

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

**ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
AGRONÒMICA I DEL MEDI NATURAL**



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Complejidad de las bases genéticas de los trastornos NBIA

Trabajo Final de Grado

Marina Guillot Fernández

Grado en Biotecnología

Curso Académico: 2018/2019

Valencia, septiembre 2019

Tutora: Dra. Carmen Espinós Armero

Co-Tutor: Dr. Vincenzo Lupo

Tutor UPV: Dr. José Javier Forment Millet



Título	Complejidad de las bases genéticas de los trastornos NBIA
Autor/a	Marina Guillot Fernández
Tutora	Dra. Carmen Espinós Armero
Co-Tutora	Dr. Vincenzo Lupo
Tutor UPV	Dr. José Javier Forment Millet
Localidad y fecha	Valencia, Septiembre de 2019

RESUMEN

Las ENACH o NBIA (*Neurodegeneration with Brain Iron Accumulation*) comprenden un grupo de enfermedades hereditarias que suelen cursar con depósitos de hierro en el cerebro, principalmente en los ganglios basales. Nuestro conocimiento sobre las bases moleculares ha mejorado enormemente gracias a la utilización de herramientas de diagnóstico basadas en *Next Generation Sequencing* (NGS), las cuales resultan útiles en el diagnóstico de enfermedades de base genética heterogénea y compleja, como es el caso de las NBIA. Sin embargo, la incapacidad para interpretar las consecuencias clínicas de algunas variantes genéticas novedades descubiertas mediante secuenciación sigue siendo un obstáculo y son necesarios estudios adicionales, como la segregación familiar, ensayos funcionales, análisis de la frecuencias poblacionales y predicciones *in silico*.

En el presente trabajo final de grado se ha realizado el filtrado de datos procedentes de secuenciación de exoma (*WES, whole exome sequencing*) de una familia (fNBIA-20) con el objetivo de determinar las bases genéticas de la enfermedad del probando. Se trata de un caso esporádico, clínicamente caracterizado por una ataxia congénita no evolutiva asociada a discapacidad intelectual con atrofia de cerebelo. Esto ha permitido seleccionar las variantes NM_015207; c. 1121G>A; (p.R374K) en el gen *OTUD3* y NM_002565.3; c. 778G>T; (p. P260T) en el gen *P2RY4* como candidatas a ser causantes del fenotipo del probando.

Por otra parte, se han llevado a cabo ensayos funcionales mediante la sobre-expresión del gen *PLEKHG2* en células HeLa para caracterizar la mutación c.158C>T (p.Thr53Ile), identificada en homocigosis en un paciente con atrofia de cerebelo. Mediante ensayo de luciferasa, se ha observado que la proteína mutada posee una actividad similar a la proteína silvestre, y en el ensayo de inmunofluorescencia no se han mostrado diferencias a nivel de localización subcelular de la proteína mutada respecto a la silvestre. Análisis adicionales de la función del gen y el cribado genético en otros pacientes con la misma u otras mutaciones en *PLEKHG2* ayudará a confirmar la posible asociación entre el fenotipo del paciente y mutaciones en el gen *PLEKHG2*.

Palabras claves: NGS (*Next Generation Sequencing*); Exoma; Diagnóstico genético; ENACH (Enfermedades Neurodegenerativas con Acumulación Cerebral de Hierro); Atrofia cerebelar; gen *PLEKHG2*.

Title	Complexity of the genetic bases of NBIA disorders
Author	Marina Guillot Fernández
Tutora	Dra. Carmen Espinós Armero
Co-Tutora	Dña. María Dolores Martínez Rubio
Tutor UPV	Dr. Vincenzo Lupo
Place and date	Valencia, September 2019

ABSTRACT

The disorders with NBIA (Neurodegeneration with Brain Iron Accumulation) comprise a heterogeneous group of hereditary diseases which are characterized by the presence of iron deposits in the brain, mainly in the basal ganglia.

Our knowledge about molecular bases has greatly improved thanks diagnostic tools based on Next Generation Sequencing (NGS), which are useful in the diagnosis of diseases with an heterogeneous and complex genetic base, as in the case of NBIA. However, sometimes it is not possible to interpret the clinical consequences of some novel genetic variants discovered through sequencing and additional studies are needed, such as family segregation, functional assays, population frequency analysis and in silico predictions.

In this end-of-degree thesis we have filtered data from whole exome sequencing (WES) of a family (fNBIA-20), in order to determine the genetic basis of the disease of the proband. This is a sporadic case, clinically characterized by a non-evolutionary congenital ataxia associated with intellectual disability and cerebellar atrophy. Filtering has allowed to select two variants as candidates for causing the phenotype of the proband: NM_015207; c. 1121G> A; (p.R374K) in the OTUD3 gene and NM_002565.3; c. 778G> T; (p. P260T) in the P2RY4 gene.

On the other hand, functional assays have been carried out by overexpression of the PLEKHG2 gene in HeLa cells to characterize the c.158C> T mutation (p.Thr53Ile), identified in homozygosis in a patient with cerebellar atrophy. By means of luciferase assay, we have observed similar activity levels in the mutant and the wild type protein. Moreover, with the immunofluorescence assay, no differences have been shown at the level of subcellular localization of the mutated protein with respect to the wild one. Additional analysis of the gene function and genetic screening in other patients with the same or other mutations in PLEKHG2 will allow us to confirm a possible association between the phenotype of the patient and mutations in the PLEKHG2 gene.

Keywords: NGS (Next Generation Sequencing); Exome; Genetic diagnosis; NBIA (Neurodegeneration with Brain Iron Accumulation); Cerebellar atrophy; PLEKHG2 gene.

ÍNDICE DE CONTENIDO

1.	INTRODUCCIÓN	1
1.1.	NBIA	1
1.2.	Diagnóstico genético de enfermedades neurodegenerativas.....	3
1.3.	Ensayos funcionales aplicados al diagnóstico genético. Gen <i>PLEKHG2</i>	5
2.	HIPÓTESIS Y OBJETIVOS	10
3.	MATERIALES Y MÉTODOS	11
3.1.	SECUENCIACIÓN DE EXOMA fNBIA-20	11
3.1.1.	Descripción de la familia.....	11
3.1.2.	Secuenciación y <i>variant calling</i>	11
3.1.3.	Filtrado y selección de variantes candidatas	11
3.1.4.	Proceso de priorización de variantes	13
3.1.5.	Validación genética mediante secuenciación Sanger	13
3.1.5.1.	<i>Diseño de cebadores</i>	13
3.1.5.2.	<i>Condiciones de PCR y secuenciación Sanger</i>	14
3.2.	ENSAYO FUNCIONAL DE PLEKHG2	16
3.2.1.	Plásmidos.....	16
3.2.1.1.	<i>Mutagénesis dirigida</i>	17
3.2.1.2.	<i>Transformación de células competentes</i>	18
3.2.1.3.	<i>Purificación de los plásmidos a pequeña y mediana escala</i>	19
3.2.1.4.	<i>Glicerinado de E.Coli</i>	19
3.2.2.	Cultivo de células eucariotas	19
3.2.2.1.	<i>Descripción de la línea celular y condiciones de mantenimiento</i>	19
3.2.2.2.	<i>PCR para la detección de micoplasma</i>	19
3.2.2.3.	<i>Descongelación de la línea celular</i>	20
3.2.3.	Transfección.....	20
3.2.4.	Medida de la actividad de PLEKHG2 mediante ensayo de luciferasa	21
3.2.5.	SDS-PAGE y <i>Western blot</i>	21
3.2.6.	Inmunofluorescencia.....	22
4.	RESULTADOS	23
4.1.	SECUENCIACIÓN DE EXOMA fNBIA-20	23
4.2.	ENSAYO FUNCIONAL DE <i>PLEKHG2</i>	27
4.2.1.	Ensayo de luciferasa	27
4.2.2.	Estudio de localización	29

5. DISCUSIÓN.....	30
5.1. SECUENCIACIÓN DE EXOMA fNBIA-20	30
5.2. ENSAYO FUNCIONAL DE PLEKHG2	31
6. CONCLUSIONES.....	33
7. REFERENCIAS	34
8. ANEXOS.....	38
Anexo I. Electroferogramas correspondientes a la secuenciación Sanger de las muestras de la familia fNBIA-20.	38

ÍNDICE DE FIGURAS

- Figura 1.** Representación esquemática de la localización celular y procesos biológicos implicados en NBIA. El hierro está representado por círculos de color azul claro; las enzimas involucradas en la síntesis de CoA (PANK2, COASY) se encuentran en amarillo; las proteínas asociadas al metabolismo de los lípidos y / o la remodelación de la membrana (PLA2G6, C19orf12, FA2H, SCP2), en azul; las proteínas relacionadas con el autofagosoma / lisosoma (WDR45, ATP13A2), en gris; las proteínas asociadas a la homeostasis del hierro (FTL, CP), en naranja; las proteínas con funciones putativas (DCAF17, GTPBP2), en verde. Los signos de interrogación identifican proteínas con roles aún desconocidos. [Imagen modificada de Schneider et al., (2019)].2
- Figura 2.** Proceso de análisis genético para el diagnóstico de pacientes con NBIA.4
- Figura 3.** Esquema para determinar si una variante se trata de una mutación clínica.5
- Figura 4.** Regulación de actividad GTPasa por GEFs y GAPs [Imagen modificada de Droppelmann et al., 2014]. Se muestra una lista representativa de efectores de GTPasas relacionados con la regulación del sistema nervioso. Las abreviaciones usadas corresponden con: ACK-1, Activated Cdc42-associated kinase 1; BICDR-1, Bicaudal-D-related protein 1; MRCK α , Myotonic dystrophy kinase-related Cdc42-binding kinase alpha; N-WASP, Neural Wiskott-Aldrich syndrome protein; PAK, P21 protein (Cdc42/Rac)-activated kinase; PKN, Serine-threonine protein kinase N; ROCK, Rho-associated coiled-coil containing protein kinase.8
- Figura 5.** Estimulación de PLEKHG2 mediante LPA. La activación de los receptores GPCRs resulta en la liberación de G α y G $\beta\gamma$. G $\beta\gamma$ induce la activación de PLEKHG2, que a su vez activa las GTPasas Rac y Cdc42, iniciando una respuesta biológica.9
- Figura 6.** Pedigrí de la familia fNBIA-20, generado mediante el software Cyrillic 3.11
- Figura 7.** Actividad basal y estimulada con LPA de WT-PLEKHG2, T53I-PLEKHG2 y control vacío (\emptyset), en células HeLa. Las barras de error muestran la desviación estándar del triplicado de un único experimento. Los valores han sido normalizados respecto a la actividad basal de PLEKHG2 WT (con valor de ratio=1).27
- Figura 8.** Western Blot de PLEKHG2 y anti-actina (control de carga). Las células HeLa fueron transfectadas, independientemente, con los plásmidos pcDNA3.1 [Myc/WT-PLEKHG2], pcDNA3.1 [Myc/T53I-PLEKHG2] y pcDNA3.1 [\emptyset]. 24 h post-transfección, las células fueron lisadas y el extracto proteico procesado mediante SDS-PAGE y Western Blot.28
- Figura 9.** Tasa de expresión de PLEKHG2 normalizada frente a la proteína actina, según Western blot.28
- Figura 10.** Actividad luciferasa frente a la cantidad de PLEKHG2. Ésta se calculó dividiendo el ratio de actividad luciferasa/renilla con el ratio de expresión de PLEKHG2/actina.28
- Figura 11.** Inmunofluorescencia de PLEKHG2-WT (A,B) y PLEKHG2-T53I (C,D) en células HeLa transfectadas. Se muestra PLEKHG2-WT, tinción nuclear (DAPI) y la superposición de ambas imágenes (merge). Barra de escala: 20 μ m.29
- Figura 12.** Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en OTUD3 c. 1121G>A (p.R374K).38

Figura 13. Electroferograma correspondiente a la secuenciación de la muestra NBIA-271 para el cambio en OTUD3 c. 1121G>A (p.R374K).....	38
Figura 14. Electroferograma correspondiente a la secuenciación de la muestra NBIA-272 para el cambio en OTUD3 c. 1121G>A (p.R374K).....	39
Figura 15. Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en VPS13B c.8788A>G (p.N2930D).....	39
Figura 16. Electroferograma correspondiente a la secuenciación de la muestra NBIA-271 para el cambio en VPS13B c.8788A>G (p.N2930D)	39
Figura 17. Electroferograma correspondiente a la secuenciación de la muestra NBIA-272 para el cambio en VPS13B c.8788A>G (p.N2930D)	40
Figura 18. Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en VPS13B c.9405G>T.....	40

ÍNDICE DE TABLAS

Tabla 1. Hipótesis de herencia para la generación de ficheros de anotación específicos. 0/0= ausencia de cambio; 0/1= cambio en heterocigosis; 1/1= cambio en homocigosis y en hemizigosis para AR y XR, respectivamente.	12
Tabla 2. Secuencias de cebadores utilizados para amplificar las regiones de interés... 14	14
Tabla 3. Componentes de la mezcla de reacción para cada muestra de PCR.....	15
Tabla 4. Programa general de PCR.	15
Tabla 5. Plásmidos y vectores utilizados.	16
Tabla 6. Cebadores utilizados para la secuenciación de pcDNA3.1 [Myc/WT- <i>PLEKHG2</i>] y pcDNA3.1 [Myc/T53I- <i>PLEKHG2</i>].....	16
Tabla 7. Cebadores utilizados para la mutagénesis dirigida de los plásmidos pcDNA3.1 [Myc/I417F- <i>PLEKHG2</i>] y pcDNA3.1 [Myc/WT- <i>PLEKHG2</i>]	17
Tabla 8. Cebadores utilizados para cada mutación dirigida.....	17
Tabla 9. Componentes de la mezcla de PCR.....	18
Tabla 10. Programa de PCR para la mutagénesis dirigida	18
Tabla 11. Cebadores utilizados para la PCR para detectar micoplasma.	20
Tabla 12. Programa de PCR para detectar micoplasma.	20
Tabla 13. Componentes de la mezcla de reacción para cada muestra de la PCR para detectar micoplasma.	20
Tabla 14. Variantes noveles tras el proceso de priorización y modo de herencia asociada.	23
Tabla 15. Variantes noveles de novo.....	25
Tabla 16. Variantes bialélicas con modo de herencia AR.....	25
Tabla 17. Variantes noveles con herencia recesiva ligada al cromosoma X	26

LISTA DE ABREVIATURAS

CNAG	Centro Nacional de Análisis Genómico (<i>Centre Nacional d'Anàlisi Genòmica</i>)
CNV	Variación en número de copias (<i>Copy Number Variation</i>)
CoA	Coenzima A
DH	Homología a Dbl (<i>Dbl Homology</i>)
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
dNTPs	Desoxirribonucleótidos
ENACH/NBIA	Enfermedades Neurodegenerativas con Acumulación Cerebral de Hierro (<i>Neurodegeneration with Brain Iron Accumulation</i>)
<i>E. Coli</i>	<i>Escherichia Coli</i>
GAP	Proteínas activadoras de GTPasa (<i>guanine activating proteins</i>)
GDP	Guanosín difosfato
GEF	Factor intercambiador de nucleótidos (<i>Guanine nucleotide exchange factor</i>)
gnomAD	<i>The Genome Aggregation Gatabase</i>
GPCR	Receptor acoplado a proteínas G (<i>G protein-coupled receptor</i>)
GTP	Guanosín trifosfato
HGVs	Sociedad de Variación del Genoma Humano (<i>Human Genome Variation Society</i>)
HSF	<i>Human Splicing Finder</i>
IGV	<i>Integrative Genomics Viewer</i>
LB	Luria Bertani
LPA	Ácido lipofosfatídico
MAF	Frecuencia del alelo menos común (<i>minor allele frequency</i>)
MLPA	Amplificación de sondas dependiente de ligandos múltiples (<i>Multiplex Ligation-dependent Probe Amplification</i>)
NGS	Secuenciación de nueva generación (<i>Next Generation Sequencing</i>)
PBS	Tampón fosfato salino (PBS, <i>phosphate buffered saline</i>)
PCR	Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>)
PFA	Paraformaldehído
PH	Homología a Pleckstrina (<i>Pleckstrin Homology</i>)
PKAN	Neurodegeneración asociada a pantotenato quinasa (<i>Pantothenate Kinase-Associated Neurodegeneration</i>)
PLAN	Neurodegeneración asociada a fosfolipasa-A2 (<i>PhosphoLipase A2 Associated Neurodegeneration</i>)
PLEKHG2	<i>Pleckstrin Homology Domain Containing, Family G, Member 2</i>
RNA	Ácido ribonucleico
SFB	Suero fetal bovino
SGGT	Servicio de Genómica y Genética Traslacional

SNC	Sistema nervioso central
SNP	Polimorfismos de un solo nucleótido (<i>single nucleotide polymorphism</i>)
SRE	Elemento de respuesta a suero (<i>serum response element</i>)
TS	Secuenciación dirigida (<i>targeted sequencing</i>)
UGGENN	Unidad de Genética y Genómica de Enfermedades Neuromusculares y Neurodegenerativas
VCF	Formato de llamada de variantes (<i>variant call format</i>)
VUS	Variante de significado incierto (<i>variants of unknown significance</i>)
WES	Secuenciación del genoma completo (<i>whole genome sequencing</i>)
WGS	Secuenciación del exoma completo (<i>whole exome sequencing</i>)
WT	Silvestre (<i>wild type</i>)
XR	Recesiva ligada al cromosoma X

1. INTRODUCCIÓN

1.1. NBIA

Las ENACH (Enfermedades Neurodegenerativas con Acumulación Cerebral de Hierro) o NBIA (*Neurodegeneration with Brain Iron Accumulation*) comprenden un grupo heterogéneo clínica y genéticamente de trastornos del movimiento, la mayoría de las cuales presentan depósitos de hierro en el cerebro y disfunción motora progresiva (Kurian y Hayflick, 2013; Tello *et al.*, 2017). Se trata de enfermedades ultra-raras, con una prevalencia estimada de 1/1.000.000 (Tello *et al.*, 2017).

La acumulación de hierro puede ser detectada *in vivo* mediante análisis de resonancia magnética y afecta principalmente a las regiones de los ganglios basales, como el globo pálido y la sustancia negra, aunque también puede afectar a otras regiones como la corteza y el cerebelo (Levi y Tiranti, 2019).

Se trata de enfermedades mendelianas o monogénicas que pueden seguir un patrón de herencia autosómico recesivo, dominante o ligado al cromosoma X (Gregory y Hayflick, 2013). Su avance es devastador e incluye dificultades progresivas en el habla, disfagia, distonía, espasticidad, parkinsonismo, acinesia pura, disfunción oculomotora, pérdida de visión y síntomas neuropsiquiátricos. Los pacientes precisan de la atención de un equipo multidisciplinario y reciben un tratamiento paliativo que se centra principalmente en reducir la distonía y la espasticidad que son altamente incapacitantes (Tello *et al.*, 2017).

Hasta la fecha se conocen 12 genes NBIA, cuyos productos están implicados en un amplio espectro de actividades biológicas (Tello *et al.*, 2017; Di Meo y Tiranti, 2018), aunque un número relevante de pacientes permanece sin diagnóstico genético. El mecanismo de la enfermedad no se conoce bien y, a pesar de que estas enfermedades se caracterizan por la acumulación de hierro en el cerebro, únicamente los genes *CP* y *FTL* están implicados directamente en el metabolismo del hierro. El resto de genes participan principalmente en la biosíntesis de coenzima A (CoA) (*PANK2* y *COASY*), autofagia (*WDR45* and *ATP13A2*), metabolismo lípido (*PLA2G6*, *FA2H* y *C19orf12*) y remodelación de membrana (*FTL* and *FA2H*). Por último, *DCAF17* codifica para una proteína nucleolar de función desconocida (Di Meo y Tiranti, 2018) (Fig. 1).

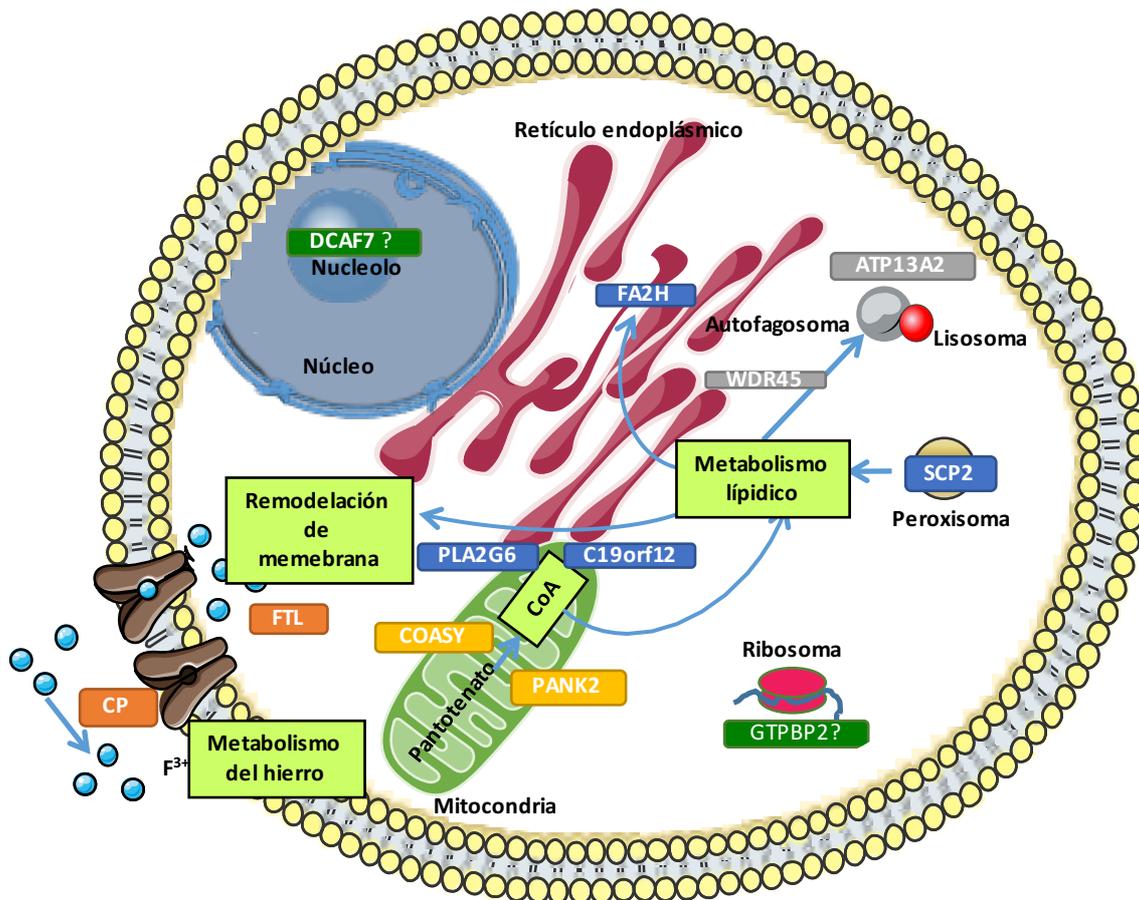


Figura 1. Representación esquemática de la localización celular y procesos biológicos implicados en NBIA. El hierro está representado por círculos de color azul claro; las enzimas involucradas en la síntesis de CoA (PANK2, COASY) se encuentran en amarillo; las proteínas asociadas al metabolismo de los lípidos y/o la remodelación de la membrana (PLA2G6, C19orf12, FA2H, SCP2), en azul; las proteínas relacionadas con el autofagosoma / lisosoma (WDR45, ATP13A2), en gris; las proteínas asociadas a la homeostasis del hierro (FTL, CP), en naranja; las proteínas con funciones putativas (DCAF17, GTPBP2), en verde. Los signos de interrogación identifican proteínas con roles aún desconocidos. [Imagen modificada de Schneider et al., (2019)].

Alrededor del 35-50% de pacientes NBIA padecen PKAN (*Pantothenate Kinase-Associated Neurodegeneration*, gen *PANK2*), mientras que la segunda forma más frecuente es PLAN (*PhosphoLipase A2 Associated Neurodegeneration*, gen *PLA2G6*) que responde a aproximadamente el 20% de los pacientes NBIA (Gregory y Hayflick, 2013). El resto de formas son poco frecuentes.

Los esfuerzos realizados en la última década en la identificación de nuevos genes asociados con NBIA han ayudado a reconocer un mayor rango de fenotipos en este grupo de enfermedades. Sin embargo, se desconoce la razón por la cual el hierro se acumula. De hecho, la deposición de hierro puede preceder al desarrollo de síntomas clínicos o no llegar a suceder. Por lo tanto, pese a que la acumulación de hierro es una característica útil en neuroimagen de NBIA, probablemente no sea necesaria ni suficiente para producir el fenotipo de la enfermedad en la mayoría de formas clínicas, dado que el grado de deposición del hierro se correlaciona de manera incompleta con los síntomas clínicos (Kruer et al., 2012).

Por otra parte, las NBIA están frecuentemente asociadas con atrofia de cerebelo. Por ejemplo, la atrofia de cerebelo es característica de PLAN y del síndrome de Kufor-Rakeb (Tello et al., 2017). El cerebelo ejerce una gran influencia en el comportamiento, dadas sus interconexiones anatómicas y funcionales con los ganglios basales y gran parte del manto cortical (Bostan et

al., 2013). Además, está involucrado en la planificación y ejecución motora, así como en un abanico de funciones emocionales y cognitivas complejas (Ito, 2006; Ramnani, 2006).

La comprensión de los mecanismos subyacentes en NBIA, junto con la identificación de los nuevos genes, es importante no solo para desarrollar estrategias terapéuticas para tratar a los pacientes con NBIA, sino también para allanar el camino en el estudio de otras enfermedades más comunes con características similares, como la demencia frontotemporal, la enfermedad de Parkinson, la enfermedad de Alzheimer, la ataxia de Friedreich y esclerosis lateral amiotrófica (Di Meo y Tiranti, 2018).

1.2. Diagnóstico genético de enfermedades neurodegenerativas

En los últimos años, nuestro conocimiento sobre las bases moleculares ha mejorado enormemente gracias a la utilización de herramientas de diagnóstico basadas en *Next Generation Sequencing* (NGS). Las tecnologías NGS resultan útiles en el diagnóstico de enfermedades de base genética heterogénea y compleja, en las que existe un gran número de genes implicados en un amplio espectro fenotípico (Di Resta *et al.*, 2018), como es el caso de las NBIA. En estos casos, las técnicas NGS permiten analizar un gran número de genes simultáneamente de forma más económica que la secuenciación Sanger (Williams y Hegde, 2013) y realizar diagnósticos genéticos en cortos periodos de tiempo (Vrijenhoek *et al.*, 2015).

En la actualidad, existen distintas opciones para el diagnóstico molecular, las cuales permiten la secuenciación seleccionada de fragmentos del genoma dependiendo de la captura realizada. Estas tres opciones comprenden la secuenciación de una selección de genes concretos (TS, *Targeted Sequencing*), del exoma (WES, *Whole Exome Sequencing*) o del genoma (WGS, *Whole Genome Sequencing*).

La utilización de TS, también conocida como paneles de genes, permite el análisis de múltiples genes potencialmente relevantes desde el punto de vista clínico. Suele ofrecer una buena cobertura para la totalidad del panel analizado (Aziz *et al.*, 2015; Johnston *et al.*, 2012) y se trata de una metodología relativamente económica. Otra de las ventajas del uso de paneles frente a WES o WGS es la reducción del número de hallazgos incidentales y/o el número de variantes de significado incierto (VUS, *Variants of Unknown Significance*), de las cuales se hablará más adelante. Sin embargo, la principal limitación de la secuenciación dirigida es la rigidez de analizar únicamente un número seleccionado de genes. Puesto que el campo de la genética evoluciona rápidamente y aparecen nuevos genes que pueden ser asociados con un fenotipo clínico, se hace necesario el re-diseño y la revalidación de los paneles en periodos de tiempo relativamente cortos para que éstos permanezcan actualizados (Williams y Hedge, 2013; Rehm *et al.*, 2013).

Por otra parte, WES analiza las regiones codificantes de proteínas en el genoma. A diferencia del panel de genes, el análisis mediante WES permite el descubrimiento de nuevos genes implicados en un fenotipo. La aplicación de WES es útil, por ejemplo, en los casos negativos tras secuenciación con panel de genes, especialmente cuando se realiza en trío (constituido por el probando –descendiente afectado- y los padres) (Di Resta *et al.*, 2018). Sin embargo, debemos tener en cuenta las limitaciones de WES. La cobertura completa de todos los exones codificantes no es factible en la práctica puesto que es posible que las sondas no hibriden a ciertos genes, o por la presencia de secuencias repetitivas o zonas ricas en GCs (Rehm, 2013). Asimismo, ciertas regiones que codifican para proteínas pueden no estar cubiertas debido a la anotación incompleta del genoma humano (Di Resta *et al.*, 2018; Goh y Choi, 2012).

Por último, WGS proporciona información sobre el genoma completo. El genoma contiene regiones no traducidas que pueden tener un papel regulador (por ejemplo, RNA no codificante o sitios de unión a la transcripción) junto con potenciales sitios codificantes de proteínas que aún no han sido anotados como genes (Chrystoja y Diamandis, 2014; Lohmann y Klein, 2014). Por lo tanto, WGS añade información sobre mutaciones en regiones reguladoras o sobre variaciones en el número de copia (CNV, *Copy Number Variation*) con mayor eficiencia que WES (Poduri *et al.*, 2014; Stavropoulos *et al.*, 2016). Se trata del enfoque más completo y se espera que se convierta en el método de elección en el futuro. No obstante, todavía existen notables limitaciones para WGS, ya que tiene un elevado coste (tanto de secuenciación como de almacenamiento de datos), y la interpretación de la enorme cantidad de datos generados supone un desafío (Olgati *et al.*, 2016)

En cuanto a la aplicación de las tecnologías NGS al diagnóstico genético de pacientes con NBIA, la Unidad de Genética y Genómica de Enfermedades Neuromusculares y Neurodegenerativas (UGGENN) inició en 2015 un proyecto financiado por la Fundació La Marató de la TV3, coordinado por la Dra. B. Pérez-Dueñas (H.U. Vall d'Hebron), para la elaboración de una escala clínica para pacientes PKAN y el diagnóstico genético de la población española NBIA con la colaboración de más de 20 hospitales españoles (Fig. 2). La serie clínica objeto de estudio comprende 350 pacientes y familiares pertenecientes a 134 familias no relacionadas.



Figura 2. Proceso de análisis genético para el diagnóstico de pacientes con NBIA.

En primer lugar, se realiza el análisis genético de *PANK2* y *PLA2G6* mediante secuenciación Sanger y amplificación de sondas dependiente de ligandos múltiples (MLPA, *Multiplex Ligation-dependent Probe Amplification*). Esto ha permitido diagnosticar el 54% de los casos.

Los pacientes que resultan negativos para estos análisis son investigados mediante un panel diseñado con tecnología *SureSelectXT* (Agilent) para Illumina, el cual comprende 498 genes implicados en NBIA y otros trastornos del movimiento (MovDisord-498; Fig. 2). Treinta y ocho pacientes han sido analizados con esta metodología y, de momento, se ha logrado el diagnóstico genético en 19 de ellos. Los resultados genéticos han revelado que el espectro clínico es más amplio del conocido hasta ahora.

Los casos que son negativos tras su estudio mediante el panel MovDisord-498 son estudiados mediante WES (Fig. 2). Se ha descrito que aproximadamente el 85% de las mutaciones patológicas se localizan en los exones (Clamp *et al.*, 2008; Choi *et al.*, 2009). Por ello, pese a que la WES sólo supone la secuenciación de alrededor del 1-2% del genoma humano (Choi *et al.*, 2009), se ha demostrado que es una herramienta que proporciona un buen rendimiento para el diagnóstico genético. En el caso de enfermedades pediátricas del neurodesarrollo, estudios comparativos han mostrado que cuando sólo se aplica WES a la muestra del probando, la tasa de éxito ronda el 21%, mientras que si se amplía el estudio a los padres (trío) ésta aumenta un 37% (Farwell *et al.*, 2015).

Es importante recalcar que la mayoría de casos son esporádicos y por tanto, es de esperar mutaciones *de novo* o bien, mutaciones con una herencia autosómica recesiva. Indudablemente, la WES permite la caracterización de nuevos genes y de nuevas mutaciones en genes conocidos, definiendo mejor la clínica y la genética asociada a estas enfermedades, lo que de por sí es extraordinariamente complejo dada la baja casuística existente.

En el marco del presente TFG, mediante WES se investigará una familia (fNBIA-20) con el objetivo de determinar las bases genéticas de la enfermedad en el probando, previamente estudiado mediante panel MovDisord-498 sin resultados de interés. Se trata de un caso esporádico clínicamente caracterizado por una ataxia congénita no evolutiva asociada a discapacidad intelectual con atrofia de cerebelo.

1.3. Ensayos funcionales aplicados al diagnóstico genético. Gen *PLEKHG2*

Los avances tecnológicos han hecho que la secuenciación rutinaria de los genomas sea cada vez más práctica, incluso en entornos clínicos. Sin embargo, nuestra incapacidad para interpretar las consecuencias clínicas de las variantes genéticas descubiertas mediante secuenciación sigue siendo un obstáculo fundamental para el progreso de la medicina de precisión (Starita *et al.*, 2017).

La mayoría de variantes interpretadas son consideradas VUS (Starita *et al.*, 2017), dado que la información sobre su asociación con la enfermedad es incompleta, y el impacto en la función del gen o proteína no es clara (Woods *et al.*, 2016). En estos casos, son necesarios estudios adicionales, como la segregación familiar, ensayos funcionales, análisis de la frecuencias poblacionales y predicciones *in silico* (Starita *et al.*, 2017; Richards *et al.*, 2015). Este proceso de análisis se encuentra recogido en la Fig. 3.

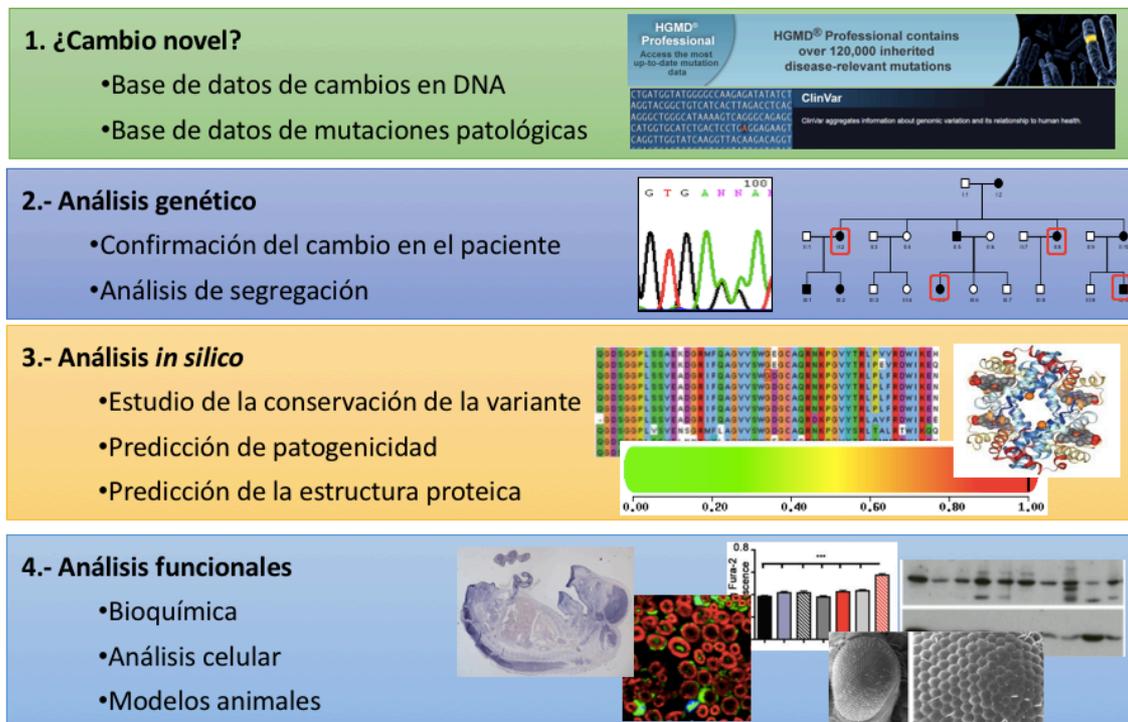


Figura 3. Esquema para determinar si una variante se trata de una mutación clínica.

Al identificar un cambio, el primer paso para determinar su posible significación biológica consiste en anotar la frecuencia del cambio en bases de datos poblacionales, lo que permite descartar polimorfismos benignos. No obstante, no podemos asumir que estas bases de datos contengan únicamente individuos sanos, y por tanto, pueden contener variantes patogénicas no identificadas como tales (Richards *et al.*, 2015). Por otra parte, las bases de mutaciones patológicas aportan información sobre si la variante de interés ha sido asociada con el fenotipo. Tras esto, es necesario confirmar el cambio en el paciente (mediante secuenciación Sanger) y la realización de análisis de segregación, para determinar si la variante co-segrega con el fenotipo de la enfermedad.

En cuanto a los estudios *in silico*, hay una gran variedad de herramientas que pueden ayudar en la interpretación de variantes. Las dos categorías principales de estas herramientas incluyen aquellas que predicen si un cambio *missense* es dañino para la función de la proteína, como SIFT, *Sorting Intolerant from Tolerant*, (<https://sift.bii.a-star.edu.sg>) o PolyPhen-2 (*Polymorphism Phenotyping v2*, <http://genetics.bwh.harvard.edu/pph2/>), y aquellas que predicen si una determinada variante tiene un efecto sobre el *splicing*, como NNSPLICE (https://www.fruitfly.org/seq_tools/splice.html) o HSF (*Human Splicing Finder*, <http://www.umd.be/HSF3/HSF.shtml>) (Richards *et al.*, 2015). No obstante, su fiabilidad es insuficiente para utilizarlos como base en la toma de decisiones clínicas, ya que tienen un carácter meramente predictivo (Weile y Roth, 2018). Son también análisis computacionales que aportan información para la interpretación de una variante, el modelaje de la estructura tridimensional de la proteína mutada y estudio de la conservación del residuo. Respecto a la primera de estas aproximaciones, cambios de aminoácido pueden causar graves alteraciones en la conformación final de la proteína alterando su función. El modelaje de la estructura tridimensional de la proteína mutada permite un análisis a nivel atómico de las posiciones proteicas que pueden ser importantes para la estabilidad de la proteína o la formación de interacciones con otras proteínas (Bhattacharya *et al.*, 2017). Y en cuanto al estudio de la conservación del aminoácido, si éste está conservado en muy diferentes organismos a lo largo de la evolución, es probable que sea porque es preciso para la actividad de la proteína.

Con el propósito de ganar evidencias que demuestren que el cambio detectado contribuye a la enfermedad, se deben llevar a cabo ensayos funcionales (Starita *et al.*, 2017; Raraigh *et al.*, 2018). La dificultad más relevante es que es preciso diseñar el estudio según la actividad del gen y puede resultar muy difícil y costoso.

En el proceso de diagnóstico genético de un caso de atrofia de cerebelo, se detectó el cambio en homocigosis *PLEKHG2* c.158C>T (p.Thr53Ile) mediante WES de un trío. El paciente había sido analizado previamente mediante el panel de genes *MovDisord-498*, mencionado anteriormente, sin resultados de interés. Además, el análisis *in silico* de patogenicidad hecho con SIFT, PolyPhen-2 y MutationTaster (<http://www.mutationtaster.org>) predicen que la variante p.T53I es deletérea.

La mutación objeto de estudio en el presente trabajo (c.158C>T, p.T53I) corresponde al segundo caso debido a mutaciones en *PLEKHG2* implicado en leucodistrofia y microcefalia, con o sin distonía (MIM 616763). La primera mutación caracterizada en *PLEKHG2*, c.610C>T (p.R204W), ha sido identificada en homocigosis en cinco pacientes de dos familias con profunda discapacidad intelectual, distonía, microcefalia postnatal y un patrón distintivo en neuroimagen. Mediante ensayo de luciferasa, Edvardson *et al.* (2016) mostró que la actividad de la proteína mutante disminuye significativamente.

El gen *PLEKHG2* codifica una proteína RhoGEF (*Rho guanine nucleotide exchange factor*) de 1387 aminoácidos, miembro de la familia Dbl (Zheng, 2001). Las proteínas de este grupo se

caracterizan por compartir dos regiones conservadas: el dominio de homología a Dbl (DH, *Dbl Homology*) y el dominio de homología a pleckstrina (PH, *Pleckstrin Homology*), los cuales se encuentran en los residuos 105-281 y 263-413, respectivamente, en la proteína *PLEKHG2* (Edvardson, 2016). El dominio DH regula la especificidad por las RhoGTPasas y la actividad intercambiadora de nucleótidos, mientras que el dominio PH tiene diversos roles, que incluyen dirigir la localización en membrana y la regulación de la actividad de factor intercambiador de nucleótidos (GEF, *guanine nucleotide exchange factor*) basal de la proteína (Rossman *et al.*, 2005; Zheng, 2001).

Las proteínas RhoGEF tienen un importante papel en la activación de las Rho GTPasas. Éstas son una familia de de veinticinco miembros de la superfamilia Ras de pequeñas GTPasas (~20 kDa). Las RhoGTPasas pueden dividirse en distintos subgrupos en función de su homología de secuencia, entre los cuales destacan los subgrupos Rac, Rho y Cdc42 por ser los más estudiados y caracterizados (Jaffe y Hall, 2005). Al igual que todas las GTPasas, las proteínas Rho actúan como interruptores binarios al alternar entre un estado conformacional inactivo -unido al guanosín difosfato (GDP)- y activo -unido al guanosín trifosfato (GTP)- (Van Aelst y D'Souza, 1997). Este mecanismo es controlado por la célula, la cual regula la interconversión y accesibilidad entre las dos formas de varias maneras (Fig. 4). Los GEFs estimulan el intercambio de GDP por GTP para generar la forma activada de las GTPasas, que luego son capaces de reconocer dianas localizadas *downstream* o proteínas efectoras (Schmidt and Hall, 2002), activando así distintas respuestas celulares. Por otra parte, las proteínas activadoras de GTPasa (GAPs, *guanine activating proteins*) catalizan la hidrólisis de GTP, inactivando las GTPasas (Schmidt y Debant, 2014).

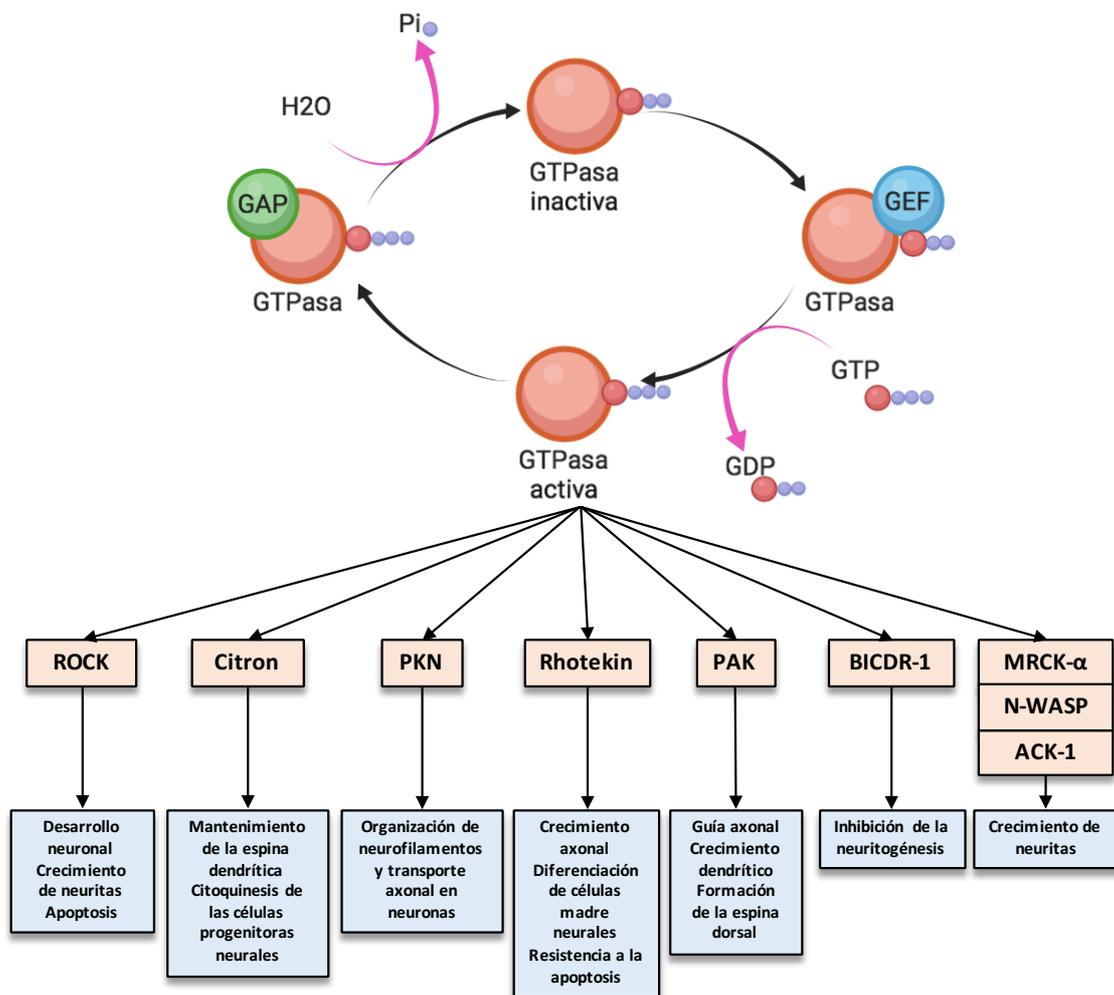


Figura 4. Regulación de actividad GTPasa por GEFs y GAPs [Imagen modificada de Droppelmann et al., 2014]. Se muestra una lista representativa de efectores de GTPasas relacionados con la regulación del sistema nervioso. Las abreviaciones usadas corresponden con: ACK-1, Activated Cdc42-associated kinase 1; BICDR-1, Bicaudal-D-related protein 1; MRCK α , Myotonic dystrophy kinase-related Cdc42-binding kinase alpha; N-WASP, Neural Wiskott-Aldrich syndrome protein; PAK, P21 protein (Cdc42/Rac)-activated kinase; PKN, Serine-threonine protein kinase N; ROCK, Rho-associated coiled-coil containing protein kinase.

Dado que las RhoGTPasas poseen una participación esencial en prácticamente todos los aspectos de la fisiología celular normal, no es sorprendente que una actividad aberrante de estas proteínas, así como de las proteínas RhoGEFs, esté asociada al cáncer y otras enfermedades (Cook *et al.*, 2014). Por ejemplo, mutaciones causantes de pérdida de función en el gen *FGD4*, el cual codifica para una GEF específica para Cdc42, han sido asociadas con la enfermedad de Charcot-Marie-Tooth tipo 4H (Delague *et al.*, 2007; Stendel *et al.*, 2007).

Por otra parte, los receptores acoplados a proteínas G (GPCRs, *G protein-coupled receptors*) son una clase de receptores de superficie celular, los cuales pueden inducir la activación de RhoGEFs a través de la interacción directa de las subunidades heterotriméricas (α , β y γ) de las proteínas G con una GEF (Runne, 2013). El ácido lipofosfatídico (LPA) es un fosfolípido soluble en agua que estimula ciertos receptores GPCRs (Lin *et al.*, 2010), induciendo un cambio conformacional que activa las subunidades heterotriméricas de las proteínas G (Rosenbaum *et al.*, 2009). En concreto, Ueda *et al.* (2008) mostraron que PLEKHG2 es activado

específicamente por las subunidades $G\beta\gamma$ (Fig. 5). La estimulación de PLEKHG2 por $G\beta\gamma$ induce la transcripción del elemento de respuesta al suero (SRE, *serum response element*) a través de la activación de Rac y Cdc42. Además, Runne y Chen (2013) demostraron que PLEKHG2 facilita la polimerización de actina y y la migración de linfocitos a través de la activación de Cdc42 y Rac.

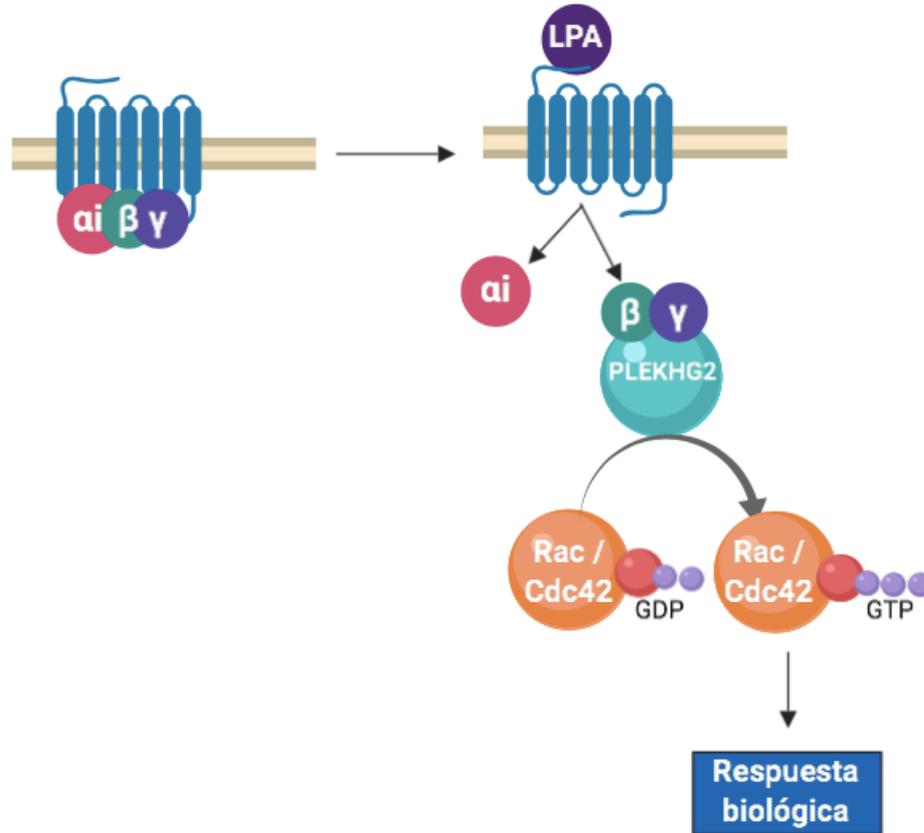


Figura 5. Estimulación de PLEKHG2 mediante LPA. La activación de los receptores GPCRs resulta en la liberación de $G\alpha$ y $G\beta\gamma$. $G\beta\gamma$ induce la activación de PLEKHG2, que a su vez activa las GTPAsas Rac y Cdc42, iniciando una respuesta biológica.

En este trabajo se llevarán a cabo ensayos funcionales para caracterizar la mutación *PLEKHG2* c.158C>T (p.Thr53Ile). Se investigará la actividad de la proteína mutada mediante ensayo de luciferasa y se realizarán estudios de localización subcelular mediante inmunofluorescencia para determinar si la proteína mutada pierde su correcta localización en la membrana celular.

2. HIPÓTESIS Y OBJETIVOS

Las bases genéticas de las enfermedades NBIA y trastornos del movimiento relacionados son complejas con cientos de genes asociados. La WES permite la identificación de variantes en genes no asociados con enfermedad, por lo que proponemos que sería posible caracterizar la mutación causal en el probando de la familia fNBIA-20 en quien ya se han descartado mutaciones en 498 genes mediante panel.

Por otra parte, la interpretación de las variantes detectadas mediante WES es habitualmente difícil. Por ello, se hacen necesarios ensayos funcionales que ayuden a entender la posible patogenicidad de una mutación. Nuestra hipótesis es que los estudios funcionales de *PLEKHG2* c.158C>T basados en ensayos de actividad y de localización subcelular, mostrarán una alteración en el perfil de la proteína mutada. Ello evidenciaría que su actividad cuando mutada es anómala y con ello, apoyará su implicación en el desarrollo de la enfermedad.

El **objetivo general** del presente TFG es ahondar en las bases moleculares de las enfermedades NBIA y trastornos relacionados.

Los **objetivos específicos** son:

- 1.- Descubrir la causa genética que subyace en la enfermedad en un paciente en quien se han descartado mutaciones en 498 genes (MovDisord-498), estudiado por WES.
- 2.- Investigar cómo la mutación novel *PLEKHG2* c.158C>T (p.Thr53Ile) identificada en un paciente en homocigosis altera la función correcta de la proteína causando enfermedad.

Palabras claves: NGS (*Next Generation Sequencing*); Exoma; Diagnóstico genético; NBIA (*Neurodegeneration with Brain Iron Accumulation*) /ENACH (Enfermedades Neurodegenerativas con Acumulación Cerebral de Hierro); Atrofia cerebelar; gen *PLEKHG2*.

3. MATERIALES Y MÉTODOS

3.1. SECUENCIACIÓN DE EXOMA fNBIA-20

3.1.1. Descripción de la familia

En el presente trabajo se ha realizado el filtrado de datos genómicos procedentes de la secuenciación de exoma de la familia fNBIA-20, en la que se estudiaron el probando (NBIA-24) y a sus progenitores (NBIA-272 y NBIA-271), representados en el pedigrí de la figura 6.

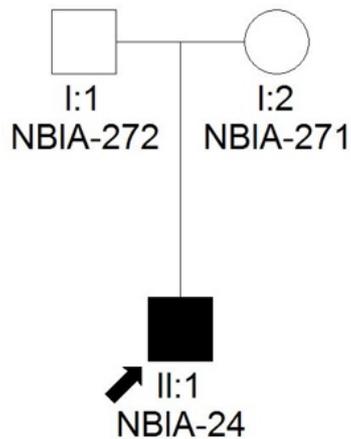


Figura 6. Pedigrí de la familia fNBIA-20, generado mediante el software Cyrillic 3.

El fenotipo del paciente corresponde con ataxia cerebelosa congénita no evolutiva, asociada a discapacidad intelectual, signos de espasticidad en miembros inferiores y estrabismo. Se trata de un caso esporádico, puesto que no existen antecedentes familiares. En las pruebas complementarias destaca atrofia de cerebelo en neuroimagen.

3.1.2. Secuenciación y *variant calling*

El primer paso en la búsqueda de posibles variantes causantes de la patología de NBIA-24 fue el análisis de la muestra de DNA del paciente mediante el panel de genes de diseño propio del laboratorio (UGGENN) Mov-Disord-498, mencionado anteriormente. El DNA fue extraído a partir de una muestra de sangre periférica. Puesto que no se identificaron variantes candidatas, se procedió a la secuenciación de exoma para continuar con el estudio. Para ello, se secuenciaron los exomas del probando y sus progenitores.

Tanto la secuenciación de exoma como la llamada de variantes (en inglés, *variant calling*) fueron llevados a cabo por el *Centre Nacional d'Anàlisi Genòmica* (CNAG, Barcelona). El CNAG proporcionó un fichero VCF (*variant call format*) anotado y, además, se generó otro fichero de anotación paralelamente en el laboratorio, utilizando la herramienta *Variant Effect Predictor* (VEP) de Ensembl (<https://www.ensembl.org/info/docs/tools/vep/index.html>). Las secuencias fueron alineadas al genoma de referencia GRCh37/hg19.

3.1.3. Filtrado y selección de variantes candidatas

Debido al gran número de variantes detectadas, se realizó un proceso de filtrado para reducir el número de éstas y priorizar las que pudieran ser candidatas a ser responsables de la enfermedad. Para ello, se utilizó un *pipeline* propio del laboratorio, que consiste en diversos pasos de filtrado y selección de variantes candidatas.

En primer lugar, se generaron cuatro ficheros de anotación en formato Excel que permitieron trabajar manualmente con los datos. Al tratarse de un caso esporádico, lo esperable es identificar mutaciones *de novo* o mutaciones bialélicas asociadas a un modo de herencia autosómica recesiva (AR) y, tratándose de un varón, no es de descartar una herencia recesiva ligada al cromosoma X (XR). Por lo tanto, para la generación de los 4 ficheros de anotación asociados al modo de herencia se tendrán en cuenta las combinaciones de genotipo que se detallan en la tabla 1.

Tabla 1. Hipótesis de herencia para la generación de ficheros de anotación específicos. 0/0= ausencia de cambio; 0/1= cambio en heterocigosis; 1/1= cambio en homocigosis y en hemizigosis para AR y XR, respectivamente.

Muestra	fNBIA	Individuo	<i>de novo</i>	AR-C-HET	AR-HOM	XR
NBIA-272	fNBIA-20	padre	0/0	0/0, 0/1	0/1	0/0
NBIA-271		madre	0/0	0/0, 0/1	0/1	0/1
NBIA-24		probando	0/1	0/1	1/1	1/1

A continuación, se procedió a la selección de variantes candidatas. Se seleccionaron únicamente las variantes etiquetadas como “*protein_coding*”, es decir, aquellas que se encuentran en transcritos anotados que codifican para proteína.

Además, se descartaron todas aquellas variantes con una frecuencia del alelo menos común (MAF, *minor allele frequency*) general superior a 0,5% en la base de datos *The Genome Aggregation Database* (gnomAD, <http://gnomad.broadinstitute.org>), donde se recogen secuencias exónicas y genómicas de individuos no relacionados secuenciados como parte de diversos estudios genéticos poblacionales y específicos para ciertas enfermedades. Tras este paso, se eliminaron también aquellas variantes con una MAF superior a 1% en alguna población en concreto.

Por otra parte, únicamente se tuvieron en cuenta las variantes localizadas dentro de la secuencia codificante e intrónica flanqueante que pudieran alterar el sitio de corte y empalme del mRNA (*splicing*).

Tras este paso de filtrado, las variantes restantes fueron manualmente clasificadas en tres grupos en función de los datos de frecuencia encontrados en gnomAD. El grupo de mayor interés está constituido por los cambios noveles, que no aparecen en la población control. El segundo grupo comprende cambios que aparecen en menos de 10 alelos, mientras que el tercer grupo agrupa aquellas variantes que aparecen entre 10 y 50 veces. Las variantes se priorizaron en función del modo de herencia que pudiera estar asociado a cada gen, es decir, las variantes noveles se consideraron mejores candidatas respecto a las que pertenecían al segundo y tercer grupo. Para el caso de dos cambios en un mismo gen y supuestamente asociadas a una herencia autosómica recesiva, se tuvieron en cuenta también variantes con frecuencia relativamente baja pero que no fueran anotadas en homocigosis en la base de datos gnomAD.

En todos los casos, se comprobó que los parámetros de calidad de la base leída tras la secuenciación fuesen óptimos y que la cobertura de la misma estuviese por encima de valores 10x. Para ello, la calidad de las variantes seleccionadas fue además revisada en el programa *Integrative Genomics Viewer* (IGV) (<https://software.broadinstitute.org/software/igv/>).

3.1.4. Proceso de priorización de variantes

Una vez realizado el proceso de selección de variables candidatas, se utilizaron diversas bases de datos para obtener más información sobre cada variante y el gen en el que se localizaban.

En primer lugar, se utilizaron las bases de datos *Online Mendelian Inheritance in Man* (OMIM, <https://www.omim.org>), ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) y *Human Gene Mutation Database* (HGMD, <http://www.hgmd.cf.ac.uk/ac/index.php>) para comprobar si los genes/variantes en cuestión habían sido relacionados previamente con alguna patología.

Además, consultamos la base de datos *The Human Protein Atlas* (<http://www.proteinatlas.org>), que muestra los patrones de expresión de la proteína de interés en distintos tejidos. Se utilizaron también las bases de datos Bgee (<https://bgee.org>), que contiene información sobre patrones de expresión génica en múltiples especies animales a partir de distintos tipos de datos (RNA-Seq, Affymetrix, hibridación in situ y datos EST), y Brain RNA Seq (https://web.stanford.edu/group/barres_lab/brain_rnaseq.html), la cual es una base de datos de RNA-Seq y *splicing* de glía, neuronas y células vasculares del córtex cerebral. Por otra parte, se buscó información del gen en modelos animales, como *The Mouse Genome Database* (MGD, <http://www.informatics.jax.org>) sobre ratón.

Finalmente, utilizamos predictores *in silico* para analizar las variantes de interés. Para las variantes *missense* se utilizaron los predictores SIFT (<https://sift.bii.a-star.edu.sg>) y PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>). El primero basa la predicción en el nivel de conservación de la secuencia, mientras que el segundo se basa en el estructura/función de la proteína, además del nivel de conservación de la secuencia. Para los cambios sinónimos, se utilizaron predictores de defectos de *splicing*. Estos fueron NNSPLICE (https://www.fruitfly.org/seq_tools/splice.html), *Human Splicing Finder* (HSF, <http://www.umd.be/HSF/>) y NetGene2 Server (<http://www.cbs.dtu.dk/services/NetGene2/>).

3.1.5. Validación genética mediante secuenciación Sanger

Tras el proceso de filtrado y priorización, se seleccionaron tres variantes como candidatas, por lo que se procedió a amplificar las regiones génicas de cada variante mediante PCR para poder llevar a cabo su validación mediante Sanger.

3.1.5.1. Diseño de cebadores

Las secuencias genómicas para el diseño de cebadores se obtuvieron de la base de datos Ensembl Genome Browser 37 (ENSEMBL, <http://www.ensembl.org>). Se tuvieron en cuenta algunos criterios para el diseño de cebadores, los cuales se indican a continuación:

- a) Los cebadores deben tener un tamaño de entre 18 a 22 bp, para evitar amplificaciones inespecíficas. Además, es preferible que la diferencia entre ambos cebadores en el número de bases sea reducida.
- b) El amplicón debe ser de 300-700 pb.
- c) Se debe evitar que los cebadores hibriden con regiones con polimorfismos, especialmente si se trata de SNPs (*single nucleotide polymorphisms*) que se encuentran con una elevada frecuencia en la población control.
- d) La temperatura de hibridación de los cebadores debe ser entre 50°C y 60°C, y la diferencia entre ambos cebadores debe ser preferiblemente menor de 2°C.
- e) Deben evitarse los cebadores con secuencias auto-complementarias, para que no se formen estructuras secundarias.

- f) Es preferible que la región 3' de los cebadores sea rica en GCs.
- g) Se deben diseñar cebadores a una distancia aproximada de 100 nucleótidos respecto de la variante de interés, dado que al secuenciar mediante el método Sanger la lectura de los primeros nucleótidos suele ser de peor calidad.

Una vez diseñados los cebadores, se comprobaron sus propiedades fisicoquímicas en la página web de Sigma-Aldrich (<https://www.sigmaaldrich.com/pc/ui/tube-home?product=standard>). Además, es recomendable comprobar la especificidad de cada pareja de cebadores mediante la aplicación informática Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

3.1.5.2. Condiciones de PCR y secuenciación Sanger

Todos los oligonucleótidos utilizados en este estudio fueron resuspendidos en H₂O libre de nucleasas a una concentración de 100 μM. Se prepararon alícuotas de uso a 10 μM. Los cebadores diseñados se recogen en la tabla 2, así como las temperaturas de hibridación y las condiciones de PCR para cada amplicón. El nombre asignado a los cebadores utilizados hace referencia al gen y, dentro de este, al exón al cual los cebadores hibridan, y a si se trata de un cebador directo (d) o reverso (r).

Las reacciones de PCR se llevaron a cabo en un termociclador *Veriti® Thermal Cycler* (Applied Biosystems). Los componentes de la mezcla de reacción se indican en la tabla 3, mientras que el programa de PCR general se describe en la tabla 4.

Tabla 2. Secuencias de cebadores utilizados para amplificar las regiones de interés.

Gen	Primer	Secuencia (5'-3')	Tamaño amplicón (pb)	Tª de hibridación (°C)	Condiciones PCR
OTUD3	Primer directo: OTUD3/E8-d	TGTTACCAACTCTTGCTGGA	465	58	Estándar
	Primer reverso: OTUD3/E8-r	TCTTACAACTCCAAGAGGC			
VPS13B	Primer directo: VPS13B/E48-d	AAGCTTAGAAATAACCAGCA	585	55	Estándar
	Primer reverso: VPS13B/E48-r	GAATCTGACAGATACATGCA			
	Primer directo: VPS13B/E51-d	GTCTTAGAAGATGTGCTTTACG	650	53	+DMSO
	Primer reverso: VPS13B/E51-r	ACTCTGTCTGGGAAACTCTG			

Tabla 3. Componentes de la mezcla de reacción para cada muestra de PCR.

Reactivo	Concentración inicial	Concentración final	Volumen (μL)	Origen
H ₂ O libre de DNAsa y RNAsa	-	-	19,5	VWR Life Science
Tampón Standard 10x con MgCl ₂	10X	1X	2,5	Biotoools
dNTPs	20 mM	0,4 mM	0,5	IBIANLab Technologies
Cebador directo	10 μM	0,2 μM	0,5	Sigma-Aldrich
Cebador reverso	10 μM	0,2 μM	0,5	Sigma-Aldrich
Taq polimerasa	1 U/ μL	0,5 U/ μL	0,5	Biotoools
Dimetilsulfóxido (DMSO)*	14,1 M	0,85 M	1,5	Sigma-Aldrich
DNA**	-	-	1	

*El dimetilsulfóxido fue utilizado únicamente para la pareja de cebadores VPS13B/E51-d y VPS13B/E51-r. En este caso, el volumen de H₂O se ajustó a 18 μL .

**Se utilizó DNA de los individuos NBIA-24, NBIA-271 y NBIA-272 a una concentración de 50-100 ng/ μL .

Tabla 4. Programa general de PCR.

Fases	Tiempo	Temperatura ($^{\circ}\text{C}$)	Número de ciclos
Inicio	5 min	95	1
Desnaturalización	30 s	95	35
Hibridación*	30 s	X	
Extensión	35 s	72	
Elongación final	7 min	72	1
Reposo	4 $^{\circ}\text{C}$	4	1

*La temperatura de hibridación varía según la pareja de cebadores, y se indica en la tabla 4.

Los productos de la PCR se analizaron mediante electroforesis en gel. Para ello, se utilizó el agente intercalante GelRed (Biotium), en geles de agarosa al 1%, preparados con TAE 1% (Tris, acetato, EDTA y H₂O) y agarosa de baja EEO (Scharlab SL). Se utilizó el marcador de peso molecular 1 Kb Plus DNA Ladder (Thermo Fisher Scientific).

La secuenciación de los amplicones fue llevada a cabo por el servicio de Servicio de Genómica y Genética Traslacional (SGGT) del Centro de Investigación Príncipe Felipe, en un secuenciador 3730xl DNA Analyzer (Applied Biosystems). Para ello, se utilizaron los cebadores OTUD3/E8-3, VPS13B/E48-r y VPS13B/E51-r.

3.2. ENSAYO FUNCIONAL DE PLEKHG2

3.2.1. Plásmidos

Los plásmidos y vectores utilizados en este trabajo se indican en la tabla 5.

Tabla 5. Plásmidos y vectores utilizados.

Plásmido	Función	Origen
pGL4.33[<i>luc2P/SRE/Hygro</i>]	Plásmido reportero de luciferasa	Promega Biotech Ibérica
pRL-TK	Plásmido reportero “control Renilla” interno en el ensayo de luciferasa	Promega Biotech Ibérica
pcDNA3.1 [<i>Myc/WT- PLEKHG2</i>]	Sobre-expresión de PLEKHG2 silvestre (WT, <i>wild type</i>)	Obtenido mediante mutagénesis dirigida a partir del plásmido pcDNA3.1 [<i>Myc/I417F- PLEKHG2</i>] cedido por Dr. Songhai Chen
pcDNA3.1 [<i>Myc/T53I- PLEKHG2</i>]	Sobre-expresión de PLEKHG2 con la mutación p.T53I	Obtenido mediante mutagénesis dirigida a partir del plásmido pcDNA3.1 [<i>Myc/WT- PLEKHG2</i>]
pcDNA3.1 [∅]	Control vacío	Invitrogen Life Technologies

Las secuencias de los constructos pcDNA3.1 [*Myc/WT-
PLEKHG2*] y pcDNA3.1 [*Myc/T53I-
PLEKHG2*] fueron verificadas mediante secuenciación Sanger, llevada a cabo por el SGGT del Centro de Investigación Príncipe Felipe, en un secuenciador 3730xl ADN Analyzer (Applied Biosystems). Para ello, se utilizaron los cebadores recogidos en la tabla 6.

Tabla 6. Cebadores utilizados para la secuenciación de pcDNA3.1 [*Myc/WT-
PLEKHG2*] y pcDNA3.1 [*Myc/T53I-
PLEKHG2*].

Cebador	Secuencia (5'-3')
pCMV-d	TTGACGCAAATGGGCGGTAGG
PLEKHG2/E3-d	GAGGTTCTGGGAGCTCCA
PLEKHG2/E3-4-r	AGGTAGTCCTCCACGATGC
PLEKHG2/E7-d	AGCTTCCTGCTGAAACCTGT
PLEKHG2/E11-d	AGAGAAGAGGCTGTGGATTC
PLEKHG2/E18-d	GACAGCGAAACCCTCACAT
PLEKHG2/E18.2-d	CTTCAGGAGGCAAGGCAG
PLEKHG2/E19-d	CTGTCTGGGTTCAAGCTGC
PLEKHG2/E19.2-d	TGTCTTGGCATGGAAGCAG
PLEKHG2/E19.3-d	TTGCCAGGTTGGAGTCTTC

En ambos plásmidos, *PLEKHG2* está fusionado a la etiqueta -myc en la región 3' y se encuentra bajo control del promotor CMV. Además, los dos plásmidos contienen un gen de resistencia a la ampicilina como marcador de selección.

Las secuencias obtenidas fueron alineadas al genoma de referencia GRCh37 (Ensembl, <https://www.ensembl.org/index.html>) mediante el programa BLAST (*Basic Local Alignment Search Tool*, <https://blast.ncbi.nlm.nih.gov>) para verificar la integridad de la secuencia.

3.2.1.1. Mutagénesis dirigida

Como se ha indicado en el apartado anterior, el plásmido pcDNA3.1 [*Myc/T53I-PLEKHG2*] fue obtenido mediante mutagénesis dirigida a partir del plásmido pcDNA3.1 [*Myc/WT-PLEKHG2*]. Para ello, se utilizó el kit *QuickChange II Site-Directed Mutagenesis Kit* (Agilent Technologies), siguiendo el protocolo proporcionado por el fabricante.

Se utilizaron los cebadores indicados en la tabla 7, los cuales fueron diseñados siguiendo las indicaciones del fabricante. Durante la verificación de la integridad del plásmido se encontró el cambio p. I417F (c.1252A>T), por lo que se realizaron dos mutagénesis dirigidas. Por una parte, se eliminó el cambio p. I417F (c.1252A>T) y, por otra parte, se introdujo el cambio p.T53I (c.158T>P), con la finalidad de obtener los plásmidos pcDNA3.1 [*Myc/WT-PLEKHG2*] y pcDNA3.1 [*Myc/T53I-PLEKHG2*].

Tabla 7. Cebadores utilizados para la mutagénesis dirigida de los plásmidos pcDNA3.1 [*Myc/I417F-PLEKHG2*] y pcDNA3.1 [*Myc/WT-PLEKHG2*]

Cebador	Secuencia (5'-3')	Tm (°C)
PLEKHG2(p.T53I)/d	TTCTGGGAGCTCCA t ATCCCTGAGCACAGTGG	80,2
PLEKHG2(p.T53I)/r	ACTGTGCTCAGGGAT a TGGAGCTCCCAGAACC	78,8
PLEKHG2(p.I417F)/d	GAGAACCACCCTGCCTCCA t CCCTGCCAAGGCAAAGC	87,4
PLEKHG2(p.I417F)/r	GCTTTGCCTTGGCAGGGAT t GGAGGCAGGGTGGTTCTC	87,4

En ambos casos se llevó a cabo una reacción de PCR para sintetizar la hebra mutante. Los cebadores y plásmidos utilizados se recogen en la tabla 8, mientras que la concentración de los reactivos se indica en la tabla 9. En la tabla 10 se detalla el programa de PCR.

Tabla 8. Cebadores utilizados para cada mutación dirigida.

Plásmido mutado	Plásmido obtenido	Cebador directo	Cebador reverso
pcDNA3.1 [<i>Myc/I417F-PLEKHG2</i>]	pcDNA3.1 [<i>Myc/WT - PLEKHG2</i>]	PLEKHG2(p.I417F)/d	PLEKHG2(p.I417F)/r
pcDNA3.1 [<i>Myc/WT - PLEKHG2</i>]	pcDNA3.1 [<i>Myc/T53I - PLEKHG2</i>]	PLEKHG2(p.T53I)/d	PLEKHG2(p.T53I)/r

Tabla 9. Componentes de la mezcla de PCR.

Reactivos	Concentración inicial	Concentración final	Volumen (μL)	Origen
H ₂ O libre de DNAsa y RNAsa	-	-	33,5	VWR Life Science
dNTPs	20 mM	0,4 mM	1	IBIANLab Technologies
Tampón Pfu Turbo	10X	1X	5	Agilent Technologies
Pfu Turbo	2.5 U/μL	0,25 U/μL	1	Agilent Technologies
Dimetilsulfóxido	14,1 M	1,41 M	5	Sigma-Aldrich
MgSO ₄	25mM	0,5 mM	1	Biotoools
Primer directo	10 mM	0,25 mM	1,25	Sigma-Aldrich
Primer reverso	10 mM	0,25 mM	1,25	Sigma-Aldrich
Plásmido de interés	20 ng/μL	0,4 ng/ μL	1	Indicado en tabla 5

Tabla 10. Programa de PCR para la mutagénesis dirigida.

Fases	Tiempo	Temperatura (°C)	Número de ciclos
Inicio	30 min	95	1
Desnaturalización	30 s	95	17
Hibridación	1 min	55	
Extensión	18 min 30 s	72	
Reposo	∞	4	1

El producto obtenido tras la PCR fue digerido añadiendo 1 μL de la enzima de restricción Dpn I (Thermo Fisher Scientific) e incubando a 37°C durante 1 h. El DNA digerido se utilizó para transformar células competentes, del modo indicado en el siguiente apartado.

3.2.1.2. Transformación de células competentes

Con la finalidad de amplificar los constructos plasmídicos descritos en el apartado 3.2.1, y como parte del proceso de la mutagénesis dirigida, se llevó a cabo la transformación de células competentes *Escherichia Coli* (*E. Coli*) *Subcloning Efficiency* DH5α (Invitrogen) con los distintos plásmidos.

Para ello se descongelaron en hielo alícuotas de 50 μL de células *E. Coli* DH5α competentes (almacenadas a 80°C). Una vez descongeladas, se pipetearon 3 μL del DNA plasmídico de interés en las bacterias competentes, tras lo que se removió el tubo cuidadosamente. Después de 30 minutos de incubación en hielo, se sometió a las células a un choque térmico a 42°C durante 25 segundos, seguido de una incubación en hielo durante 2 minutos. Posteriormente, se añadieron 250 μL de medio líquido Luria Bertani (LB) (Scharlau), pre-calentado a 37°C, a la mezcla de transformación, y se incubaron las células durante 1 h a 37°C y en agitación (200 rpm) para permitir la recuperación de las células del choque térmico. A continuación, se centrifugaron a 6000 rpm durante 60 s y se retiró parte del sobrenadante para concentrar las solución de medio con células. Finalmente, se sembraron las células en placas de agar LB

(Scharlau), pre-calentadas a 37°C y suplementadas con ampicilina a 100 µg/mL. La placas se incubaron durante la noche a 37°C y al día siguiente fueron guardadas en la nevera.

3.2.1.3. Purificación de los plásmidos a pequeña y mediana escala

El DNA plasmídico fue aislado a partir de 5 ml de cultivo crecido durante la noche. Para ello, se inocularon colonias aisladas en 5 ml de medio líquido LB con ampicilina a 100 µg/mL y se dejaron crecer durante la noche a 37°C y en agitación (200 rpm).

La extracción de ácidos nucleicos fue llevada a cabo mediante dos procedimientos. Por una parte, se realizaron *minipreps* (extracciones a pequeña escala) mediante el kit *High Pure Plasmid Isolation* (Roche) y, por otra parte, se realizaron *midipreps* (extracciones a mediana escala) mediante el kit *PureLink™ HiPure Plasmid Midiprep* (Thermo Fisher Scientific), en ambos casos siguiendo las instrucciones del fabricante. Se cuantificó la concentración de ácidos nucleicos mediante espectroscopia, utilizando el espectrofotómetro NanoDrop® 1000 (Thermo Fisher Scientific).

3.2.1.4. Glicerinado de E.Coli

Para el almacenamiento de un stock de las células transformadas con los plásmidos de interés en glicerol, se mezclaron 750 µL de un cultivo de *E. Coli* DH5α transformado con el plásmido de interés con 250 µL de glicerol (VWR chemicals) en un tubo de 1,5 mL, y se almacenaron a -80°C. Los cultivos se obtuvieron de la manera descrita en el apartado anterior.

3.2.2. Cultivo de células eucariotas

3.2.2.1. Descripción de la línea celular y condiciones de mantenimiento

Se utilizó la línea celular HeLa, línea inmortalizada derivada de una muestra de cáncer cervicouterino (ECAC nº 93921913). Este tipo celular es adherente, por lo que las células se cultivaron en monocapa, en frascos de poliestireno (BioLite Cell Culture Products, Thermo Scientific)

Las células fueron cultivadas en *Dulbecco's modified Eagle's Medium* (DMEM) – *high glucose* suplementado con suero fetal bovino (SFB) (10%), penicilina-estreptomicina al 1% (100 unidades de penicilina y 0,1 mg de estreptomicina/mL) y solución de L-Glutamina al 1% (2mM). Todos los compuestos mencionados fueron obtenidos de Sigma-Aldrich. Las células se mantuvieron en incubador a 37°C y al 5% de CO₂.

Se realizaron pases de mantenimiento regularmente al alcanzar la confluencia del cultivo. Para separar las células del sustrato se empleó la solución tripsina-EDTA Gibco® (Thermo Fisher Scientific). Para lavar las células se utilizó tampón fosfato salino (PBS, *phosphate buffered saline*) (Sigma-Aldrich).

3.2.2.2. PCR para la detección de micoplasma

Se realizó una PCR para detectar posibles contaminaciones por micoplasma. Para ello, se utilizaron los cebadores indicados en la tabla 11. El programa de PCR se encuentra en la tabla 12, mientras que los componentes de la mezcla de reacción se encuentran en la tabla 13.

Tabla 11. Cebadores utilizados para la PCR para detectar micoplasma.

Primer	Secuencia (5'-3')	Tamaño amplicón (pb)	T _m (°C)
Myco1-f	GGCGAATGGGTGAGTAACAGG	500	58
Myco2-r	CGGATAACGCTTGC GACTATG		

Tabla 12. Programa de PCR para detectar micoplasma.

Fases	Tiempo	Temperatura (°C)	Número de ciclos
Inicio	5 min	95	1
Desnaturalización	30 s	95	35
Hibridación	30 s	X	
Extensión	45 s	72	
Elongación final	7 min	72	1
Reposo	4 °C	4	1

Tabla 13. Componentes de la mezcla de reacción para cada muestra de la PCR para detectar micoplasma.

Reactivo	Concentración inicial	Concentración final	Volumen (µL)
H₂O libre de DNAsa y RNAsa	-	-	19
Tampón Standard 10x con MgCl₂	10X	1X	2,5
dNTPs	20 mM	0,4 mM	0,5
Myco1-f	10 µM	0,2 µM	0,5
Myco2-r	10 µM	0,2 µM	0,5
Taq polimerasa	1 U/µL	0,5 U/µL	0,5
Muestra*	-	-	1,5

* La se obtuvo a partir del medio en el que se habían cultivado las células durante 3 días. Éste fue hervido a 100°C durante 5 minutos antes de ser utilizado para la PCR.

3.2.2.3. Descongelación de la línea celular

Se descongelaron las células HeLa en baño de 37°C. Una vez descongeladas, la suspensión celular fue transferida a un tubo Falcon de 15 mL y se centrifugó durante 5 min a 1200 rpm. Tras esto, se desechó el sobrenadante y se resuspendieron las células en 500 µL de medio. Éstas se dejaron crecer durante la noche en una placa p100 (diámetro de 100mm) y, al día siguiente, fueron transferidas a un frasco t75 (superficie de crecimiento de 75 cm²). A partir de este momento, se llevó a cabo el mantenimiento regular del cultivo celular, explicado en el apartado 3.2.2.1.

3.2.3. Transfección

Para lograr la transfección transitoria de las células, se utilizó el reactivo Fugene HD (Promega), siguiendo el protocolo del fabricante. Se utilizó un ratio 3:1 de reactivo Fugene HD (µL):DNA(µG).

Para los ensayos de luciferasa, la cantidad de DNA transfectado por pocillo fue de 1 µg del plásmido pcDNA3.1 [*Myc/WT-PLEKHG2*] o pcDNA3.1 [*Myc/T53I-PLEKHG2*], 200 ng de pGL4.33[*luc2P/SRE/Hygro*] y 100 ng del plásmido control pRL-TK. Para la inmunofluorescencia, se transfectó con 1 µg del plásmido pcDNA3.1 [*Myc/WT-PLEKHG2*] o pcDNA3.1 [*Myc/T53I-PLEKHG2*].

Las células fueron sembradas el día anterior de la transfección en placas de 6 pocillos (1,6 cm² por pocillo). Al alcanzar una confluencia del 60-70% aproximadamente, fueron transfectadas con las cantidades de DNA indicadas previamente.

Seis horas post-transfección, el medio fue cambiado a medio DMEM - *High glucose* suplementado con penicilina-estreptomicina (1%) y solución de L-Glutamina (1%), sin SFB, con o sin LPA. Para ello, se retiró el medio, se lavó con PBS dos veces, y se añadió el medio sin SFB fresco. Además, para aumentar la inducción de la actividad de PLEKHG2 mediante LPA, la mitad de las células transfectadas fueron estimuladas con LPA 10 µM en el medio sin SFB. El LPA fue obtenido resuspendiendo sal de sodio del ácido lisofosfatídico (Sigma-Aldrich) en PBS con albúmina de suero bovino (Sigma-Aldrich)(1mg/ml).

3.2.4. Medida de la actividad de PLEKHG2 mediante ensayo de luciferasa

Para medir la activación de PLEKHG2 WT y mutante, realizamos un ensayo de luciferasa. En este ensayo, la activación de las RhoGTPasas, y consecuentemente de las RhoGEFs, induce la transcripción de luciferasa en la secuencia posterior al promotor SRE. De esta forma, se puede usar la luminiscencia como medida de la actividad de PLEKHG2.

Para medir la actividad luciferasa en células HeLa, se utilizó el kit *Dual Luciferase Reporter Assay System* (Promega Co.). 24 horas post-transfección, las células HeLa fueron lavadas dos veces con PBS. Tras esto, fueron despegadas de las placas por rascado mediante raspador de células (Sarstedt) en presencia de 100 µL de *Passive Lysis Buffer* (PLB) 1X del kit mencionado. Los lisados fueron centrifugados para remover los residuos celulares (13,200 rpm durante 30 s).

La actividad de la luciferasa de los extractos se midió siguiendo el protocolo establecido por el fabricante del kit mencionado, en un lector de microplacas Perkin Elmer Wallac 1420 Victor 2TM. La actividad del gen reportero pGL4.33[*luc2P/SRE/Hygro*] se normalizó contra la actividad del vector control pRL-TK.

3.2.5. SDS-PAGE y Western blot

Se llevó a cabo un SDS-PAGE y un *Western Blot* para detectar la expresión de la proteína PLEKHG2 sobre-expresada. Para ello, se cuantificó la concentración de proteína total de los lisados obtenidos en el apartado anterior mediante el kit PierceTM *BCA Protein Assay* (Thermo Fisher Scientific), para así analizar cantidades equivalentes de proteína. En concreto, las muestras se prepararon de forma que contuvieran 25 µg de proteína. Las medidas de cantidad de proteína se obtuvieron en un lector de microplacas Perkin Elmer Wallac 1420 Victor 2TM.

Las muestras a analizar se separaron mediante SDS-PAGE en un gel de poliacrilamida al 8%. Se utilizó el marcador *Precision Plus ProteinTM All Blue Prestained Protein standards* (BioRad). Posteriormente, las proteínas fueron transferidas a membranas de difluoruro de polivinilideno (Thermo Fisher Scientific) durante 2 horas a 100 V, utilizando un aparato de transferencia *Mini Trans-Blot[®] Electroforetic transfer cell* (BioRad), de acuerdo al protocolo del fabricante. Tras esto, las membranas fueron lavadas con agua destilada y teñidas con Ponceau S. (Sigma

Aldrich) para verificar que la transferencia se hubiera llevado a cabo correctamente. A continuación, las membranas fueron lavadas con TBST (Tris·HCl 50mM, pH 7.5, NaCl 150 mM, Tween-20 0.1%) hasta eliminar la tinción, y luego bloqueadas con 5% de leche en polvo disuelta en TBST durante 1 hora. Posteriormente, fueron incubadas con el anticuerpo primario diluido en la misma solución de bloqueo, durante toda la noche a 4°C. Se utilizaron los anticuerpos primarios Myc-Tag (9B11) *Mouse* (Cell Signalling Technology) (dilución 1:1000) y anticuerpo anti-actina producido en ratón (Sigma-Aldrich) (dilución 1:2000). Este último fue utilizado como control de carga.

Tras la incubación con el anticuerpo primario, las membranas fueron lavadas 3 veces durante 10 minutos con TBST, incubadas con el anticuerpo secundario *Goat anti-Mouse IgG (H+L)* conjugado a peroxidasa de rábano (1 hora y media a temperatura ambiente, 1:10,000 en la misma solución de bloqueo) y finalmente lavadas con TBST (3 lavados 10 minutos)

Finalmente, las proteínas fueron detectadas por quimioluminiscencia usando el kit *Amersham ECL Plus Western Blotting Detection Reagents* (GE Healthcare Life Sciences). Los *blots* fueron expuestos sobre películas de rayos X y se generaron las imágenes mediante una máquina Kodak. Se cuantificaron las bandas mediante el software ImageJ.

3.2.6. Inmunofluorescencia

Para llevar a cabo el marcaje fluorescente de PLEKHG2, las células HeLa crecidas sobre cubreobjetos fueron fijadas con paraformaldehído (PFA) al 4% en PBS durante 10 minutos. Tras esto, se lavaron las células tres veces con PBS y se llevó a cabo la permeabilización celular con Tritón 0,5%/PBS durante 15 minutos. A continuación, se realizaron 3 lavados de 5 minutos con PBS, y se incubaron las células con suero equino al 5%/PBS durante 1 hora para bloquear. Tras esto, se añadió el anticuerpo primario Myc-Tag (9B11) *Mouse* (Cell Signalling Technology) a la dilución 1:8000 y se dejó durante toda la noche a 4°C. Posteriormente, se realizaron 3 lavados de 5 minutos con PBS y las células fueron incubadas con el anticuerpo secundario marcado con Alexa Fluor™ 488 *goat anti-mouse* (Invitrogen), a la dilución 1:250, durante 1 h y media. Tras esto, se volvieron a realizar 3 lavados de 5 minutos con PBS. Todos los pasos, excepto la incubación con el anticuerpo primario, se llevaron a cabo a temperatura ambiente. Los cubreobjetos fueron colocados sobre los portaobjetos con medio de montaje, y se almacenaron a 4°C en la oscuridad. Las imágenes de fluorescencia fueron adquiridas con un microscopio confocal espectral con escaneado láser Leica TCS SP2, utilizando un objetivo 63X para inmersión.

4. RESULTADOS

4.1. SECUENCIACIÓN DE EXOMA fNBIA-20

Tras obtener el fichero de anotación de variantes procedente de la secuenciación del exoma del trío (progenitores y descendiente afectado) de la familia fNBIA-20, se procedió al filtrado y priorización de variantes. Tratándose de un caso esporádico y, siendo el paciente un varón, se plantearon las siguientes hipótesis: cambios *de novo*; herencia autosómica recesiva; herencia recesiva ligada al cromosoma X. Tras el filtrado de variantes, se seleccionaron 5 variantes noveles de interés que segregaban con la enfermedad en 4 genes (Tabla 14).

Tabla 14. Variantes noveles tras el proceso de priorización y modo de herencia asociada.

	NBIA-272 (Padre)	NBIA-271 (Madre)	NBIA-24 (Probando)	Herencia asociada
OTUD3 c.1121G>A (p.R374K)	-/-	-/-	+/-	<i>de novo</i>
VPS13B c.8788A>G (p.N2930D)	-/-	+/-	+/-	AR
VPS13B c.9405G>T	-/-	-/-	+/-	<i>de novo</i>
DMD c.6029G>C (p. A2010G)	-/-	+/-	+/+	XR
P2RY4 c.778G>T (p. P260T)	-/-	+/-	+/+	XR

En las tablas 15-17, se compila información sobre el tipo de cambio, patologías asociadas al gen (base de datos OMIM) y predictores *in silico* de la patogenicidad de la variante: PolyPhen-2 y SIFT para cambios de aminoácidos, y BGD, HSF y NetGene2 para los cambios que pueden afectar a la región de *splicing*. La nomenclatura de las variantes se reporta siguiendo las guías de HGVS (*Human Genome Variation Society*). Además, se indica la frecuencia alélica en población control, consultada en la base de datos gnomAD.

Entre las variantes noveles *de novo*, destaca el cambio c.1121G>A (p.R374K) en *OTUD3* (*OTU Deubiquitinase 3*) (Tabla 15), gen que codifica para una enzima deubiquitinasa. Según el predictor de *splicing* HSF, la mutación podría crear un nuevo sitio aceptor, lo que alteraría potencialmente el procesado del mRNA. A nivel de proteína, según los predictores SIFT y PolyPhen-2, el cambio *missense* tendría un efecto benigno y tolerado, respectivamente. El residuo al cual afecta la mutación (p.Arg374) se encuentra en una región (C-terminal) donde no parece haber un dominio funcional importante. *OTUD3* se expresa en distintas células del sistema nervioso central (SNC), con una mayor expresión en neuronas y astrocitos fetales y maduros (Zhang *et al.*, 2016). También se ha detectado expresión en distintos estadios de desarrollo en regiones del SNC tales como cerebelo, el córtex del cíngulo anterior o córtex prefrontal (Bastian *et al.*, 2008).

Por otra parte, se identificaron dos variantes en heterozigosis compuesta en el gen *VPS13B* (*vacuolar protein sorting 13 homolog B*) (Tabla 16), también conocido como *COH1* por su

asociación con el síndrome de Cohen (MIM 607817). La variante c.8788A>G (p.N2930D) genera un cambio tolerado y benigno, según los predictores *in silico* de patogenicidad SIFT y Polyphen, respectivamente, mientras que, a nivel del mRNA, la variante c.9405G>T podría generar un nuevo sitio aceptor de *splicing*. Se ha descrito alta expresión de la proteína en las neuronas del SNC (Mochida *et al.*, 2004), así como en glía y tejido endotelial del SNC (Zhang *et al.*, 2016). No obstante, al realizar la secuenciación mediante Sanger, corroboramos que el cambio c.9405G>T resultó ser un falso positivo.

Finalmente, para la hipótesis de una herencia recesiva ligada al cromosoma X, se detectaron dos variantes de interés en los genes *DMD* (*dystrophin*) y *P2RY4* (*pyrimidinergic receptor P2Y4*) (Tabla 17). En cuanto al cambio en *DMD*: c.6029G>C (p. A2010G), gen que codifica para distrofina, las herramientas de predicción SIFT y PolyPhen-2 lo señalan como tolerado y benigno, respectivamente. Mutaciones en este gen se asocian a distrofia muscular de Becher y distrofia muscular de Duchenne, ambas con herencia recesiva ligada al cromosoma X. Para *DMD*, se ha descrito expresión en células del SNC, siendo mayor en astrocitos (Zhang *et al.*, 2016). Además, se ha descrito expresión en cerebro en distintos estadios de desarrollo humano (Bastian *et al.*, 2008).

Por último, cabe destacar el cambio en el gen *P2RY4*: c. 778G>T (p. P260T). Este gen codifica para un receptor pirimidinérgico que, hasta la fecha, no ha sido asociado con fenotipos clínicos. Los predictores SIFT y PolyPhen estiman que la variante podría tener un efecto deletéreo y probablemente perjudicial, respectivamente. No obstante, no se ha descrito expresión del gen en tejido nervioso.

La calidad de secuenciación de las variantes seleccionadas se comprobó mediante el empleo del software IGV. Ésta fue buena para todos los cambios, excepto para el cambio c.9405G>T en *VPS13B*. Las variantes en *OTUD3* y *VPS13B* fueron secuenciadas mediante Sanger con la finalidad de validarlas (Anexo I). El cambio c.9405G>T en *VPS13B* resultó ser un falso positivo.

Tabla 15. Variantes novedosas de novo.

Gen	HGVS	Frecuencia	Tipo de cambio	Predictores			OMIM	
OTUD3*	NM_015207 c. 1121G>A p.R374K	Novel	Cambio de aminoácido Región de splice	BDGP		HSF	NetGene2	-
				Nuevo sitio aceptor Nuevo sitio silenciador de <i>splicing</i> exónico Alteración potencial del <i>splicing</i>			-	
				SIFT		PolyPhen-2		
				Tolerado		Benigno		

* En el caso de *OTUD3*, se estudió tanto el efecto del cambio de aminoácido sobre la proteína, como el efecto sobre el *splicing*, ya que la mutación podría afectar también al *splicing* por su ubicación.

Tabla 16. Variantes bialélicas con modo de herencia AR.

Gen	HGVS	Frecuencia	Tipo de cambio	Predictores			OMIM
VPS13B/COH1	NM_017890.4 c.8788A>G p.N2930D	Novel	Cambio de aminoácido	SIFT		PolyPhen-2	Síndrome de Cohen
				Tolerado		Benigno	
	NM_017890.4 c.9405G>T	Novel	Región de <i>splice</i>	BDGP	HSF	NetGene2	
				Nuevo sitio aceptor de <i>splicing</i>	Alteración de sitio silenciador de <i>splicing</i> exónico Probablemente no altera el <i>splicing</i>	Nuevo sitio aceptor de <i>splicing</i>	

Tabla 17. Variantes novedales con herencia recesiva ligada al cromosoma X.

Gen	HGVS	Frecuencia	Tipo de cambio	Predictores		OMIM
DMD	NM_004006.2 c.6029G>C (p. A2010G)	Novel	Cambio de aminoácido	SIFT	PolyPhen-2	Distrofia muscular de Becher, cardiomiopatía, distrofia muscular de Duchenne
				Tolerado	Benigno	
P2RY4	NM_002565.3 c. 778G>T (p. P260T)	Novel	Cambio de aminoácido	SIFT	PolyPhen-2	-
				Deletéreo	Probablemente perjudicial	

4.2. ENSAYO FUNCIONAL DE PLEKHG2

4.2.1. Ensayo de luciferasa

Para determinar si la mutación p.T53I podría afectar la función de PLEKHG2, se realizó un ensayo de luciferasa para comparar los niveles de activación del plásmido reportero de luciferasa, entre PLEKHG2 mutante (p.T53I) y WT. El resultado de un único ensayo, realizado por triplicado, muestra que p.T53I presenta mayores niveles de actividad luciferasa frente al WT (Fig. 7). La inducción a nivel celular por parte de LPA, revela una mayor actividad de PLEKHG2 comparado con la condición basal (sin LPA), y además muestra que la diferencia observada entre p.T53I y el WT es mayor en presencia de LPA (Fig. 7).

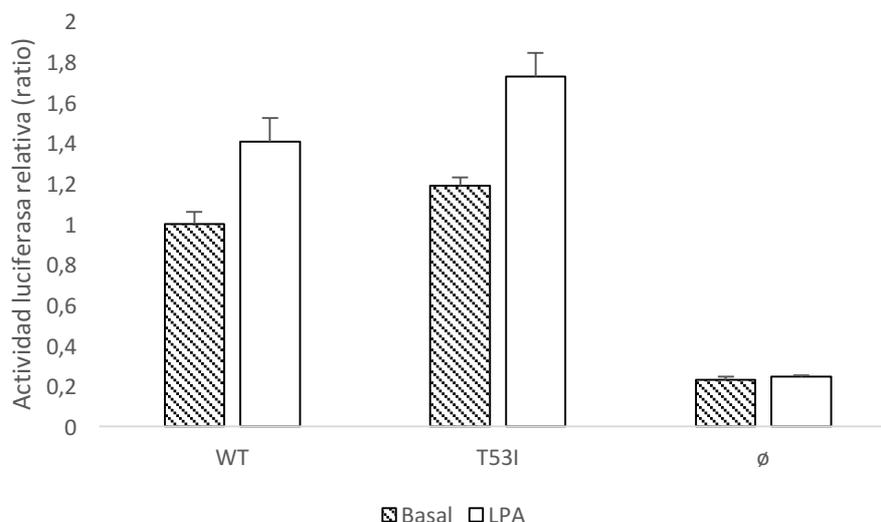


Figura 7. Actividad basal y estimulada con LPA de WT-*PLEKHG2*, T53I-*PLEKHG2* y control vacío (\emptyset), en células HeLa. Las barras de error muestran la desviación estándar del triplicado de un único experimento. Los valores han sido normalizados respecto a la actividad basal de *PLEKHG2* WT (con valor de ratio=1).

Para determinar los niveles de sobre-expresión de la proteína mutante y WT, así como la integridad de las mismas, se realizó un experimento de *Western Blot* a partir de los lisados celulares tratados con LPA empleados en el ensayo de luciferasa. Con dicho ensayo, se detectó una banda de aproximadamente 150kDa correspondiente a la proteína *PLEKHG2*, tanto cuando se sobre-expresó la forma WT -pcDNA3.1 [*Myc*/*WT-PLEKHG2*]-, como la mutante -pcDNA3.1 [*Myc*/*T53I-PLEKHG2*]-. En el caso del control -el vector vacío pcDNA3.1 [\emptyset]-, se confirmó la ausencia de dicha banda (Fig. 8). La normalización de los valores de *PLEKHG2* con la proteína actina mostró niveles de expresión similares entre la forma WT y mutante p.T53I (Fig. 9), aunque ligeramente inferiores en el caso de la proteína WT en algunas réplicas.

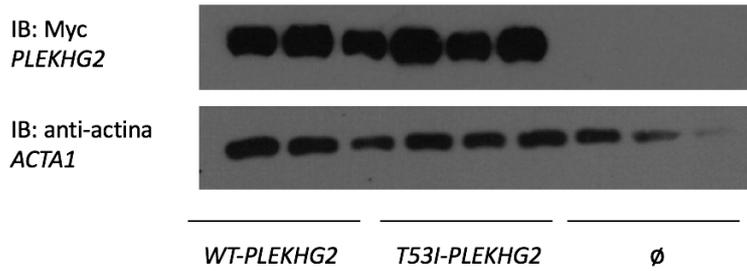


Figura 8. Western Blot de PLEKHG2 y anti-actina (control de carga). Las células HeLa fueron transfectadas, independientemente, con los plásmidos pcDNA3.1 [Myc/WT-PLEKHG2], pcDNA3.1 [Myc/T53I-PLEKHG2] y pcDNA3.1 [ø]. 24 h post-transfección, las células fueron lisadas y el extracto proteico procesado mediante SDS-PAGE y Western Blot.

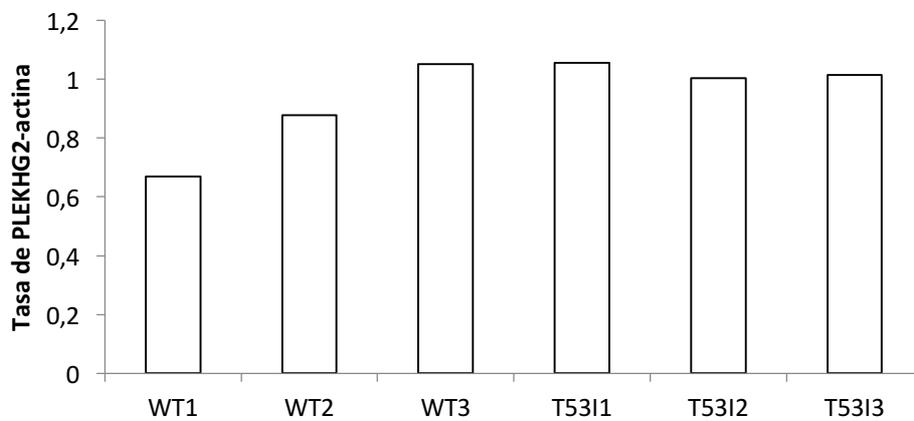


Figura 9. Tasa de expresión de PLEKHG2 normalizada frente a la proteína actina, según Western blot.

Debido a que la actividad luciferasa es dependiente de la cantidad de PLEKHG2 presente a nivel intracelular, normalizamos los valores de la actividad luciferasa con los niveles de expresión de PLEKHG2 WT y p.T53I. Con dicho análisis, la diferencia observada entre p.T53I y PLEKHG2 WT disminuye, llegando a ser muy similar entre ambos constructos (Fig. 10).

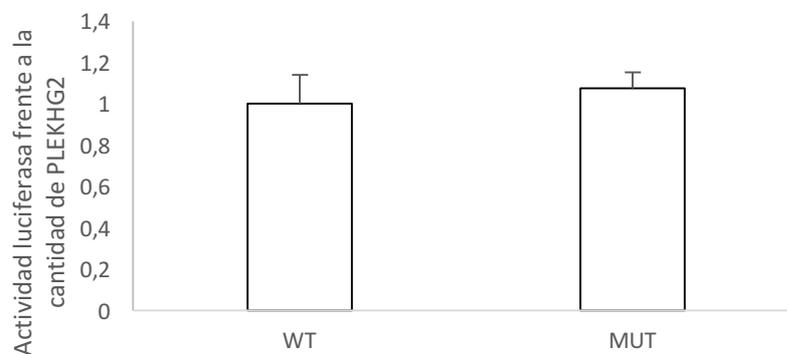


Figura 10. Actividad luciferasa frente a la cantidad de PLEKHG2. Ésta se calculó dividiendo el ratio de actividad luciferasa/renilla con el ratio de expresión de PLEKHG2/actina.

4.2.2. Estudio de localización

Para investigar posibles cambios en la localización celular de la proteína PLEKHG2 mutada, realizamos un ensayo de inmunofluorescencia en células HeLa que sobre-expresaban la forma WT o mutada p.T53I. Como se muestra en las figura 11, el patrón de distribución de PLEKHG2 en membrana plasmática y en citoplasma en la línea celular HeLa aparece tanto al sobre-expresar la proteína WT como la mutante. Este resultado sugiere que la mutación p.T53I no altera la localización de PLEKHG2.

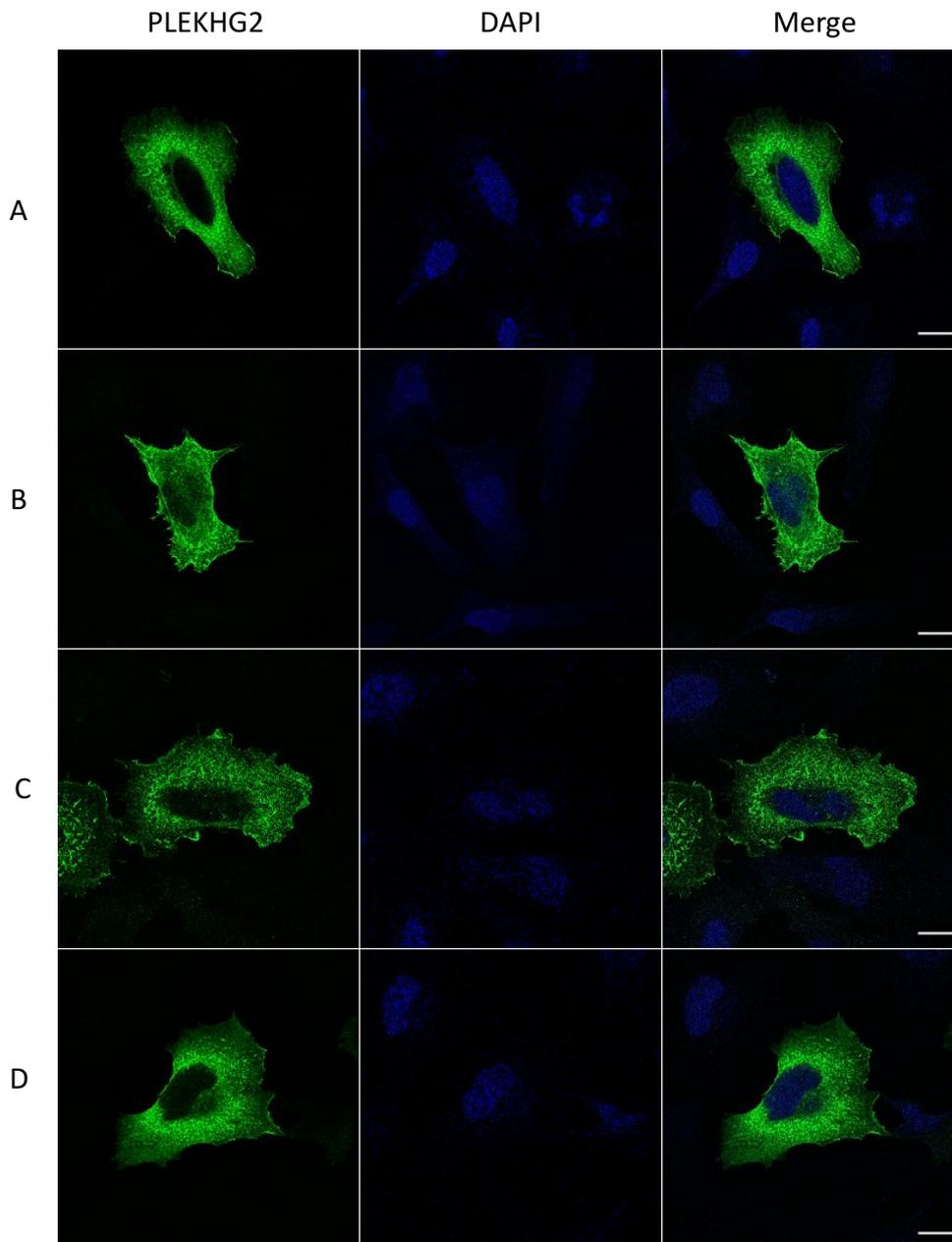


Figura 11. Inmunofluorescencia de PLEKHG2-WT (A,B) y PLEKHG2-T53I (C,D) en células HeLa transfectadas. Se muestra PLEKHG2-WT, tinción nuclear (DAPI) y la superposición de ambas imágenes (merge). Barra de escala: 20 μ m.

5. DISCUSIÓN

5.1. SECUENCIACIÓN DE EXOMA fNBIA-20

En el presente trabajo se ha realizado la secuenciación del exoma de un paciente y sus progenitores. El fenotipo del probando se corresponde con ataxia cerebelosa congénita no evolutiva, asociada a discapacidad intelectual, signos de espasticidad en miembros inferiores y estrabismo, junto con atrofia de cerebelo en neuroimagen en las pruebas complementarias. Tras el filtrado de variantes genéticas, seleccionamos 5 cambios noveles en 4 genes candidatos, entre los que podría resultar complicado determinar la variante causante del fenotipo. Los estudios de validación mediante secuenciación Sanger nos ayudaron a descartar el gen *VPS13B* como candidato a ser responsable de la enfermedad en el paciente de la fNBIA-20. Una de las dos variantes identificadas en este gen, la c.9405G>T, resultó ser un falso positivo. Además, el síndrome de Cohen –asociado a mutaciones en *VPS13B*– se relaciona con un fenotipo característico de microcefalia, ausente en el paciente bajo estudio.

La mutación c.6029G>C (p. A2010G) en *DMD* fue también descartada como candidata, ya que las casi 4000 mutaciones patológicas descritas en este gen se han asociado a distrofia muscular de Duchenne (MIM 310200), distrofia muscular de Becker (MIM 300376) y cardiomiopatía (MIM 302045), fenotipos no presentados en el paciente.

Por otro lado, la variante en el gen *P2RY4* (c. 778G>T, p. P260T) resulta interesante. Pese a que no se ha descrito expresión de *P2RY4* en el SNC humano, Harata *et al.* (2019) mostraron que el gen mencionado se expresa en la región del organizador de cabeza y en el neuroectodermo prospectivo en el estadio temprano de gástrula en *Xenopus laevis*, y que tiene alta expresión en los componentes de la cabeza. Además, la alteración de la función de *P2RY4* da como resultado el fenotipo de cabeza pequeña y una expresión reducida de los genes marcadores específicos para el neuroectodermo y el borde neural en una etapa temprana de la neurula. Por lo tanto, el gen *P2RY4* podría también estar relacionado con el desarrollo del SNC en humanos.

Finalmente, la variante en el gen *OTUD3* (c.1121G>A; p.Arg374Lys) es la única candidata *de novo* identificada en la secuenciación del trío de la fNBIA-20. La proteína OTUD3 ha sido descrita como de-ubiquitinasa de PTEN (Yuan *et al.*, 2015), ya que interactúa, de-poliubiquitina y estabiliza PTEN. A su vez, *PTEN* se ha relacionado con la patogénesis de trastornos neurodegenerativos como la enfermedad de Alzheimer, la enfermedad de Parkinson y la esclerosis lateral amiotrófica (Wong and Cuervo, 2010). Se ha detectado una elevada frecuencia de mutaciones asociadas con glioblastoma multiforme y trastornos neurológicos en pacientes con mutaciones en la línea germinal de *PTEN* (Liaw *et al.*, 1997; Nelen *et al.*, 1997), lo que señala su importante papel en el SNC. Chang *et al.* (2007) demostraron que *PTEN* juega un rol crítico en la regulación del desarrollo neuronal, el crecimiento y la regeneración axonal, la plasticidad sináptica y la supervivencia y diferenciación neuronal. Además, existen fenotipos clínicos con *PTEN* asociados a lesión cerebelosa con herencia autosómica dominante: el síndrome de Cowden (MIM 158350), generalmente asociado con el síndrome de Lhermitte-Duclos (MIM 158350).

Los genes NBIA conocidos permiten el diagnóstico del 80% de casos aproximadamente (Gregory y Hayflick, 2013). La idea de que una enfermedad es consecuencia de mutaciones en un único gen está cambiando, y sabemos que las patologías son resultado de complejas interacciones y alteraciones que afectan a un gran número de genes y, consecuentemente, a las proteínas para las que éstos codifican (Tello *et al.*, 2017). Por ello, la búsqueda de nuevas

rutas biológicas que conecten grupos de genes NBIA y otros genes asociados a enfermedades neurológicas puede proporcionar información sobre los mecanismos de las NBIA, que ayuden en la comprensión de este grupo de enfermedades y en el desarrollo de nuevas estrategias terapéuticas (Bettencourt *et al.*, 2016). En ese sentido, la interacción mencionada entre *P TEN* y *OTUD3* refuerza la hipótesis de que la variante en este último gen sea un buena candidata como responsable del desarrollo de la enfermedad en el probando de la familia fNBIA-20.

Como se ha explicado con anterioridad, para determinar la patogenicidad de una variante son necesarios análisis genéticos, *in silico* y funcionales. Los análisis genéticos en muestras de DNA de familiares resultan cruciales, pues permiten confirmar o descartar aquellas variantes que segregan o no con la enfermedad. Se estima que la tasa de éxito en la identificación de genes mediante NGS en un entorno clínico es del 25% (Yang *et al.*, 2013) y puede aumentar en gran medida en enfermedades de segregación mendeliana al analizar más miembros de la familia (Rehman *et al.*, 2015). Por lo tanto, como próximo abordaje experimental sería conveniente extraer muestras de más familiares para analizar la segregación de los cambios detectados en *OTUD3* y *P2RY4*.

5.2. ENSAYO FUNCIONAL DE PLEKHG2

En el presente trabajo se han realizado ensayos funcionales para investigar el posible efecto patogénico de la mutación p.T53I sobre la proteína PLEKHG2, una RhoGEF de 1387 aminoácidos, miembro de la familia Dbl (Zheng, 2001). Dicha mutación había sido identificada en homocigosis en un paciente de nuestra serie clínica con atrofia de cerebelo.

Otras proteínas RhoGEFs han sido implicadas en la patogénesis de la neurodegeneración. Por ejemplo, RGNEF/p190RhoGEF ha sido relacionada con esclerosis lateral amiotrófica (ELA) (Droppelman *et al.*, 2014). Gracias a la comprensión cada vez mayor del papel clave de los GEF en la modulación de la función de las GTPasas en la supervivencia celular, se ha sugerido que los GEFs tienen un papel crítico como mediadores del equilibrio entre la citotoxicidad y la neuroprotección que, cuando se altera, contribuye a la pérdida neuronal (Droppelman *et al.*, 2014).

Edvardson *et al.* (2016) llevaron a cabo un ensayo de luciferasa con la proteína PLEKHG2 mutante p.R204W y mostraron que ésta tiene menor actividad catalítica que la proteína WT. Por otro lado, Runne y Chen (2013) demostraron que PLEKHG2 es capaz de autoinhibirse mediante la unión de su extremo C-terminal a la región N-terminal que contiene los dominios DH-PH. Dada la importancia de las RhoGTPasas y, por lo tanto, de las RhoGEFs en multitud de procesos celulares, es plausible que tanto la hiperactivación como una actividad deficiente de PLEKHG2 sea causa –al menos, en parte- de neurodegeneración. Debido a que la mutación identificada en nuestro laboratorio, p.T53I, también se localiza en el dominio N-terminal, decidimos estudiar su efecto sobre la actividad de PLEKHG2 mediante el ensayo de luciferasa realizado en los trabajos mencionados anteriormente.

No obstante, nuestros resultados preliminares para la mutación p.T53I no confirman que ésta altere la funcionalidad de la proteína PLEKHG2. Mediante ensayo de luciferasa, en principio observamos que la proteína PLEKHG2 mutante poseía mayor actividad RhoGEF que la proteína WT. Sin embargo, cuando normalizamos los valores de luciferasa/Renilla con la expresión de la proteína PLEKHG2/actina, el ratio resultante nos indica que no hay casi diferencias entre WT y mutante. Por otra parte, el experimento por *Western blot* no revela cambios, en cuanto al tamaño e integridad, entre la proteína sobre-expresada WT y mutada (p.T53I).

El estudio de localización realizado mediante ensayo de inmunofluorescencia mostró que la proteína PLEKHG2, tanto WT como mutante, se localiza en membrana celular y citoplasma, por lo que la mutación p.T53I no interfiere en la localización subcelular. El mecanismo por el cual PLEKHG2 se localiza en la membrana celular, donde se localizan las proteínas Rac y Cdc42 en su forma inactiva (Runne, 2013), se desconoce.

Para poder concluir con mayor exactitud si existe una correlación entre la mutación p.T53I y el fenotipo del probando, sería necesario realizar un mayor número de repeticiones del ensayo luciferasa. Además, el análisis de la función enzimática directamente de muestras biológicas del paciente, como una biopsia de fibroblastos o sangre, o la utilización de modelos animales, podría aportar pruebas más sólidas que la expresión de la proteína *in vitro* (Richards *et al.*, 2015).

Por otra parte, el estudio genético de PLEKHG2 en otros pacientes de la serie clínica con el mismo fenotipo o similar, y la identificación en ellos de otras (o la misma variante) en *PLEKHG2*, ayudaría a asociar la variante con la enfermedad.

Los resultados obtenidos en el presente trabajo aportan información sobre posibles nuevas bases moleculares subyacentes en las NBIA, lo que establece una base para el desarrollo de nuevas estrategias terapéuticas, allanando el camino hacia la medicina personalizada.

6. CONCLUSIONES

En este trabajo se ha profundizado sobre la complejidad de las bases genéticas subyacentes en NBIA, un grupo heterogéneo clínica y genéticamente de trastornos del movimiento.

Por una parte, la secuenciación mediante WES de un trío (ambos progenitores y descendiente afectado) y su posterior procesamiento bioinformático ha permitido seleccionar dos variantes candidatas a ser causantes del fenotipo del probando fNBIA-24. Estas son las mutaciones en el gen *OTUD3*: NM_015207; c. 1121G>A; (p.R374K) y en el gen *P2RY4*: NM_002565.3; c. 778G>T; (p. P260T). La validación del diagnóstico propuesto requiere estudios adicionales, los cuales incluyen principalmente estudios de segregación y ensayos funcionales.

Por otra parte, mediante la sobre-expresión del gen *PLEKHG2* en células HeLa se han llevado a cabo ensayos funcionales con la finalidad de caracterizar la mutación c.158C>T (p.Thr53Ile), identificada en homocigosis en un paciente con atrofia de cerebelo. Nuestros estudios preliminares en células HeLa, no han mostrado diferencias sustanciales entre la proteína mutada y WT, tanto mediante el ensayo de luciferasa, como mediante el estudio de la localización celular realizada con la técnica de inmunofluorescencia.

7. REFERENCIAS

- AZIZ, N.; ZHAO, Q.; BRY, L.; DRISCOLL, D. K.; FUNKE, B.; GIBSON, J. S.; ... & MERKER, J. D. (2014). College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. *Archives of Pathology and Laboratory Medicine*, 139(4), 481-493.
- BA, W.; VAN DER RAADT, J. & KASRI, N. N. (2013). Rho GTPase signaling at the synapse: implications for intellectual disability. *Experimental cell research*, 319(15), 2368-2374.
- BASTIAN, F.; PARMENTIER, G.; ROUX, J.; MORETTI, S.; LAUDET, V. & ROBINSON-RECHAVI, M. (2008). Bgee: integrating and comparing heterogeneous transcriptome data among species. In *International Workshop on Data Integration in the Life Sciences, Lecture Notes in Computer Science*. (pp. 124-131). Springer, Berlin, Heidelberg.
- BETTENCOURT, C.; FORABOSCO, P.; WIETHOFF, S.; HEIDARI, M.; JOHNSTONE, D. M.; BOTÍA, J. A.; ... & HOULDEN, H. (2016). Gene co-expression networks shed light into diseases of brain iron accumulation. *Neurobiology of disease*, 87, 59-68.
- BHATTACHARYA, R.; ROSE, P. W.; BURLEY, S. K. & PRLIĆ, A. (2017). Impact of genetic variation on three dimensional structure and function of proteins. *PLoS one*, 12(3), e0171355.
- BOSTAN, A. C.; DUM, R. P. & STRICK, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in cognitive sciences*, 17(5), 241-254.
- CHANG, N.; EL-HAYEK, Y. H.; GOMEZ, E. & WAN, Q. (2007). Phosphatase PTEN in neuronal injury and brain disorders. *Trends in neurosciences*, 30(11), 581-586.
- CHOI, M.; SCHOLL, U. I.; JI, W.; LIU, T.; TIKHONOVA, I. R.; ZUMBO, P.; ... & NELSON-WILLIAMS, C. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences*, 106(45), 19096-19101.
- CHRYSOJA, C. C. & DIAMANDIS, E. P. (2014). Whole genome sequencing as a diagnostic test: challenges and opportunities. *Clin. Chem.* 60, 724–733. doi: 10.1373/clinchem.2013.20921
- CLAMP, M.; FRY, B.; KAMAL, M.; XIE, X.; CUFF, J.; LIN, M. F.; ... & LANDER, E. S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences*, 104(49), 19428-19433.
- COOK, D. R.; ROSSMAN, K. L.; & DER, C. J. (2014). Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. *Oncogene*, 33(31), 4021.
- DELAGUE, V.; JACQUIER, A.; HAMADOUCHE, T.; POITELON, Y.; BAUDOT, C.; BOCCACCIO, I.; ... & GRID, D. (2007). Mutations in FGD4 encoding the Rho GDP/GTP exchange factor FRABIN cause autosomal recessive Charcot-Marie-Tooth type 4H. *The American Journal of Human Genetics*, 81(1), 1-16.
- DI MEO, I. & TIRANTI, V. (2018). Classification and molecular pathogenesis of NBIA syndromes. *European journal of paediatric neurology*, 22(2), 272-284.
- DI RESTA, C.; GALBIATI, S.; CARRERA, P. & FERRARI, M. (2018). Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. *Ejifcc*, 29(1), 4.
- DROPPELMANN, C. A.; CAMPOS-MELO, D.; VOLKENING, K. & STRONG, M. J. (2014). The emerging role of guanine nucleotide exchange factors in ALS and other neurodegenerative diseases. *Frontiers in cellular neuroscience*, 8, 282.
- EDVARDSON, S.; WANG, H.; DOR, T.; ATAWNEH, O.; YAACOV, B.; GARTNER, J., ... & ELPELEG, O. (2016). Microcephaly-dystonia due to mutated PLEKHG2 with impaired actin polymerization. *neurogenetics*, 17(1), 25-30.

- FARWELL, K. D.; SHAHMIRZADI, L.; EL-KHECHEN, D.; POWIS, Z.; CHAO, E. C.; DAVIS, B. T.; ... & GANDOMI, S. K. (2015). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genetics in Medicine*, 17(7), 578.
- GOH, G. & CHOI, M. (2012). Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics & informatics*, 10(4), 214.
- GREGORY, A. & HAYFLICK, S. (2013). Neurodegeneration with brain iron accumulation disorders overview. In *GeneReviews®[Internet]*. University of Washington, Seattle.
- HARATA, A.; HIRAKAWA, M.; SAKUMA, T.; YAMAMOTO, T. & HASHIMOTO, C. (2019). Nucleotide receptor P2RY4 is required for head formation via induction and maintenance of head organizer in *Xenopus laevis*. *Development, Growth & Differentiation*, 61(2), 186-197.
- ITO, M. (2006). Cerebellar circuitry as a neuronal machine. *Progress in neurobiology*, 78(3-5), 272-303.
- JAFFE, A. B. & HALL, A. (2005). Rho GTPases: biochemistry and biology. *Annu. Rev. Cell Dev. Biol.*, 21, 247-269.
- JOHNSTON, J. J.; RUBINSTEIN, W. S.; FACIO, F. M.; NG, D.; SINGH, L. N.; TEER, J. K.; ... & BIESECKER, L. G. (2012). Secondary variants in individuals undergoing exome sequencing: screening of 572 individuals identifies high-penetrance mutations in cancer-susceptibility genes. *The American Journal of Human Genetics*, 91(1), 97-108.
- KRUEER, M. C.; BODDAERT, N.; SCHNEIDER, S. A.; HOULDEN, H.; BHATIA, K. P.; GREGORY, A.; ... & HAYFLICK, S. J. (2012). Neuroimaging features of neurodegeneration with brain iron accumulation. *American Journal of Neuroradiology*, 33(3), 407-414.
- KURIAN, M. A.; & HAYFLICK, S. J. (2013). Pantothenate kinase-associated neurodegeneration (PKAN) and PLA2G6-associated neurodegeneration (PLAN): review of two major neurodegeneration with brain iron accumulation (NBIA) phenotypes. In *International review of neurobiology* (Vol. 110, pp. 49-71). Academic Press.
- LEVI, S.; & TIRANTI, V. (2019). Neurodegeneration with brain iron accumulation disorders: valuable models aimed at understanding the pathogenesis of iron deposition. *Pharmaceuticals*, 12(1), 27.
- LIAW, D.; MARSH, D. J.; LI, J.; DAHIA, P. L.; WANG, S. I.; ZHENG, Z.; ... & ENG, C. (1997). Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature genetics*, 16(1), 64.
- LIN, M. E.; HERR, D. R.; & CHUN, J. (2010). Lysophosphatidic acid (LPA) receptors: signaling properties and disease relevance. *Prostaglandins & other lipid mediators*, 91(3-4), 130-138.
- LOHMANN, K.; & KLEIN, C. (2014). Next generation sequencing and the future of genetic diagnosis. *Neurotherapeutics* 11, 699–707. doi: 10.1007/s13311-014-0288-8
- MOCHIDA, G. H.; RAJAB, A.; EYALD, W.; LU, A.; AL-NOURI, D.; KOSAKI, K.; ... & APSE, K. (2004). Broader geographical spectrum of Cohen syndrome due to COH1 mutations. *Journal of medical genetics*, 41(6), e87-e87.
- NELEN, M. R.; VAN STAVEREN, W. C.; PEETERS, E. A.; BEN HASSEL, M.; GORLIN, R. J.; HAMM, H.; ... & MARIMAN, E. C. (1997). Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. *Human molecular genetics*, 6(8), 1383-1387.
- OLGIATI, S.; QUADRI, M.; & BONIFATI, V. (2016). Genetics of movement disorders in the next-generation sequencing era. *Movement Disorders*, 31(4), 458-470.

- PODURI, A.; SHEIDLEY, B. R.; SHOSTAK, S.; & OTTMAN, R. (2014). Genetic testing in the epilepsies-developments and dilemmas. *Nat. Rev. Neurol.* 10, 293–299. doi: 10.1038/nrneurol.2014.60
- RARAIGH, K. S.; HAN, S. T.; DAVIS, E.; EVANS, T. A.; PELLICORE, M. J.; MCCAGUE, A. F.; ... & SHERIDAN, M. B. (2018). Functional assays are essential for interpretation of missense variants associated with variable expressivity. *The American Journal of Human Genetics*, 102(6), 1062-1077.
- RAMNANI, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature reviews neuroscience*, 7(7), 511.
- REHM, H. L. (2013). Disease-targeted sequencing: a cornerstone in the clinic. *Nature Reviews Genetics*, 14(4), 295.
- REHM, H. L.; BALE, S. J.; BAYRAK-TOYDEMIR, P.; BERG, J. S.; BROWN, K. K.; DEIGNAN, J. L.; ... & LYON, E. (2013). ACMG clinical laboratory standards for next-generation sequencing. *Genetics in medicine*, 15(9), 733.
- REHMAN, A. U.; SANTOS-CORTEZ, R. L. P.; DRUMMOND, M. C.; SHAHZAD, M.; LEE, K.; MORELL, R. J.; ... & RIAZUDDIN, S. (2015). Challenges and solutions for gene identification in the presence of familial locus heterogeneity. *European Journal of Human Genetics*, 23(9), 1207.
- RICHARDS, S.; AZIZ, N.; BALE, S.; BICK, D.; DAS, S.; GASTIER-FOSTER, J.; ... & VOELKERDING, K. (2015). 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*, 17(5), 405.
- ROSENBAUM, D. M.; RASMUSSEN, S. G. & KOBILKA, B. K. (2009). The structure and function of G-protein-coupled receptors. *Nature*, 459(7245), 356.
- ROSSMAN, K. L.; DER, C. J. & SONDEK, J. (2005). GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nature reviews Molecular cell biology*, 6(2), 167.
- RUNNE, C. M. (2013). Function and Activation Mechanism of PLEKHG2, A Novel G Beta Gamma-Activated RhoGEF in Leukemia Cells. PhD (Doctor of Philosophy) thesis, University of Iowa, USA. <https://ir.uiowa.edu/etd/4907>.
- RUNNE, C. & CHEN, S. (2013). PLEKHG2 promotes heterotrimeric G protein $\beta\gamma$ -stimulated lymphocyte migration via Rac and Cdc42 activation and actin polymerization. *Molecular and cellular biology*, 33(21), 4294-4307.
- SCHMIDT, A. & HALL, A. (2002). Guanine nucleotide exchange factors for Rho GTPases: turning on the switch. *Genes & development*, 16(13), 1587-1609.
- SCHMIDT, S. & DEBANT, A. (2014). Function and regulation of the Rho guanine nucleotide exchange factor Trio. *Small GTPases*, 5(4), e983880.
- SCHNEIDER, S.A.; ESPINÓS, C. & PÉREZ-DUEÑAS, B (2019). Syndromes of Neurodegeneration with Brain Iron Accumulation, en: *Inherited Metabolic Movement Disorders: Recognition, Understanding, Improving Outcomes*. Cambridge University Press. Cambridge (UK) (en prensa).
- STARITA, L. M.; AHITUV, N.; DUNHAM, M. J.; KITZMAN, J. O.; ROTH, F. P.; SEELIG, G.; ... & FOWLER, D. M. (2017). Variant interpretation: functional assays to the rescue. *The American Journal of Human Genetics*, 101(3), 315-325.
- STAVROPOULOS, D. J.; MERICCO, D.; JOBLING, R.; BOWDIN, S.; MONFARED, N.; THIRUVAHINDRAPURAM, B.; ... & MARSHALL C.R. (2016). Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. *Npj Genomic Med.* 1:15012. doi: 10.1038/npjgenmed.2015.12

- STENDEL, C.; ROOS, A.; DECONINCK, T.; PEREIRA, J.; CASTAGNER, F.; NIEMANN, A.; ... & PARMAN, Y. (2007). Peripheral nerve demyelination caused by a mutant Rho GTPase guanine nucleotide exchange factor, frabin/FGD4. *The American Journal of Human Genetics*, 81(1), 158-164.
- TELLO, C.; DARLING, A.; LUPO, V.; PÉREZ-DUEÑAS, B. & ESPINÓS, C. (2017). On the complexity of clinical and molecular bases of neurodegeneration with brain iron accumulation. *Clinical genetics*, 93(4), 731-740.
- UEDA, H.; NAGAE, R.; KOZAWA, M.; MORISHITA, R.; KIMURA, S.; NAGASE, T.; ... & ASANO, T. (2008). Heterotrimeric G protein $\beta\gamma$ subunits stimulate FLJ00018, a guanine nucleotide exchange factor for Rac1 and Cdc42. *Journal of Biological Chemistry*, 283(4), 1946-1953.
- VAN AELST, L. & D'SOUZA-SCHOREY, C. (1997). Rho GTPases and signaling networks. *Genes & development*, 11(18), 2295-2322.
- VRIJENHOEK, T.; KRAAIJEVELD, K.; ELFERINK, M.; DE LIGT, J.; KRANENDONK, E.; SANTEN, G.; ... & DORLIJN, W. (2015). Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects. *European Journal of Human Genetics*, 23(9), 1142.
- WEILE, J. & ROTH, F. P. (2018). Multiplexed assays of variant effects contribute to a growing genotype–phenotype atlas. *Human genetics*, 137(9), 665-678.
- WILLIAMS, E. S. & HEGDE, M. (2013). Implementing genomic medicine in pathology. *Advances in anatomic pathology*, 20(4), 238-244.
- WONG, E. & CUERVO, A. M. (2010). Autophagy gone awry in neurodegenerative diseases. *Nature neuroscience*, 13(7), 805.
- WOODS, N. T.; BASKIN, R.; GOLUBEVA, V.; JHURANEY, A.; DE-GREGORIIS, G.; VACLOVA, T.; ... & MONTEIRO, A. N. (2016). Functional assays provide a robust tool for the clinical annotation of genetic variants of uncertain significance. *NPJ genomic medicine*, 1, 16001.
- YANG, Y.; MUZNY, D. M.; REID, J. G.; BAINBRIDGE, M. N.; WILLIS, A.; WARD, P. A.; ... & HARDISON, M. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *New England Journal of Medicine*, 369(16), 1502-1511.
- YUAN, L.; LV, Y.; LI, H.; GAO, H.; SONG, S.; ZHANG, Y.; ... & ZHOU, T. (2015). Deubiquitylase OTUD3 regulates PTEN stability and suppresses tumorigenesis. *Nature cell biology*, 17(9), 1169.
- ZHANG, Y.; SLOAN, S. A.; CLARKE, L. E.; CANEDA, C.; PLAZA, C. A.; BLUMENTHAL, P. D.; ... & DUNCAN III, J. A. (2016). Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron*, 89(1), 37-53.
- ZHENG, Y. (2001). Dbl family guanine nucleotide exchange factors. *Trends in biochemical sciences*, 26(12), 724-732.

8. ANEXOS

Anexo I. Electroferogramas correspondientes a la secuenciación Sanger de las muestras de la familia fNBIA-20.

OTUD3, c. 1121G>A (p.R374K)

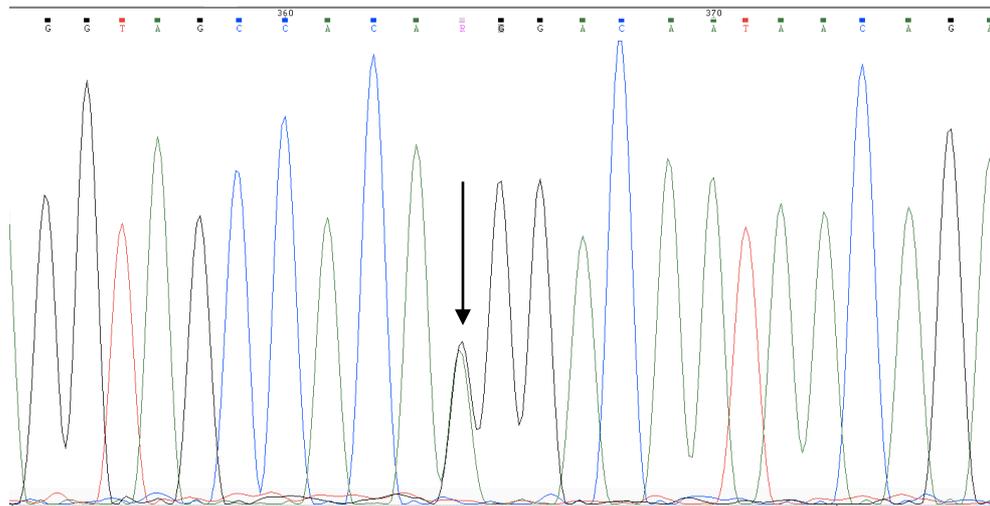


Figura 12. Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en OTUD3 c. 1121G>A (p.R374K).

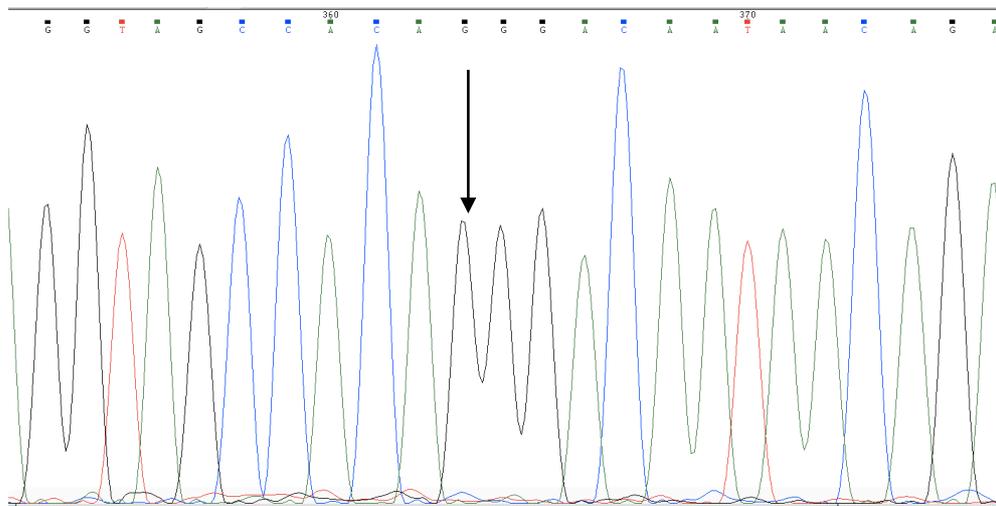


Figura 13. Electroferograma correspondiente a la secuenciación de la muestra NBIA-271 para el cambio en OTUD3 c. 1121G>A (p.R374K).

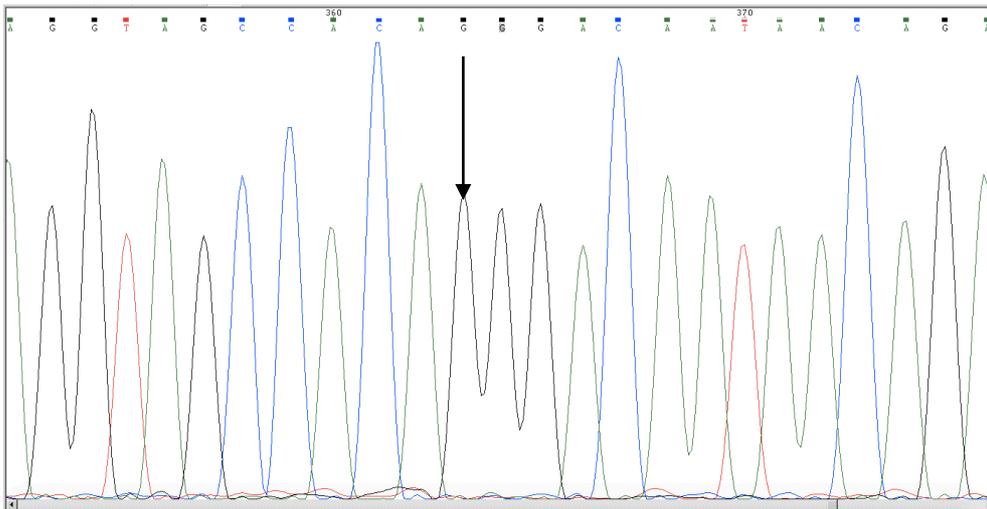


Figura 14. Electroferograma correspondiente a la secuenciación de la muestra NBIA-272 para el cambio en OTUD3 c. 1121G>A (p.R374K).

VPS13B, c.8788A>G (p.N2930D)

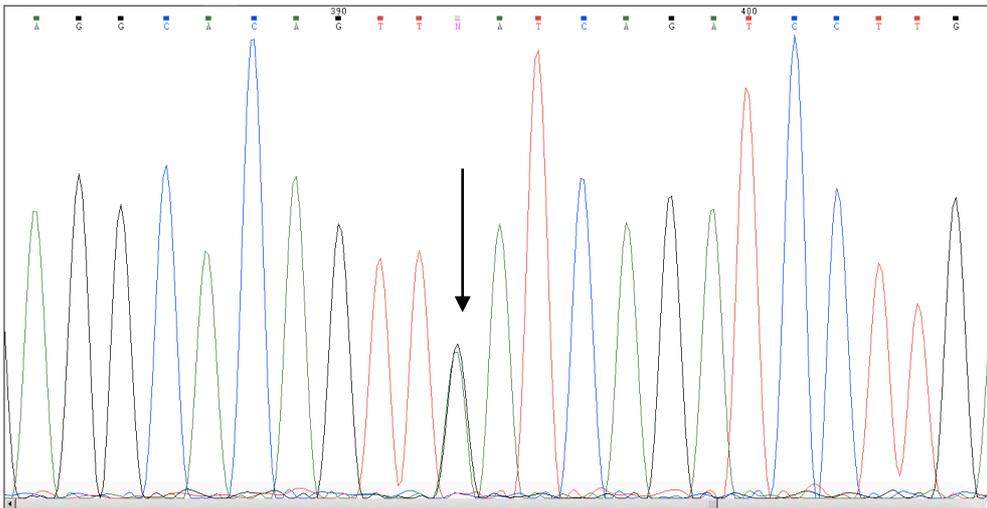


Figura 15. Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en VPS13B c.8788A>G (p.N2930D).

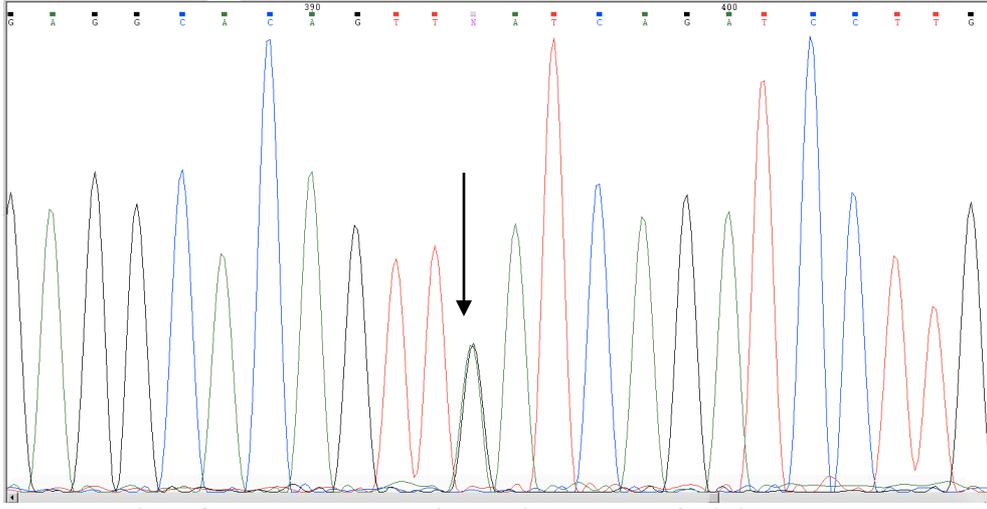


Figura 16. Electroferograma correspondiente a la secuenciación de la muestra NBIA-271 para el cambio en VPS13B c.8788A>G (p.N2930D).

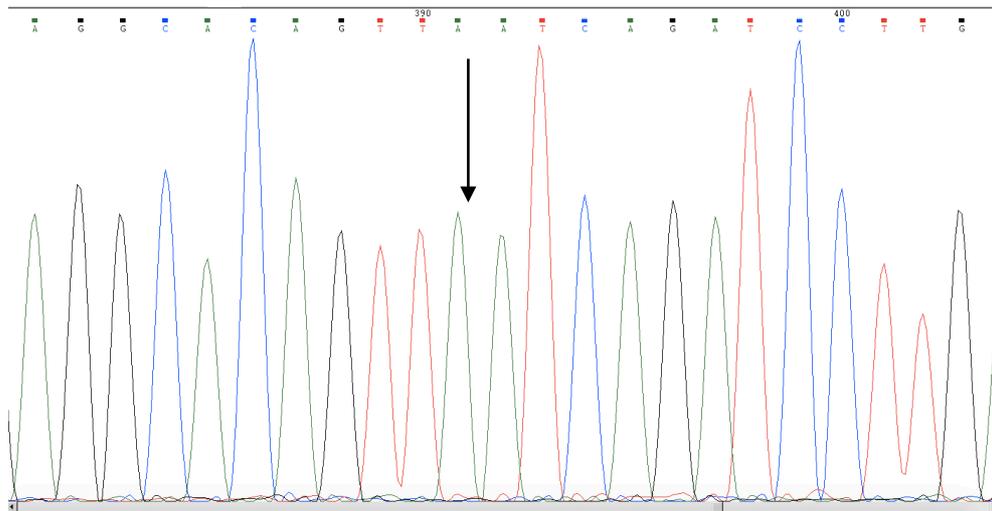


Figura 17. Electroferograma correspondiente a la secuenciación de la muestra NBIA-272 para el cambio en *VPS13B* c.8788A>G (p.N2930D).

VPS13B, c.9405G>T

En el caso del cambio c.9405G>T en *VPS13B* se muestra únicamente el electroferograma correspondiente a la muestra 24, dado que se trata de un falso positivo.

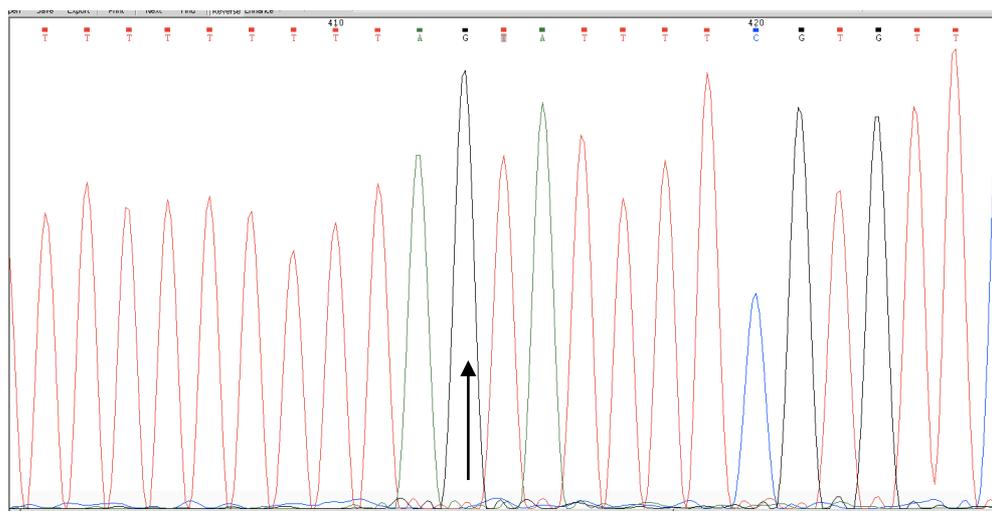


Figura 18. Electroferograma correspondiente a la secuenciación de la muestra NBIA-24 para el cambio en *VPS13B* c.9405G>T.