. Therefore, we used 6 functional prediction scores (SIFT, Polyphen-2, MutationTaster, MutationAssesor, LRT, and FATHMM), 4 conservation scores (GERP++, PhyloP, PhastCons, and Grantham Score), and three ensemble-based prediction scores (CADD, LR, and RadialSVM). A threshold was set for each tool above in which the variant was classified as damaging (Suppl. Table 1). In addition, all variants were analyzed using GWAVA, CADD, and the splicing predictions from the Batch version of the annotation software package Alamut (Interactive Biosoftware, Rouen, France), that integrates 5 sources: Splice Site Finder, MaxEntScan, NNSplice, GeneSplicer and Human Splicing Finder (Suppl. Table 1). We also considered the dbcsnv11 database for predicting the splicing impact by Ada Boost and Random Forest (Suppl. Table 1). Moreover, non-coding variants were prioritized when ReMM-score >0.8 and IW-score with an associated P -value <0.05 .

3) Score from Regulome-DB of 1.

4) Co-occurrence with pathogenic variants.

SUPPLEMENTARY FIGURE 1



Supplementary Table 1. Summary of deleteriousness prediction methods used in our study.

Name	Category	Threshold
SIFT	Function prediction	< 0.05
Polyphen-2	Function prediction	> 0.85
Mutation Taster	Function prediction	Disease causing o Disease causing automatic
PhyloP	Conservation score	> 0.7
CADD	Ensembl score	> 10
Grantham Score	Conservation score	> 100
PhastCons	Conservation score	> 0.7
GERP++	Conservation score	> 3
LRT	Function prediction	Deleterious
Mutation Assessor	Function prediction	High o Medium
LR	Ensembl score	Deleterious
FATHMM	Function prediction	Deleterious
RadialSVM	Ensembl score	Deleterious
SSF	Effect on splicing	> 70
MaxEnt	Effect on splicing	> 0
NNSPLICE	Effect on splicing	> 0.4
GeneSplicer	Effect on splicing	> 0
HSF	Effect on splicing	> 65
dbscSNV_RF_Score	Effect on splicing	< 0.6
dbscSNV_RF_Score	Effect on splicing	< 0.6

Supplementary Table 2. Description of preselected candidate pathogenic variants after applying bioinformatics filters. Nonsense variants and those missense variants which were predicted as likely pathogenic by at least 53'9% (seven of thirteen) of the bioinformatics tools.

Gene	Chr: position	Amino acid change	Nucleotide change	Zygoty	SIFT		PolyPhen-2		MutationTaster		PhyloP	CADD	Grantham Score	Phast Cons	GERP++	MutationAssessor		LRT	FATHMM	RadialSVM	LR	ACMG	ClinVar	N
					Score	Prediction	Score	Prediction	Score	Prediction						Score	Prediction							
<i>ATPI3A2</i>	1:17328585	p.G217S	c.G649A	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	5,577	22,8	56	0,956	4,89	3,41	Medium	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>ATPI3A2</i>	1:17318246	p.R745H	c.G2234A	Hetero	0,03	Damaging	1	Probably damaging	1	Disease causing	7,265	21,1	29	1	4,93	2,315	Medium	Deleterious	Tolerated	Deleterious	Deleterious	VOUS	ND	1
<i>DNAJC6</i>	1:65831737	p.S77F	c.C230T	Hetero	0,09	Tolerated	1	Probably damaging	1	Disease causing	7,247	29,1	155	1	5,51	2,62	Medium	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>DNAJC6</i>	1:65854097	p.K341E	c.A1021G	Hetero	0	Damaging	0,992	Probably damaging	1	Disease causing	6,548	20,4	56	1	4,9	2,44	Medium	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>DNAJC6</i>	1:65877062	p.V865M	c.G2593A	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	9,657	29,6	21	1	5,03	3,03	Medium	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	1
<i>EIF4G1</i>	3:184040997	p.G693C	c.G2077T	Hetero	0,05	Damaging	0,999	Probably damaging	1	Disease causing	0,872	18,54	159	1	5,16	0,895	Low	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	1
<i>EIF4G1</i>	3:184045661	p.L1282H	c.T3845A	Hetero	0,56	Tolerated	1	Probably damaging	1	Disease causing	7,698	31	99	1	6,17	1,74	Low	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	2
<i>EIF4G1</i>	3:184046529	p.M1362T	c.T4085C	Hetero	0	Damaging	0,998	Probably damaging	1	Disease causing	7,698	18,64	81	1	4,54	2,48	Medium	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	1
<i>GBA</i>	1:155207997	p.V230G	c.T689G	Hetero	0,25	Tolerated	0,011	Benign	1	Disease causing	5,309	4,868	109	0,987	2,5	0,77	Non-functional	Neutral	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>GBA</i>	1:155206220	p.I347T	c.T1040C	Hetero	0,11	Tolerated	0,937	Possibly damaging	1	Disease causing	8,274	17,27	89	0,984	3,67	1,365	Low	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>GBA</i>	1:155206167	p.E365K	c.G1093A	Hetero	0,43	Tolerated	0,03	Benign	0	Disease causing	0,838	11,06	56	0,845	3,67	1,01	Low	Neutral	Deleterious	Deleterious	Deleterious	VOUS	P / RF	1
<i>GBA</i>	1:155206052	p.S403N	c.G1208A	Hetero	0,1	Tolerated	0,996	Probably damaging	0,554	Polymorphism	5,361	20,3	46	1	3,57	2,045	Medium	Deleterious	Deleterious	Deleterious	Deleterious	LP	P	1
<i>GIGYF2</i>	2:233712207	p.R1225C	c.C3673T	Hetero	0	Damaging	0,946	Possibly damaging	1	Disease causing	1,725	17,35	180	0,997	5,4	0,625	Non-functional	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	1
<i>LRRK2</i>	12:40704236	p.R1441C	c.C4321T	Hetero	0,14	Tolerated	1	Probably damaging	1	Disease causing	3,654	23,5	180	0,082	4,74	1,97	Medium	Deleterious	Tolerated	Deleterious	Deleterious	P	P	1
<i>LRRK2</i>	12:40713845	p.R1628H	c.G4883A	Hetero	0,22	Tolerated	1	Probably damaging	1	Disease causing	8,853	33	29	1	5,54	1,43	Low	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	VOUS	2
<i>LRRK2</i>	12:40734202	p.G2019S	c.G6055A	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	9,527	36	56	0,966	5,69	3,765	High	Deleterious	Deleterious	Deleterious	Deleterious	LP	P	5
<i>PARK2</i>	6:162864434	p.K27X	c.A79T	Homo	1	Tolerated	.	.	1	Disease causing	6,231	38	NA	1	5,63	.	.	Deleterious	.	.	.	LP	ND	1
<i>PARK2</i>	6:162864388	p.R70P	c.G209A	Hetero	0,15	Tolerated	0,726	Possibly damaging	1	Disease causing	0,727	21,2	103,29	1	4,74	3,47	Medium	Deleterious	Tolerated	Deleterious	Deleterious	VOUS	VOUS	1
<i>PARK2</i>	6:162394349	p.T240M	c.C635T	Hetero	0,05	Damaging	0,991	Probably damaging	0,983	Polymorphism	6,594	16,31	81	0,299	5,14	1,78	Low	Neutral	Deleterious	Deleterious	Deleterious	LP	P	1
<i>PARK2</i>	6:162206909	p.R256C	c.C766T	Hetero	0,02	Damaging	1	Probably damaging	0,999	Disease causing	8,289	26,9	180	0,06	5,75	2,3	Medium	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>PARK2</i>	6:161781200	p.R402H	c.G1205A	Hetero	0,01	Damaging	1	Probably damaging	1	Disease causing	5,469	21,2	29	0,686	4,89	2,56	Medium	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>PARK2</i>	6:161771195	p.W445X	c.G1334A	Hetero	1	Tolerated	.	.	1	Disease causing	5,476	34	NA	1	5,27	.	.	Deleterious	.	.	.	P	ND	1
<i>PINK1</i>	1:20972133	p.L347P	c.T1040C	Homo	0	Damaging	1	Probably damaging	1	Disease causing	6,959	15,71	98	0,775	4,9	4,13	High	Deleterious	Tolerated	Deleterious	Deleterious	LP	P	1
<i>PLA2G6</i>	22:38508553	p.R745Q	c.G2234A	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	6,995	28,7	43	1	4,54	2,64	Medium	Deleterious	Tolerated	Deleterious	Deleterious	VOUS	ND	1
<i>PLA2G6</i>	22:38509628	p.V690I	c.G2068A	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	4	17,29	29	0,962	4,36	1,85	Low	Deleterious	Tolerated	Deleterious	Deleterious	VOUS	VOUS	1
<i>SMPD1</i>	11:6414486	p.R377C	c.C1129T	Hetero	0	Damaging	1	Probably damaging	1	Disease causing	4,64	22,4	180	1	5,41	3,785	High	Deleterious	Deleterious	Deleterious	Deleterious	VOUS	ND	1
<i>SMPD1</i>	11:6415566	p.R541Q	c.G1622A	Hetero	0,14	Tolerated	0,981	Probably damaging	0,75	Disease causing	1,982	13,8	43	0,919	2,94	2,255	Medium	Neutral	Deleterious	Deleterious	Tolerated	VOUS	VOUS	1
<i>SYNJ1</i>	21:34060625	p.A320V	c.C959T	Hetero	0,02	Damaging	1	Probably damaging	1	Disease causing	9,061	36	64	1	6,16	2,11	Medium	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	VOUS	1
<i>SYNJ1</i>	21:34003786	p.G1453E	c.G4358A	Hetero	0,01	Damaging	0,989	Probably damaging	1	Disease causing	4,989	22,9	NA	0,988	5,13	1,845	Low	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	B	4
<i>SYNJ1</i>	21:34060664	p.R307H	c.G920A	Hetero	0,64	Tolerated	1	Probably damaging	1	Disease causing	7,193	18,67	29	1	6,16	1,145	Low	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	1
<i>VPS35</i>	16:46696262	p.A654S	c.G1960T	Hetero	0,08	Tolerated	0,999	Probably damaging	1	Disease causing	7,818	34	99	1	5,49	2,87	Medium	Deleterious	Tolerated	Deleterious	Deleterious	VOUS	ND	1
<i>VPS35</i>	16:46717460	p.Q21L	c.A62T	Hetero	0,32	Tolerated	0	Benign	1	Disease causing	7,972	15,92	113	1	5,42	0,64	Non-functional	Deleterious	Tolerated	Tolerated	Tolerated	VOUS	ND	2

Supplementary Table 3. Putative pathogenic variants found in our PD-population from central Spain

Gene (MI)	Transcript	Exon (E) / Intron (I)	Nucleotide change	Amino acid change	dbSNP ID	Carriers (n=117)		maf			Previously Reported	Effect
						Het	Hom	Our study	GO-ESP ^{&}	ExAc*		
ATP13A2 (AD)	NM_022089.3	E8	c.649G>A	p.Gly217Ser	rs199961048	1	0	4.27e-3	ND	1.62e-05	Yes	MS
		E20	c.2234G>A	p.Arg745His	rs140455242	1	0	4.27e-3	1.16e-04	1.5e-05	Yes	MS
		I24	c.2762+21T>C	-	rs560897844	1	0	4.27e-3	ND	1.5e-05	Yes	SR
DNAJC6 (AR)	NM_001256864	E4	c.401C>T	p.Ser134Phe	-	1	0	4.27e-3	ND	ND	No	MS
		E9	c.1192A>G	p.Lys398Glu	-	1	0	4.27e-3	ND	ND	No	MS
		E18	c.2764G>A	p.Val922Met	-	1	0	4.27e-3	ND	ND	No	MS
EIF4G1 (AD)	NM_001194946	E15	c.2077G>T	p.Gly693Cys	rs112019125	1	0	4.27e-3	ND	9e-05	Yes	MS/SR
		E28	c.4085T>C	p.Met1362Thr	rs144059151	1	0	4.27e-3	8.14e-04	5.8e-04	Yes	MS
GBA (AD-RF)	NM_001005741.2	E9	c.1040T>C	p.Ile347Thr	-	1	0	4.27e-3	ND	ND	No	MS
		E9	c.1208G>A	p.Ser403Asn	-	1	0	4.27e-3	ND	ND	No	MS
GIGYF2 (AD-S)	NM_001103147	E29	c.3673C>T	p.Arg1225Cys	rs149440697	1	0	4.27e-3	1.16e-04	4.47e-05	Yes	MS
LRRK2 (AD)	NM_198578.3	E30	c.4229C>T	p.Thr1410Met	rs72546327	1	0	4.27e-3	0	1.2e-04	Yes	MS
		E31	c.4321C>T	p.Arg1441Cys	rs33939927	1	0	4.27e-3	ND	1.5e-05	Yes	MS
		I31	c.4536+3A>G	-	rs41286476	1	0	4.27e-3	5.81e-04	4.38e-04	Yes	SR
		E41	c.6055G>A	p.Gly2019Ser	rs34637584	5	0	0.022	5.8e-04	6.29e-04	Yes	MS
PARK2 (AR)	NM_004562.2	2	c.79A>T	p.Lys27Stop	-	0	1	8.55e-3	ND	ND	No	SG
		2	c.125G>C	p.Arg42Pro	rs368134308	1	0	4.27e-3	1.16e-04	4.5e-05	Yes	MS
		2	c.155delA	p.Asn52Metfs	rs754809877	3	7	0.073	6.06e-04	3.15e-04	Yes	FS
		6	c.719C>T	p.Thr240Met	rs137853054	1	0	4.27e-3	2.33e-04	1.35e-04	Yes	MS
		7	c.766C>T	p.Arg256Cys	rs150562946	1	0	4.27e-3	4.65e-04	5.14e-04	Yes	MS
		11	c.1205G>A	p.Arg402His	rs766915327	1	0	4.27e-3	ND	0	Yes	MS
		12	c.1334G>A	p.Trp445Stop	-	1	0	4.27e-3	ND	ND	No	SG
		3	Del	Del	-	0	1	8.55e-3	-	-	-	DEE
		4	Del	Del	-	1	0	4.27e-3	-	-	-	DEE
		3, 4, 5, 6	Del	Del	-	1	2	0.021	-	-	-	DEE
7	Del	Del	-	4	0	0.017	-	-	-	DEE		
PINK1 (AR)	NM_032409.2	5	c.1040T>C	p.Leu347Pro	rs28940285	0	1	8.55e-3	ND	0	Yes	MS
PLA2G6 (AR)	NM_003560.2	E15	c.2068G>A	p.Val690Ile	rs141777179	1	0	4.27e-3	8.14e-04	9.17e-04	Yes	MS
		E16	c.2234G>A	p.R745Q	rs368514303	1	0	4.27e-3	1.2e-04	0	Yes	MS

Gene (MI)	Transcript	Exon (E) / Intron (I)	Nucleotide change	Amino acid change	dbSNP ID	Carriers (n=117)		maf			Previously Reported	Effect
						Het	Hom	Our study	GO-ESP ^{&}	ExAc [*]		
SMPD1 (RF)	NM_000543.4	E2	c.441G>A	p.V147V	rs148944108	2	0	8.55e-3	0	1.67e-04	Yes	SN / SR
		E3	c.1132C>T	p.R378C	rs369088417	1	0	4.27e-3	0	0	Yes	MS
SYNJ1 (AR)	NM_003895.3	E7	c.959C>T	p.A320V	-	1	0	4.27e-3	ND	ND	No	MS
VPS35 (AD)	NM_018206.5	E15	c.1960G>T	p.A654S	-	1	0	4.27e-3	No	ND	No	MS

MI: Mode of Inheritance; AD: Autosomal dominant; AR: Autosomal recessive; S: Susceptibility; RF: Risk factor; N: Number of carriers; Het: Heterozygosity; Hom: Homozygosity; SG: Stop gain; MS: Missense; FS: Frameshift; DEE: Deletion of the entire exon; SR: Splice region; SN: Synonymous; MAF: Minor allele frequency; &: European Americans; *: Non-Finnish European.

Supplementary Table 4. Variants putatively affecting splicing

Gene	Intron / Exon	Chr position	Nucleotide change	Strand	Splice Site	Distance to Splice Site	ID	N	Frequency		SSF		MaxEnt		NNSplice		GeneSplicer		HSF		dnCSNV scores		Putative consequence
									1000G EUR	Affected	Score	variation	Score	variation	Score	variation	Score	variation	Score	variation	ada	rf	
ATP13A2	E10	1: 17326767	c.881G>A	-	5'	27	rs56367069*	5	0.018	0.0214	79.69	(+12.7%)	6.46	(+55.9%)	0.79	appear	6.76	(+142%)	87.86	(+5.9%)	-	-	appear a putative new 5' donor site in c.882
	I24	1: 17314796	c.2762+21T>C	-	5'	19	rs560897844*	1	0.0002 (all)	0.0043	81.66	(+5.3%)	8.48	(+25.7%)	0.54	appear	5.18	(+146.3%)	92.05	(+2.2%)	-	-	stronger cryptic 5' donor site
	I26	1: 17313454	c.3084-3C>T	-	3'	1	rs7531163	68	0.25	0.325	76.52	(-7.5%)	9.90	(+6.5%)	0.96	(+4.4%)	13.19	(+14.3%)	84.49	(-8.3%)	0.0011	0.1	weaker 3' acceptor site
	I27	1: 17313283	c.3235+17G>A	-	5'	15	rs56146840	5	0.019	0.0214	70.14	appear	3.78	appear	-	-	1.96	appear	83.47	appear	-	-	appear a new 3' acceptor site in c.3235+18
DNAJC6	E9	1: 65854097	c.1192A>G	+	5'	2	-	1	-	0.0043	78.90	(-10.2)	10.15	(+0.7%)	0.99	(-0.2%)	4.94	(+21.9%)	84.56	(-5.4%)	0.9934	0.96	weaker 5' donor site in c.1193
EIF4G1	I11	3: 184040606	c.1796-3C>T	+	3'	1	rs190378563	1	0.001	0.0043	87.71	(-6.6%)	11.84	(-12.84%)	0.99	(+0.2%)	10.07	(-12.9%)	87.91	(-8.0%)	0.0018	0.196	weaker 3' acceptor site
	E13	3: 184040997	c.2056G>T	+	5'	33	rs112019125*	1	-	0.0043	80.14	appear	5.28	appear	0.92	appear	2.52	appear	87.15	appear	-	-	appear a putative new 5' donor site in c.2054
	I29	3: 184049399	c.4398+5G>A	+	5'	3	rs16858676	1	0.014 (all)	0.0043	74.71	(-14.0%)	6.04	(-37.4%)	0.88	(-9.3%)	6.17	(-37.3%)	77.73	(-13.5%)	0.9556	0.656	weaker 5' donor site
GIGYF2	I4	2: 233568203	c.-44+4A>G	+	5'	2	rs746723892	1	-	0.0043	89.92	(-10.1%)	10.28	(-5.3%)	0.96	(-3.4%)	5.73%	(-16.2%)	91.66	(-8.3%)	0.3339	0.44	weaker 5' donor site
LRRK2	I31	12: 40704454	c.4536+3A>G	+	5'	1	rs41286476	1	-	0.0043	77.31	(-5.3%)	5.19	(-41.8%)	desappear	-	desappear	-	86.21	(-11.4%)	0.9712	0.86	weaker 5' donor site
SMPD1	E2	11: 6412736	c.441G>A	+	3'	120	rs148944108*	2	0.002	0.0086	81.75	(+5.6%)	8.98	(+100.6%)	0.75	appear	4.85	appear	87.50	(+1.3%)	-	-	appear a new 5' donor site in c.438
SYNJ1	I1	21: 34100124	c.-23+24G>A	-	5'	22	rs755252765	2	-	0.0086	70.57	appear	5.10	(+79.3%)	-	-	4.73	(+60.9%)	82.29	(+1.0%)	-	-	appear a putative new 5' donor site (referred to isoforms 3 and 4)
UCHL1	I2	4: 41259149	c.45+6T>C	+	5'	4	rs11556273	21	0.09	0.098	77.42	(-6.9%)	5.89	(-29.9%)	0.90	(-8.3%)	4.22	(-34.4%)	85.13	(-2.5%)	0.8166	0.724	weaker 5' donor site

Chr: chromosome; E: Exon; I: Intron; N: number of carriers; SSF: Splicing Sequences Finder; MaxEnt: Maximum Entropy Scan; NNSplice: Splice site prediction by Neural Network; HSF: Human Splicing Finder; ADA: adaptive boosting, RF: random forest; *: Predicted as likely pathogenic by all 5 tools with high scores

1.1.2. Resumen

Desde que se describió la primera mutación causativa de EP, se han realizado numerosos estudios con el fin de dilucidar la complejidad genética que se esconde tras esta enfermedad. Genes de herencia autosómica dominante como *SNCA*, *LRRK2* o *VPS35* se sabe que son causa de EP, aunque se han descrito otros como *EIF4G1*, *GIGYF2* o *UCHL1* cuya implicación continúa bajo estudio. También se han identificado genes causativos de EP de herencia autosómica recesiva como *PARK2*, *PINK1*, *DJ-1*, *ATP13A2*, *PLA2G6*, *FBXO7*, *DNAJC6* y *SYNJ*. Sin embargo, las mutaciones monogénicas no son la causa más común de la EP. Lo más frecuente en la población son variaciones genéticas comunes de riesgo moderado. El gen *GBA* es el factor de riesgo descrito más importante de EP idiopático, y otros genes como *SMPD1* o *GIGYF2* se han descrito también de riesgo para la EP, aunque su implicación continúa bajo estudio. A pesar de todos los factores genéticos descritos hasta la fecha, la mayor proporción de la heredabilidad del Parkinson sigue siendo desconocida.

En los pacientes de EPIT hay una mayor contribución de los factores genéticos en el desarrollo de la enfermedad, por lo que el objetivo de este trabajo fue estudiar genes descritos causativos y de riesgo relacionados con la EP, en una cohorte de 117 pacientes de EPIT procedentes de distintos hospitales de la Comunidad Autónoma de Madrid (España). Como la tecnología de secuenciación masiva se espera que contribuya a identificar nuevas variaciones genéticas, se realizó la secuenciación dirigida de 17

genes causativos y de riesgo para la EP, y de genes no confirmados (*SNCA*, *LRRK2*, *PINK1*, *PARK2*, *ATP13A2*, *FBXO7*, *VPS35*, *DJI*, *PLA2G6*, *EIF4G1*, *DNAJC6*, *HTRA2*, *SMPD1*, *SNYJ1*, *UCHL1*, *GIGYF2* y *GBA*).

Además se llevó a cabo un estudio de dosis génica mediante MLPA.

En total se detectaron 224 variaciones categorizadas como probablemente patogénicas, patogénicas o con significado incierto (VOUS) de acuerdo con las guías internacionales del ACMG (Figura 1). De éstas, 82 estaban en regiones codificantes y 142 en regiones no codificantes.

Se consideraron variaciones con potencial significado clínico aquellas predichas como patogénicas por el 69% de las herramientas bioinformáticas, así como las descritas como patogénicas en ClinVar (Tabla suplementaria 3).

En el gen *PARK2*, 3 variaciones se consideraron directamente de pérdida de función ya que introducían un codón de *stop* o provocaba una pauta abierta de lectura de la proteína. Además, se detectaron 4 deleciones en este gen: 3 deleciones de un solo exón y 1 afectando a 4 exones consecutivos (Tabla suplementaria 3).

Otras 13 variaciones afectaban al sitio de *splicing* según las herramientas bioinformáticas (Tabla suplementaria 4), y de ellas 4 (rs560897844, rs112019125, rs148944108 y rs41286476) fueron clasificadas como posiblemente patogénicas por todas las herramientas de predicción bioinformáticas empleadas y presentaban frecuencias alélicas mayores que las descritas en las principales bases de datos, considerándolas candidatas a ser deletéreas (Tabla suplementaria 4).

Finalmente, se escogieron aquellas variaciones que se encontraban en heterocigosis en genes dominantes y en homocigosis o en heterocigosis

compuesta en los recesivos. Así, 34 de los 117 pacientes (29,06%) tenían alguna variación potencialmente patogénica en 8 genes (Figura 2; Tabla 2).

27 pacientes (23,08%) tenían variaciones potencialmente patogénicas para la EP en *PARK2*, *LRRK2*, *PINK1* o *GBA*, genes bien descritos relacionados con la EP. Concretamente, 16 de ellos (13,68%) tenían estas variaciones en el gen *PARK2*, tratándose en 8 de los casos de deleciones de exones completos (Figura 2; Tabla2). La variación más frecuente fue la p.N52fs, también en *PARK2*. Además, 8 pacientes de los 27 (6,84%) tenían variaciones potencialmente patogénicas en el gen *LRRK2*. No se encontraron diferencias significativas ($p= 0,205$) en la edad de inicio entre los pacientes que tenían variaciones potencialmente patogénicas en el gen *LRRK2* y los que tenían variaciones de este tipo en el gen *PARK2*.

Los genes *PINK1* y *VPS35* han resultado ser una causa rara de EP en la población estudiada, ya que sólo se localizó una variación potencialmente patogénica en cada gen.

Finalmente, 2 pacientes tenían 2 variaciones no descritas previamente en *GBA* (p. I347T y p.S403S).

Un total de 8 variaciones potencialmente patogénicas se localizaron en otros genes relacionados con la EP: *ATP13A2* (p.G217S, p.R754H y c.2762+21T>C), *EIF4G1* (p.G693C y p.M1362T), *GIGYF2* (p.R1225C), y *SMPD1* (p.V.147V y p.R378C) (Tabla 2).

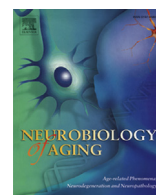
1.2. Sub-objetivo 2: Estudiar la implicación en la fisiopatogenia de la EP de *CHCHD2*, nuevo gen propuesto causativo de EP, en población del sur de España.

1.2.1. Publicación: Genetic analysis of CHCHD2 in a southern Spanish population. Tejera-Parrado C, Jesús S, Huertas-Fernández I, Bernal-Bernal I, Bonilla-Toribio M, Córdoba-Tevar I, Abreu-Rodríguez I, Carrillo F, Bernal-Escudero M, Vargas-González L, Carballoa M, Gómez-Garrea P, Mir P.



Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Negative results

Genetic analysis of *CHCHD2* in a southern Spanish population

Cristina Tejera-Parrado^a, Silvia Jesús^a, Ismael Huertas-Fernández^a,
 Inmaculada Bernal-Bernal^a, Marta Bonilla-Toribio^a, Isabel Córdoba-Tevar^a,
 Irene Abreu-Rodríguez^a, Fátima Carrillo^a, Maravilla Bernal-Escudero^a,
 Laura Vargas-González^a, Manuel Carballo^a, Pilar Gómez-Garre^{a,b,*}, Pablo Mir^{a,b,*}

^aUnidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología Clínica, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla, Spain

^bCentro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain

ARTICLE INFO

Article history:

Received 13 September 2016

Accepted 15 October 2016

Keywords:

CHCHD2

Parkinson's disease

Genetic risk

ABSTRACT

Researching genetic factors involved in Parkinson's disease (PD) is crucial to increase our knowledge about the pathophysiology of the disorder. A missense mutation has recently been reported within *CHCHD2*, a gene newly associated with autosomal dominant PD. Subsequent studies in different ethnic populations have not reached any conclusive result about the role of *CHCHD2* in PD. Therefore, the aim of this study was to investigate the implication of this gene for a PD population from southern Spain (including 536 PD patients and 518 unrelated control subjects). We studied all 4 exons of *CHCHD2* and their exon-intron boundary regions. Four variants were observed in non-coding regions. No significant differences were observed in the allele frequencies of these variants between patients and controls. Thus, our study suggests that *CHCHD2* is probably not involved in the etiopathogenesis of PD in our population.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Parkinson's disease (PD) is a complex disorder where environmental and genetic factors are involved. More than 18 genes, such as *PARK2*, *LRRK2*, and *SNCA*, among others, are well documented as causative genes or risk factors for PD (Trinh and Farrer, 2013). However, study of some of these genes is still needed to understand their implications in the pathophysiology of PD.

The *CHCHD2* gene has been a focus of research since 2015, when Funayama et al. described the missense mutation Thr61Ile to be a cause of autosomal dominant PD in a Japanese family (Funayama et al., 2015). In addition, they found other putative mutations (Arg145Gln and 300+5G>A) and 2 single nucleotide variations (-9T>G and 5C>T) with different frequencies in control subjects (CS) and sporadic PD patients (Funayama et al., 2015). To elucidate these findings, *CHCHD2* has been studied in different populations. In Taiwanese, Han Chinese, and Canadian populations, neither causative mutation nor risk factors for PD have been reported in *CHCHD2* (Fan et al., 2016; Lu et al., 2016; Zhang et al., 2016). Simultaneously, other authors have confirmed variants in *CHCHD2*

as a risk factor for PD in the Japanese population and other data sets of the Chinese population (Foo et al., 2015; Funayama et al., 2015; Shi et al., 2016). In the European population, the first-reported 3 putative mutations have been assessed and were not present in a large cohort. Consequently, a possible Asian-specific link between these variants and the risk of PD has been proposed (Jansen et al., 2015). Our aim was therefore to elucidate the involvement of *CHCHD2* in a PD population from southern Spain.

2. Methods

We included a total of 536 PD patients (304 males and 232 females, mean age 70 ± 12 years; mean age at onset 56 ± 13 years) from the Movement Disorders Unit of the Hospital Universitario Virgen del Rocío (Seville, Spain) and 518 unrelated CS (291 males and 227 females; mean age 62 ± 16 years). The demographic characteristics of the PD patients and CS are shown in Supplementary Table 1. PD was diagnosed using the criteria of the United Kingdom Parkinson's Disease Society Brain Bank (Gibb and Lees, 1988). Seventy-six patients had familial PD and 101 displayed an early-onset PD (EOPD; <45 year old; Supplementary Table 1). Signed consent forms were obtained from each individual studied. Genomic DNA was isolated from peripheral blood from each subject, according to established protocols with standard or automated methods (Maxwell 16 System, Promega Corporation).

* Corresponding authors at: Unidad de Trastornos del Movimiento, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío, Avda. Manuel Sirot s/n, 41013 Sevilla, Spain. Tel.: +34-955923039; fax: +34-955923101.

E-mail addresses: mgomez-ibis@us.es (P. Gómez-Garre), pmir@us.es (P. Mir).

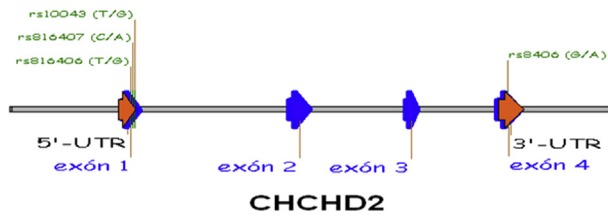


Fig. 1. CHCHD2 gene schematic representation. Exonic regions are shown by blue arrows and UTR regions by orange arrows. SNVs studied are indicated by a slight orange line and their rsID.

All 4 exons and their exon-intron boundary regions were analyzed using high-resolution melting analysis in a LightCycler 480 (Roche). Samples showing abnormal melting profiles were sequenced using an ABI 3500 Genetic Analyzer (Applied Biosystems). The primers used for amplification cover all coding regions of the *CHCHD2* gene and its 5' and 3' untranslated regions. A variant call of sequenced samples was made using Variant Reporter software v1.1. The variants detected by the first screening using high-resolution melting and sequencing samples were genotyped for all patients and controls, using TaqMan SNP (Applied Biosystems) and SimpleProbe chemistry (Roche). The case-control association was analyzed using the χ^2 -test with PLINK software v1.07. Results were considered statistically significant when *p*-values were lower than 0.05.

3. Results

We detected a total of 4 genetic variants in our study population: 3 in 5' UTR region (c.–9T>G [rs10043], c.–34C>A [rs816407], c.–65T>G [rs816406]) and 1 in 3' UTR region (c.*125G>A [rs8406]; Fig. 1).

All variants were in Hardy–Weinberg equilibrium in the control population. Total genotyping rate in individuals was higher than 95%. Around 5% of the samples were genotyped twice or thrice to check genotyping consistency, showing 100% identical genotypes. Minor allele frequency and allelic association results between PD cases and controls are shown in Table 1. No statistically significant differences were observed between cases and controls for any of the variants identified. Separate study of familial cases and EOPD against CS, respectively, yielded the same results (Supplementary Table 2).

4. Discussion

In this study, we have reported 4 non-exonic variants, c.–9T>G (rs10043), c.–34C>A (rs816407), c.–65T>G (rs816406), and c.*125G>A (rs8406), with no significant association in PD cases compared with CS. In recent years, variants rs10043, rs816407, and rs8406 have been detected in other studies on PD patients, but no difference has been found between their frequency in PD cases and controls (Fan et al., 2016; Lu et al., 2016; Shi et al., 2016; Wu et al., 2016; Zhang et al., 2016). Only Funayama et al. observed an association between the variant rs10043 and the risk of PD in their Japanese population (Funayama et al., 2015). However, in our study, we did not find such as observation. The other variant, c.–65T>G (rs816406), has not been reported by any other study into *CHCHD2* in PD. This could be explained by its low frequency, 0.01141 in our population and 0.0187 in 1000 genomes in the Iberian population. No exonic variants were found in this study. These results suggest that potential causal variants in *CHCHD2* are not common and

Table 1

Allelic association analyses of the 4 genetic variants found in our study population

SNP	Alleles	MAFs		Chi-square	Allelic association	
		PD cases	CS		<i>p</i> -value	Odds ratio
rs10043	T>G	0.234	0.254	1.145	0.285	0.896
rs816407	C>A	0.234	0.257	1.488	0.223	0.881
rs816406	T>G	0.011	0.006	1.880	0.170	1.965
rs8406	G>A	0.221	0.252	2.490	0.115	0.841

Key: CS, control subjects; MAFs, minor allele frequencies; PD, Parkinson's disease; SNP, single nucleotide polymorphism.

subsequently they may not act as a risk factor for PD, at least in our population from southern Spain.

Disclosure statement

The authors have no conflicts of interest to disclose.

Acknowledgements

The authors would like to thank the patients and healthy individuals who participated in this study. This research was conducted using samples from the *HUVR-IBiS Biobank (Andalusian Public Health System Biobank and ISCIII-Red de Biobancos PT13/0010/0056)*. The authors thank the donors and the *HUVR-IBiS Biobank (Andalusian Public Health System Biobank and ISCIII-Red de Biobancos PT13/0010/0056)* for the human specimens used in this study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.neurobiolaging.2016.10.019>.

References

- Fan, T.S., Lin, H.I., Lin, C.H., Wu, R.M., 2016. Lack of CHCHD2 mutations in Parkinson's disease in a Taiwanese population. *Neurobiol. Aging* 38, 218.e1–218.e2.
- Foo, J.N., Liu, J., Tan, E.K., 2015. CHCHD2 and Parkinson's disease. *Lancet Neurol.* 14, 681–682.
- Funayama, M., Ohe, K., Amo, T., Furuya, N., Yamaguchi, J., Saiki, S., Li, Y., Ogaki, K., Ando, M., Yoshino, H., Tomiyama, H., Nishioka, K., Hasegawa, K., Saiki, H., Satake, W., Mogushi, K., Sasaki, R., Kokubo, Y., Kuzuhara, S., Toda, T., Mizuno, Y., Uchiyama, Y., Ohno, K., Hattori, N., 2015. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. *Lancet Neurol.* 14, 274–282.
- Gibb, W.R., Lees, A.J., 1988. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 51, 745–752.
- Jansen, I.E., Bras, J.M., Lesage, S., Schulte, C., Gibbs, J.R., Nalls, M.A., Brice, A., Wood, N.W., Morris, H., Hardy, J.A., Singleton, A.B., Gasser, T., Heutink, P., Sharma, M., IPDGC, 2015. CHCHD2 and Parkinson's disease. *Lancet Neurol.* 14, 678–679.
- Lu, Q., Deng, X., Song, Z., Guo, Y., Yang, Y., Deng, H., 2016. Mutation analysis of the CHCHD2 gene in Chinese Han patients with Parkinson's disease. *Parkinsonism Relat. Disord.* 29, 143–144.
- Shi, C.H., Mao, C.Y., Zhang, S.Y., Yang, J., Song, B., Wu, P., Zuo, C.T., Liu, Y.T., Ji, Y., Yang, Z.H., Wu, J., Zhuang, Z.P., Xu, Y.M., 2016. CHCHD2 gene mutations in familial and sporadic Parkinson's disease. *Neurobiol. Aging* 38, 217.e9–217.e13.
- Trinh, J., Farrer, M., 2013. Advances in the genetics of Parkinson disease. *Nat. Rev. Neurol.* 9, 445–454.
- Wu, H., Lu, X., Xie, F., Cen, Z., Zheng, X., Luo, W., 2016. Genetic analysis of the CHCHD2 gene in a cohort of Chinese patients with Parkinson disease. *Neurosci. Lett.* 629, 116–118.
- Zhang, M., Xi, Z., Fang, S., Ghani, M., Sato, C., Moreno, D., Liang, Y., Lang, A.E., Rogaeva, E., 2016. Mutation analysis of CHCHD2 in Canadian patients with familial Parkinson's disease. *Neurobiol. Aging* 38, 217.e7–217.e8.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Demographic characteristics of PD patients and controls

		Sex (M/F)	Age (y)	AO (y)	n
Control subjects		291/227	62 ± 16	-	518
Parkinson disease	Total	304/232	70 ± 12	56 ± 13	536
	Familial PD	38/38	69 ± 10	52 ± 14	76
	EOPD	63/38	56 ± 10	37 ± 8	101

PD: Parkinson's disease; y, years; AO: age at onset; n: number of samples; EOPD: early onset Parkinson's disease.

Data are presented as the mean ± standard deviation.

Supplementary Table 2. Allelic association analyses of the four genetic variants found in familial PD and early onset PD groups

Association study of familial PD cases vs control						
SNP	Alleles	MAFs		Chi-Square	Allelic Association	
		Familial PD cases	CS		P-value	Odds ratio
rs10043	T>G	0.2727	0.2544	0.2349	0.6279	1.0990
rs816407	C>A	0.2877	0.2570	0.6230	0.4299	1.1680
rs816406	T>G	0.0128	0.0058	0.9844	0.3211	2.2120
rs8406	G>A	0.2462	0.2521	0.0214	0.8836	0.9688
Association study of EOPD cases vs control						
SNP	Alleles	MAFs		Chi-Square	Allelic Association	
		EOPD cases	CS		P-value	Odds ratio
rs10043	T>G	0.2644	0.2544	0.0772	0.7811	1.0530
rs816407	C>A	0.2805	0.2570	0.4049	0.5246	1.1270
rs816406	T>G	0.0169	0.0058	2.4860	0.1148	2.9200
rs8406	G>A	0.2595	0.2521	0.0393	0.8429	1.0400

PD: Parkinson's disease; MAFs: Minor allele frequency; CS: control subjects; SNP: single nucleotide polymorphism; EOPD: early onset Parkinson's disease.

1.2.2. Resumen

En este trabajo se evaluó la existencia de variaciones genéticas en el gen *CHCHD2*, propuesto como gen involucrado en la fisiopatogenia de la EP desde 2015 por Funayama y col. En su trabajo, Funayama y col. describen por primera vez una mutación causativa de EP de herencia autosómica dominante en el gen *CHCHD2*, p.T61I, en una familia japonesa. Posteriores trabajos identifican también variaciones de riesgo asociada a la EP en este gen, aunque su papel en la enfermedad sigue siendo controvertido ya que en otros estudios no se corroboran estos resultados. Con el fin de dilucidar la implicación de *CHCHD2* en la EP, el objetivo de este trabajo fue analizar este gen en una población del sur de España compuesta por 518 sujetos sanos y 536 pacientes de EP de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla, España), de los cuales 101 tenían EPIT. En total se identificaron 4 variaciones genéticas en la población estudiada: 3 en la región 5'-UTR (c.9T>G, c.34C>A, c.65T>G) y 1 en la región 3'-UTR (c.*125G>A).

Ninguna de las variaciones genéticas se encontraba asociada a la EP en este estudio, ni tampoco cuando se analizaron por separado los 101 casos de EPIT frente a los controles sanos. El gen *CHCHD2* no parece actuar como un factor de riesgo para la EP, al menos en la población de estudio.

1.3. Sub-objetivo 3: Estudiar la implicación en la fisiopatogenia de la EP de *TMEM230*, nuevo gen propuesto causativo de EP, en población del sur de España.

1.3.1. Publicación: TMEM230 in Parkinson's disease in a southern Spanish population. Tejera-Parrado C, Jesús S, López-Ruíz A, Buiza-Rueda D, Bonilla-Toribio M, Bernal.Bernal I, Periñán MT, Vargas-González L, Gómez-Garre P, Mir P.

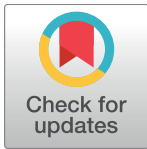
RESEARCH ARTICLE

TMEM230 in Parkinson's disease in a southern Spanish population

Cristina Tejera-Parrado¹, Silvia Jesús¹, Adrián López-Ruiz¹, Dolores Buiza-Rueda¹, Marta Bonilla-Toribio¹, Inmaculada Bernal-Bernal¹, María Teresa Perrián¹, Laura Vargas-González¹, Pilar Gómez-Garre^{1,2*}, Pablo Mir^{1,2*}

1 Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología Clínica, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla, España, **2** Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, España

* pmir@us.es (PM); mgomez-ibis@us.es (PG)



OPEN ACCESS

Citation: Tejera-Parrado C, Jesús S, López-Ruiz A, Buiza-Rueda D, Bonilla-Toribio M, Bernal-Bernal I, et al. (2018) *TMEM230* in Parkinson's disease in a southern Spanish population. PLoS ONE 13(5): e0197271. <https://doi.org/10.1371/journal.pone.0197271>

Editor: Patrick Lewis, University College London Institute of Neurology, UNITED KINGDOM

Received: February 6, 2018

Accepted: April 30, 2018

Published: May 17, 2018

Copyright: © 2018 Tejera-Parrado et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was supported by grants from the Spanish Ministry of Economy and Competitiveness [PI14/01823 to Pilar Gómez-Garre, PI16/01575 to Pablo Mir] co-founded by the Instituto de Salud Carlos-III (ISCIII) and by Fondo Europeo de Desarrollo Regional (FEDER), the Consejería de Economía, Innovación, Ciencia y Empleo de la Junta de Andalucía [CVI-02526, CTS-

Abstract

TMEM230 has been associated with autosomal dominant Parkinson's disease (PD). Subsequent studies have remained negative, and none of previous described mutation has been reported anymore. We investigated the implication of this gene in the PD in a population of 703 PD patients and 695 unrelated healthy controls from southern Spain. Thirteen variants were found, twelve of them observed only in controls or in patients and controls, and one (c.190A>G) observed only in one patient. Subsequent analysis of this variant indicates that probably it is not pathogenic. In addition, we found a variation in the 3'-UTR (rs183551373) and related with the miRNA hsa-miR-4299 but it was observed only in healthy controls. Our results suggest that variants in *TMEM230* gene are not associated with the development of PD.

Introduction

Parkinson's disease (PD) is a complex pathology in which both environmental and genetic factors are involved. More than 18 genes, like *PARK2*, *LRRK2* or *SNCA*, among others, are well documented as causative genes or risk factors for PD [1]. However, the study of other described genes is still needed for understanding their implication in the pathophysiology of PD.

Recently, a missense mutation (p.Arg141Leu) in *TMEM230* gene has been described to be a cause of autosomal dominant PD in a large family from North American [2]. Two other mutations (p.Tyr92Cys and p.*184Trpext5*) were also identified in two young onset unrelated PD patients in that study. Furthermore, in a Chinese population, a new mutation in *TMEM230* (p.*184ProGlyext5*) was identified as being related with PD in 7 familial cases. Several additional studies in different populations have failed to find neither causative mutation nor risk factors in *TMEM230* associated with PD [3–16] (S1 Table).

In order to complete the picture of the involvement of *TMEM230* gene in PD, we have analyzed this gene in a PD population from southern Spain.

7685], the Consejería de Salud y Bienestar Social de la Junta de Andalucía [PI-0437-2012, PI-0471-2013], the Sociedad Andaluza de Neurología to Pablo Mir, the Jacques and Gloria Gossweiler Foundation to Pablo Mir, the Fundación Alicia Koplowitz to Pablo Mir, and the Fundación Mutua Madrileña. Pilar Gómez-Garre was supported by the "Miguel Servet" (from ISCIII-FEDER) and "Nicolás Monardes" (from Andalusian Ministry of Health) programmes. Silvia Jesús was supported by the "Juan Rodés" programme (from ISCIII-FEDER). Cristina Tejera-Parrado was supported by VPPI-US from the Universidad de Sevilla. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interest exist.

Materials and methods

The study was approved by the Ethics Committee of the Hospital Virgen del Rocío and was conducted according to the Declaration of Helsinki. All subjects, both patients and controls, were informed (orally and in writing) of the nature of the study and signed a written consent before blood collection.

We included a total of 703 PD patients from the Movement Disorders Unit of the Hospital Universitario Virgen del Rocío (Seville, Spain), and 695 healthy controls (HC). One hundred forty-eight patients had familial PD (Table 1). PD was diagnosed using the criteria of the United Kingdom Parkinson's Disease Society Brain Bank [17]. A positive family history was defined as the presence of Parkinson's disease in at least other first-degree and/or second-degree relative. Signed consent forms were obtained for each individual studied and the study was approved by the local ethic committee. Genomic DNA was isolated from peripheral blood from each subject, according to established protocols by standard or automated methods (MagNA Pure LC, Roche Diagnostics, Indianapolis, IN). To detect mutations, all the 5 exons of *TMEM230* (NM_001009923 and NM_001009925) and their intron boundary regions were analyzed using high-resolution melting analysis in a LightCycler 480-II (Roche). Samples showing abnormal melting profiles were sequenced in an ABI3500 Genetic Analyzer (Applied Biosystem). The identified variants were annotated according to the longest isoform (NM_001009923). A variant call of sequenced samples was made using Variant Reporter software v1.1. (Applied Biosystem).

Results

We detected a total of 13 genetic variants (two not reported before) in our population: two at 3'-untranslated region, one in 5'-untranslated region, one intronic variant, four synonymous variants, and five nonsynonymous variants (Table 2). Ten were rare variants (minor allele frequency <0.01). From those, four variants were observed in patients and controls (rs141394228, rs147693982, rs148033002, and c.*670T>C) and five were observed only in controls (rs368686615, rs186628284, rs776037148, rs143571424, and rs183551373). One novel variant (c.190A>G) was present in one patient only. This variant results in the loss of the start codon in the most common isoform of *TMEM230*. However, another variant found in the present study, c.191T>C, carrying the same effect was present in two healthy controls and in one PD patient. Another novel variant was identified in 3'UTR region, c.*670T>C but it was present in patients and control subjects. None of these variations were identified in the familial cases. Neither of the known variants previously related with PD have been found in this study.

Three frequent variants (rs10221980, rs6107576, and rs6116651) were present in PD case cohort and control cohort with similar frequencies.

Discussion

In this study we have reported 13 variants in *TMEM230* but neither of them seems to be related with the development of PD, since they are present only in HC or in both patients and HC. The exception has been c.190A>G, which is present in a single patient. This variation causes

Table 1. Demographic characteristics of studied population.

Subjects (n)		% males (M/F)	Mean age (y)	Mean AO (y)
Healthy controls (695)		55.4 (385 / 310)	57 ± 15	-
PD	Total (703)	54.6 (384 / 319)	64 ± 12	56 ± 13
	Familial (148)	45.9 (68 / 80)	62.5 ± 11	53 ± 14

PD: Parkinson's disease; y: years; AO: age at onset; n: number of samples

<https://doi.org/10.1371/journal.pone.0197271.t001>

Table 2. Variants identified in *TMEM230* in our population.

Chromosome position ^{&}	Variant [§]			Maf			PD population (Pres/Abs)
	Nucleotide change	Aminoacid change	rs number	1000 genomes [#]	ExAc	ESP	
Chr20: 5093685	c.-11C>T	-	rs368686615	NI	0	0	Abs
Chr20: 5092233	c.87G>A	p.Ser29Ser	rs186628284	NI	NI	NI	Abs
Chr20: 5092141	c.174+5G>C	-	rs10221980	0.103	0.3214	NI	Pres
Chr20: 5090076	c.190A>G	p.Met64Val	-	NI	NI	NI	Pres
Chr20: 5090075	c.191T>C	p.Met64Thr	rs141394228	0.004	0.0018	0.0015	Pres
Chr20: 5090057	c.209A>G	p.Asn70Ser	rs776037148	NI	1.5 e-05	NI	Abs
Chr20: 5086939	c.306T>A	p.Pro102Pro	rs6116651	0.0905	0.1114	0.1093	Pres
Chr20: 5086918	c.327C>T	p.Ile109Ile	rs147693982	0.0169	0.0162	0.0181	Pres
Chr20: 5086915	c.330A>G	p.Ala110Ala	rs6107576	0.0099	0.0126	0.0136	Pres
Chr20: 5086870	c.375A>G	p.Ile125Met	rs148033002	0.001	0.0012	0.0008	Pres
Chr20: 5081478	c.511C>T	p.Arg171Cys	rs143571424	0	0.0039	0.0020	Abs
Chr20: 5080819	c.*618A>G	-	rs183551373	0	0	NI	Abs
Chr20: 5079983	c.*670T>C	-	-	NI	NI	NI	Pres

Chr: chromosome; &: GRCh37/hg19)

§: referred to the longest *TMEM230* isoform (NM_001009923, NP_001009923)

#: European population

*ExAC: Exome Aggregation Consortium (Non-Finnish European)

ESP: Exome Sequencing Project (European American); NI: No information. Pres/Abs: present or absent in our analyzed PD population

<https://doi.org/10.1371/journal.pone.0197271.t002>

the loss of the ATG start codon in the isoform 2 of *TMEM230* (NP_001009925; c.1A>G, p.Met1Val), therefore, it could be disease causing. However, the variant c.191T>C, with similar consequences in the protein (NP_001009925; c.2A>G, p.M1Thr), is present in healthy controls in this study and in previous ones [3]. In addition, the subsequent amino acid of *TMEM230* protein is also methionine, which would act as an alternative start codon to generate a novel protein that lacks only the first amino acid. So, it is expected that this variation does not cause a significant alteration of the protein. In addition, a variation in the 3'-UTR (rs183551373) has been related with the miRNA hsa-miR-4299 but it was found only in healthy controls in our population and, thereupon, the pathogenicity of this variant seems not be probable.

Conclusions

Our findings suggest that the incidence of pathogenic variations in *TMEM230* is very low and, therefore, *TMEM230* do not play a major role in familial and sporadic PD patients in southern Spanish population which can have important implication in clinical investigation.

Supporting information

S1 Table. Summary of characteristics of previously reported studies for *TMEM230* (up to April 2018).
(PDF)

Acknowledgments

We thank the patients and healthy individuals who participated in this study for their cooperation. This research has been conducted using samples from the *HUVR-IBiS Biobank (Andalusian Public Health System Biobank and ISCIII-Red de Biobancos PT13/0010/0056)*.

Author Contributions

Conceptualization: Pilar Gómez-Garre, Pablo Mir.

Data curation: Cristina Tejera-Parrado, Silvia Jesús.

Formal analysis: Cristina Tejera-Parrado, Silvia Jesús, María Teresa Periñán.

Funding acquisition: Pilar Gómez-Garre, Pablo Mir.

Investigation: Cristina Tejera-Parrado, Adrián López-Ruíz, Dolores Buiza-Rueda, Marta Bonilla-Toribio, Inmaculada Bernal-Bernal.

Resources: Cristina Tejera-Parrado, Silvia Jesús, Laura Vargas-González.

Supervision: Pilar Gómez-Garre, Pablo Mir.

Visualization: Silvia Jesús.

Writing – review & editing: Cristina Tejera-Parrado, Silvia Jesús, Pilar Gómez-Garre, Pablo Mir.

References

1. Trinh J, Farrer M. Advances in the genetics of Parkinson disease. *Nature reviews Neurology*. 2013; 9(8):445–54. <https://doi.org/10.1038/nrneurol.2013.132> PMID: 23857047
2. Deng HX, Shi Y, Yang Y, Ahmeti KB, Miller N, Huang C, et al. Identification of TMEM230 mutations in familial Parkinson's disease. *Nature genetics*. 2016; 48(7):733–9. <https://doi.org/10.1038/ng.3589> PMID: 27270108
3. Buongarzone G, Monfrini E, Franco G, Trezzi I, Borellini L, Frattini E, et al. Mutations in TMEM230 are rare in autosomal dominant Parkinson's disease. *Parkinsonism & related disorders*. 2017; 39:87–8.
4. Giri A, Mok KY, Jansen I, Sharma M, Tesson C, Mangone G, et al. Lack of evidence for a role of genetic variation in TMEM230 in the risk for Parkinson's disease in the Caucasian population. *Neurobiology of aging*. 2017; 50:167 e11–e13.
5. Ibanez L, Dube U, Budde J, Black K, Medvedeva A, Davis AA, et al. TMEM230 in Parkinson's disease. *Neurobiology of aging*. 2017; 56:212 e1–e3.
6. Quadri M, Breedveld GJ, Chang HC, Yeh TH, Guedes LC, Toni V, et al. Mutations in TMEM230 are not a common cause of Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society*. 2017; 32(2):302–4.
7. Shi CH, Li F, Shi MM, Yang ZH, Mao CY, Zhang SY, et al. Genetic analysis of the TMEM230 gene in Chinese Han patients with Parkinson's disease. *Scientific reports*. 2017; 7(1):1190. <https://doi.org/10.1038/s41598-017-01398-9> PMID: 28446760
8. Wu H, Zheng X, Cen Z, Xie F, Chen Y, Lu X, et al. Genetic analysis of the TMEM230 gene in Chinese patients with familial Parkinson disease. *Parkinsonism & related disorders*. 2017; 36:105–6.
9. Yan W, Tang B, Zhou X, Lei L, Li K, Sun Q, et al. TMEM230 mutation analysis in Parkinson's disease in a Chinese population. *Neurobiology of aging*. 2017; 49:219 e1–e3.
10. Yang X, An R, Xi J, Zheng J, Chen Y, Huang H, et al. Sequencing TMEM230 in Chinese patients with sporadic or familial Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society*. 2017; 32(5):800–2.
11. Baumann H, Wolff S, Munchau A, Hagenah JM, Lohmann K, Klein C. Evaluating the role of TMEM230 variants in Parkinson's disease. *Parkinsonism & related disorders*. 2017; 35:100–1.
12. Conedera SA, Li Y, Funayama M, Yoshino H, Nishioka K, Hattori N. Genetic analysis of TMEM230 in Japanese patients with familial Parkinson's disease. *Parkinsonism & related disorders*. 2018; 48:107–8.
13. Fan TS, Lin CH, Lin HI, Chen ML, Wu RM. Lack of TMEM230 mutations in patients with familial and sporadic Parkinson's disease in a Taiwanese population. *American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics*. 2017; 174(7):751–6.
14. He YC, Huang P, Li QQ, Sun Q, Li DH, Wang T, et al. TMEM230 stop codon mutation is rare in parkinson's disease and essential tremor in eastern China. *Movement disorders: official journal of the Movement Disorder Society*. 2017; 32(2):301–2.

15. Ma D, Foo JN, Yulin Ng E, Zhao Y, Liu JJ, Tan EK. Screening for TMEM230 mutations in young-onset Parkinson's disease. *Neurobiology of aging*. 2017; 58:239 e9–e10.
16. Wei Q, Ou R, Zhou Q, Chen Y, Cao B, Gu X, et al. TMEM230 Mutations Are Rare in Han Chinese Patients with Autosomal Dominant Parkinson's Disease. *Molecular neurobiology*. 2018; 55(4):2851–5. <https://doi.org/10.1007/s12035-017-0542-2> PMID: 28455698
17. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *Journal of neurology, neurosurgery, and psychiatry*. 1988; 51(6):745–52. PMID: 2841426

Supplementary Table 1. Summary of characteristics of previously reported studies for TMEM230 (up to April 2018).

Reference	Population	N PD (F) / HC	Ancestry	Position (GRCh37; chr20)	Exon	Variation (NM_001009923; NP_001009923)			Frequency			
						Change of Nt	Change of Aa	rs ID	PD	FPD?	HC	
Deng et al. 2016	North America	832 (433) / ~1000	-	5090075	3	c.191T>C	p.Met64Thr	rs141394228	2 / 832	ND	5 / 1081	
				5089991	3	c.275A>G	p.Tyr92Cys	rs1056737920	1 / 832	No	0 / 1081	
				5086870	4	c.375A>G	p.Ile125Met	rs148033002	2 / 832	ND	3 / 1028	
				5081567	5	c.422G>T	p.Arg141Leu	None	1 PD	Yes	0 / 1238	
				5081478	5	c.511C>T	p.Arg171Cys	rs143571424	1 / 832	ND	3 / 1238	
	5081438	5	c.551A>G	p.*184Trpext*5	-	1 / 832	Yes	0 / 1238				
	Chinese	574 (225) / ~528	-	5081437	5	c.550_552delTAGinsCCCGGG	p.*184ProGlyext*5	-	7 / 574	Yes	0 / 528	
Yan et al. 2017	Chinese (western)	1235 (192) / 1252	-	5086888	4	c.357G>A	p.Gly119Gly	rs749201963	1 / 1235	Yes	0 / 1252	
				5081434	5	c.*3G>T	3'-UTR	-	2 / 1235	No	0 / 1252	
Giri et al. 2017	Caucasian	1450 / 2267	North America	5086929	4	c.316T>C	p.Tyr106His	rs374122606	1 / 1450	No	0 / 2267	
				5081505	5	c.484A>G	p.Ile162Val	rs368707598	1 / 1450	No	0 / 2267	
	European	5081502	5	c.487G>A	p.Ala163Thr	rs374122606	1 [‡] / 29	Yes	#			
He et al. 2017 ^{&}	Chinese Han (eastern)	414 (207) / 400	-	5081439	5	Negative for p.*184Trpext*5 (c.551A>G) and p.*184ProGlyext*5 (c.550_552delTAGinsCCCGGG)			0 / 414	-	0 / 400	
Baumann et al. 2017	Caucasian	53 (53) / 0	Ukrainian	5090063	3	c.203C>T	p.Arg68His	rs780460399	1 / 53	Yes	#	
			German	5081478	5	c.511C>T	p.Arg171Cys	rs143571424	1 / 53	Yes	Described by Deng et al	
Wu et al. 2017	Chinese (eastern)	122 (122) / 0	-	5093629	1	c.46G>A	p.Glu16Arg	rs1272596579	1 [‡] / 122	Yes	NA	
				5092135	2	c.174+11C>T	(intronic)	rs144736056	1 / 122	Yes	#	
Quadri et al. 2017	Taiwanese [%]	815 (98) / 417	-	5081595	5	c.494A>G	p.Tyr165Cys	rs758033952	1 / 815	No	417 [#]	
				5081426	5	c.*11A>G	(3'-UTR)	rs750802038	1 / 815	No	417 [#]	
	Dutch [%]	90 (31) / 0	-	-	5	None			0 / 90	-	NA	
	White	266 (266) / 0	-	Italian	5093674	1	c.1A>G	p.Met1Val	rs768390203	1 / 266	Yes	#
				Brazilian	5092251	2	c.69A>G	p.Arg23Arg	-	1 / 266	Yes	NA
					5090075	3	c.191T>C	p.Met64Thr	rs141394228	2 / 266	Yes	Described by Deng et al
				5086870	4	c.375A>G	p.Ile125Met	rs148033002	2 / 266	Yes	Described by Deng et al	
	5081478	5	c.511C>T	p.Arg171Cys	rs143571424	2 / 266	Yes	Described by Deng et al				
Boungarzone et al. 2017	Caucasian (Italian)	86 (86) / 0	-	5092141	2	c.174+5G>C	(intronic)	rs10221980	3 / 86	Yes	#	
				5090075	3	c.191T>C	p.Met64Thr	rs141394228	1 / 86	Yes	Described by Deng et al [#]	
				5086939	4	c.306T>A	p.Pro102Pro	rs6116651	13.37%	Yes	#	
				5086936	4	c.309G>A	p.Lys103Lys	rs763383477	1 / 86	Yes	#	
				5086915	4	c.330A>G	p.Ala110Ala	rs6107576	3 / 87	Yes	#	
				5081621	5	c.412-44G>A	(intronic)	rs201276122	2 / 86	Yes	#	
Yang et al. 2017	Chinese (southwestern)	366 (11) / 0	-	5093629	1	c.46G>T	p.Gly16Trp	rs1272596579	1 / 366	No	Also described by Wei et al	
				5093607	1	c.68G>A	p.Arg23Gln	rs765701142	3 / 366	Yes	Also described by	

										Fan et al	
				5092253	1	c.69-3T>A	(intronic)	rs1034386774	1 / 366	No	-
				5086963	4	c.289-7C>T	(intronic)	rs767174424	1 / 366	No	#
				5086917	4	c.328G>A	p.Ala110Thr	rs1055543199	2 / 366	No	-
Shi et al. 2017	Chinese Han	550 (0) / 560	-	5093425	1	c.68+182G>A	(intronic)	rs149865687	maf 0.032	No	maf 0.018
				5092242	2	c.78A>G	p.Leu26Leu	rs745443202	3 / 550	No	1 / 560
				5092135	2	c.174+11C>T	(intronic)	rs144736056	MAF 0.0172	No	MAF 0.0143
				5081426	5	c.*11A>G	(3'-UTR)	rs750802038	5 / 550	No	4 / 560
Wei et al. 2018	Chinese Han (southwestern)	120 (120) / 650	-	5093629	1	c.46G>T	p.Gly16Trp	rs1272596579	1 / 120	Yes	0 / 650
				5081560	5	c.429delT	p.Val143ValfsX4	-	0 / 120	-	1 / 650
Ibanez et al. 2017	American European	499 / 294	WUSTL	Negative for those variants previously described by Deng et al					0 / 499	-	0 / 294
		340 / 140	PPMI	5093490	1	c.68+117C>A	(intronic)	rs41282112	ND	ND	ND
				5093410	1	c.68+19C>G	(intronic)	rs41282110			
				5090075	3	c.191T>C	p.Met64Thr	rs141394228	ND	ND	Described by Deng et al
				5086939	4	c.306T>A	p.Pro102Pro	rs6116651	ND	ND	Also described by Boungarzone et al.
				5086935	4	c.310A>C	p.Ile104Leu	rs139824737			#
				5086918	4	c.327C>T	p.Ile109Ile	rs147693982			#
				5086915	4	c.330A>G	p.Ala110Ala	rs6107576	ND	ND	Also described by Boungarzone et al.
				5081478	5	c.511C>T	p.Arg171Cys	rs143571424			ND
				5081276	5	c.*161C>T	(3'-UTR)	rs1249078621	ND	ND	ND
				5081049	5	c.*388A>G	(3'-UTR)	rs2715			
				5080819	5	c.618*A>G	(3'-UTR)	rs183551373			
				5080691	5	c.*746G>A	(3'-UTR)	rs45610034			
				5080631	5	c.*806A>T	(3'-UTR)	-			
5080560	5	c.*877A>C	(3'-UTR)	rs540970582							
Ma et al. 2017	Chinese (southeastern)	99 (23) / 99	-	5086883	4	c.362T>C	p.Phe121Ser	rs1393260987	0 / 99	-	1 / 99
Fan et al. 2017	Taiwanese	680 (180) / 992	-	5093607	1	c.68G>A	p.Arg23Gln	rs765701142	3 / 680	Yes	3 / 992
				5093573	1	c.68+34C>T	(intronic)	rs999822627	Yes	ND	ND
				5093498	1	c.68+109C>T	(intronic)	rs565123586	Yes	ND	ND
				5093425	1	c.68+182G>A	(intronic)	rs149865687	Yes	ND	Also described by Shi et al
				5092135	2	c.174+11C>T	(intronic)	rs144736056	Yes	ND	Also described by Shi et al
Conedera et al. 2018	Japanese	182 (182) / 0	-	Negative (for coding variations)					0 / 182	-	ND

This table summarizes the data from different studies in which variants in TMEM230 have been tested in relation with Parkinson disease. Articles appear in the same order that they appear in PubMed. ND: No data. Population: it is referred to the origin of the studied population. Ancestry: the ancestry of those PD patients carrying the variations (when the populations are from different countries). N: number of screened samples. Nt: nucleotide. Aa: Amino acid. rs ID: reference SNP ID number PD (F): PD cases (familial cases included). HC: Healthy controls. FPD: Familial PD cases. ¥: This variant does not segregate with PD in the family. #: This variant is present in populations from Exome Aggregation Consortium (ExAC). &: This study is focused in

detect variations affecting to the stop (TAG) codon. %: This study is focused on the study of the exon 5 of the gene. WUSTL: GWASs data and DNA from the Washington University in Saint Louis (This cohort was only analyzed for those variants previously described by Deng et al). PPMI: GWASs and Whole Exome Sequencing data from the Parkinson's Progression Markers Initiative.

1.3.2. Resumen

Recientemente se ha descrito el gen *TMEM230* como causa de EP de herencia autosómica dominante. Deng y col., describen por primera vez en 2016 una mutación (p.R141L) en una familia de Norteamérica y otras dos en dos pacientes de EPIT no relacionados (p.Y92C y p.*184Wext*5). En otro estudio en población china, se localizó otra mutación en este gen asociada a la EP en 7 casos familiares. Sin embargo, en sucesivos estudios en diferentes poblaciones este gen no parece ser causativo de EP ni tampoco factor de riesgo. Para comprender mejor el papel de *TMEM230* en la fisiopatología de la EP, en este trabajo se analiza la presencia de mutaciones en este gen en población española comprendida por 703 pacientes de EP de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla, España) y 695 controles sanos no relacionados.

En total se identificaron 13 variaciones en nuestra población de estudio: 2 en la región 3'-UTR, 1 en la región 5'-UTR, 1 intrónica, 4 sinónimas y 5 no sinónimas. Diez de estas variaciones son variaciones raras con una frecuencia menor de 0,01. De éstas, 4 estaban presentes en pacientes y controles (p.M64T, p.I109I, p.I125M y c.*670T>C), 5 sólo se observaron en controles (c.-11C>T, p.S29S, p.N70S, p.R171C y c.*618A>G) y 1 variación sólo se encontraba en un paciente. Esta variación genética no estaba descrita previamente (p.M64V) y produce la pérdida del codón de inicio en la isoforma más común del gen *TMEM230* (NM_001009923). Sin embargo, otra variación de este estudio, la p.M64T, produce el mismo efecto en la

secuencia y está presente en dos controles y un paciente, por lo que parece que su efecto no es deletéreo.

Se identificó además otra variación no descrita en la región 3'-UTR (c.*670T>C), pero estaba presente tanto en controles como en pacientes. Ninguna de las variaciones identificadas se encontraban en pacientes de EP familiar. Tampoco se localizaron variaciones descritas previamente relacionadas con la EP. En cuanto a las variaciones más frecuentes, su presencia era similar en los dos grupos de estudio (c.174+5G>C, p.A110A y p.P102P). Los resultados derivados de este trabajo sugieren que el gen *TMEM230* no está implicado en la EP, al menos en la población de este estudio.

2. Objetivo 2: Replicación de 5 loci de riesgo para la EP, previamente asociados a una cohorte de EP de Granada, en otra cohorte independiente del sur de España.

2.1. Manuscrito en vías de publicación: A replication study of GWAS-genetic risk variants associated with Parkinson's disease in a Spanish population: a case-control study and meta-analysis. Tejera-Parrado C, Jesús S, Perrián MT, Buiza-Rueda D, Oliva-Ariza G, Adarmes-Gómez AD, Macías-García D, Gómez-Garre P, Mir P.

A replication study of GWAS-genetic risk variants associated with Parkinson's disease in a Spanish population: a case-control study and meta-analysis.

Cristina Tejera-Parrado¹, Silvia Jesús^{1,2}, María T Periñán^{1,2}, Dolores Buiza-Rueda^{1,2}, Guillermo Oliva-Ariza¹, Astrid D Adarmes-Gómez^{1,2}, Daniel Macías-García^{1,2}, Pilar Gómez-Garre^{1,2*}, Pablo Mir^{1,2*}.

¹Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología Clínica, Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla, Sevilla, España.

²Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, España

*** Corresponding authors:**

Dr. Pilar Gómez-Garre and Dr. Pablo Mir

Unidad de Trastornos del Movimiento

Instituto de Biomedicina de Sevilla (IBiS)

Hospital Universitario Virgen del Rocío

Avda. Manuel Siurot s/n. 41013 Seville, Spain

Fax: +34-955923101

Telephone: +34-955923039

e-mail: mgomez-ibis@us.es / pmir@us.es

Abstract

Recently, 5 previously Parkinson's disease (PD)-related loci: *ACMSD/TMEM163*, *STK39*, *MIR4697*, *SREBF/RAI1PD* and *MAPT*, have been associated to PD in a Southern Spanish population. However, due to the small sample size of the cohort, this association did not reach genome wide significance. Our aim was to investigate the robustness of this association in a larger and independent cohort from the South of Spain. Variants were genotyped employing TaqMan SNP Genotyping Assay and high resolution melting analysis in 738 PD patients and 1138 healthy controls. Furthermore, a meta-analysis study was carried out with both cohorts. In the replication analysis, only two loci (*ACMSD/TMEM16* and *MAPT*) were replicated with a Bonferroni significance level. In the meta-analysis study no loci reached a genome-wide significance level ($P < 5 \times 10^{-8}$), but a suggestive association (P -value = 1.04×10^{-6}) between rs6430538 (*ACMSD/TMEM163*) and an increased risk of PD was found. In addition, rs9468 (*MAPT*) was associated with a decreased risk of PD (P -value = 5.70×10^{-7}). Our results add further support for the genetic involvement of these two loci in the susceptibility to PD in population from the South of Spain. We believe that our findings will be very useful for future genetic studies on PD.

Keywords: Parkinson's disease, Genome wide association studies, GWAS, genetic risk, Spanish population

1. Introduction

Parkinson's disease (PD) is the second most frequent neurodegenerative disorder affecting approximately 1% of the population over the age of 60 years. However, the etiology underlying the development of PD remains unclear. Broad efforts are being made to identify genes involved in the development of PD. While high penetrant pathogenic mutations causative of dominant or recessive forms of the disease are rare in the population, most PD cases are sporadic [5]. A combination of low penetrant variants, but also common variants and environmental factors, seems to be the cause of sporadic parkinsonism [13]. Last years, many Genome-Wide Association Studies (GWAS) have identified risk loci to contribute to PD [2, 4, 6, 8, 10, 11]. In 2014, Nalls et al. reported 24 risk loci for PD in a large scale meta-analyses study of all PD GWAS of European ancestry existing so far [7]. It is well known that replication studies in other populations are essential to check the reliability of these GWAS. Therefore, in 2016, Bandrés-Ciga et al. reported a marginal association (nominal P -value <0.05) of 5 of those 24 previously reported PD loci (*ACMSD/TMEM163*, *MIR4697*, *SREBF/RAI1*, *MAPT*, and *STK39*) through a GWAS in an independent Spanish cohort from Granada [1]. However, results from that study were limited because of the small sample size of the studied cohort (240 PD cases and 192 healthy controls).

Consequently, the purpose of the present study was to investigate whether there was a link between those variants in *ACMSD/TMEM163*, *MIR4697*, *SREBF/RAI1*, *MAPT*, and *STK39* loci, and PD in a larger independent PD cohort also from the South of Spain. Furthermore, we performed a meta-analysis that combined our data with those of Bandrés-Ciga et al, to understand the role of these risk variants in our population.

2. Methods

2.1 Patients

We included a total of 738 PD patients (432 males and 306 females, mean age 65.4 ± 11.5 years, mean age of onset 55.3 ± 12.6 years) from the Movement Disorders Unit of the Hospital Universitario Virgen del Rocío (Seville, Spain) and 1138 healthy controls (HC) (646 males and 492 females, mean age 58.4 ± 15.3 years). Demographic characteristics are shown in Table 1. PD was diagnosed using the criteria of the United Kingdom Parkinson's Disease Society Brain Bank [3]. Patients were clinically evaluated by movement disorders specialists and an extensive set of clinical features were obtained. Peripheral blood samples were obtained from each subject.

2.2 Ethical Statements

The study was approved by the ethic committee of the Hospital Virgen del Rocío de Sevilla and conducted according to the principles expressed in the Helsinki Declaration. Signed consent forms were obtained from each individual studied (both patients and healthy controls) before to blood withdrawal.

2.3 Genetic analysis

2.3.1. DNA isolation

Genomic DNA was isolated from peripheral blood from each subject according to established protocols with standard or automated methods (Magna Pure LC, Roche Applied Science, Indianapolis, IN, USA). DNA quantification was determined by Qubit 3.0 fluorometer (Invitrogen) and DNA purity was assessed by the NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Florida, USA). In addition, a quality control was established checking sex inconsistencies.

2.3.2. Genotyping

Five risk variants with nominal uncorrected P -value <0.05 from Bandrés-Ciga's study were selected for genotyping in our population. Those were rs6430538 (in *ACMSD/TMEM163* loci), rs329648 (in *MIR4697* loci), rs11868035 (in *SREBF/RAI1* loci), rs17649553 (in *MAPT* loci) and rs1955337 (in *STK39* loci). The first four variants were genotyped using TaqMan SNP Genotyping Assays (Applied Biosystems) in a LightCycler480 (Roche Applied Science). The SNP rs1955337 was analyzed by high-resolution melting analysis and subsequent DNA sequencing for those samples showing abnormal melting profiles.

2.3.3 Association analysis

Genotyped samples were excluded if they failed in more than one genotyped variant. An association analysis of the variants was carried out with PLINK software [9]. Allelic association between cases and controls and Hardy-Weinberg equilibrium in the control group were tested by Chi-square test. The Bonferroni method was used to adjust the significance level for multiple comparisons. Therefore, statistically significance threshold was established in nominal P -value < 0.01 . Quality controls were applied as follow: remove SNPs with $P < 0.01$ in a Hardy-Weinberg equilibrium test, genotype call rate less than 90%, or MAF less than 1%.

2.3.4. Meta-analysis

The meta-analysis with the Granada cohort [1] and our cohort from Seville was conducted using PLINK software. Heterogeneity between the studies was assessed using the Q-statistic and its derived metric I^2 . When no heterogeneity was observed, that is P -value of Q test (P_Q) greater than 0.1 and I^2 less than 50%, the pooled OR would be estimated using fixed-effect models. Otherwise, random-effect models would be applied for the analyses.

3. Results

Demographic characteristics of both Spanish populations are described in Table 1. Our cohort was showed to be similar overall to the cohort from Granada, but the average age at onset of

the affected subjects from our cohort was statistically younger than those from the Granada study (P -value = $1.4E-10$).

Table 1. Demographic characteristics of both Spanish populations

	Seville					Granada		
	Whole	PD			HC	Whole	PD	HC
		All	Familial	Non-familial				
N (% males)	1876 (57.46)	738 (58.54)	193 (59.07)	545 (58.35)	1138 (56.77)	432 (55.56)	240 (64.56)	192 (44.77)
Age (Y \pm SD)	61,2 \pm 14,4	65.4 \pm 11.5	62.6 \pm 11.7	66.5 \pm 11.3	58.4 \pm 15.3	63.14 \pm 13	67.4 \pm 11	58.07 \pm 14
AAO (Y \pm SD)	-	55.3 \pm 12.6	51.6 \pm 13.2	56.6 \pm 12.1	-	-	61.9 \pm 13.8	-

PD: Parkinson's Disease; HC: Healthy Controls; N: Number of samples; Y: Years; SD: Standard Deviation; AAO: Age At Onset.

3.1 Association analysis in Seville cohort

Each locus was genotyped in more than 95% of individuals. Around 4.5% of the samples were genotyped twice or thrice to check genotyping consistency, showing 100% identical genotypes. All variants were in Hardy-Weinberg equilibrium in the control individuals and all of them had minor allele frequency greater than 5%. The main genotypic features are shown in Table 2. The SNP rs17649553, in *MAPT*, had to be excluded because of assay failure. Therefore, it was replaced by the SNP rs9468, which was also in *MAPT* and in perfect linkage disequilibrium (LD) with the SNP rs17649553.

Table 2. Association analyses of the genotyped SNPs with PD in our population from Seville.

Loci	SNP	Chr	BP	Alleles		MAF		Allelic Association Analysis			N Genotypes (A ¹ A ¹ /A ¹ A ² /A ² A ²)		Genotypic Association Analysis (Adj)		
				A ¹ >A ²		PD	HC	χ^2	P_{Adj}	OR _{Adj} (95% CI)	PD	HC	$P_{genotypic}$	P_{dom}	P_{rec}
<i>MAPT</i>	rs9468*	17	43994648	T>C		0.236	0.306	21.4	6.9 e-6	0.7 (0.6 - 0.82)	433/260/44	532/473/102	2.64 e-5	7.24 e-6	0.0203
<i>ACMSD/TMEM163</i>	rs6430538	2	135539967	T>C		0.478	0.396	24.3	1.3 e-5	1.36 (1.18 - 1.56)	208/341/176	409/545/174	1.72 e-5	0.0049	7.38 e-6
<i>STK39</i>	rs1955337	2	169129145	G>T		0.169	0.132	9.66	0.0011	1.37 (1.14 - 1.66)	499/203/21	840/255/20	0.0048	0.0019	0.0798
<i>SREBF/RAI1</i>	rs11868035	17	17715101	G>A		0.275	0.312	5.75	0.0413	0.85 (0.73 - 0.99)	381/301/51	522/495/101	0.1044	0.1073	0.072
<i>MIR4697</i>	rs329648	11	133765367	C>T		0.434	0.418	0.93	0.2428	1.09 (0.95 - 1.25)	238/347/142	372/568/187	0.2323	0.726	0.0902

SNP: Single Nucleotide Polymorphism; *: In perfect linkage disequilibrium with rs17649553; Chr: Chromosome; BP: Base position in the chromosome; A¹: Reference allele; A²: Alternate allele; MAF: Minor Allele Frequency; PD: Parkinson's disease patients; HC: Healthy Controls, N: Number of samples; OR: Odds Ratio; CI: Confidence Interval; Adj: Adjusted for age; P_{geno} : Genotypic P -value; P_{dom} : P -value for dominant model; P_{rec} : P -value for recessive model.

The genotype distribution of three SNPs, rs9468 (*MAPT*), rs6430538 (*ACMSD/TMEM16*) and rs1955337 (*STK39*), were significantly associated to PD (adjusted P -value < 0.01) in our

population (Table 2). Interestingly, the SNP rs9468 was associated with a decreased risk of PD in the dominant model (P -value = 7.24E-6; OR = 0.64; CI: 0.53 - 0.78), while the SNP rs6430538 was associated with an increased risk of PD in the recessive model (P -value = 7.38E-6; OR = 1.74; CI: 1.4 - 2.2). The SNP rs1955337 did not reach this high association with the PD risk, but a P -value < 0.05 was observed in the dominant model (P -value = 1.9E-3; OR = 1.4; CI: 1.13 – 1.74). The other SNPs failed the association with PD in our population from Seville (Table 2).

3.2 Meta-analysis with Granada cohort

Finally, 944 PD patients (206 from the Granada cohort and 738 from our cohort) and 1310 healthy controls (172 from the Granada cohort and 1138 from our cohort) were included in the meta-analysis study. All variants showed similar minor allele frequency in both Spanish populations (Supplementary Table). A certain degree of heterogeneity was observed only for allele contribution OR of the SNP rs329648. Our analysis revealed that no variant was PD related at genome-wide significance level ($P < 5.0E-8$) in the Spanish population. However, rs6430538 (*ACMSD/TMEM163*) and rs9468 (*MAPT*) showed a suggestive adjusted for age P -values. This study also suggests that loci *SREBF/RAI1*, *STK39* and *MIR4697* are not related with PD, at least in population from the South of Spain (Table 3).

Table 3. Results from meta-analysis with Southern Spanish cohorts

Locus	SNP	P_Q	I^2	Het obs?	P -value	OR (95% CI)
<i>MAPT</i>	rs9468*	0.6475	0.00	No	5.70E-7	0.69 (0.6 – 0.8)
<i>ACMSD/TMEM163</i>	rs6430538	0.5405	0.00	No	1.04E-6	1.38 (1.21 – 1.57)
<i>STK39</i>	rs1955337	0.1963	40.10	No	2.43E-4	1.41 (1.18 – 1.70)
<i>SREBF/RAI1</i>	rs11868035	0.2451	25.97	No	9.10E-3	0.83 (0.72 – 0.95)
<i>MIR4697</i>	rs329648	0.1292	56.56	Yes	0.1818	1.2 (0.92 – 1.56)

SNP, Single Nucleotide Polymorphism; P_Q , P -value of Cochran's Q test; Het obs, Heterogeneity between the studies observed; OR, Odds Ratio; CI, Confidence Interval.

The forest plots of the meta-analysis for the allelic model in all SNPs are shown in Supplementary Figures 1, 2, 3, 4, and 5 (for rs9468, rs6430538, rs1955337, rs11868035, and rs329648, respectively).

4. Discussion

In this study we have analyzed the association of 5 loci, previously reported PD related in a cohort from Granada, in the South of Spain [1]. The sample size in that previous study was the limitation of the analysis. Therefore, we have included more than double the PD cases from the same area of Spain, which solves this limitation. In addition, the ancestry is homogeneous in both cohorts, since both are from nearby areas of southern Spain.

Here we have reported that two of the previously reported loci related to PD in the Bandrés-Ciga cohort (*SREBF/RAI1* and *MIR4697*) failed the association analysis in our bigger cohort. However, we replicate the association of the loci *ACMSD/TMEM163* and *MAPT* with a strong significance level (Bonferroni corrected P -value < 0.05). *STK39* showed a moderate association with an uncorrected P -value < 0.05 .

In the meta-analysis study with both cohorts from the south of Spain, no loci reached a genome-wide significance level, but *ACMSD/TMEM163* and *MAPT* loci still showed a suggestive uncorrected P -value $< 5.0E-6$. *STK39* showed a moderate association level, and the last two loci also failed the association analysis with PD.

These findings reinforce the complex genetic architecture of PD risk and that replication studies in large independent cohorts are necessary to support association reports from population studies.

Here, we have replicated (using an independent technique) only 2 of those 5 loci previously associated with PD in a similar cohort from Spain.

From our point of view, the association of PD with rs6430538, in proximity of *ACMSD* is particularly interesting. *ACMSD* is involved in tryptophan metabolism and mutations in its gene

have been associated with neurological disease with features of parkinsonism. In addition, it has been shown that *ACMSD* looks to exert a critical regulatory effect on excitotoxic and inflammatory pathways in the CNS and has been proposed as a novel therapeutic target in Parkinson's disease [12]. Therefore, additional studies need to be followed in order to confirm the implication of this gene in the pathogenesis of PD.

5. Conclusion

Results from our study suggest the association of two loci (*ACMSD/TMEM16* and *MAPT*) with PD in population from the south of Spain, and that *MAPT* gene polymorphism might be a protective factor against PD, while *ACMSD* gene polymorphism would be a risk factor for PD. Our results provide powerful information for tracking the most viable gene candidates to take into account in order to establish cumulative genetic risk scores specific for populations.

Acknowledgments

We would like to thank the patients and healthy individuals who participated in this study. This research has been conducted using samples from the *HUVR-IBiS Biobank (Andalusian Public Health System Biobank and ISCIII-Red de Biobancos PT13/0010/0056)*. We also thank Ismael Huertas for his valuable comments and suggestions.

Author Contributions

PGG and PM conceived, designed, and supervised the study. CTP and PGG wrote the paper. SJ, ADAG, and DMG performed the clinical analysis. DBR and GOA performed the genotyping experiments. MTP analyzed the data.

Conflict of interests:

The authors declare no conflict of interest.

Financial disclosure

This study was supported by grants from the Spanish Ministry of Economy and Competitiveness [PI14/01823, PI16/01575, PI18/01898], co-founded by ISCIII (Subdirección General de Evaluación y Fomento de la Investigación) and by Fondo Europeo de Desarrollo Regional (FEDER), the Consejería de Economía, Innovación, Ciencia y Empleo de la Junta de Andalucía [CVI-02526, CTS-7685], the Consejería de Salud y Bienestar Social de la Junta de Andalucía [PI-0437-2012, PI-0471-2013], the Sociedad Andaluza de Neurología, the Jacques and Gloria Gossweiler Foundation, the Fundación Alicia Koplowitz, the Fundación Mutua Madrileña. Pilar Gómez-Garre was supported by the "Miguel Servet" (from ISCIII-FEDER) and "Nicolás Monardes" (from Andalusian Ministry of Health) programmes. Silvia Jesús Maestre was supported by the "Juan Rodés" programme and Daniel Macías-García was supported by the "Río Hortega" programme (both from ISCIII-FEDER). Cristina Tejera-Parrado was supported by VPPI-US from the Universidad de Sevilla.

References

- [1] S. Bandres-Ciga, T.R. Price, F.J. Barrero, F. Escamilla-Sevilla, J. Pelegriña, S. Arepalli, D. Hernandez, B. Gutierrez, J. Cervilla, M. Rivera, A. Rivera, J.H. Ding, F. Vives, M. Nalls, A. Singleton, R. Duran, Genome-wide assessment of Parkinson's disease in a Southern Spanish population, *Neurobiol Aging* 45 (2016) 213 e213-213 e219.
- [2] T.L. Edwards, W.K. Scott, C. Almonte, A. Burt, E.H. Powell, G.W. Beecham, L. Wang, S. Zuchner, I. Konidari, G. Wang, C. Singer, F. Nahab, B. Scott, J.M. Stajich, M. Pericak-Vance, J. Haines, J.M. Vance, E.R. Martin, Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease, *Ann Hum Genet* 74 (2010) 97-109.

- [3] W.R. Gibb, A.J. Lees, The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease, *J Neurol Neurosurg Psychiatry* 51 (1988) 745-752.
- [4] D.G. Hernandez, M.A. Nalls, P. Ylikotila, M. Keller, J.A. Hardy, K. Majamaa, A.B. Singleton, Genome wide assessment of young onset Parkinson's disease from Finland, *PLoS One* 7 (2012) e41859.
- [5] C. Klein, A. Westenberger, Genetics of Parkinson's disease, *Cold Spring Harb Perspect Med* 2 (2012) a008888.
- [6] C.M. Lill, J.T. Roehr, M.B. McQueen, F.K. Kavvoura, S. Bagade, B.M. Schjeide, L.M. Schjeide, E. Meissner, U. Zauft, N.C. Allen, T. Liu, M. Schilling, K.J. Anderson, G. Beecham, D. Berg, J.M. Biernacka, A. Brice, A.L. DeStefano, C.B. Do, N. Eriksson, S.A. Factor, M.J. Farrer, T. Foroud, T. Gasser, T. Hamza, J.A. Hardy, P. Heutink, E.M. Hill-Burns, C. Klein, J.C. Latourelle, D.M. Maraganore, E.R. Martin, M. Martinez, R.H. Myers, M.A. Nalls, N. Pankratz, H. Payami, W. Satake, W.K. Scott, M. Sharma, A.B. Singleton, K. Stefansson, T. Toda, J.Y. Tung, J. Vance, N.W. Wood, C.P. Zabetian, C. and Me Genetic Epidemiology of Parkinson's Disease, C. International Parkinson's Disease Genomics, G.C. Parkinson's Disease, C. Wellcome Trust Case Control, P. Young, R.E. Tanzi, M.J. Khoury, F. Zipp, H. Lehrach, J.P. Ioannidis, L. Bertram, Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database, *PLoS Genet* 8 (2012) e1002548.
- [7] M.A. Nalls, N. Pankratz, C.M. Lill, C.B. Do, D.G. Hernandez, M. Saad, A.L. DeStefano, E. Kara, J. Bras, M. Sharma, C. Schulte, M.F. Keller, S. Arepalli, C. Letson, C. Edsall, H. Stefansson, X. Liu, H. Pliner, J.H. Lee, R. Cheng, C. International Parkinson's Disease Genomics, G.I. Parkinson's Study Group Parkinson's Research: The Organized, and Me, GenePd, C. NeuroGenetics Research, G. Hussman Institute of Human, I. Ashkenazi Jewish Dataset, H. Cohorts for, E. Aging Research in Genetic, C. North American Brain

- Expression, C. United Kingdom Brain Expression, C. Greek Parkinson's Disease, G. Alzheimer Genetic Analysis, M.A. Ikram, J.P. Ioannidis, G.M. Hadjigeorgiou, J.C. Bis, M. Martinez, J.S. Perlmutter, A. Goate, K. Marder, B. Fiske, M. Sutherland, G. Xiromerisiou, R.H. Myers, L.N. Clark, K. Stefansson, J.A. Hardy, P. Heutink, H. Chen, N.W. Wood, H. Houlden, H. Payami, A. Brice, W.K. Scott, T. Gasser, L. Bertram, N. Eriksson, T. Foroud, A.B. Singleton, Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease, *Nat Genet* 46 (2014) 989-993.
- [8] N. Pankratz, G.W. Beecham, A.L. DeStefano, T.M. Dawson, K.F. Doheny, S.A. Factor, T.H. Hamza, A.Y. Hung, B.T. Hyman, A.J. Iverson, D. Krainc, J.C. Latourelle, L.N. Clark, K. Marder, E.R. Martin, R. Mayeux, O.A. Ross, C.R. Scherzer, D.K. Simon, C. Tanner, J.M. Vance, Z.K. Wszolek, C.P. Zabetian, R.H. Myers, H. Payami, W.K. Scott, T. Foroud, P.G. Consortium, Meta-analysis of Parkinson's disease: identification of a novel locus, *RIT2*, *Ann Neurol* 71 (2012) 370-384.
- [9] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. de Bakker, M.J. Daly, P.C. Sham, PLINK: a tool set for whole-genome association and population-based linkage analyses, *Am J Hum Genet* 81 (2007) 559-575.
- [10] W. Satake, Y. Nakabayashi, I. Mizuta, Y. Hirota, C. Ito, M. Kubo, T. Kawaguchi, T. Tsunoda, M. Watanabe, A. Takeda, H. Tomiyama, K. Nakashima, K. Hasegawa, F. Obata, T. Yoshikawa, H. Kawakami, S. Sakoda, M. Yamamoto, N. Hattori, M. Murata, Y. Nakamura, T. Toda, Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease, *Nat Genet* 41 (2009) 1303-1307.

- [11] M. Sharma, J.P. Ioannidis, J.O. Aasly, G. Annesi, A. Brice, C. Van Broeckhoven, L. Bertram, M. Bozi, D. Crosiers, C. Clarke, M. Facheris, M. Farrer, G. Garraux, S. Gispert, G. Auburger, C. Vilarino-Guell, G.M. Hadjigeorgiou, A.A. Hicks, N. Hattori, B. Jeon, S. Lesage, C.M. Lill, J.J. Lin, T. Lynch, P. Lichtner, A.E. Lang, V. Mok, B. Jasinska-Myga, G.D. Mellick, K.E. Morrison, G. Opala, P.P. Pramstaller, I. Pichler, S.S. Park, A. Quattrone, E. Rogaeva, O.A. Ross, L. Stefanis, J.D. Stockton, W. Satake, P.A. Silburn, J. Theuns, E.K. Tan, T. Toda, H. Tomiyama, R.J. Uitti, K. Wirdefeldt, Z. Wszolek, G. Xiromerisiou, K.C. Yueh, Y. Zhao, T. Gasser, D. Maraganore, R. Kruger, G.-P. Consortium, Large-scale replication and heterogeneity in Parkinson disease genetic loci, *Neurology* 79 (2012) 659-667.
- [12] K. Thirtamara-Rajamani, P. Li, M.L. Escobar Galvis, V. Labrie, P. Brundin, L. Brundin, Is the Enzyme ACMSD a Novel Therapeutic Target in Parkinson's Disease?, *J Parkinsons Dis* 7 (2017) 577-587.
- [13] A. Verstraeten, J. Theuns, C. Van Broeckhoven, Progress in unraveling the genetic etiology of Parkinson disease in a genomic era, *Trends Genet* 31 (2015) 140-149.

Supplementary figures captions

Supplementary Figure 1. Forest plot of rs9468* meta-analysis on Parkinson's disease. Results of meta-analysis for allelic model.

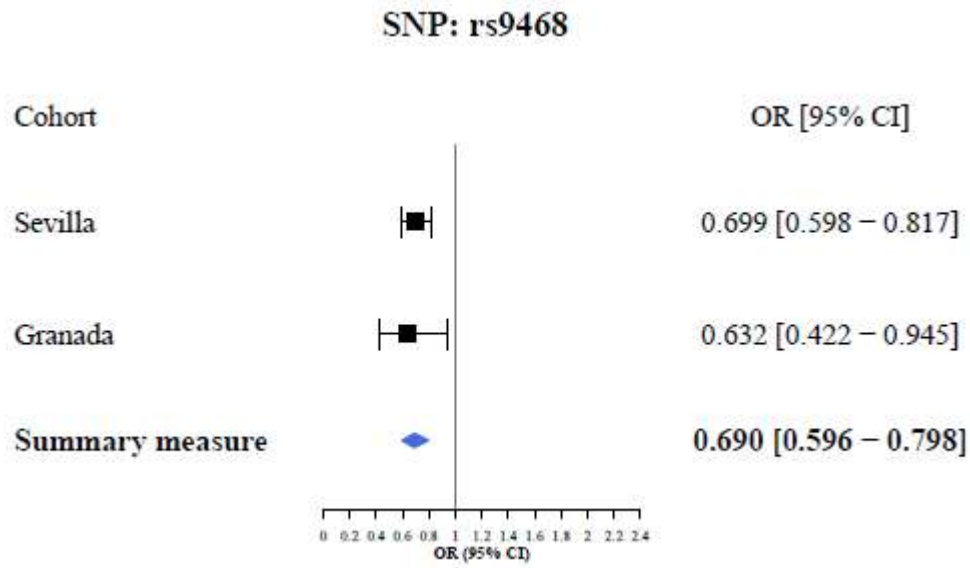
Supplementary Figure 2. Forest plot of rs6430538 meta-analysis on Parkinson's disease. Results of meta-analysis for allelic model.

Supplementary Figure 3. Forest plot of rs1955337 meta-analysis on Parkinson's disease. Results of meta-analysis for allelic model.

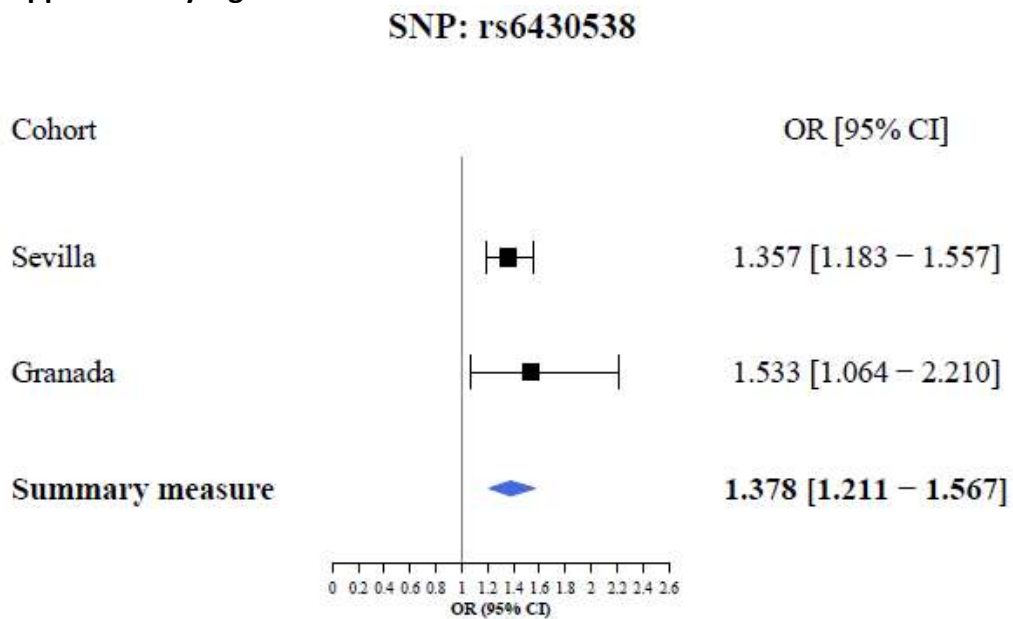
Supplementary Figure 4. Forest plot of rs11868035 meta-analysis on Parkinson's disease. Results of meta-analysis for allelic model.

Supplementary Figure 5. Forest plot of rs329648 meta-analysis on Parkinson's disease. Results of meta-analysis for allelic model.

Supplementary Figure 1.

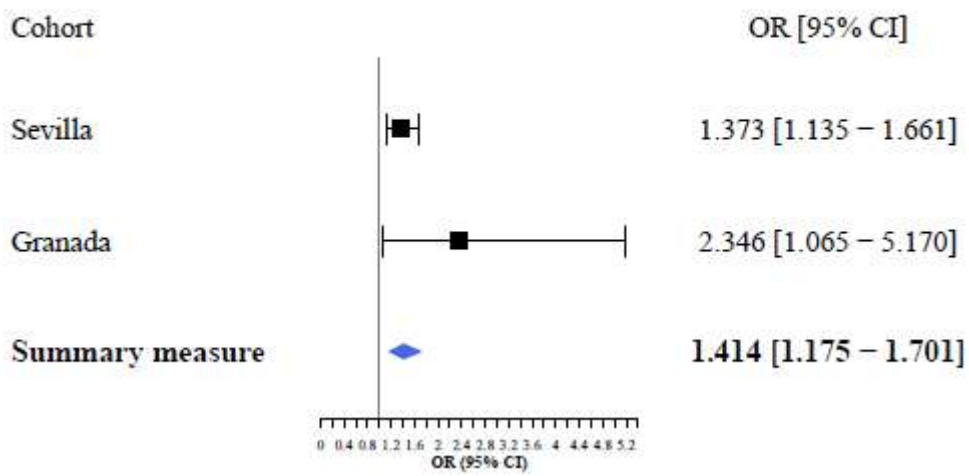


Supplementary Figure 2.



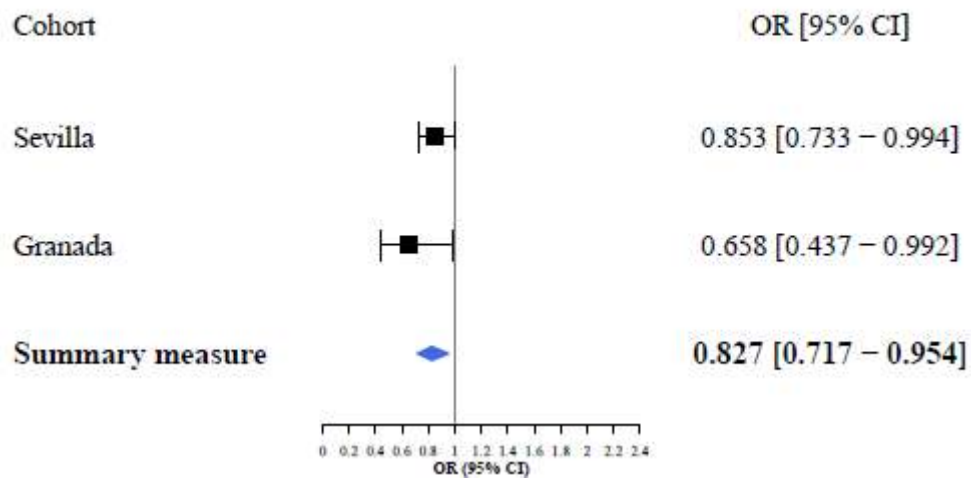
Supplementary Figure 3.

SNP: rs1955337



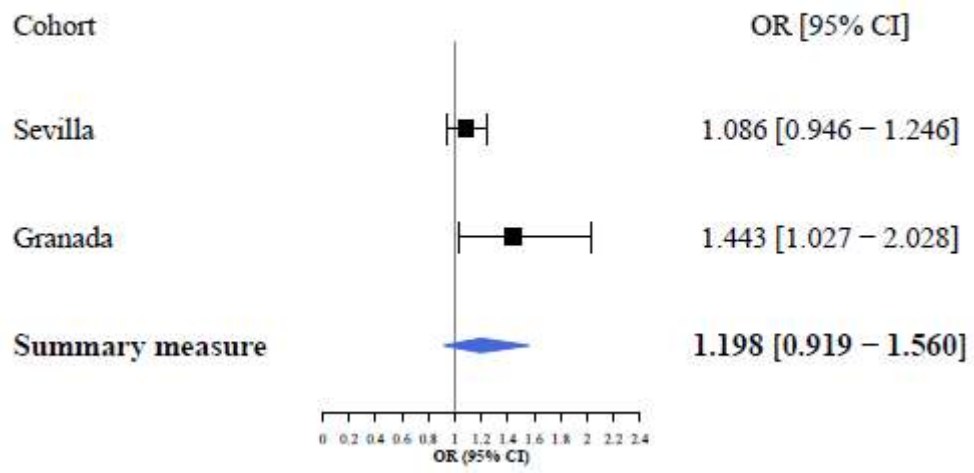
Supplementary Figure 4.

SNP: rs11868035



Supplementary Figure 5.

SNP: rs329648



Supplementary Table. Data from Association analysis in Seville cohort, Granada cohort and Nalls' meta-analysis cohort.

Locus	SNP	Seville cohort			Granada cohort			Nalls' meta-analysis cohort		
		M A	MAF	OR	M A	MAF	OR	M A	MAF	OR
<i>MAPT</i>	rs1764955 3	C*	0.27 8	0.69 9	T	0.30 7	0.63 2	T	0.22 6	0.76 9
<i>ACMSD/TMEM16</i> 3	rs6430538	C	0.42 8	1.35 7	C	0.41 1	1.53 3	T	0.43 0	0.87 5
<i>STK39</i>	rs1955337	T	0.14 7	1.37 3	T	0.17 9	2.34 6	T	0.12 8	1.21 4
<i>SREBF/RAI1</i>	rs1186803 5	A	0.29 7	0.85 3	A	0.29 2	0.65 8	A	0.29 8	0.93 9
<i>MIR4697</i>	rs329648	T	0.42 4	1.08 6	T	0.43 4	1.44 3	T	0.35 4	1.10 5

SNP, Single Nucleotide Polymorphism; MA, Minor Allele; MAF, Minor Allele Frequency; OR, Odds Ratio;

*Due to problems with the genotyping of the SNP rs17649553, in our study it was replaced by rs9468, in perfect LD.

2.2. Resumen

En los últimos años, muchos estudios GWA han identificado *loci* de riesgo que contribuyen al desarrollo de la EP. Pero no fue hasta 2014 que Nalls y col. realizaron un extenso meta-análisis de todos los estudios GWA publicados hasta la fecha en población de descendencia europea, logrando identificar 24 *loci* de riesgo para la EP. Para confirmar estos resultados es necesario que se realicen sucesivos estudios de replicación en diferentes cohortes y grupos étnicos. Bandrés-Ciga y col. en 2016 estudiaron los factores genéticos de riesgo asociados a la EP en población de Granada, en el sur de España, a través de un GWAS. De los 24 *loci* identificados por Nalls y col., encontraron que sólo 5 de ellos se mantenían asociados a la EP (p nominal <0,05). Sin embargo, los resultados de este estudio son limitados debido al pequeño tamaño de la cohorte analizada (240 pacientes de EP y 192 controles).

Por ello, el propósito de este trabajo fue investigar la asociación de estos 5 *loci* de riesgo identificados en la población de Bandrés-Ciga y col. (*ACMSD/TMEM163*, *MIR4697*, *SREBF/RAI1*, *MAPT* y *STK39*) en una población mayor de EP, también del sur de España. Además, se realizó un meta-análisis combinando los resultados de ambas cohortes con el fin de comprender mejor el papel de estos *loci* en población del sur de España. Para ello se cuenta con una cohorte de 738 pacientes de EP de la Unidad de Trastornos del Movimiento del Hospital Universitario Virgen del Rocío (Sevilla, España), y 1138 controles sanos.

Tres SNPs se asociaron significativamente a la EP en nuestra población: rs6430538 (*ACMSD/TMEM16*), rs9468 (*MAPT*) y rs1955337 (*STK39*), con una *p* ajustada por edad < 0,01. El SNP rs9468 fue asociado a una disminución de riesgo de EP siguiendo un modelo dominante ($p=7,24 \times 10^{-6}$; OR=0,64; CI, 0,53 - 0,78), mientras que el SNP rs6430538 se asoció a un incremento de riesgo de EP siguiendo un modelo recesivo ($p=7,38 \times 10^{-6}$; OR=1,74; CI, 1,4 - 2,2). El SNP rs1955337 no alcanzó una asociación tan significativa, pero sí mostró un patrón de herencia dominante con una $p < 0,05$ ($p=0,0019$; OR= 1,4; CI, 1,13 - 1,74). Los dos SNPs restantes no mostraron estar asociados a la EP en este estudio.

Para realizar el meta-análisis se contaron finalmente con 944 pacientes de EP (206 de Bandrés-Ciga y col. y 738 de nuestra cohorte) y 1310 controles sanos (172 de Bandrés-Ciga y col. y 1138 de nuestra cohorte). Ninguno de los SNPs analizados mostraron estar asociados a la EP a un nivel de significación de genoma completo ($p < 5 \times 10^{-8}$), aunque los SNPs rs6430538 (*ACMSD/TMEM16*) y rs9468 (*MAPT*) sí mostraron una *p* ajustada por edad sugestiva ($p < 5 \times 10^{-6}$).

Se ha conseguido replicar en este estudio la asociación de los *loci ACMSD/TMEM16* y *MAPT*. El *loci STK39* mostró una asociación moderada. Y en cuanto a los *loci MIR4697* y *SREBF/RAI1*, no se ha replicado su asociación con la EP, al menos en población del sur de España.

IV. DISCUSIÓN GENERAL

El esclarecimiento de las bases genéticas que subyacen tras las enfermedades complejas es una cuestión de gran interés, objeto de una intensa investigación por parte de la comunidad científica. Esta tesis doctoral se ha centrado en el estudio de factores genéticos que se encuentran involucrados en la predisposición a padecer la segunda enfermedad neurodegenerativa más frecuente en la población, la enfermedad de Parkinson. Y es que, a pesar de los avances logrados en los últimos años, la etiopatogenia de la enfermedad continúa sin conocerse en gran medida. Además, en las enfermedades neurodegenerativas en las que la enfermedad va avanzando a lo largo del tiempo, resulta de crucial importancia su detección temprana cuando el proceso neurodegenerativo ya ha comenzado pero no hay síntomas que lleven al diagnóstico de la enfermedad. Identificar la enfermedad en esta ventana temporal resulta de gran importancia para poder poner en práctica una terapia neuroprotectora. La aplicación temprana de este tipo de terapia, podría disminuir considerablemente el impacto de la enfermedad y su evolución. La mejora del conocimiento de las bases genéticas de las enfermedades complejas ayuda a la comprensión de las bases moleculares y de los mecanismos fisiopatológicos subyacentes. Con este tipo de estudios se puede contribuir a la identificación de nuevos marcadores útiles para el diagnóstico temprano, mejor pronóstico, tratamiento personalizado e incluso constituir nuevas dianas terapéuticas contra las que diseñar nuevos fármacos. Este trabajo se ha enfocado por una parte en el estudio de la contribución de genes descritos causativos o de riesgo para la EP, y por otra parte, en el estudio de replicación de 5 *loci* de riesgo previamente asociados con la EP en población del sur de España. Con ello se pretende contribuir a esclarecer los factores genéticos involucrados en la EP.

1. Objetivo 1: Estudiar la implicación en la fisiopatogenia de la EP de genes causativos de formas monogénicas de EP.

1.1. Sub-objetivo 1: Estudiar exhaustivamente 17 genes relacionados con la EP y otros parkinsonismos en pacientes de inicio temprano de la Comunidad de Madrid.

En este trabajo se han examinado 17 genes previamente relacionados con la EP mediante secuenciación masiva y MLPA en 117 pacientes de EP procedentes de la región central de España. 27 pacientes de los 117 estudiados (23,08%) tenían variaciones potencialmente patogénicas en genes robustamente asociados con la EP.

Las variaciones patogénicas en el gen *PARK2* son la causa más frecuente de EPIT de herencia autosómica recesiva, con una prevalencia variable entre los diferentes estudios (Macedo y col., 2009; Koziorowski y col., 2013; Alcalay y col., 2010). En el presente estudio, la causa más frecuente de EP fueron las mutaciones en este gen (13,68%, 16 pacientes de 117), siendo una variación conocida, la p.N52fs, la más frecuente (6,4%, 15/234; Spataro y col 2017). En un trabajo previo en 72 pacientes de EPIT del País Vasco sólo se localizó esta variación en homocigosis en 1 paciente (1,4%, 2/244; Gorostidi y col., 2016). Sí que fue similar la frecuencia con la que se identificó en un estudio previo del sur de España de 26 pacientes de EPIT (5,8%; 3/52; Bandrés-Ciga y col., 2016). Estos resultados sugieren que aunque esta mutación está altamente extendida por todo el mundo, es menos prevalente en población vasca y podría estar asociada al efecto de un fundador y la consiguiente dispersión de la misma.

Resulta interesante un paciente (EP-17) que tenía la mutación p.N52fs en homocigosis en el gen *PARK2*, y también era portador de una variación genética en *ATP13A2*

(c.2762+21T>C), considerada patogénica por las herramientas bioinformáticas afectando al sitio de *splicing*. La edad de inicio de este paciente fue menor que la del resto de pacientes portadores de la mutación p.N52fs, lo que podría indicar un efecto aditivo entre ambas variaciones. Variaciones en el gen *ATP13A2* se han descrito anteriormente relacionadas con ciertos tipos de parkinsonismo de inicio temprano (Ramírez y col., 2006). Aunque el papel de estas variaciones en heterocigosis continúa bajo estudio, se ha descrito su efecto como factor de riesgo, y se ha sugerido también que las neuronas dopaminérgicas que sobreexpresan la proteína ATP13A2 son menos susceptibles a la muerte celular (Ysselstein y col., 2019). Sin embargo, debido a la baja frecuencia de estas variaciones, para confirmar estas hipótesis es necesario secuenciar poblaciones mayores.

Por otra parte, las mutaciones en *LRRK2*, una proteína multidominio con función GTPasa y kinasa, es la causa genética más frecuente de EP de herencia autosómica dominante. Se han identificado más de 100 variaciones asociadas a la EP en este gen. La más frecuente es la p.G2019S y su prevalencia es muy variable según la población (Bandrés-Ciga y col., 2016, Lill, 2016, Hernández y col., 2016). En caucásicos su prevalencia está en torno al 4%, que es similar a la que se ha encontrado en este estudio (5 pacientes de 117, 4,27%). En total el 6,4% de los pacientes de este estudio tenían variaciones potencialmente patogénicas en este gen, sin diferencia en la edad de inicio respecto a aquellos que tenían variaciones en el gen *PARK2*. Nuestros resultados muestran el papel tan importante de *LRRK2* en EPIT en población española.

Las variaciones en el gen *GIGYF2* se han relacionado anteriormente con EP familiar (Lautier y col., 2008), aunque no se ha demostrado en todos los casos (Bras y col., 2009; Guella y col., 2011; Bonetti y col., 2009). Un paciente de este estudio (EP-33) que tenía la variación c.4536+3A>G en *LRRK2*, que afecta al sitio de *splicing*, también era

portador de una variación en *GIGYF2* clasificada como patogénica (p.R1225C). El paciente presentaba deterioro cognitivo con edad de inicio tardía, similar a los síntomas que presentaba una familia española en otro estudio, también con una variación potencialmente patogénica en este gen (Ruíz-Martínez y col., 2015). Se sabe que ambos genes están implicados en el proceso de autofagia en la célula (Kim y col., 2015). Esta co-ocurrencia de variaciones en ambos genes se ha descrito previamente y se ha asociado con EPIT, sugiriendo la acción de efectos aditivos (Lautier y col., 2008). Sin embargo, son necesarios más estudios para esclarecer el papel de estas variaciones en la EP.

En cuanto a los factores de riesgo, el gen *GBA* es el mejor conocido en la EP, y está presente en múltiples étnias (Jesús y col., 2016; Sidransky y col., 2009). Las mutaciones de riesgo en el gen *GBA* se heredan de manera autosómica dominante, tienen una penetrancia alta y se relacionan con EPIT (Ferreira y Mossano, 2016). Las variaciones patogénicas más frecuentes son la L444P y la N370S (Zhao y col., 2016), pero ninguna se ha localizado en este estudio. Esto puede deberse a la dificultad para secuenciar *GBA* empleando secuenciación dirigida ya que existe un pseudo gen, con secuencia similar. Además, *GBA* tiene distinto impacto en la EP según su patogenicidad (puede ser deletéreo o benigno). Puesto que hemos seleccionado aquellas variaciones con posible efecto patogénico y VOUS, hemos podido subestimar la cantidad de variaciones en *GBA* relacionadas con la EP en nuestra población.

Este trabajo se ha enfocado en el estudio de variaciones exónicas e intrónicas, sin embargo, la mayoría de las variaciones localizadas se encontraban en otras regiones del gen (principalmente 3'-UTR). De hecho, hemos identificado algunos pacientes con variaciones potencialmente patogénicas en heterocigosis en estas regiones no

traducidas. El análisis exhaustivo de estas variaciones y su contribución al desarrollo de la EP queda pendiente para un futuro trabajo.

Este estudio es el primero en analizar sistemáticamente la variabilidad genética de genes relacionados con la EPIT en pacientes de España, teniendo en cuenta el número de pacientes incluidos, el número de genes analizados y las técnicas empleadas. Con nuestros resultados se pretende contribuir a completar los resultados de estudios previos para comprender mejor el perfil genético que hay detrás del desarrollo de la EP a edad temprana. Además, se evidencia una posible interacción de los genes *ATP13A2* y *PARK2*, de manera que un paciente que tenía variaciones potencialmente patogénicas en ambos genes parece desarrollar EP mucho antes que los que tenían variaciones en uno sólo de los genes. Por último, se demuestra que el papel de *LRRK2* en la EPIT podría ser mayor en población española que el que se creía hasta ahora. Nuestros resultados por tanto, podrían ser útiles para su aplicación en el consejo genético.

1.2. Sub-objetivo 2: Estudiar la implicación en la fisiopatogenia de la EP de *CHCHD2*, nuevo gen propuesto causativo de EP, en población del sur de España.

El papel de *CHCHD2* en la EP continua bajo estudio ya que desde su identificación en 2015 por Funayama y col. como gen causativo de EP de herencia autosómica dominante, se han publicado tanto trabajos en los que se corrobora esta hipótesis (Foo y col., 2015; Funayama y col., 2015; Shi y col., 2016), como trabajos en los que *CHCHD2* no se muestra como gen causativo de EP, ni tampoco portador de variaciones de riesgo (Fan y col., 2016; Lu y col., 2016; Zhang y col., 2016). Aquí, en este trabajo hemos identificado 4 variaciones genéticas no exónicas (c.-9T>G, c.-34C>A, c.-65T>G,

y c.*125G>A) no asociadas a la EP en la población de este estudio procedente del sur de España. Estas variaciones se habían identificado previamente en otras cohortes sin observarse tampoco diferencia en sus frecuencias en pacientes de EP frente a controles (Fan y col., 2016; Lu y col., 2016; Shi y col., 2016; Wu y col., 2016; Zhang y col., 2016, Ogaki y col., 2015). Funayama y col. en cambio, observaron las variaciones c.-9T>G y c.*125G>A asociadas a la EP en población japonesa (Funayama y col., 2015), lo que podría deberse a diferencias en las frecuencias alélicas entre diferentes etnias. De hecho, en un estudio multicéntrico en población caucásica en el que se examinaron más de 2000 casos de EP y de demencias con cuerpos de Lewy, la frecuencia de estas variaciones fueron mucho más altas (23%) que en población japonesa (4%), lo que sugiere la existencia de un haplotipo diferente entre ambas poblaciones. La asociación observada con la EP en el estudio de Funayama y col. podría por tanto ser específica de población japonesa (Ogaki y col., 2015). En nuestra población de estudio, de acuerdo con otras cohortes, parece que variaciones en el gen *CHCHD2* no suponen un factor de riesgo para la EP.

Aunque el papel del gen *CHCHD2* en la EP no está aún claro, estudios funcionales de este gen señalan su implicación en la función mitocondrial y el estrés oxidativo, ambos procesos relacionados con la pérdida neuronal en la EP. La proteína codificada por este gen se encuentra en el espacio intermembrana de la mitocondria y parece que su función está relacionada con la producción de energía a través de la fosforilación oxidativa. En este proceso, el citocromo c situado en el espacio intermembrana, es el encargado de transferir electrones desde el complejo III al IV, que finalmente reduce moléculas de oxígeno cargadas de electrones a agua, evitando la creación de especies oxidativas. *CHCHD2* podría estabilizar al citocromo c en el espacio intermembrana para mantener

el flujo de la cadena de electrones, ya que cuando hay una deficiencia de este gen en moscas *knockout*, el citocromo c se ve afectado, la cadena de transporte de electrones se ve comprometida y el exceso de electrones reacciona con el oxígeno molecular produciendo superóxido y peróxido de hidrógeno, la mayor causa de estrés oxidativo (Meng y col., 2017). Esto conlleva a la pérdida de neuronas dopaminérgicas y al desarrollo de disfunciones motoras con la edad en estas moscas. La expresión en estos animales del gen *CHCHD2* portando mutaciones descritas patogénicas en humanos, no revierten del todo los defectos producidos por la pérdida del gen, lo que señala la patogenicidad de dichas mutaciones. Así mismo, parece que este gen está también relacionado con los genes *PINK1* y *PARK2* ya que cuando se sobreexpresan aumenta la mitofagia reconociendo el daño mitocondrial existente (Meng y col., 2017).

Con todo esto, a pesar de que los resultados poblacionales no son concluyentes sobre la asociación de variaciones en el gen *CHCHD2* y la EP, o si ésta es dependiente de la etnia, parece que los estudios funcionales apuntan a que *CHCHD2* juega un papel fundamental en la función mitocondrial, evitando el daño oxidativo.

1.3. Sub-objetivo 3: Estudiar la implicación en la fisiopatogenia de la EP de *TMEM230*, nuevo gen propuesto causativo de EP, en población del sur de España.

En este estudio se han localizado 13 variaciones genéticas en el gen *TMEM230* en nuestra población procedente del sur de España, aunque ninguna parece estar asociada a la EP, ya que se encuentran sólo en controles sanos o en controles sanos y casos de Parkinson en similar proporción. La excepción es la variación p.M1V que provoca la pérdida del codón de inicio en la isoforma más abundante y se encuentra en un solo

paciente de EP. Sin embargo, otra variación identificada en este estudio provoca el mismo cambio en la proteína (p.M1T), y sin embargo es un conocido polimorfismo descrito previamente presente en controles (Buongarzone y col., 2017). Además, el siguiente aminoácido es una metionina que puede actuar de codón de inicio alternativo, por lo que se perdería sólo un aminoácido en la proteína final. La implicación de *TMEM230* en la EP, por tanto, debe ser mejor explorada, en ésta y otras poblaciones.

La primera vez que este gen fue identificado como causa de EP de herencia autosómica dominante en una familia canadiense de 15 miembros afectados de EP, la cosegregación con la enfermedad se cumplió en todos sus miembros (Deng y col., 2016). Además, la mutación causativa propuesta, p.R141L, se encontraba en una región altamente conservada de la proteína. Otra mutación fue propuesta previamente como causa de EP en esta familia, p.N855S en el gen *DNAJC13*, sin embargo esta mutación no se encontraba presente en 3 miembros afectados y se encontraba en un familiar sano, lo que pone en duda su implicación en la EP (Vilariño-Güell y col., 2014).

A nivel celular, *TMEM230* se localiza en la red del trans Golgi y en la vía de reciclaje de vesículas, por lo que parece que esté involucrado en la vía de reciclaje de la red trans Golgi hacia los endosomas, el cual es mediado precisamente por la proteína VPS35, relacionada con la EP. De hecho, algunos estudios muestran que estas dos proteínas parcialmente colocalizan (Vilariño-Güell y col., 2011, Zimprich y col., 2011). Alteraciones en el tráfico y reciclaje de vesículas son procesos bien conocidos implicados en el desarrollo de la EP, a través de genes vinculados a esta patología como *LRRK2*, *SNCA*, *SYNJ1* y *DNAJC6*. En otros estudios se ha visto además que *DNAJC13* también puede estar implicado en esta ruta (Seaman y col., 2014), originando todavía más confusión con su papel en la EP.

En 15 estudios publicados hasta la fecha en este gen, aproximadamente el 0,28% de los pacientes de EP examinados contenían variaciones genéticas potencialmente patogénicas en *TMEM230*, localizándose la mayoría en la isoforma 2, la más abundante y conservada (Deng y col., 2018). Sin embargo, en muchos otros estudios publicados, el gen *TMEM230* no se encuentran relacionados con la EP (Yan y col., 2017, Wu y col., 2017, Fan y col., 2017, He y col., 2017, Shi y col., 2017, Ma y col., 2017, Buongarzone y col., 2017, Conedera y col., 2018, Ibanez y col., 2017). Además, algunas variaciones como la p.*184PGext*5 localizada en 7 casos familiares independientes procedentes de China, fueron localizadas tanto en homocigosis como en heterocigosis (Deng y col., 2016). Esto podría explicarse si su penetrancia fuera incompleta, pero entonces debería estar presente también en controles, y sin embargo no se encontró en ninguno de los 10434 controles sanos de China estudiados.

La inconsistencia en los estudios sobre el papel de *TMEM230* sugiere que, o bien *TMEM230* no es un gen relacionado con la EP, o que su penetrancia es incompleta y su efecto depende del fondo genético de la persona o de otros factores epigenéticos o ambientales, como podría ser una posible interacción con el gen *DNAJC13*. Sin embargo, con los datos que se tienen hasta ahora no es posible concluir su papel en la EP.

2. Objetivo 2: Replicación de 5 loci de riesgo para la EP, previamente asociados a una cohorte de EP de Granada, en otra cohorte independiente del sur de España.

En este estudio se ha analizado la asociación con la EP de 5 loci previamente descritos relacionados con la enfermedad, en una población de Granada, en el sur de España

(Bandrés-Ciga y col., 2016). Sin embargo, el pequeño tamaño de la cohorte en este primer estudio fue una limitación de sus resultados. Para solventar este problema, en este trabajo se cuenta con una nueva población con más del doble de pacientes de EP también del sur de España, con la que pretendemos contribuir a entender mejor la implicación de estos 5 *loci* en la EP.

En nuestra población de estudio, los *loci SREBF/RAII* y *MIR4697*, que previamente se han descritos asociados a la EP en la cohorte de Bandrés-Ciga y col., no parecen estar relacionados con la enfermedad. Sin embargo, dos de los *loci* previamente relacionados con la EP, *ACMSD/TMEM163* y *MAPT*, si muestran una asociación fuerte en este estudio (p corregido por Bonferroni $< 0,05$). En cuanto al *loci STK39*, su asociación con la EP es más moderada (p sin corregir $< 0,05$). En el meta-análisis realizado con los resultados de ambas cohortes, ninguno de los *loci* muestran estar asociados a la EP con un nivel de significación de genoma completo ($p < 5 \times 10^{-8}$), sin embargo si alcanzaron una asociación sugestiva ($p < 5 \times 10^{-6}$) los *loci ACMSD/TMEM163* y *MAPT*. *STK39* sigue mostrando una asociación moderada, pero los otros dos *loci* no parecen estar asociados a la EP.

Estos resultados muestran la complejidad de la arquitectura genética que hay detrás de la EP y la necesidad de replicar en poblaciones independientes los resultados de los estudios de asociación poblacionales. De hecho, en una población de similares características ancestrales que la previamente analizada por Bandrés-Ciga y col., sólo se replican 2 *loci* de los 5 previamente asociados a la EP, utilizando una técnica independiente.

Es de resaltar la asociación del SNP rs6430538 en este estudio, cercano al gen *ACMSD*. Este gen está implicado en el metabolismo del triptófano y sus mutaciones han sido relacionadas previamente con enfermedades neurológicas con características de

parkinsonismo. Además, *ACMSD* parece estar implicada en las vías de excitotoxicidad y de respuesta inflamatoria en el sistema nervioso central, por lo que ha sido propuesto como una nueva diana terapéutica en la EP (Thirumara-Rajamani y col., 2017). Sin embargo, éste y otros genes deben ser mejor estudiados para confirmar su papel en la EP.

Por tanto en este estudio se ha conseguido replicar la asociación de dos *loci* con la EP: *MAPT*, que parece tener un papel protector en el desarrollo de la enfermedad con un patrón de herencia dominante; y *ACMSD/TMEM16*, que puede ser un factor de riesgo con un patrón de herencia recesiva. La asociación con el *loci STK39* no parece ser tan robusta, mientras que los *loci SREBF/RAI1* y *MIR4697* no parecen estar asociados a la EP. Estos datos contribuyen a esclarecer los principales genes candidatos que deben ser tenidos en cuenta a la hora de establecer una puntuación del riesgo acumulado para el desarrollo de la EP en poblaciones específicas.

Los resultados de esta tesis doctoral ponen de manifiesto la complejidad genética que hay detrás de la EP. Genes robustamente asociados a la EP, se presentan con diferente frecuencia según la población de estudio. De hecho, la variación p.N52fs localizada en éste y otros estudios, parece variar su frecuencia dentro de nuestro país. Otros genes también ampliamente asociados a la EP, como *PINK1* y *VPS35*, son, sin embargo, poco frecuentes en la población estudiada. En cuanto a los genes nuevos que se identifican implicados en el desarrollo de distintas enfermedades, deben ser analizados en cohortes independientes para confirmar su verdadero papel en ellas. En este sentido, en esta tesis, los nuevos genes propuestos, *CHCHD2* y *TMEM230*, no se han mostrado asociados a la EP, a pesar de que han sido identificados en estudios previos como causativos o de riesgo para la EP. Por otra parte, se ha analizado una población de similares

característica a la del estudio previo donde se identificaron 5 *loci* de riesgo y sólo se han replicado 2, poniendo de manifiesto la importancia de estos estudios para no llegar a conclusiones erróneas. Sólo la realización de sucesivos estudios puede confirmar la robustez de estos hallazgos. Los estudios funcionales también pueden contribuir a esta confirmación de resultados, pero éstos son muy complicados en el contexto de la EP debido a la complejidad del sistema nervioso y a la falta de modelos *in vivo* que realmente reproduzcan la enfermedad. Aún así, los genes propuestos causativos se pueden estudiar en modelos simplificados. Pero en el caso de los *loci* de riesgo, su estudio aún es más complejo, puesto que el gen responsable es desconocido. Por otra parte, la penetrancia incompleta de algunas mutaciones dificulta aún más la resolución de la red genética que hay tras estas enfermedades, y es lo que ha podido ocurrir en este estudio con el gen *GBA*. Para complicar más el panorama, se han descrito efectos aditivos entre genes en ésta y otras enfermedades complejas, por lo que también hay que tener en cuenta la relación que puede haber entre las diferentes variaciones genéticas que se identifican. En este estudio, se ha propuesto que los genes *PARK2* y *ATP13A2* podrían estar actuando conjuntamente, así como *GIGYF2* y *LRRK2*.

A pesar de todos los esfuerzos realizados en esta materia, todavía queda mucho por conocer sobre la etiopatogenia de la EP. Por un lado, las cohortes de estudios GWA son cada vez mayores, sin embargo, la secuenciación del genoma completo todavía es muy limitada. A medida que aumente la aplicación de esta última tecnología, se podrán identificar mutaciones que todavía se nos escapan en otros tipos de análisis. Además, aumentar el número de estudios fuera de población caucásica también contribuirá a la localización de otras variaciones que no se dan con la misma frecuencia o que son específicas de otras poblaciones. Asociada a toda esta maquinaria, es necesaria una buena red de análisis de datos masivos, ya que la interpretación de los datos arrojados

por las nuevas tecnologías es un proceso complejo y crucial para ampliar nuestro conocimiento de la EP y otras enfermedades complejas.

V. CONCLUSIONES

1. 27 de los 117 pacientes estudiados (23,08%) tienen variaciones potencialmente patogénicas en los genes *PARK2*, *LRRK2*, *PINK1* o *GBA*.
2. El gen en el que se localizaron más variaciones potencialmente patogénicas fue *PARK2*, siendo la variación p.N52fs en este gen la causa más frecuente de EP en nuestra población.
3. Se ha identificado la co-ocurrencia de variaciones en los genes *ATP13A2* y *PARK2* y entre *LRRK2* y *GIGYF2*, sugiriendo efectos aditivos en ambos casos.
4. El 6,84% de los pacientes tenía variaciones potencialmente patogénicas en el gen *LRRK2*, sugiriendo que este gen juega un papel importante en la EP de inicio temprano en población española.
5. No se encontraron diferencias significativas en la edad de inicio en los pacientes que tenían variaciones potencialmente patogénicas en *LRRK2* comparado con los que tenían este tipo de variaciones en *PARK2*.
6. No se han localizado variaciones potencialmente patogénicas en los genes *SNCA*, *FBXO7*, *PARK7*, *HTRA2*, *DNAJC6*, *PLA2G6* y *UCHL1*.
7. Se han identificado 4 variaciones no exónicas en el gen *CHCHD2* en población del sur de España, pero ninguna parece ser causativa o factor de riesgo para la EP, por lo que el gen *CHCHD2* no parece estar involucrado en el desarrollo de la EP, al menos en población del sur de España.
8. Se han identificado 13 variaciones genéticas en el gen *TMEM230* en nuestra población de estudio, pero ninguna parece ser causativa o factor de riesgo para el desarrollo de la EP. Estos resultados sugieren que el gen *TMEM230* no parece

estar involucrado en el desarrollo de la EP, al menos en población del sur de España.

9. Se ha replicado la asociación de los *loci ACMSD/TMEM16* y *MAPT* con la EP, de manera que polimorfismos en *MAPT* parecen jugar un papel protector frente a la EP, mientras que variaciones en *ACMSD/TMEM16* pueden ser un factor de riesgo para la enfermedad.

VI. CONCLUSIONS

1. We have identified 27 of 117 studied patients (23,08%) carrying pathogenic variants in *PARK2*, *LRRK2*, *PINK1* or *GBA*.
2. The most frequently mutated gene was *PARK2*, and p.N52Mfs in *PARK2* was the most frequent cause of PD in the studied population.
3. Co-occurrence of mutations involving two genes suggests additive effects of *ATP13A2* and *PARK2* genes, as well as of *LRRK2* and *GIGYF2* genes.
4. 6,84% of patients carried some potential pathogenic variants in *LRRK2*, which suggests an important role for *LRRK2* in EOPD in Spanish population.
5. No significant differences relating to age at onset were observed between patients with pathogenic variants in *LRRK2* in comparison with those with pathogenic variants in *PARK2*.
6. Pathogenic variants were not observed in genes *SNCA*, *FBXO7*, *PARK7*, *HTRA2*, *DNAJC6*, *PLA2G6*, and *UCHL1*.
7. We have reported 4 non-exonic variants in *CHCHD2*, with no significant association in PD cases compared with control subjects. These results suggest that potential causal variants in *CHCHD2* are not common in our population, and subsequently they may not act as a risk factor, at least in population from southern Spain.
8. We have reported 13 variants in *TMEM230* but neither of them seems to be related with the development of PD. These results suggest that *TMEM230* may not play a mayor role in the development of PD, at least in population from southern Spain.

9. We have replicated the association with PD of two *loci* *ACMSD/TMEM16* and *MAPT*, and suggested that *MAPT* gene polymorphisms might be a protective factor against PD while *ACMSD/TMEM16* gene polymorphisms could be a risk factor for PD.

VII. BIBLIOGRAFÍA

Abou-Sleiman PM, Healy DG, Quinn N, Lees AJ, Wood NW. The role of pathogenic *DJ-1* mutations in Parkinson's disease. *Ann Neurol*. 2003; 54: 283-286.

Aharon-Peretz J, Rosenbaum H, Gershoni-Baruch R. Mutations in the glucocerebrosidase gene and Parkinson's disease in Ashkenazi Jews. *N Engl J Med*. 2004; 351: 1972-1977.

Alcalay RN, Caccappolo E, Mejia-Santana H, Tang MX, Rosado L, Ross BM, et al. Frequency of known mutations in early-onset Parkinson disease: implication for genetic counseling: the consortium on risk for early onset Parkinson disease study. *Arch. Neurol*. 2010; 67: 1116-1122.

Alvarez-Erviti L, Couch Y, Richardson J, et al. Alpha-synuclein release by neurons activates the inflammatory response in a microglial cell line. *Neurosci Res*. 2011; 69: 337-342.

Annesi G, Savettieri G, Pugliese P, DiAmelio M, Tarantino P, Ragonese P, et al. *DJ-1* mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex. *Ann Neurology*. 2005; 58:803-807.

Athanassiadou A, Voutsinas G, Psiouri L, Leroy E, Polimeropoulos MH, Maniatis GM, Papapetropoulos T. Genetic analysis of families with Parkinson disease that carry the Ala53Thr mutation in the gene encoding alpha-synuclein. *Am J Hum Genet*. 1999; 65: 555-8.

Bandres-Ciga S, Mencacci NE, Duran R, Barrero FJ, Escamilla-Sevilla F, Morgan S, et al. Analysis of the genetic variability in Parkinson's disease from Southern Spain. *Neurobiol Aging*. 2016; 37: 210 e1-5.

Baumann H, Wolff S, Munchau A, Hagenah JM, Lohmann K, Klein C. Evaluating the role of *TMEM230* variants in Parkinson's disease. *Parkinsonism Relat Disord*. 2017; 35: 100-101.

Benito-León J, Bermejo-Pareja F, Rodríguez J, Molina JA, Gabriel R, Morales JM. Prevalence of PD and other types of Parkinsonism in three elderly populations of central Spain. *Mov Disord* 2003; 18: 267-274.

Billingsley KJ, Bandres-Ciga S, Saez-Atienzar S, Singleton AB. Genetic risk factors in Parkinson's disease. *Cell Tissue Res*. 2018; 373: 9-20.

Bonetti M, Ferraris A, Petracca M, Bentivoglio AR, Dallapiccola B, Valente EM. *GIGYF2* variants are not associated with Parkinson's disease in Italy. *Mov Disord*. 2009; 24: 1867-8.

Bonifati V, Rizzu P, Squitieri F, Krieger E, Vanacore N, van Swieten JC, et al. *DJ-1(PARK7)*, a novel gene for autosomal recessive, early onset parkinsonism. *Neurol. Sci*. 2003; 24: 159-160.

Bonifati V, Rizzu P, Van Baren MJ, Schaap O, Breedveld GJ, Krieger R, et al. Mutations in the *DJ-1* gene associated with autosomal recessive early-onset parkinsonism. *Science*. 2003; 299: 256-259.

Bonifati V, Rohé CF, Breedveld GJ, Fabrizio E, De Mari M, Tassorelli C, et al. Early-onset parkinsonism associated with *PINK1* mutations: frequency, genotypes, and phenotypes. *Neurology*. 2005; 65: 87-95.

Bonifati V. Genetics of Parkinson's disease - state of the art, 2013. *Parkinsonism Relat Disord*. 2014; 20: S23-S28.

Bras J, Simon-Sanchez J, Federoff M, Morgadinho A, Januario C, Ribeiro M, et al. Lack of replication of association between *GIGYF2* variants and Parkinson disease. *Hum Mol Genet*. 2009;18: 341-6.

Bras JM, Singleton A. Genetic susceptibility in Parkinson's disease. *Biochim Biophys Acta*. 2009; 1792: 597-603.

Buongarzone G, Monfrini E, Franco G, Trezzi I, Borellini L, Frattini E, et al. Mutations in *TMEM230* are rare in autosomal dominant Parkinson's disease. *Parkinsonism Relat Disord*. 2017; 39: 87-88.

Cacabelos R. Parkinson's disease: from pathogenesis to pharmacogenomics. *Int. J Mol Sci*. 2017; 18: E551.

Carlsson A, Lindqvist M, Magnusson T. 3,4-Dihydroxyphenylalanine and 5-hydroxytryptophan as reserpine antagonists. *Nature*. 1957; 180: 1200.

Chang D, Nalls MA, Hallgrímsdóttir IB, Hunkapiller J, van der Brug M, Cai F, et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat Genet*. 2017; 49: 1511-1516.

Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S, et al. Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet*. 2004; 364: 1167-1169.

Chartier-Harlin MC, Dachsel JC, Vilarino-Guell C, Lincoln SJ, LePrete F, Hulihan MM, et al. Translation initiator *EIF4G1* mutations in familial Parkinson disease. *Am J Hum Genet*. 2011; 89: 398-406.

Choi JM, Woo MS, Ma HI, Kang SY, Sung YH, Yong SW, et al. Analysis of *PARK* genes in a Korean cohort of early-onset Parkinson disease. *Neurogenetics*. 2008; 9: 263-9.

Clark LN, Wang Y, Karlins E, Saito L, Mejia-Santana H, Harris J, et al. Frequency of *LRRK2* mutations in early- and late-onset Parkinson disease. *Neurology*. 2006; 67: 1786-1791.

Claveria LE, Duarte S, Sevillano MD, Perez-Sempere A, Cabezas C, Rodriguez F, et al. Prevalence of Parkinson's disease in Cantalejo, Spain: a door-to-door survey. *Mov Disord* 2002; 17: 242-249.

Conedera SA, Li Y, Funayama M, Yoshino H, Nishioka K, Hattori N. Genetic analysis of *TMEM230* in Japanese patients with familial Parkinson's disease. *Parkinsonism & Relat disord.* 2018; 48: 107-8.

Dagan E, Schlesinger I, Ayoub M, Mory A, Nassar M, Kurolap A, et al. The contribution of Niemann-Pick *SMPDI* mutations to Parkinson disease in Ashkenazi Jews. *Parkinsonism Relat Disord.* 2015; 21: 1067-71.

Deng HX, Shi Y, Yang Y, Ahmeti KB, Miller N, Huang C, et al. Identification of *TMEM230* mutations in familial Parkinson's disease. *Nat Genet.* 2016; 48: 733-739.

Deng H, Fan K, Jankovic J. The role of *TMEM230* gene in Parkinson's disease. *J Parkinsons Dis.* 2018; 8: 469-477

Dexter DT, Wells FR, Agid F, Agid Y, Lees AJ, Jenner P, et al. Increased nigral iron content in postmortem parkinsonian brain. *Lancet.* 1987; 2: 1219-20.

Di Fonzo A, Wu-Chou YH, Lu CS, van Doeselaar M, Simons EJ, Rohé CF et al. A common missense variant in the *LRRK2* gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan. *Neurogenetics.* 2006; 7: 133-138.

Di Fonzo A, Dekker MC, Montagna P, Baruzzi A, Yonova EH, Correia Guedes L, et al. *FBXO7* mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. *Neurology.* 2009;72(3):240-5.

Dorsey ER, Constantinescu R, Thompson JP, Biglan KM, Holloway RG, Kieburtz K, et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology.* 2007. 68: 384-386.

Double KL, Gerlach M, Schünemann V, Trautwein AX, Zecca L, Gallorini M, et al. Iron-binding characteristics of neuromelanin of the human substantia nigra. *Biochem Pharmacol.* 2003; 66: 489-94.

Edvardson S, Cinnamon Y, Ta-Shma A, Shaag A, Yim YI, Zenvirt S, et al. A deleterious mutation in *DNAJC6* encoding the neuronal-specific clathrin-uncoating Co-chaperone auxilin, is associated with juvenile parkinsonism. *PLoS One.* 2012; 7: 4-8.

Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, 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; 74:97-109.

Eiberg H, Hansen L, Korbo L, Nielsen IM, Svenstrup K, Bech S, et al. Novel mutation in *ATP13A2* widens the spectrum of Kufor-Rakeb syndrome (*PARK9*). *Clin Genet.* 2012; 82: 256-63.

Elbaz A, Bower JH, Maraganore DM, McDonnell SK, Peterson BJ, Ahlskog JE, et al. Risk tables for parkinsonism and Parkinson's disease. *J Clin Epidemiol.* 2002; 55: 25-31.

Elbaz A, Carcaillon L, Kab S, Moisan F. Epidemiology of Parkinson's disease. *Rev Neurol.* 2016; 172: 14-26.

Enciclopedia Médica A.D.A.M. [Internet]. Atlanta (GA): A.D.A.M., Inc.; ©2005. Substancia nigra and Parkinson disease; [actualizado 19 ene 2005; consulta 26 mar 2005]; [aprox.1p.]. Disponible en: <https://medlineplus.gov/ency/imagepages/19515.htm>.

Fan TS, Lin HI, Lin CH, Wu RM. Lack of *CHCHD2* mutations in Parkinson's disease in a Taiwanese population. *Neurobiol Aging.* 2016; 38: 218.e1e218.e2.

Fan TS, Lin CH, Lin HI, Chen ML, Wu RM. Lack of *TMEM230* mutations in patients with familial and sporadic Parkinson's disease in a Taiwanese population. *Am J Med Genet B Neuropsychiatr Geneti.* 2017; 174: 751-6.

Ferreira M, Massano J. An updated review of Parkinson's disease genetics and clinicopathological correlations. *Acta Neurol Scand.* 2017; 135: 273-284.

Farrer M, Kachergus J, Forno L, Lincoln S, Wang DS, Hulihan M, et al. Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. *Ann. Neurol.* 2004; 55: 174-179.

Farrer MJ, Stone JT, Lin CH, Dächsel JC, Hulihan MM, Haugarvoll K et al. *Lrrk2* G2385R is an ancestral risk factor for Parkinson's disease in Asia. *Parkinsonism Relat Disord.* 2007; 13: 89-92.

Fleming SM. Mechanisms of Gene-Environment Interactions in Parkinson's Disease. *Curr Environ Health Rep.* 2017; 4: 192-199.

Foo JN, Liany H, Bei JX, Yu XQ, Liu J, Au WL, et al. Rare lysosomal enzyme gene *SMPD1* variant (p.R591C) associates with Parkinson's disease. *Neurobiol Aging.* 2013; 34: 2890.e13-5.

Foo JN, Liany H, Tan LC, Au WL, Prakash KM, Liu J, et al. *DNAJ* mutations are rare in Chinese Parkinson's disease patients and controls. *Neurobiol Aging.* 2014; 35: 935.e1-935.e2.

Foo JN, Liu J, Tan EK. *CHCHD2* and Parkinson's disease. *Lancet Neurol.* 2015; 14: 681.e682.

Frank-Cannon TC, Tran T, Ruhn K a, et al. Parkin deficiency increases vulnerability to inflammation-related nigral degeneration. *J Neurosci.* 2008; 28: 10825-10834.

Fuchs J, Nilsson C, Kachergus J, Munz M, Larsson EM Schüle B, et al. Phenotypic variation in a large Swedish pedigree due to *SNCA* duplication and triplication. *Neurology.* 2007; 68: 916-22.

Fujita KA, Ostaszewski M, Matsuoka Y, Ghosh S, Glaab E, Trefois C, et al. Integrating Pathways of Parkinson's Disease in a Molecular Interaction Map. *Mol Neurobiol*. 2014; 49 :88-102.

Fung HC, Xiomerisiou G, Gibbs JR, Wu YR, Eerola J, Gurbali V, et al. Association of tau haplotype-tagging polymorphisms with Parkinson's disease in diverse ethnic Parkinson's disease cohorts. *Neurodegener Dis*. 2006; 3: 327-333.

Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, et al. *CHCHD2* mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. *Lancet Neurol*. 2015; 14: 274-82.

Gaig C, Ezquerro M, Marti MJ, Munoz E, Valldeoriola F, Tolosa E. *LRRK2* mutations in Spanish patients with Parkinson disease: frequency, clinical features, and incomplete penetrance. *Arch Neurol*. 2006; 63: 377-382.

Gan-Or Z, Ozelius LJ, Bar-Shira A, Saunders-Pullman R, Mirelman A, Kornreich R, et al. The p.L302P mutation in the lysosomal enzyme gene *SMPD1* is a risk factor for Parkinson disease. *Neurology*. 2013; 80: 1606-10.

Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 1988; 51: 745-52.

Gillardon F, Schmid R, Draheim H. Parkinson's disease-linked leucine-rich repeat kinase 2 (R1441G) mutation increases proinflammatory cytokine release from activated primary microglial cells and resultant neurotoxicity. *Neuroscience*. 2012; 208: 41-48.

Giri A, Mok KY, Jansen I, Sharma M, Tesson C, Mangone G, et al. Lack of evidence for a role of genetic variation in *TMEM230* in the risk for Parkinson's disease in the Caucasian population. *Neurobiol Aging*. 2017; 50: 167 e11-167 e13.

Gorell JM, Ordidge RJ, Brown GG, Deniau JC, Buderer NM, Helpert JA. Increased iron-related MRI contrast in the substantia nigra in Parkinson's disease. *Neurology*. 1995; 45: 1138-43.

Gorostidi A, Martí-Masó JF, Bergareche A, Rodríguez-Oroz MC, López de Munain A, Ruiz-Martínez J. Genetic mutation analysis of Parkinson's disease patients using multigene next-generation sequencing panels. *Mol Diagn Ther*. 2016; 20: 481-491.

Guella I, Pistocchi A, Asselta R, Rimoldi V, Ghilardi A, Sironi F, et al. Mutational screening and zebrafish functional analysis of *GIGYF2* as a Parkinson-disease gene. *Neurobiol Aging*. 2011; 32: 1994-2005.

Gustavsson EK, Trinh J, Guella I, Vilarino-Guell C, Appel-Cresswell S, Stoessl AJ, et al. *DNAJC13* genetic variants in parkinsonism. *Mov Disord*. 2015; 30: 273-278.

Halperin A, Elstein D, Zimran A. Increased incidence of Parkinson disease among relatives of patients with Gaucher disease. *Blood Cells Mol Dis*. 2006; 36: 426-428.

Haugarvoll K and Wszolek ZK. Clinical features of *LRRK2* parkinsonism. *Parkinsonism Relat. Disord.* 2009; 15: S205-S208.

He YC, Huang P, Li QQ, Sun Q, Li DH, Wang T, et al. *TMEM230* stop codon mutation is rare in Parkinson's disease and essential tremor in eastern China. *Mov Disord.* 2017; 32: 301-302.

Healy DG, Falchi M, O'Sullivan SS, Bonifati V, Bressman S, Brice A, et al. International *LRRK2* Consortium. Phenotype, genotype, and worldwide genetic penetrance of *LRRK2*-associated Parkinson's disease: a case-control study. *Lancet Neurol.* 2008; 7: 583-590.

Hernàn MA, Takkouche B, Caaman~o-Isorna F, Gestal-Otero JJ. A meta-analysis of coffee drinking, cigarette smoking, and the risk of Parkinson's disease. *Ann Neurol.* 2002; 52: 276-284.

Hernandez DG, Nalls MA, Ylikotila P, Keller M, Hardy JA, Majamaa K, Singleton AB. Genome wide assessment of young onset Parkinson's disease from Finland. *PLoS ONE.* 2012; 7: e41859.

Hernandez DG, Reed X, Singleton AB. Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. *J Neurochem.* 2016; 139: 59-74.

Ibanez P, Bonnet AM, Debarges B, Lohmann E, Tison F, Pollak P, Agid Y, Durr A, Brice A. Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet.* 2004; 364: 1169-1171.

Ibanez L, Dube U, Budde J, Black K, Medvedeva A, Davis AA, et al. *TMEM230* in Parkinson's disease. *Neurobiol Aging.* 2017; 56: 212 e1-e3.

Ikeuchi T, Kakita A, Shiga A, Kasuga K, Kaneko H, Tan CF, et al. Patients homozygous and heterozygous for *SNCA* duplication in a family with parkinsonism and dementia. *Arch. Neurol.* 2008; 65: 514-519.

Jaberi E, Rohani M, Shahidi GA, Nafissi S, Arefian E, Soleimani M, et al. Mutation in *ADORA1* identified as likely cause of early-onset parkinsonism and cognitive dysfunction. *Mov Disord.* 2016; 31: 1004-11.

Jansen IE, Bras JM, Lesage S, Schulte C, Gibbs JR, Nalls MA, et al. *CHCHD2* and Parkinson's disease. *Lancet Neurol.* 2015;14: 678e679.

Jesús S, Pérez I, Cáceres-Redondo MT, Carrillo F, Carballo M, Gómez-Garre P and Mir O. Low serum uric acid concentration in Parkinson's disease in southern Spain. *European Journal of Neurology.* 2013; 20: 208-210.

Jesus S, Huertas I, Bernal-Bernal I, Bonilla-Toribio M, Caceres-Redondo MT, Vargas-Gonzalez L, et al. *GBA* Variants Influence Motor and Non-Motor Features of Parkinson's Disease. *PLoS One.* 2016; 11: e0167749.

Karimi-Moghadam A, Charsouei S, Bell B, Jabalameli MR. Parkinson disease from mendelian forms to genetic susceptibility: new molecular insights into the neurodegeneration process. *Cell. Mol. Neurobiol.* 2018; 38: 1153-1178.

Keller MF, Saad M, Bras J, Bettella F, Nicolaou N, Simon-Sanchez J, et al. Using genome-wide complex trait analysis to quantify ‘missing heritability’ in Parkinson’s disease. *Hum Mol Genet.* 2012; 21: 4996-5009.

Khodadadi H, Azcona LJ, Aghamollai V, Omrani MD, Garshasbi M, Taghavi S, et al. *PTRHDI* (C2orf79) mutations lead to autosomal-recessive intellectual disability and parkinsonism. *Mov Disord.* 2017; 32: 287-91.

Ki CS, Stavrou EF, Lee WY, Chung EJ, Kim JY, Athanassiadou A. The Ala53Thr mutation in the alpha-synuclein gene in a Korean family with Parkinson disease. *Clin Genet.* 2007; 71: 471-3.

Kim M, Semple I, Kim B, Kiers A, Nam S, Park HW, et al. *Drosophila* Gyf/GRB10 interacting GYF protein is an autophagy regulator that controls neuron and muscle homeostasis. *Autophagy.* 2015; 11: 1358-1372.

Kirola L, Behari M, Shishir C, Thelma BK. Identification of a novel homozygous mutation Arg459Pro in *SYNJ1* gene of an Indian family with autosomal recessive juvenile parkinsonism. *Parkinsonism Relat Disord.* 2016; 31: 124-128.

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

Klein RJ, Zeiss C, Chew EY, Sackler RS, Haynes C, Henning AK, et al. Complement factor H polymorphism in age-related macular degeneration. *Science.* 2005; 308: 385-389.

Köroğlu Ç, Baysal L, Cenkaya M, Karasoy H, Tolun A. *DNAJC6* is responsible for juvenile parkinsonism with phenotypic variability. *Parkinsonism Relat Disord.* 2013; 19: 320-324.

Koziorowski D, Hoffman-Zacharska D, Sławek J, Jamrozik Z, Janik P, Potulska-Chromik A, et al. Incidence of mutations in the *PARK2*, *PINK1*, *PARK7* genes in Polish early-onset Parkinson disease patients. *Neurol Neurochir Pol.* 2013; 47: 319-324.

Krebs CE, Karkheiran S, Powell JC, Cao M, Makarov V, Darvish H, et al. The Sac1 domain of *SYNJ1* identified mutated in a family with early-onset progressive parkinsonism with generalized seizures. *Hum Mutat.* 2013; 34: 1200-7.

Kruger R, Kuhn W, Muller T, Woitalla D, Graeber M, Kosel S, Przuntek H, Epplen JT, Schols L and Riess O. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson’s disease. *Nat. Genet.* 1998; 18: 106-108.

Kumar KR, Djarmati-Westenberger A, Grunewald A. Genetics of Parkinson’s disease. *Semin. Neurol.* 2011; 31: 433-440.

Kumar KR, Weissbach A, Heldmann M, Kasten M, Tunc S, Sue CM, et al. Frequency of the D620N mutation in *VPS35* in Parkinson disease. *Arch Neurol*. 2012; 69: 1360-1364.

Kumazawa R, Tomiyama H, Li Y, Imamichi Y, Funayama M, Yoshino H, et al. Mutation analysis of the *PINK1* gene in 391 patients with Parkinson disease. *Arch Neurol*. 2008; 65: 802-808.

Kwakye GF, Paoliello MM, Mukhopadhyay S, Bowman AB, Aschner M. Manganese-induced parkinsonism and Parkinson's disease: shared and distinguishable features. *Int J Environ Res Public Health*. 2015; 12: 7519-40.

Kwok JB, Teber ET, Loy C, Hallupp M, Nicholson G, Mellick GD, et al. Tau haplotypes regulate transcription and are associated with Parkinson's disease. *Ann Neurol*. 2004; 55: 329-334.

Lang AE, Obeso JA. Time to move beyond nigrostriatal dopamine deficiency in Parkinson's disease. *Ann Neurol* 2004; 55: 761-765.

de Lau L. M., Breteler M. M. Epidemiology of Parkinson's disease. *The Lancet Neurology*. 2006; 5: 525-535.

Lautier C, Goldwurm S, Durr A, Giovannone B, Tsiaras WG, Pezzoli G, et al. Mutations in the *GIGYF2* (TNRC15) gene at the *PARK11* locus in familial Parkinson disease. *Am J Human Genet*. 2008; 82: 822-33.

Leroy E, Boyer R, Auburger G, Leube B, Ulm G, Mezey G, et al. The ubiquitin pathway in Parkinson's disease. *Nature*. 1998; 395: 451-452.

Lesage S, Condroyer C, Klebe S, Lohmann E, Durif F, Damier P, et al. *EIF4G1* in familial Parkinson's disease: pathogenic mutations or rare benign variants?. *Neurobiol Aging* 2012; 33: 2233.e1-2233.e5.

Lesage S, Bras J, Cormier-Dequaire F. Loss-of-function mutations in *RAB39B* are associated with typical early-onset Parkinson disease. *Neurol Genet*. 2015; 1: e9.

Lesage S, Drouet V, Majounie E, Deramecourt V, Jacoupy M, Nicolas A, et al. Loss of *VPS13C* function in autosomal-recessive parkinsonism causes mitochondrial dysfunction and increases *PINK1/parkin*-dependent mitophagy. *Am J Hum Genet*. 2016; 98: 500-513.

Levran O, Desnick RJ, Schuchman EH. Niemann-Pick Disease: a frequent missense mutation in the acid sphingomyelinase gene of Ashkenazi Jewish type A and B patients. *Proc. Natl. Acad. Sci. USA*. 1991; 88: 3738-3752.

Levran O, Desnick RJ, Schuchman EH. Niemann-Pick Type B Disease: Identification of a single codon deletion in the acid sphingomyelinase gene and genotype/phenotype correlations in type A and B patients. *J. Clin. Invest*. 1991; 88: 806-810.

Levran O, Desnick RJ, Schuchman EH. Identification and expression of a common

missense mutation (L302P) in the acid sphingomyelinase gene of Ashkenazi Jewish Type A Niemann-Pick disease patients. *Blood*. 1992; 8: 2081-2087.

Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BMM, et al. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: the PDGene database. *PLoS Genet*. 2012; 8: e1002548.

Lill CM. Genetics of Parkinson's disease. *Mol Cell Probes*. 2016; 30: 386-96.

Liou HH, Tsai MC, Chen CJ, Jeng JS, Chang YC, Chen SY, et al. Environmental risk factors and Parkinson's disease: a case-control study in Taiwan. *Neurology*. 1997; 48: 1583-8.

Lorenzo-Betancor O, Ogaki K, Soto-Ortolaza AI, Labbe C, Walton RL, Strongosky AJ, et al. *DNAJC13* p.Asn855Ser mutation screening in Parkinson's disease and pathologically confirmed Lewy body disease patients. *Eur J Neurol*. 2015; 22: 1323-1325.

Lu, Q, Deng X, Song Z, Guo Y, Yang Y, Deng H. Mutation analysis of the *CHCHD2* gene in Chinese Han patients with Parkinson's disease. *Parkinsonism Relat. Disord*. 2016; 29: 143e144.

Lucking CB, Dür A, Bonifati V, Vaughan J, De Michele G, Gasser T, et al. Association between early-onset Parkinson's disease and mutations in the parkin gene. *N Engl J Med*. 2000; 342: 1560-7.

Ma D, Foo JN, Yulin Ng E, Zhao Y, Liu JJ, Tan EK. Screening for *TMEM230* mutations in young-onset Parkinson's disease. *Neurobiol Aging*. 2017; 58: 239e9-e10.

Macedo MG, Verbaan D, Fang Y, van Rooden SM, Visser M, Anar B, et al. Genotypic and phenotypic characteristics of Dutch patients with early onset Parkinson's disease. *Mov. Disord*. 2009; 24: 196-203.

Mackenzie I. The pathology of Parkinson's disease. *BC Med J*. 2001; 43: 142-147.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461: 747-753.

Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP, Krüger R, et al. Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. *JAMA*. 2006; 296: 661-670.

Martínez-Suárez MM, Blázquez-Menes B. Estimation of the prevalence of Parkinson's disease in Asturias, Spain. A pharmacoepidemiological study of the consumption of antiparkinson drugs. *Rev Neurol*. 2000; 31: 1001-1006.

Meng H, Yamashita C, Shiba-Fukushima K, Inoshita T, Funayama M, Sato S, et al. Loss of Parkinson's disease-associated protein *CHCHD2* affects mitochondrial crista structure and destabilizes cytochrome c. *Nat Commun*. 2017; 8: 15500.

Mitsui J, Mizuta I, Toyoda A, Ashida R, Takahashi Y, Goto J, et al. Mutations for Gaucher disease confer high susceptibility to Parkinson disease. *Arch Neurol*. 2009; 66: 571-576.

Morgan NV, Westaway SK, Morton JE, Gregory A, Gissen P, Sonek S, et al. *PLA2G6*, encoding a phospholipase A2, is mutated in neurodegenerative disorders with high brain iron. *Nat Genet*. 2006; 38: 752-4.

Mortiboys H, Johansen KK, Aasly JO, Bandmann O. Mitochondrial impairment in patients with Parkinson disease with the G2019S mutation in *LRRK2*. *Neurol*. 2010; 75: 2017-2020.

Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet*. 2014; 46: 989-993.

Neudorfer O, Giladi N, Elstein D, Abrahamov A, Turezkite T, Aghai E, et al. Occurrence of Parkinson's syndrome in type I Gaucher disease. *QJM*. 1996; 89: 691-694.

Nichols WC, Kissell DK, Pankratz N, Pauciulo MW, Elsaesser VE, Clark KA, et al. Parkinson Study Group-PROGENI Investigators. Variation in *GIGYF2* is not associated with Parkinson disease. *Neurology*. 2009; 72: 1886-92.

Nishioka K, Hayashi S, Farrer MJ, Singleton AB, Yoshino H, Imai H, et al. Clinical heterogeneity of alpha-synuclein gene duplication in Parkinson's disease. *Ann. Neurol*. 2006; 59: 298-309.

Nuytemans K, Meeus B, Crosiers D, Brouwers N, Goossens D, Engelborghs S, Theuns J. Relative contribution of simple mutations vs. copy number variations in five Parkinson disease genes in the Belgian population. *Hum Mutat*. 2009; 30: 1054-1061.

Nuytemans K, Theuns J, Cruts M, Van Broeckhoven C. Genetic etiology of Parkinson disease associated with mutations in the *SNCA*, *PARK2*, *PINK1*, *PARK7*, and *LRRK2* genes: a mutation update. *Hum Mutat*. 2010; 31: 763-780.

Obeso JA, Stamelou M, Goetz CG, Poewe W, Lang AE, Weintraub D, et al. Past, present, and future of Parkinson's disease: a special essay on the 200th anniversary of the shaking palsy. *Mov Disord*. 2017; 32: 1264-1310.

Ogaki K, Koga S, Heckman MG, Fiesel FC, Ando M, Labbe C, et al. Mitochondrial targeting sequence variants of the *CHCHD2* gene are a risk for Lewy body disorders. *Neurology*. 2015; 85: 2016-25.

Okubadejo NU, Bower JH, W, Rocca WA, Maraganore DM. Parkinson's disease in Africa: a systematic review of epidemiologic and genetic studies. *Mov Disord*. 2006; 21: 2150-2156.

Oligiati S, De Rosa A, Quadri M, Criscuolo C, Breedveld GJ, Picillo M, et al. *PARK20* caused by *SYNJ1* homozygous Arg258Gln mutation in a new Italian family. *Neurogenetics*. 2014; 15: 183-8.

- Olgiami S, Quadri M, Fang M, Rood JP, Saute JA, Chein HF, et al. *DNAJC6* mutations associated with early-onset Parkinson's disease. *Ann Neurol*. 2015; 65: 1353- 1357.
- Paisán-Ruiz C, Bhatia KP, Li A, Hernandez D, Davis M, Wood NW, et al. Characterization of *PLA2G6* as a locus for dystonia-parkinsonism. *Ann Neurol*. 2009; 65: 19-23.
- Palacios N, Gao X, McCullough ML, et al. Caffeine and risk of Parkinson's disease in a large cohort of men and women. *Mov Disord* 2012; 27: 1276-1282.
- Park JS, Blair NF, Sue CM. The role of *ATP13A2* in Parkinson's disease: Clinical phenotypes and molecular mechanisms. *Mov Disord*. 2015; 30: 770-9.
- Parikshak NN, Gandal MJ, Geschwind DH. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nat Rev Genet*. 2015; 16: 441-458.
- Parkinson J. An essay on the shaking palsy. London: Writting-ham and Rowland Sherwood, Neely and Jones 1817.
- Pihlstrøm L, Axelsson G, Bjørnara KA, Dizdar N, Fardell C, Forsgren L, et al. Supportive evidence for 11 loci from genome-wide association studies in Parkinson's disease. *Neurobiol Aging*. 2013; 34: 1708.e7-13.
- Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, Schrag AE, Lang AE. Parkinson disease. *Nat Rev Dis Primers*. 2017; 3: 17013.
- Polymeropoulos MH, Hurko O, Hsu F, Rubenstein J, Basnet S, Lane K, et al. Linkage of the locus for cerebral cavernous hemangiomas to human chromosome 7q in four families of Mexican-American descent. *Neurology*. 1997; 48: 752-7.
- Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Deheija A, Dutra A, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*. 1997; 276: 2045-7.
- Pont-Sunyer C, Hotter A, Gaig C, Seppi K, Compta Y, Katzenschlager R, et al. The onset of non motor symptoms in Parkinson's disease (The ONSET PD Study). *Mov Disord*. 2015; 30: 229-237.
- Pringsheim N, Jette A, Frolkis TD. Steeves. The prevalence of Parkinson's disease: a systematic review and meta-analysis. *Mov Disord*. 2014; 29: 583-1590.
- Puschmann A, Ross OA, Vilarino-Guell C, Lincoln SJ, Kachergus JM, Cobb SA, et al. A Swedish family with de novo α -synuclein A53T mutation: Evidence for early cortical dysfunction. *Parkinsonism Relat Disord*. 2009; 15: 627-632.
- Puschmann A. Monogenic Parkinson's disease and parkinsonism: clinical phenotypes and frequencies of known mutations. *Parkinsonism Relat Disord*. 2013; 19: 407-415.

Quadri M, Fang M, Picillo M, Olgiati S, Breedveld GJ, Graafland J, et al. Mutation in the *SYNJ1* gene associated with autosomal recessive, early-onset parkinsonism. *Hum Mutat.* 2013; 34: 1208-1215.

Quadri M, Breedveld GJ, Chang HC, Yeh TH, Guedes LC, Toni V, et al. Mutations in *TMEM230* are not a common cause of Parkinson's disease. *Mov Disord.* 2017; 32: 302-304.

Ramachandiran S, Hansen JM, Jones DP, Richardson JR, Miller GW. Divergent mechanisms of paraquat, MPP+, and rotenone toxicity: oxidation of thioredoxin and caspase-3 activation. *Toxicol Sci.* 2007; 95: 163-71.

Ramirez A, Heimbach A, Grundemann J, Stiller B, Hampshire D, Cid LP, et al. Hereditary parkinsonism with dementia is caused by mutations in *ATP13A2*, encoding a lysosomal type 5 P-type ATPase. *Nat Genet.* 2006; 38: 1184-91.

Rauschendorf MA, Jost M, Stock F, Zimmer A, Rosler B, Rijntjes M, et al. Novel compound heterozygous synaptotagmin-1 mutation causes l-dopa-responsive dystonia-parkinsonism syndrome. *Mov Disord.* 2017; 32: 478-480.

Ritz B, Ascherio A, Checkoway H, et al. Pooled analysis of tobacco use and risk of Parkinson disease. *Arch Neurol* 2007; 64: 990-997.

Ruiz-Martinez J, Krebs CE, Makarov V, Gorostidi A, Martí-Massó JF, Paisán-Ruiz C. *GIGYF2* mutation in late-onset Parkinson's disease with cognitive impairment. *J. Hum. Genet.* 2015; 60: 637-640.

de Rijk MC, Tzourio C, Breteler MMB, Dartigues JF, Amaducci L, Lopez Pousa S, et al. Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON collaborative study. *J Neurol Neurosurg Psychiatry.* 1997; 62: 10-15.

Rocha EM, De Miranda B, Sanders LH. α -synuclein: pathology, mitochondrial dysfunction and neuroinflammation in Parkinson's disease. *Neurobiol. Dis.* 2018; 109: 249-257.

Ross OA, Soto AI, Vilariño-güell C, Heckman MG, Diehl NN, Hulihan MM, et al. Genetic variation of *Omi/HtrA2* and Parkinson's disease. *Parkinsonism Relat Disord.* 2008; 14: 539-543.

Ross OA, Spanaki C, Griffith A, Lin CH, Kachergus J, Haugarvoll K, et al. Haplotype analysis of *Lrrk2* R1441H carriers with parkinsonism. *Parkinsonism Relat Disord* 2009; 15: 466-467.

Ross OA, Soto-Ortolaza AI, Heckman MG, Aasly JO, Abahuni N, Annesi G, et al. Association of *LRRK2* exonic variants with susceptibility to Parkinson's disease: a case-control study. *Lancet Neurol.* 2011; 10: 898-908.

Ross JP, Dupre N, Dauvilliers Y, Strong S, Ambalavanan A, Spiegelman D, et al. Analysis of *DNAJC13* mutations in French-Canadian/French cohort of Parkinson's disease. *Neurobiol Aging*. 2016; 45: 212.e13-212.e17.

Ross JP, Dupre N, Dauvilliers Y, Strong S, Dionne-Laporte A, Dion PA, et al. *RIC3* variants are not associated with Parkinson's disease in French-Canadians and French. *Neurobiol Aging*. 2017; 53: 194e9-194e11.

Spataro N, Roca-Umbert A, Cervera-Carles L, Vallès M, Anglada R, Pagonabarraga J, et al. Detection of genomic rearrangements from targeted resequencing data in Parkinson's disease patients. *Mov. Disord*. 2017; 32:165-169.

Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet*. 2009; 41: 1303-1307.

Schapira, A. H. Mitochondrial dysfunction in Parkinson's disease. *Cell Death Differ*. 2007; 14: 1261-1266.

Schoenberg BS, Osuntokun BO, Adeuja AO, Bademosi O, Nottidge V, Anderson DW, et al. Comparison of the prevalence of Parkinson's disease in black populations in the rural United States and in rural Nigeria: door-to-door community studies *Neurology*. 1988; 38 (4): 645-646.

Schrag A, Horsfall L, Walters K, Noyce A, Petersen I. Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *Lancet Neurol*. 2015; 14: 57-64.

Seaman M, Freeman CL. Analysis of the Retromer complex- WASH complex interaction illuminates new avenues to explore in Parkinson disease. *Commun Integr Biol* 2014; 7: e29483.

Semchuk KM, Love EJ, Lee RG. Parkinson's disease and exposure to agricultural work and pesticide chemicals. *Neurology*. 1992; 42: 1328-35.

Sheerin UM, Charlesworth G, Bras J, Guerreiro R, Bhatia K, Foltynie T, et al. Screening for *VPS35* mutations in Parkinson's disease. *Neurobiol Aging*. 2012; 33: 838.e1-838.e5.

Shi CH, Mao CY, Zhang SY, Yang J, Song B, Wu P, et al. *CHCHD2* gene mutations in familial and sporadic Parkinson's disease. *Neurobiol Aging*. 2016; 38: 217.e9e217.e13.

Shi CH, Li F, Shi MM, Yang ZH, Mao CY, Zhang SY, et al. Genetic analysis of the *TMEM230* gene in Chinese Han patients with Parkinson's disease. *Sci Rep*. 2017; 7: 1190.

Shojaee S, Sina F, Banihosseini SS, Kazemi MH, Kalhor R, Shahidi GA, et al. Genome-wide linkage analysis of a Parkinsonian-pyramidal syndrome pedigree by 500 K SNP arrays. *Am J Hum Genet*. 2008; 82: 1375-84.

Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N Engl J Med.* 2009; 361: 1651-1661.

Siitonen A, Majounie E, Federo M, Ding J, Majamaa K, Singleton AB. Mutations in *EIF4G1* are not a common cause of Parkinson's disease. *Eur J Neurol.* 2013; 20: e59.

Simon-Sanchez J, Singleton AB. Sequencing analysis of *OMI/HTRA2* shows previously reported pathogenic mutations in neurologically normal controls. *Hum Mol Genet.* 2008; 17: 1988-93.

Simón-Sánchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat Genet.* 2009; 41: 1308-1312.

Sina F, Shojaee S, Elahi E, Paisan-Ruiz C. R632W mutation in *PLA2G6* segregates with dystonia-parkinsonism in a consanguineous Iranian family. *Eur J Neurol.* 2009; 16: 101-4.

Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, et al. Alpha-Synuclein locus triplication causes Parkinson's disease. *Science.* 2003; 302: 841.

Skipper L, Wilkes K, Toft M, Baker M, Lincoln S, Hulihan M, et al. Linkage disequilibrium and association of *MAPT* H1 in Parkinson disease. *Am J Hum Genet.* 2004; 75: 669-677.

Sofic E, Riederer P, Heinsen H, Beckmann H, Reynolds GP, Hebenstreit G, et al. Increased iron (III) and total iron content in post mortem substantia nigra of parkinsonian brain. *J Neural Transm.* 1988; 74: 199-205.

Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. *Nature.* 1997; 388: 839-840.

Spira PJ, Sharpe DM, Halliday G, Cavanagh J, Nicholson GA. Clinical and pathological features of a Parkinsonian syndrome in a family with an Ala53Thr alpha-synuclein mutation. *Ann Neurol.* 2001; 49: 313-9.

Strauss KM, Martins LM, Plun-Favreau H, Marx FP, Kautzmann S, Breg D, et al. Loss of function mutations in the gene encoding *Omi/HtrA2* in Parkinson's disease. *Hum Mol Genet.* 2005; 14: 2099-2111.

Sudhama S, Muthane UB, Behari M, Govindappa ST, Juyal RC, Thelma BK. Evidence of mutations in *RIC3* acetylcholine receptor chaperone as a novel cause of autosomal-dominant Parkinson's disease with non-motor phenotypes. *J Med Genet.* 2016; 53: 559-566.

Sudhama S, Prasad K, Behari M, Muthane UB, Juyal RC, Thelma BK. Discovery of a frameshift mutation in podocalyxin-like (*PODXL*) gene, coding for a neural adhesion molecule, as causal for autosomal-recessive juvenile parkinsonism. *J Med Genet.* 2016; 53: 450-456.

Sveinbjornsdottir S. The clinical symptoms of Parkinson's disease. *Journal of Neurochemistry*. 2016; 318-324.

Taghavi S, Chaouni R, Tafakhori A, Azcona LJ, Firouzabadi SG, Omrani MD, et al. A clinical and molecular genetic study of 50 families with autosomal recessive parkinsonism revealed known and novel gene mutations. *Mol Neurobiol*. 2018; 55: 3477-3489.

Tan EK, Schapira AH. Uniting Chinese across Asia: the *LRRK2* Gly2385Arg risk variant. *Eur J Neurol*. 2008; 15: 203-204.

Tan EK, Zhao Y, Tan L, Lim H-Q, Lee J, Yuen Y, et al. Analysis of *LRRK2* Gly2385Arg genetic variant in non-Chinese Asians. *Mov Disord*. 2007; 22: 1816-181.

Taylor KS, Cook JA, Counsell CE. Heterogeneity in male to female risk for Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2007; 78: 905-906.

Thirtamara-Rajamani K, Li P, Escobar Galvis ML, Labrie V, Brundin P, Brundin L. Is the enzyme ACMSD a novel therapeutic target in Parkinson's disease?. *J Parkinsons Dis*. 2017; 7: 577- 587.

Troiano AR, Cazeneuve C, Le Ber I, Bonnet AM, Lesage S, Brice A. Alpha-synuclein gene duplication is present in sporadic Parkinson disease. *Neurology*. 2008; 71:1295.

Tucci A, Charlesworth G, Sheerin UM, Plagnol V, Wood NW, Hardy J. Study of the genetic variability in a Parkinson's disease gene: *EIF4G1*. *Neurosci Lett*. 2012; 518: 19-22.

Valente EM, Bentivoglio AR, Dixon PH, Ferraris A, Ialongo T, Frontali M, Albanese A, Wood NW. Localization of a novel locus for autosomal recessive early-onset parkinsonism, *PARK6*, on human chromosome 1p35-p36. *Am. J. Hum. Genet*. 2001; 68: 895-900.

Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S, et al. Hereditary early-onset Parkinson's disease caused by mutations in *PINK1*. *Science*. 2004; 304: 1158-60.

Vilarino-Guell C, Wider C, Ross OA, Dachsel JC, Kachergus JM, Lincoln SJ, et al. *VPS35* mutations in Parkinson disease. *Am. J. Hum. Genet*. 2011; 89: 162-167.

Vilarino-Guell C, Rajput A, Milnerwood AJ, Shah B, Szu-Tu C, Trinh J, et al. *DNAJC13* mutations in Parkinson disease. *Hum Mol Genet*. 2014; 23: 1794-1801.

Vines JJ. Incidence of idiopathic and secondary Parkinson disease in Navarre. Population-based case registry. *Neurologia*. 1999; 14: 16-22.

Waak J, Weber SS, Waldenmaier A, et al. Regulation of astrocyte inflammatory responses by the Parkinson's disease-associated gene *DJ-1*. *FASEB J*. 2009; 23: 2478-2489.

- Weisskopf MG, O'Reilly E, Chen H, Schwarzschild MA, Ascherio A. Plasma urate and risk of Parkinson's disease. *Am J Epidemiol* 2007; 166: 561-567.
- Wilson GRR, Sim JC, McLean C, Giannandrea M, Galea CA, Riseley JR, et al. Mutations in *RAB39B* cause X-linked intellectual disability and early-onset parkinson disease with α -synuclein pathology. *Am J Hum Genet.* 2014; 95: 729-735.
- Wooten GF, Currie LJ, Bovbjerg VE, Lee JK, Patrie J. Are men at greater risk for Parkinson's disease than women? *J Neurol Neurosurg Psychiatry.* 2004; 75: 637-639.
- Wu H, Lu X, Xie F, Cen Z, Zheng X, Luo W. Genetic analysis of the *CHCHD2* gene in a cohort of Chinese patients with Parkinson disease. *Neurosci. Lett.* 2016; 629: 116e118.
- Wu H, Zheng X, Cen Z, Xie F, Chen Y, Lu X, et al. Genetic analysis of the *TMEM230* gene in Chinese patients with familial Parkinson disease. *Parkinsonism Relat Disord.* 2017; 36: 105-106.
- Yamada-Fowler N, Soferkvist P. Coffee, genetic variants, and Parkinson's disease: gene-environment interactions. *J Caffeine Res.* 2015; 5: 3-10.
- Yan W, Tang B, Zhou X, Lei L, Li K, Sun Q, et al. *TMEM230* mutation analysis in Parkinson's disease in a Chinese population. *Neurobiol Aging.* 2017; 49: 219 e1-219 e3.
- Yang X, An R, Xi J, Zheng J, Chen Y, Huang H, et al. Sequencing *TMEM230* in Chinese patients with sporadic or familial Parkinson's disease. *Mov Disord.* 2017; 32:8 00-802.
- Yarnall A, Archibald Neil, Burn D. Parkinson's disease. *Mov Disord.* 2012; 40: 529-535.
- Ysselstein D, Shulman JM, Krainc D. Emerging links between pediatric lysosomal storage diseases and adult parkinsonism. *Mov Disord.* 2019.
- Zabetian CP, Hutter CM, Factor SA, Nutt JG, Higgins DS, Griffith A, et al. Association analysis of *MAPT* H1 haplotype and subhaplotypes in Parkinson's disease. *Ann Neurol.* 2007; 62: 137-144.
- Pont-Sunyer C, Hotter A, Gaig C, et al. The onset of nonmotor symptoms in Parkinson's disease (The ONSET PD Study). *Mov. Disord.* 2015; 30: 229-237
- Zarranz JJ, Alegre J, Gomez-Esteban JC, Lezcano E, Ros R, Ampuero I, et al. The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Ann. Neurol.* 2004; 55: 164-173.
- Zhang W, Wang T, Pei Z, et al. Aggregated alpha-synuclein activates microglia: a process leading to disease progression in Parkinson's disease. *FASEB J.* 2005; 19: 533-542.

Zhang Y, Sun QY, Yu RH, Guo JF, Tang BS, Yan XX. The contribution of *GIGYF2* to Parkinson's disease: a meta-analysis. *Neurol Sci.* 2015; 36: 2073-9.

Zhang M, Xi Z, Fang S, Ghani M, Sato C, Moreno D, Liang Y, Lang AE, Rogaeva E. Mutation analysis of *CHCHD2* in Canadian patients with familial Parkinson's disease. *Neurobiol Aging.* 2016; 38: 217.e7e217.e8.

Zhao F, Bi L, Wang W, Wu X, Li Y, Gong F, et al. Mutations of glucocerebrosidase gene and susceptibility to Parkinson's disease: An updated meta-analysis in a European population. *Neuroscience.* 2016; 320: 239-246.

Zimprich A, Benet-Pages A, Struhal W, et al. A mutation in *VPS35*, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am J Hum Genet* 2011; 89: 168-175.