



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Biotecnología

Bases moleculares de la enfermedad de Wilson y
enfermedades Wilson-like

Trabajo Fin de Máster

Máster Universitario en Biotecnología Biomédica

AUTOR/A: Vidal Garcia, Clara

Tutor/a: Galindo Orozco, Máximo Ibo

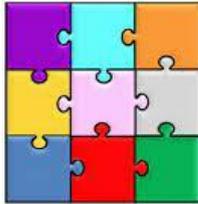
Cotutor/a externo: ESPINOS ARMERO, CARMEN ANGELES

Director/a Experimental: SANCHEZ MONTEAGUDO, ANA

CURSO ACADÉMICO: 2021/2022

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

DEPARTAMENTO DE BIOTECNOLOGÍA



CIPF – Lab of
RARE
NEURODEGENERATIVE
DISEASES



PRINCIPE FELIPE
CENTRO DE INVESTIGACION



BASES MOLECULARES DE LA ENFERMEDAD DE WILSON Y ENFERMEDADES *WILSON-LIKE*

TRABAJO FIN DE MÁSTER EN BIOTECNOLOGÍA BIOMÉDICA

ALUMNA: Clara Vidal Garcia

TUTOR UPV: Máximo Ibo Galindo Orozco

TUTORAS CIPF: Carmen Espinós Armero y Ana Sánchez Monteagudo

Curso Académico 2021-2022

VALENCIA, JUNIO DE 2022

RESUMEN

La enfermedad de Wilson (EW) es un trastorno hereditario autosómico recesivo, causado por mutaciones en *ATP7B*. Se trata de un trastorno multisistémico, siendo lo más frecuente las manifestaciones hepáticas, neurológicas y psiquiátricas. La evaluación clínica en los pacientes se basa en la escala de Leipzig, que engloba manifestaciones clínicas y hallazgos de pruebas bioquímicas, pero es el resultado del análisis genético de *ATP7B* lo que permite establecer un diagnóstico certero. Algunos pacientes pueden llegar a ser diagnosticados con EW en base a esta escala, pero no se logran identificar mutaciones en *ATP7B* causantes de la patología. Estos cuadros clínicos, que mimetizan la EW al presentar signos similares, pero con distinta etiología, se denominan fenotipos *Wilson-like*. Identificar los genes implicados en estas enfermedades es importante para mejorar el diagnóstico y tratamiento que reciben los pacientes.

En este trabajo se han analizado datos de secuenciación de exoma completo en un caso familiar con fenotipo *Wilson-like*, con la finalidad de identificar mutaciones candidatas en otros genes que pudieran explicar el cuadro clínico observado en los pacientes. Mediante el filtrado y priorización de variantes obtenidas, se identificaron en la familia fMD-218 cambios candidatos en *NOTCH2*, *CARD10*, *CBFA2T3* y *COG5*. El análisis genético de las variantes detectadas en familiares reveló que únicamente los cambios en *NOTCH2* y *CARD10* segregarían con la enfermedad, si bien estos hallazgos no son suficientes para alcanzar un diagnóstico definitivo. Uno de los desafíos que existe actualmente en el diagnóstico molecular es la interpretación y caracterización de las mutaciones que permita establecer la variante identificada como la mutación causal del fenotipo observado. Se han llevado a cabo estudios funcionales de las variantes c.704+2T>G en *INSIG1* (identificada en la familia fEW-18) y c.5003-23T>C en *NOTCH2* (identificada en la familia fMD-218). Para ambas se realizó un análisis de mRNA en sangre periférica del paciente y un ensayo de minigenes con los que se descartó que la variante en *NOTCH2* se trate de una mutación patológica. En cambio, para la variante en *INSIG1* no se obtuvieron resultados concluyentes, lo que hace necesario ampliar su estudio para caracterizarla correctamente.

El seguimiento de los pacientes con EW incluye pruebas bioquímicas y de imagen que resultan insuficientes para establecer el diagnóstico y asegurar la cumplimentación terapéutica. Por ello, existe la necesidad de identificar nuevos marcadores que permitan anticipar la progresión de la enfermedad. Los microRNAs (miRNAs), moléculas liberadas en biofluidos como respuesta a procesos fisiológicos, son ampliamente investigados como biomarcadores en patologías con muy diferente etiología. En este trabajo se han analizado mediante PCR cuantitativa los niveles en circulación de miR-146a-5p, miR-340-3p y miR-485-3p, previamente identificados en el laboratorio como significativamente desregulados en plasma de pacientes EW, en plasma de ratones *Atp7b*^{-/-} y *wild-type*. Los tres miRNAs fueron detectados en las muestras estudiadas, observándose una tendencia a disminuir para miR-146a-5p, miR-340-3p y un incremento significativo de miR-485-3p en plasma de ratones *Atp7b*^{-/-} con respecto al grupo *wild-type*. Es necesario ampliar la cohorte de muestras para determinar si las diferencias entre grupos son estadísticamente significativas.

Palabras clave: Enfermedad de Wilson; diagnóstico molecular; secuenciación de nueva generación; fenotipo *Wilson-like*; miRNAs circulantes

ABSTRACT

Wilson's disease (WD) is an autosomal recessive inherited disorder, caused by mutations in *ATP7B*. It is a multisystemic disorder, with hepatic, neurologic and psychiatric manifestations being the most common. Clinical evaluation of patients is based on the Leipzig scale, which encompasses clinical manifestations and biochemical test findings; however, it is the result of the genetic analysis of *ATP7B* that allows a definitive WD diagnosis. Some patients are diagnosed of WD according to this score, but no disease-causing mutations in *ATP7B* are identified in them. These clinical pictures, which mimic WD by presenting similar signs but with different aetiology, are called Wilson-like phenotypes. Identifying the genes involved in these diseases is important to improve the diagnosis and treatment that patients receive.

In this work, whole exome sequencing data have been analyzed in a familial case with Wilson-like phenotype, in order to identify candidate mutations in other genes that could explain the clinical picture observed in patients. By filtering and prioritizing the obtained variants, candidate changes in *NOTCH2*, *CARD10*, *CBFA2T3* and *COG5* were identified in the fMD-218 family. Genetic analysis of the variants detected in relatives revealed that only changes in *NOTCH2* and *CARD10* would segregate with the disease, although these findings are not enough to reach a definitive diagnosis. One of the challenges that currently exists in molecular diagnosis is the interpretation and characterization of the mutations that allow establishing the variant identified as the causal mutation of the observed phenotype. Functional studies of the variants c.704+2T>G in *INSIG1* (identified in the fEW-18 family) and c.5003-23T>C in *NOTCH2* (identified in the fMD-218 family) have been carried out. For both, an analysis of mRNA in the patient's peripheral blood and a minigene assay were performed, in order to rule out that the variant in *NOTCH2* is a pathological mutation. In contrast, for the variant in *INSIG1*, no conclusive results were obtained, which makes it necessary to extend its study to characterize it correctly.

Monitoring of patients with WD includes biochemical and imaging tests that are insufficient to establish the diagnosis and ensure therapeutic compliance. Therefore, there is a need to identify new markers that allow anticipating the progression of the disease. MicroRNAs (miRNAs), molecules released in biofluids in response to physiological processes, are broadly studied as biomarkers for pathologies with disparate etiology. In this work, the circulating levels of miR-146a-5p, miR-340-3p and miR-485-3p, previously identified in the laboratory as significantly deregulated in plasma from WD patients, have been analyzed using quantitative PCR in plasma from mice *Atp7b*^{-/-} and wild-type. The three miRNAs were detected in the studied samples, observing a tendency to decrease for miR-146a-5p, miR-340-3p and a significant increase of miR-485-3p in plasma of *Atp7b*^{-/-} mice with respect to the wild-type group. It is necessary to increase the sample cohort to determine if the differences between groups are statistically significant.

Keywords: Wilson disease; molecular diagnosis, next-generation sequencing, Wilson-like phenotype; circulating miRNAs

ABREVIATURAS

1000GP	Proyecto de los 1000 genomas
AD	Autosómica dominante
AR	Autosómica recesiva
CDG2I	Trastorno congénito de glicosilación tipo Iii (<i>Congenital Disorder of Glycosylation type Iii</i>)
cDNA	DNA complementario
CNAG	Centro Nacional de Análisis Genómico
COG	Complejo oligomérico de Golgi (<i>Oligomeric Golgi Complex</i>)
Ct	Umbral de ciclo (<i>cycle threshold</i>)
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico (<i>deoxyribonucleic acid</i>)
EW	Enfermedad de Wilson
HPO	Ontología del fenotipo humano (<i>Human Phenotype Ontology</i>)
iPSC	Células madre pluripotentes inducidas
kDa	Kilo-Dalton
KF	Kayser-Fleischer
MAF	Frecuencia del alelo menos frecuente (<i>Minor Allele Frequency</i>)
miRNA	Micro RNA
MLPA	Amplificación de sonda dependiente de ligadura múltiple (<i>Multiplex Ligation dependent Probe Amplification</i>)
mRNA	RNA mensajero
NGS	Secuenciación de próxima generación (<i>Next generation sequencing</i>)
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa (<i>Polymerase chain reaction</i>)
qPCR	PCR cuantitativa (<i>quantitative PCR</i>)
rAAV8	Virus adenoasociado recombinante 8 (<i>Recombinant Adeno-associated virus 8</i>)
RNA	Ácido ribonucleico (<i>ribonucleic acid</i>)
RT-PCR	Reacción en cadena de la polimerasa con transcriptasa inversa (<i>Reverse transcription polymerase chain reaction</i>)
SNP	Polimorfismo de un único nucleótido (<i>Single nucleotide polymorphism</i>)
UTR	Región no traducida (<i>Un-Translated Regions</i>)
WES	Secuenciación de exoma completo (<i>Whole-Exome Sequencing</i>)
WGS	Secuenciación de genoma completo (<i>Whole-Genome Sequencing</i>)

OMIM	Herencia mendeliana en el hombre (<i>Online Mendelian Inheritance in Man</i>)
ESE	Secuencia exónica que potencia el <i>splicing</i> (<i>Exonic Splicing Enhancer</i>)
ESS	Secuencia exónica silenciadora de <i>splicing</i> (<i>Exonic Splicing Silencer</i>)
MCS	Sitio de clonación múltiple (<i>Multiple Cloning Site</i>)
KO	<i>Knockout</i> .
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
FBS	Suero fetal bovino (<i>Fetal Bovine serum</i>)
RNasas	Ribonucleasas
HGVS	<i>Human Genome Variation Society</i>
NBIA	Síndromes de neurodegeneración con acumulación cerebral de hierro (<i>neurodegeneration with brain iron accumulation</i>)

GENES

ATP7B	<i>ATPase copper transporting beta</i>
ATP7A	<i>ATPase copper transporting alfa</i>
JAG1	<i>Jagged-1</i>
NOTCH2	<i>Notch receptor 2</i>
CCDC115	<i>Coiled-Coil Domain Containing 115</i>
TMEM199	<i>Transmembrane protein 199</i>
CARD10	<i>Caspase Recruitment Domain Family Member 10</i>
COG5	<i>Component of Oligomeric Golgi complex 5</i>
CBFA2T3	<i>CBFA2/RUNX1 partner transcriptional co-repressor 3</i>

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. Enfermedad de Wilson (EW)	3
1.1.1. Sintomatología de la enfermedad	3
1.1.2. Diagnóstico y tratamiento	4
1.2. <i>ATP7B</i> : caracterización molecular y su función en la homeostasis del cobre	6
1.3. miRNAs como biomarcadores en la EW	8
2. HIPÓTESIS Y OBJETIVOS.....	11
3. MATERIAL Y MÉTODOS.....	15
3.1. Sujetos del estudio	17
3.2. Estudio genético de la familia fMD-218 mediante secuenciación de exoma	18
3.2.1. Filtrado y priorización de variantes	18
3.2.2. Validación de variantes candidatas y secuenciación Sanger	21
3.2.2.1. Diseño de cebadores y amplificación por PCR	21
3.2.2.2. Análisis de secuencias.....	22
3.3. Análisis funcional de variantes de <i>splicing</i>	23
3.3.1. Caracterización de variantes de <i>splicing</i> mediante ensayo de minigenes	23
3.3.2. Caracterización de variantes de <i>splicing</i> en mRNA de sangre periférica.....	25
3.4. Estudio de miRNA circulantes en plasma de ratón.....	26
3.4.1. Extracción de RNA total circulante	26
3.4.2. Transcripción reversa a cDNA.....	26
3.4.3. PCR cuantitativa (qPCR) de los miRNA de interés.....	27
3.4.4. Análisis estadístico de los datos de qPCR.....	27
4. RESULTADOS	29
4.1. Variantes candidatas identificadas en la familia fMD-218	31
4.2. Análisis funcional de variantes de <i>splicing</i>	33
4.2.1. Variante de <i>splicing</i> c.704+2T>G en <i>INSIG1</i>	33
4.2.2. Variante de <i>splicing</i> c.5003-23T>C en <i>NOTCH2</i>	35
4.3. Estudio de miRNAs circulantes en plasma como biomarcadores de la EW	37
5. DISCUSIÓN	39
5.1. Diagnóstico genético.....	41
5.2. Biomarcadores en la EW	46
6. CONCLUSIONES	49
7. BIBLIOGRAFÍA.....	53
ANEXOS	61

ÍNDICE DE FIGURAS

Figura 1.1. Sintomatología causada por la acumulación de cobre en la enfermedad de Wilson.....	4
Figura 1.2. Diagrama de flujo de los pasos a seguir para el diagnóstico genético de la EW y enfermedades con fenotipos Wilson-like.....	7
Figura 1.3. Topología de la proteína ATP7B	8
Figura 3.1. Árbol genealógico de las familias fEW-18 y fMD-218	17
Figura 3.2. Esquema del procedimiento de filtrado y priorización de variantes registradas en secuenciación de exoma de la familia fMD-218.	19
Figura 3.3. Mapa del vector reportero pSPL3	23
Figura 3.4. Esquema de la construcción de minigenes a partir del vector pSPL3 utilizada en ambos estudios	24
Figura 4.1. Resultados del análisis del mRNA en sangre periférica de la variante de splicing c.704+2T>G en <i>INSIG1</i>	34
Figura 4.2. Resultados del ensayo de minigenes de la variante de splicing c.704+2T>G en <i>INSIG1</i>	35
Figura 4.3. Resultados del análisis del mRNA en sangre periférica de la variante de splicing c.5003-23T>C en <i>NOTCH2</i>	36
Figura 4.4. Resultados del ensayo de minigenes de la variante de splicing c.5003-23T>C en <i>NOTCH2</i>	37
Figura 4.5. Análisis comparativo entre grupos de los niveles de miR146a-5p, miR340-3p y miR485-3p mediante qPCR	38

ÍNDICE DE TABLAS

Tabla 1.1 Escala diagnóstica desarrollada en la 8ª Reunión Internacional sobre la EW (Leipzig 2001).....	5
Tabla 3.1. A) Concentraciones y componentes de la Master Mix empleada en la reacción de PCR. B) Programa de PCR configurado en el termociclador para las amplificaciones	22
Tabla 3.2. Sondas <i>TaqMan Advanced</i> utilizadas para la amplificación de miRNAs circulantes mediante qPCR.....	27
Tabla 4.1. Variantes candidatas priorizadas en el WES de la familia fMD-218	31

1. INTRODUCCIÓN

1.1. Enfermedad de Wilson (EW)

La degeneración hepatolenticular progresiva, o enfermedad de Wilson (EW), es un trastorno raro que afecta al metabolismo del cobre. El neurólogo Samuel Alexander Kinner Wilson describió por primera vez en 1912 varios casos como una enfermedad nerviosa de naturaleza desconocida en aquel momento. No fue hasta 40 años después que se estableció la EW como una enfermedad autosómica recesiva (Bearn, 1953) y ya en 1994 se caracterizaron mutaciones en el gen *ATP7B* (ATPasa transportadora de cobre) como causantes de la enfermedad. En 1984 se calculó la prevalencia de esta enfermedad como 1:30.000, lo que correspondería a una frecuencia de portadores de 1:90-1:150. Estos datos son similares a los calculados recientemente, dónde se estima una prevalencia de 1:29.000 – 1:40.000, la cual es mayor en poblaciones con alta consanguinidad como en India o la isla de Cerdeña (Sandahl *et al.*, 2020).

1.1.1. Sintomatología de la enfermedad

La EW puede aparecer a cualquier edad, pero la mayoría de los pacientes se diagnostican entre los 5 y 35 años. Se trata de un trastorno multisistémico con síntomas variables caracterizados por el depósito anormal de cobre, siendo los más frecuentes los trastornos hepáticos, neurológicos y psiquiátricos (Sánchez-Monteagudo *et al.*, 2021).

Cualquier tipo de enfermedad hepática puede desarrollarse en pacientes con EW. Más concretamente, los síntomas hepáticos van desde asintomáticos con anormalidades bioquímicas hasta hepatitis crónica, fibrosis y cirrosis. Incluso puede cursar con insuficiencia hepática aguda, a veces asociada con anemia hemolítica de Coombs negativa, pudiendo preceder estos síntomas a los neuropsiquiátricos hasta en 10 años. Por otra parte, los signos neurológicos son los síntomas clínicos más frecuentes tras los hepáticos (en torno al 50% de los pacientes presentan enfermedad hepática), teniendo mayor susceptibilidad a la toxicidad por acumulación de cobre en los ganglios basales y el tronco encefálico. La presentación neurológica puede ser muy sutil e interrumpida durante muchos años, pero también puede desarrollarse con gran rapidez, alcanzándose en poco tiempo una incapacidad completa. Los signos neurológicos pueden cursar con diversos trastornos del movimiento como la ataxia y apraxia de la marcha, llegando en los casos más extremos a distonía, temblores, parkinsonismo, convulsiones y coreoacantocitosis. Asimismo, sobre un 10-20% de los pacientes presentan trastornos psiquiátricos como un aumento de la impulsividad, tendencias autolesivas, cambios de personalidad y exhibicionismo sexual (Czlonkowska *et al.*, 2018). En el caso de niños con la EW se puede observar un menor rendimiento escolar, impulsividad o cambios en la personalidad que inicialmente pueden confundirse con problemas de comportamiento asociados a la pubertad (*Guías de Práctica Clínica de la EASL*, 2012).

Finalmente, se distinguen síntomas adicionales como la presencia del anillo de Kayser-Fleischer (KF) como consecuencia de la deposición de cobre en la membrana de Descemet en la córnea, signo

presente en más de la mitad de pacientes sin síntomas neurológicos y en el 95% de los que presentan síntomas neurológicos, además de la aparición de complicaciones renales o síntomas de enfermedad articular (Kasztelan-Szczerbinska and Cichoz-Lach, 2021). La Figura 1.1 esquematiza el cuadro clínico de la EW.

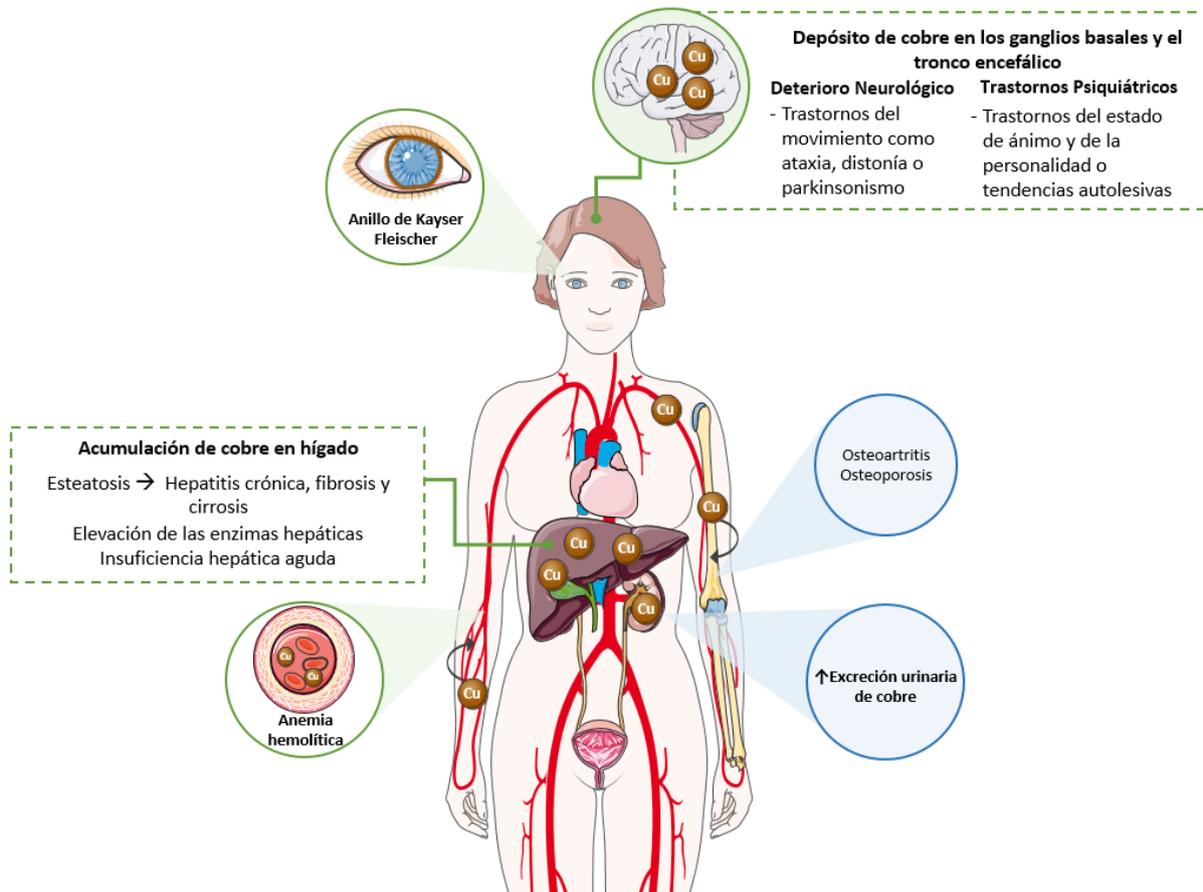


Figura 1.1. Sintomatología causada por la acumulación de cobre en la enfermedad de Wilson. En verde se indican los principales síntomas y signos de la enfermedad, mientras que en azul podemos observar características secundarias que pueden presentar los pacientes (adaptado de Sánchez-Monteagudo *et al.*, 2021)

1.1.2. Diagnóstico y tratamiento

El diagnóstico temprano de la EW no es sencillo. Este es importante para evitar futuras complicaciones, principalmente neurológicas, ya que este trastorno del metabolismo del cobre dispone de terapia para evitar la progresión de la enfermedad. Las manifestaciones de la enfermedad son diversas, no siempre están presentes y los marcadores indicativos de la acumulación del cobre no son del todo sensibles ni específicos de la enfermedad. Actualmente el diagnóstico está basado en la escala de Leipzig, desarrollada en la 8ª Reunión Internacional sobre la EW (Tabla 1.1) (*Guías de Práctica Clínica de la EASL*, 2012). Esta escala tiene en cuenta los síntomas clínicos, la histopatología, la bioquímica y la genética, necesitando una puntuación final de 4 o más para confirmar el diagnóstico de EW, siendo el test genético suficiente para diagnosticar

la enfermedad; como se recogen en la Tabla 1.1, la identificación de dos mutaciones patológicas en *ATP7B* se corresponde con una puntuación de 4. Algunos pacientes pueden ser diagnosticados en una fase inicial como enfermos de Wilson basándose en esta escala, pero tras el estudio genético adecuado no se logran identificar mutaciones patológicas en *ATP7B*. Estas enfermedades, que cursan con un fenotipo similar al descrito para la EW, se denominan enfermedades *Wilson-like*, ya que mimetizan el cuadro clínico de la EW, si bien su etiología es distinta. Por ello, es crucial llevar a cabo el estudio genético de *ATP7B* para alcanzar un diagnóstico certero y poder proporcionar un tratamiento adecuado a los pacientes.

Tabla 1.1: Escala diagnóstica desarrollada en la 8ª Reunión Internacional sobre la EW (Leipzig 2001)

Síntomas y signos clínicos típicos		Otros test	
Anillos de Kayser-Fleischer		Cobre hepático	
Presente	2	>5x LSN (>4 μmol/g)	2
Ausente	0	0,8-4 umol/g	1
Síntomas neurológicos**		Normal (<0,8μmol/g)	-1
Severos	2	Gránulos rodamina positivos*	1
Moderados	1	Cobre en orina 24h (sin hepatitis aguda)	
Ausentes	0	Normal	0
Ceruloplasmina sérica		1-2x LSN	1
Normal (> 0,2 g/L)	0	>2x LSN	2
0,1-0,2 g/L	1	Normal, pero >5x LSN con D-penicilamina	2
<0,1 g/L	2	Estudio genético <i>ATP7B</i>	
Anemia hemolítica Coombs negativa		2 mutaciones en <i>trans</i>	4
Presente	1	1 mutación	1
Ausente	0	Negativo	0
PUNTUACIÓN TOTAL		Resultado evaluación:	
4 puntos o más		Diagnóstico confirmado	
3 puntos		Diagnóstico posible, se necesitan más tests	
<2 puntos		Diagnóstico poco probable	

LSN: límite superior de la normalidad; **Alternativamente, anomalías típicas en RMN cerebral

*Si no hay cuantificación de cobre intrahepático disponible

Sin la terapia adecuada, la EW no tiene buen pronóstico, siendo la mayoría de muertes de los pacientes por enfermedad hepática y en menor número por complicaciones neurológicas. Afortunadamente, hoy en día se dispone de distintos tratamientos, que en tanto que la enfermedad es crónica, son de por vida, como la D-penicilamina y la trientina, cuyos mecanismos de acción implican la quelación del cobre circulante para su excreción en la orina. Con otro modo de acción, las sales de zinc que inducen la síntesis de metalotioneínas en enterocitos, que se unirán al cobre para impedir su absorción en la circulación portal, excretándose posteriormente en las heces; el tetratiomolibdato de amonio que forma un complejo con proteínas y cobre en los hepatocitos, que posteriormente se excretará en la bilis. Por último, el trasplante de hígado que se realiza en un 5%

de los casos de EW que presentan insuficiencia hepática aguda; en estos casos el metabolismo del cobre se normaliza y ya no se requiere de medicación para evitar su acumulación (Kasztelan-Szczerbinska and Cichoz-Lach, 2021).

Hasta la fecha, los tratamientos se han basado en eliminar el exceso de cobre, en evitar su absorción o, en última instancia, en el trasplante hepático. En la actualidad, existen ensayos clínicos en marcha de nuevas terapias centradas en restablecer la vía para la correcta eliminación del cobre en el hígado. Un ejemplo de estudio que supondría la cura de la enfermedad es la terapia génica. Una de las iniciativas surgida en 2016 va dirigida al hígado, en el que mediante el uso de un vector viral recombinante adeno-asociado (rAAV8), se introduce una copia normal del gen *ATP7B* para suplir la función del gen mutado del paciente. Esta aproximación ha mostrado buenos resultados en ratones (Murillo *et al.*, 2016), y se está llevando a cabo el ensayo clínico en fase I/II con el que pretenden comprobar la seguridad, tolerabilidad y actividad farmacológica de dosis únicas del fármaco (Ref. NCT04537377). Otro enfoque es el trasplante de hepatocitos. Para ello, se obtienen iPSC a partir de fibroblastos de pacientes donde se corrige la mutación causante de la enfermedad con un vector lentiviral, estos hepatocitos corregidos podrían usarse posteriormente para llevar a cabo un trasplante autólogo (Lucena-Valera *et al.*, 2021).

1.2. *ATP7B*: caracterización molecular y su función en la homeostasis del cobre

El gen *ATP7B* se encuentra en el brazo corto del cromosoma 13 y su transcrito canónico, que abarca alrededor de 8 kb, contiene 21 exones (NM_000053). Ya se han descrito más de 900 variantes patológicas o probablemente patológicas a lo largo del gen (Tang *et al.*, 2019; HGMD Professional versión 2022.3), y se ha comprobado que el 96,8% de las mutaciones se pueden detectar mediante la secuenciación de los exones y sus secuencias intrónicas flanqueantes. Fuera de esta ratio de detección encontramos grandes deleciones/duplicaciones en el gen que pueden identificarse mediante el uso de la técnica MLPA (*Multiplex Ligation-dependent Probe Amplification*) y, además, se han descrito algunas mutaciones intrónicas profundas que no podrían detectarse con este tipo de aproximaciones. También es frecuente estudiar las muestras con paneles de genes asociados a alteraciones hepáticas o paneles de trastornos del movimiento (según la clínica del paciente). En el caso de pacientes de determinado origen geográfico donde tienen mayor prevalencia ciertas mutaciones, es común estudiar primero estas por agilizar el proceso, como es el ejemplo de la deleción de 15 nucleótidos en la región promotora (c.-441_-427del), mutación prevalente en la isla de Cerdeña (Hahn, 2014).

Los pacientes pueden ser homocigotos, portando una misma mutación en ambos alelos, o heterocigotos compuestos portando dos mutaciones distintas patológicas/probablemente patológicas en *ATP7B* que causen la enfermedad. En un estudio genético completo, incluyendo regiones codificantes e intrónicas flanqueantes, promotor y UTRs (*Un-Translated Regions*), la ratio de detección de mutaciones causales en series clínicas de casos con cuadro clínico compatible con la EW está en torno a 77-98% (Mukherjee *et al.*, 2014; Ferenci *et al.*, 2019).

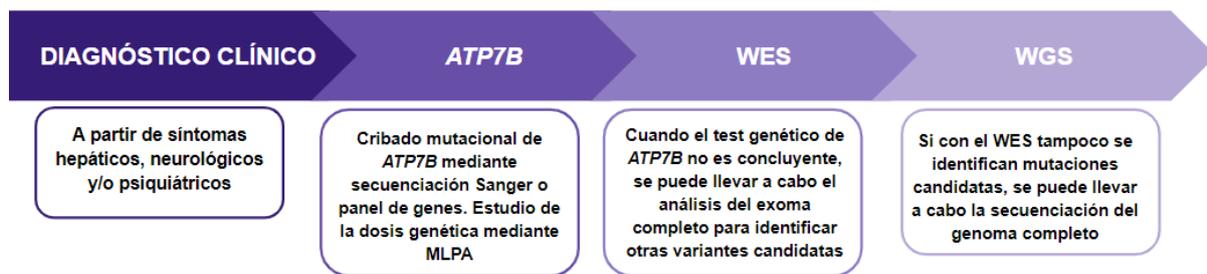


Figura 1.2. Diagrama de flujo de los pasos a seguir para el diagnóstico genético de la EW y enfermedades con fenotipos *Wilson-like*. Inicialmente se cuenta con un diagnóstico clínico en base al fenotipo del paciente, que se confirma con el cribado mutacional de *ATP7B*. Si el resultado del análisis no es concluyente, se pueden contemplar otras aproximaciones NGS (*Next Generation Sequencing*) como WES (*Whole-Exome Sequencing*) o WGS (*Whole-Genome Sequencing*) con las que podemos identificar otras posibles variantes causantes de fenotipos *Wilson-like*. MLPA: *Multiplex Ligation-dependent Probe Amplification*

Si en el cribado mutacional de *ATP7B* no se identifican variantes deletéreas en este gen y la sintomatología es clara, se podría plantear que el paciente sufre otra enfermedad que mimetiza los síntomas de la EW. Algunas enfermedades con fenotipos *Wilson-like* son el síndrome de Alagille, causado por mutaciones en los genes *JAG1* o *NOTCH2*, trastornos congénitos de glicosilación asociados a alteraciones en los genes *CCDC115* o *TMEM199*, o la enfermedad de Menkes, causada por mutaciones en el gen *ATP7A*. Llegados a este punto, identificar los genes implicados en una enfermedad *Wilson-like* es importante para poder mejorar el tratamiento recibido por los pacientes. Por este motivo se amplían los estudios usando técnicas NGS (*Next Generation Sequencing*) como el exoma completo o WES (*Whole-Exome Sequencing*), o incluso de genoma completo o WGS (*Whole-Genome Sequencing*) (Espinós and Ferenci, 2020). El WES se basa en el enriquecimiento empleando sistemas de captura de las regiones codificantes junto con las intrónicas flanqueantes del genoma para su posterior secuenciación masiva. En cambio, el WGS permitiría el estudio mediante secuenciación masiva de todo el genoma, incluidas las regiones no codificantes. Estas nuevas tecnologías se están utilizando cada vez más en el ámbito clínico a medida que se van mejorando las técnicas y se abaratan los costes. Aunque su uso no asegure detectar las variantes causantes de la enfermedad en todos los pacientes estudiados, sí permiten aumentar la tasa de éxito. En el caso del diagnóstico genético de enfermedades raras en niños mediante WES se ha visto que estudiando sólo el probando se consigue una tasa de éxito de alrededor al 28% de los casos, aumentando esta tasa hasta un 40% cuando también se lleva a cabo la secuenciación de los padres del paciente (estudio de un trío), lo cual mejora el rendimiento diagnóstico (Wright *et al.*, 2018). Otros estudios han mostrado un aumento en el rendimiento del diagnóstico genético en un amplio espectro de enfermedades raras mediante WGS, aumentándose el ratio diagnóstico hasta el 35% (Smedley *et al.*, 2021).

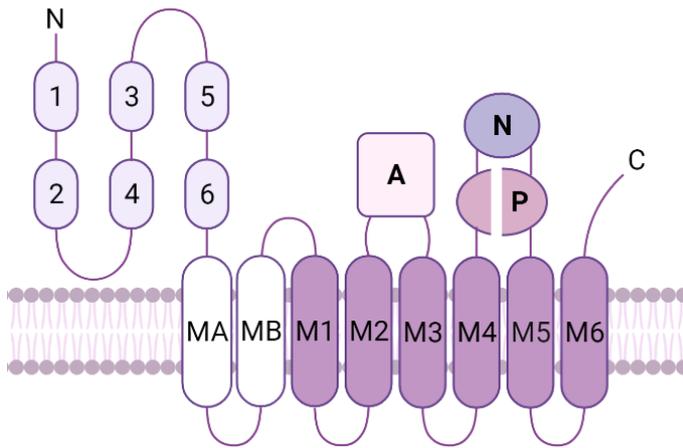


Figura 1.3. Topología de la proteína ATP7B (adaptado de Bitter *et al.*, 2022)

La proteína ATP7B consta de 1465 aminoácidos y tiene un peso molecular de aproximadamente 160 kDa. Esta proteína es un transportador de cobre transmembrana, localizado en el aparato de Golgi en condiciones fisiológicas, que pertenece a la subfamilia P1B de ATPasas. Su estructura consta de una región N-terminal citoplasmática formada por 6 dominios de unión a metales; seguida de un dominio transmembrana compuesto por un total de 8 hélices,

siendo las dos primeras (MA y MB) específicas de la subfamilia P1B, mientras que las 6 hélices restantes (M1 a M6) son compartidas con el resto de ATPasas tipo P; además posee un dominio activador (dominio A), uno fosforilador (dominio P) y otro de unión a ATP (dominio N) (Figura 1.3) (Bitter *et al.*, 2022).

En los hepatocitos, por una parte, la proteína ATP7B transporta el cobre presente en el citosol hacia el aparato de Golgi y media su incorporación a la ceruloplasmina (seis átomos de cobre por cada molécula de ceruloplasmina), que cargada con el cobre pasa a ser holoceruloplasmina y es secretada a la circulación. Por otra parte, también se encarga de mediar en la excreción biliar del cobre cuando este se encuentra en exceso en los hepatocitos. Para ello, ATP7B participa en la eliminación de vesículas con alto contenido de cobre, que son desplazadas desde la red *trans* Golgi al canalículo biliar (Kerker and Roberts, 2018). Defectos en la proteína resultan en una pérdida de función que dará lugar a la ausencia de holoceruloplasmina y a la incapacidad de eliminar correctamente el cobre a través de la bilis, lo que se traduce en la acumulación de cobre en los hepatocitos y, en última instancia, a la circulación de cobre libre, no unido a ceruloplasmina o fijado por glutatión o las metalotioneínas, depositándose en diferentes órganos y tejidos generando toxicidad, lo que caracteriza el cuadro clínico descrito en los pacientes para esta patología.

1.3. miRNAs como biomarcadores en la EW

Los microRNAs (miRNAs) son pequeños RNA, de unos 22 nucleótidos, no codificantes que participan en la regulación de la expresión génica a nivel post-transcripcional. Se ha demostrado que son esenciales para el desarrollo normal del individuo ya que están relacionados con una amplia variedad de procesos biológicos importantes. Aunque normalmente se encuentran en el citoplasma de las células, los miRNAs pueden liberarse en los fluidos y llevar a cabo su función en otras zonas

del organismo. Estos miRNAs extracelulares se han propuesto como marcadores mínimamente invasivos que permiten la detección temprana de enfermedades, ya que pueden estudiarse en biofluidos como suero o plasma y, a diferencia de las demás especies de RNA, son altamente estables y resisten a la degradación por endonucleasas (O'Brien *et al.*, 2018). No obstante, se ha visto que los niveles de miRNAs en biofluidos depende de la edad y del sexo, consecuentemente influyendo en la fisiopatología de la enfermedad. La determinación de firmas moleculares basadas en la combinación de varios miRNAs permite que éstos puedan ser analizados como biomarcadores, lo que permite una mayor especificidad para el diagnóstico de la enfermedad (Backes *et al.*, 2016).

Actualmente existen unas 2600 secuencias que se corresponden con miRNAs maduros de origen humano anotadas en la miRBase (*miRBase 22.1*, 2018). Además, también está disponible la base de datos que recoge asociaciones entre enfermedades y desregulación de miRNAs en humanos (*HMDD v3.2*, 2019). En su última actualización en 2019, se incluyen 32.281 entradas que describen asociaciones entre miRNAs y enfermedades experimentalmente respaldadas, incluyendo 1102 miRNAs y 850 enfermedades extraídas de 17.412 artículos (Huang *et al.*, 2019). En el caso de la EW, el estudio de miRNAs que permitan un diagnóstico temprano de la enfermedad es necesario, puesto que no se conocen biomarcadores específicos y los marcadores pronósticos utilizados en la actualidad, como el estudio de la ceruloplasmina y de la cuantificación de cobre intrahepático o en orina no son del todo efectivos o suponen someter al paciente a pruebas invasivas (Ryan *et al.*, 2019).

2. HIPÓTESIS Y OBJETIVOS

HIPÓTESIS

- 1- Diagnóstico genético: El análisis genético de *ATP7B* permite establecer un diagnóstico certero en la EW. Esta enfermedad dispone de un tratamiento específico, por lo que un diagnóstico temprano permite evitar el desarrollo de la enfermedad hasta la aparición de daños irreversibles. Algunos pacientes pueden ser diagnosticados con EW en base a la escala de Leipzig, sin embargo, tras el estudio genético completo de *ATP7B*, no se logran identificar mutaciones causantes de la patología. Dichos casos deben ser investigados con estudios genéticos más amplios. La secuenciación de exoma o WES (*Whole-Exome Sequencing*) es relativamente económica y permite un análisis más detallado de pacientes bien caracterizados clínicamente con el fin de identificar los genes implicados en el fenotipo observado. En el caso de identificar variantes candidatas no asociadas previamente a enfermedad, son necesarios estudios funcionales con el fin de caracterizarlas que ayude a alcanzar un diagnóstico definitivo.
- 2- Biomarcadores para la EW: Los miRNAs se encuentran altamente conservados evolutivamente y son liberados en biofluidos como respuesta a procesos fisiológicos. Su desregulación puede estar involucrada en la patogénesis de enfermedades y las alteraciones en su perfil de expresión pueden servir como biomarcadores de evolución de la enfermedad y de pronóstico. En estudios previos del laboratorio, mediante miRNA-seq, se identificaron distintos miRNAs de interés significativamente desregulados en plasma de pacientes con EW. El análisis de estos miRNAs circulantes en el modelo murino *knockout* de la EW (*Atp7b*^{-/-}) aportaría información acerca de la relación entre las diferentes fases de la patología y su desregulación. Ello permitirá la investigación de los miRNAs de interés como biomarcadores que puedan ser utilizados para la monitorización de nuevos tratamientos o el descubrimiento de nuevas dianas terapéuticas.

OBJETIVOS

- 1- Determinar las bases moleculares de la enfermedad en la familia fMD-218 con un fenotipo *Wilson-like* mediante el análisis de WES.
- 2- Caracterización funcional de variantes candidatas a ser causantes de la enfermedad identificadas mediante el análisis de WES y WGS en las familias fMD-218 y fEW-18, respectivamente.
- 3- Detección mediante PCR cuantitativa (qPCR) de miR-146a-5p, miR-340-3p y miR-485-3p en muestra de plasma de ratones *Atp7b*^{-/-} y *wild-type*, y su análisis comparativo.

3. MATERIAL Y MÉTODOS

3.1. Sujetos del estudio

El paciente EW27, probando de la familia fEW-18, y los pacientes MD489 y MD492, probando e hijo probablemente afecto respectivamente de la familia fMD-218 (Figura 3.1), fueron diagnosticados clínicamente con EW por parte de los especialistas de hepatología que los supervisan, al presentar síntomas compatibles con el cuadro clínico habitual: alta concentración de cobre intrahepático ($\geq 250 \mu\text{g/g}$ tejido hepático seco) e hipertransaminasemia persistente.

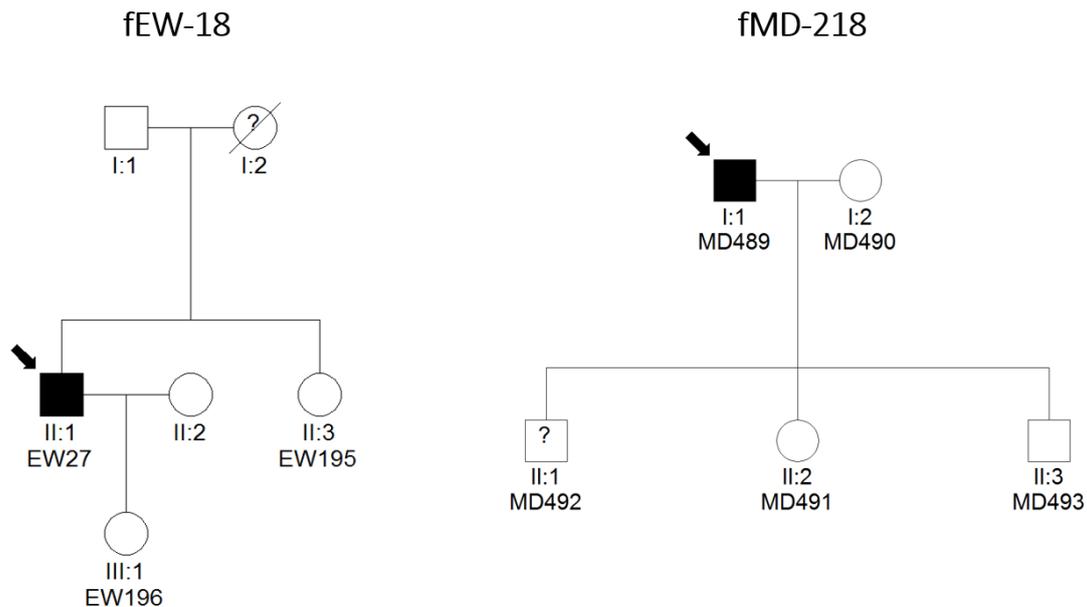


Figura 3.1. Árbol genealógico de las familias fEW-18 y fMD-218. Se indica con una flecha el probando de cada familia.

En el estudio genético de las muestras EW27, MD489 y MD492, mediante secuenciación Sanger de *ATP7B* incluyendo promotor, regiones codificantes e intrónicas flanqueantes, no se detectaron variantes patológicas asociadas a la EW. Tampoco se identificaron grandes duplicaciones/delecciones mediante MLPA (*Multiplex Ligation-dependent Probe Amplification*). De forma complementaria, se descartó que otras regiones de *ATP7B* excluidas del estudio habitual albergaran variantes causantes del fenotipo observado. Para ello, se realizó el análisis completo del gen incluyendo todos los exones, intrones, regiones 5' y 3' UTR así como la región promotora, empleando una herramienta diagnóstica de diseño propio del laboratorio basada en enriquecimiento dirigido y secuenciación masiva (Sánchez-Monteagudo *et al.*, 2020).

Dados los resultados negativos, ambas familias se estudiaron mediante secuenciación de exoma. Esta aproximación ya se incluyó en la publicación anteriormente mencionada Sánchez-Monteagudo *et al.* (2020) para la familia fEW-18, sin obtener resultados positivos. Por este motivo, el análisis continuó con la secuenciación de genoma completo en muestra del probando EW27, además de EW195 y EW196, hermana e hija sanas, respectivamente (Figura 3.1). De entre las variantes

candidatas identificadas, destaca el cambio c.704+2T>G en *INSIG1*, detectado en heterocigosis únicamente en EW27. En este TFM se incluye el estudio funcional de esta variante. Finalmente, en cuanto a la familia fMD-218, en este TFM se incluye la secuenciación de exoma, tal y como se describe a continuación.

3.2. Estudio genético de la familia fMD-218 mediante secuenciación de exoma

Con la finalidad de identificar variantes en otros genes que puedan explicar el fenotipo observado, se llevó a cabo la secuenciación de exoma, a partir de muestra de DNA genómico aislado desde sangre periférica, del probando MD489, así como de su hijo probablemente afecto, MD492, y de su hija sana, MD491 (Figura 3.1). Tanto la obtención de librerías de DNA genómico, el enriquecimiento de regiones exónicas mediante captura y la secuenciación masiva mediante tecnología Illumina, se llevaron a cabo en el Centro Nacional de Análisis Genómico (CNAG). El kit de captura empleado fue *KAPA HyperExome Probes* (Roche).

Los datos de secuenciación de exoma obtenidos por el CNAG fueron depositados en la plataforma *RD-Connect*, en la que está implementado un *pipeline* propio de análisis primario para la identificación de variantes (*variant calling*), a partir del alineamiento de lecturas o *reads* frente al genoma humano de referencia versión 37 (GRCh37) (Laurie *et al.*, 2022). Además, esta plataforma ha integrado herramientas bioinformáticas que facilitan la anotación y filtrado de las variantes registradas, su interpretación, así como el análisis colaborativo entre investigadores de distintas partes del mundo (*RD-Connect Dashboard*, 2021).

3.2.1. Filtrado y priorización de variantes

Para poder estudiar de forma eficiente la gran cantidad de variantes detectadas en la secuenciación de exoma, haciendo uso de la plataforma *RD-Connect* se aplicaron distintas estrategias de filtrado y priorización que se muestran en el esquema de la Figura 3.2. Las variantes resultantes de aplicar cada estrategia de filtrado se recogen en ficheros en formato CSV generados por la propia plataforma, lo que permite continuar su estudio fuera de ella. De forma general, se tuvieron en cuenta todas las variantes con profundidad de lectura mínima de 10x y un valor de calidad de genotipado igual o superior a 30 en escala Phred.

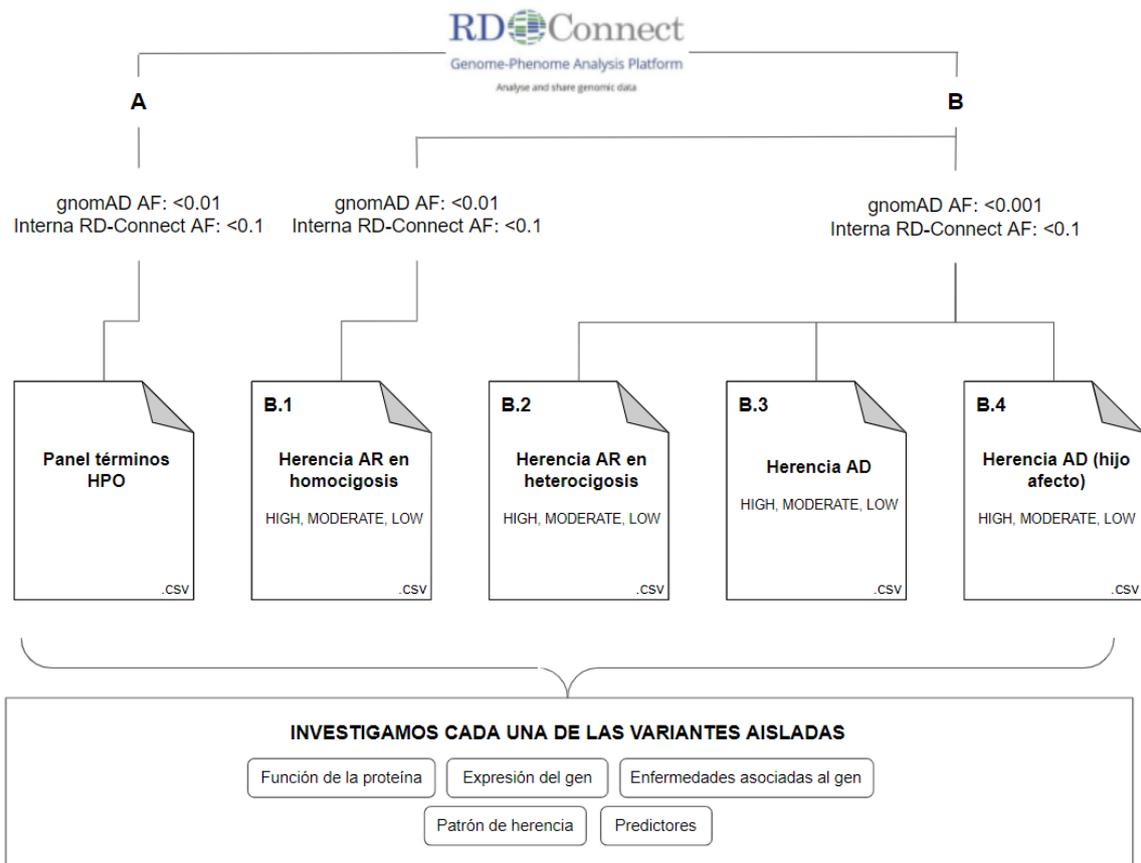


Figura 3.2. Esquema del procedimiento de filtrado y priorización de variantes registradas en secuenciación de exoma de la familia fMD-218. **A)** Filtrado de variantes según términos HPO (*Human Phenotype Ontology*) simulando un panel de genes **B)** Filtrado de variantes aplicando distintas hipótesis de herencia: autosómica recesiva (AR) en homocigosis (B.1) o heterocigosis compuesta (B.2), herencia autosómica dominante (AD) considerando únicamente a MD489 como afecto (B.3) o también a su hijo, MD492 (B.4).

En una primera fase de estudio, se aplicó un filtrado de variantes mediante términos HPO (*Human Phenotype Ontology*) que habían sido previamente asignados al probando en función de los síntomas y signos identificados en la descripción clínica aportada por el especialista que lo supervisa (Figura 3.2A). La plataforma devuelve todas las variantes localizadas en genes asociados a dichos términos HPO, lo que permite simular un panel de genes a investigar confeccionado a partir del fenotipo descrito en el paciente. El listado de genes que se incluyeron en este análisis se recoge en el Anexo I. Se tuvieron en cuenta sólo variantes con una frecuencia alélica menor o MAF (*Minor Allele Frequency*) <1% en la base de datos de frecuencias poblacionales gnomAD versión 2.1.1 y con una frecuencia interna en la plataforma *RD-Connect* <10%. Dado que se trata de genes previamente asociados a patología, se consideraron todos los tipos de cambios posibles anotados en función del efecto que causan sobre la proteína, el procesamiento y/o expresión del mRNA.

En la segunda fase de estudio, se filtraron y priorizaron todas las variantes detectadas en la secuenciación de exoma, según los distintos tipos de herencia de la enfermedad posibles en base

al árbol genealógico (Figura 3.1) y la información disponible de las muestras estudiadas (Figura 3.2B). Los siguientes tres filtrados se realizaron asumiendo como afecto sólo al probando MD489 y a sus hijos MD491 y MD492 sanos: B.1) Considerando una enfermedad autosómica recesiva (AR) en homocigosis; B.2) Considerando una enfermedad AR en heterocigosis compuesta; y B.3) Considerando una enfermedad autosómica dominante (AD). Para el primer tipo de herencia (AR en homocigosis) se examinaron las variantes potencialmente deletéreas por el efecto que causan en la proteína o sobre el procesamiento y/o expresión del mRNA, esto es con un impacto *HIGH* [cambio de pauta de lectura y/o generación de un codón de parada temprano, alteración del sitio de *splicing* 5' (*splicing donor site*) o 3' (*splicing acceptor site*) cercano al exón], *MODERATE* (cambio de aminoácido, inserciones/deleciones en pauta), y *LOW* (cambios sinónimos, alteración de sitios de *splicing* 5'-3' algo más alejados del exón) con un MAF <1% en gnomAD versión 2.1.1 y con una frecuencia interna en *RD-Connect* <10%. En las dos siguientes (AR en heterocigosis compuesta y AD), al obtener un número superior de variantes candidatas, se bajó el valor de MAF a filtrar en gnomAD a 0,1%, dejando el resto de parámetros igual que en el anterior filtrado.

Por último, teniendo en cuenta la información clínica del resto de miembros de la familia fMD-218, se realizó un filtrado considerando también como afecto a MD492, hijo del probando MD489 (Figura 3.2B4). En este filtrado se consideró un tipo de herencia AD, teniendo en cuenta variantes con un impacto *HIGH*, *MODERATE* y *LOW*, con un MAF <0,1% en gnomAD, y frecuencia interna inferior a 10% en *RD-Connect*.

Una vez obtenidos los ficheros correspondientes a las distintas estrategias de filtrado, se priorizaron los cambios identificados con el fin de reducir el listado para determinar variantes candidatas. Para ello, se profundizó en la búsqueda de información en bibliografía y bases de datos como *OMIM* (*Online Mendelian Inheritance in Man*), *PubMed*, *NCBI Gene* o *Human Protein Atlas* (Uhlén *et al.*, 2015), para conocer acerca de la función de la proteína que codifica el gen, si este ha sido previamente asociado a patología y su expresión en los órganos de interés. Para evaluar la posible patogenicidad de variantes de cambio de aminoácido, se utilizaron los predictores MutationTaster (Schwarz *et al.*, 2010), PolyPhen2 (Adzhubei *et al.*, 2010), CADD (Rentzsch *et al.*, 2021, p.), SIFT (Kumar *et al.*, 2009) y PROVEAN (Choi *et al.*, 2012). Para las variantes que pudiesen afectar al *splicing* del mRNA, se utilizaron predictores de alteración de sitios de reconocimiento canónicos (secuencias consenso de los extremos 5' y 3' de exones e intrones) de la maquinaria que participa en este proceso (spliceosoma) o de elementos activadores o silenciadores (ESE y ESS, respectivamente) como son NetGene2 (Brunak *et al.*, 1991), FSplice (<http://www.softberry.com/berry.phtml?topic=fssplice&group=programs&subgroup=gfind>), y NNSplice (Reese *et al.*, 1997).

3.2.2. Validación de variantes candidatas y secuenciación Sanger

Aquellas variantes candidatas como posibles causantes de la patología se validaron mediante amplificación por PCR y secuenciación Sanger en muestra de los individuos de la familia fMD-218 estudiados mediante secuenciación de exoma, y se llevaron a cabo estudios de segregación en las muestras de DNA disponible del resto de familiares (MD490 y MD491, Figura 3.1).

3.2.2.1. Diseño de cebadores y amplificación por PCR

Se diseñaron parejas de cebadores que permitiesen amplificar y secuenciar las regiones que albergan los cambios de interés. Las secuencias de referencia de los transcritos canónicos frente a los que están anotadas las variantes de interés a estudiar se obtuvieron de Ensembl GRCh37. Los cebadores se diseñan a partir de la región a amplificar siguiendo unos criterios establecidos:

- La longitud del amplificado recomendable es de mínimo de 300 pb y máximo 800 pb.
- Longitud de los cebadores entre 18 y 25 nucleótidos, para proporcionar suficiente especificidad.
- El porcentaje de bases GC ha de situarse entre el 40-60%
- La temperatura de *annealing* debe estar en torno a 62-65°C, y esta no debe diferir en más de 2°C entre ambos cebadores.

Además, se comprobó que las secuencias de cebadores seleccionadas no presentasen SNPs frecuentes en la población y que no formasen estructuras secundarias y/o dímeros entre sí. Por último, se confirmó la especificidad de los cebadores, es decir que no hibridasen en otras zonas del genoma humano, mediante PCR *in-silico* con la herramienta *Primer-BLAST* del NCBI (*National Center for Biotechnology Information*).

Para la puesta a punto de las condiciones de amplificación por PCR con cada una de las parejas de cebadores, se emplearon distintos gradientes de temperatura de *annealing* utilizando como molde DNA de individuos control. En algunos casos se recurrió al uso de aditivos como DMSO o MgCl₂ para favorecer la reacción de PCR. Una vez fijadas las condiciones de PCR óptimas para cada región de interés, se pasó a amplificar las mismas a partir de DNA genómico de los individuos objeto de estudio. Las condiciones tanto de las concentraciones de los componentes empleados en la reacción de PCR como de los ciclos y temperaturas del termociclador, se recogen en la Tabla 3.1. Los cebadores utilizados para las validaciones y sus condiciones de amplificación se encuentran en la Tabla 1 del Anexo II.

Tabla 3.1. A) Concentraciones y componentes de la Master Mix empleada en la reacción de PCR. B) Programa de PCR configurado en el termociclador para las ampliificaciones.

A) Master Mix reacción PCR		B) Programa de PCR			
Componentes	Concentración final	Fase	Tª	Tiempo	Ciclos
Cebador 5' (10 µM)	0,2 µM	Desnaturalización inicial	95 °C	5'	1
Cebador 3' (10 µM)	0,2 µM	Desnaturalización	95 °C	30''	35
Mix dNTPs (20 mM)	0,4 mM	<i>Annealing</i>	*	30''	
Tampón polimerasa 10X	1X	Extensión	72 °C	30''	
DNA polimerasa de BIOTOOLS	0,5 U totales	Extensión final	72 °C	7'	1
DNA molde	~ 50 ng	Conservación	4 °C	∞	
Agua libre de DNAsas	Hasta 25 µL de volumen final				

*Las diferentes temperaturas de *annealing* empleadas se encuentran anotadas en las tablas del Anexo II.

Para comprobar que la reacción de PCR ha funcionado correctamente, se testa una pequeña parte del producto obtenido mediante electroforesis en gel de agarosa al 1% con tinción de *GelRed* (Biotium), empleando un patrón de tamaños moleculares (*1 kb Plus DNA Ladder*, Invitrogen) y tampón TAE 1x (30 minutos a 100V). Posteriormente, el gel se revela en un transiluminador con luz UV acoplado a una cámara fotográfica (Biorad).

3.2.2.2. Análisis de secuencias

Los productos de PCR se secuenciaron mediante secuenciación automática basada en el método de Sanger, en un equipo ABI 3730xl (Applied Biosystems) en el Servicio de Genómica del Centro de Investigación Príncipe Felipe (CIPF). Las secuencias obtenidas en formato FASTA se alinearon frente a la secuencia de referencia del amplicón a estudiar, obtenida de Ensembl GRCh37 con ayuda de la herramienta *blast2seq* del NCBI. En paralelo, se revisaron los cromatogramas, en formato ab1, con el programa Chromas para evaluar la calidad de las secuencias e identificar las variantes de interés en las muestras problema tras descartar errores o artefactos.

3.3. Análisis funcional de variantes de *splicing*

3.3.1. Caracterización de variantes de *splicing* mediante ensayo de minigenes

Se investigó mediante ensayo de minigenes la posible implicación de c.704+2T>G en *INSIG1* y c.5003-23T>C en *NOTCH2* en el procesamiento del mRNA haciendo uso del vector reportero pSPL3, cuyas características se recogen en la Figura 3.3.

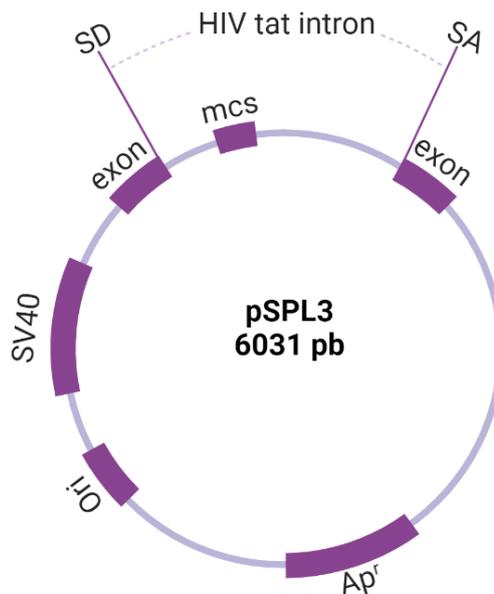


Figura 3.3. Mapa del vector reportero pSPL3. El vector contiene secuencias que permiten su replicación en *E.coli* así como el promotor SV40 para expresión constitutiva en células huésped de mamífero. También presenta el gen de resistencia a ampicilina para ser utilizado como marcador de selección en bacteria.

Se diseñaron cebadores para obtener fragmentos de DNA que incluyeran el exón en el que se localiza la variante y secuencia intrónica flanqueante, siguiendo las indicaciones descritas en el apartado 3.2.2.1. y cuyas secuencias se recogen en la Tabla 2 del Anexo II. La variante c.704+2T>G en *INSIG1* se localiza en el sitio de *splicing* 5' (*splicing donor site*), zona de unión entre exón e intrón 4, consecuentemente se incluyeron en el diseño el exón 4 y secuencia flanqueante a este (625 pb en total) (Figura 3.4A). En cuanto al cambio c.5003-23T>C en *NOTCH2*, este se localiza en secuencia flanqueante al exón 27, por lo que los exones 27 y 28 y secuencia intrónica flanqueante de ambos incluyendo también intrón 27 completo (963 pb en total) se incluyeron en el diseño del amplicón (Figura 3.4B).

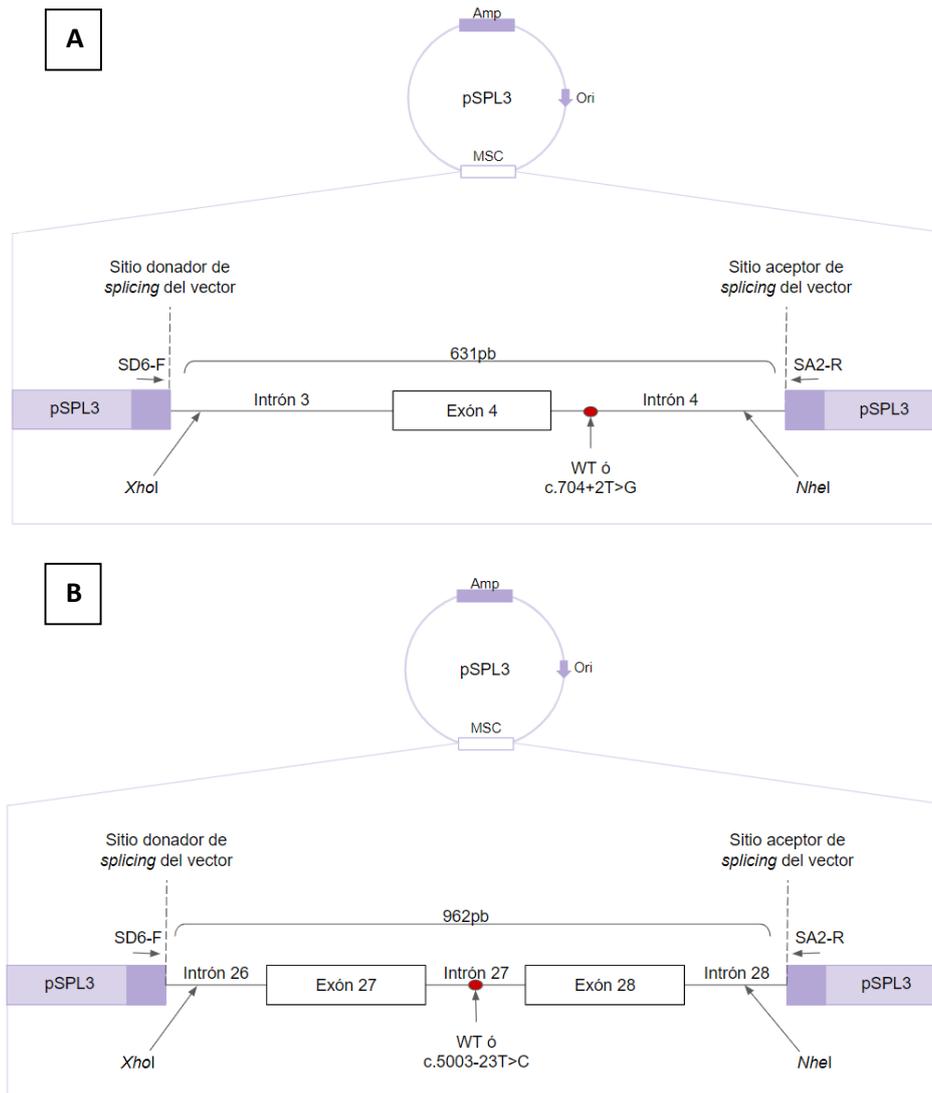


Figura 3.4. Esquema de la construcción de minigenes a partir del vector pSPL3 utilizada en ambos estudios. A) Construcción realizada para la variante c.704+2T>G del gen *INSIG1*. **B)** Construcción realizada para la variante c.5003-23T>C del gen *NOTCH2*. Las cajas moradas representan los exones SD6 y SA2 del vector, las cajas blancas representan los exones a estudiar, que se subclonan en el vector junto con las secuencias intrónicas flanqueantes (representadas como líneas). En cada constructo se indica la mutación de interés además del tamaño del inserto a clonar.

Para la reacción de PCR se empleó como molde el DNA del paciente portador de la variante en heterocigosis y la enzima de alta fidelidad *Pfu Ultra II Fusion HS DNA Polymerase* (Agilent Technologies) siguiendo el protocolo del fabricante. El fragmento obtenido se subclonó en el vector pSPL3 entre los sitios de corte *XhoI* y *NheI*. Para ello, los productos de PCR se purificaron con el kit *NucleoSpin® Gel and PCR Clean-up* (Macherey-Nagel), para posteriormente llevar a cabo la digestión tanto de los insertos como del vector pSPL3 con ambas enzimas *overnight* a 37°C. La ligación se realizó con la enzima *T4 DNA Ligase* (Thermo Scientific) con ratios inserto:vector de 3:1

y 6:1 *overnight* a 4°C. Las ligaciones se transformaron por choque térmico utilizando bacterias competentes de la cepa *DH5α* de *E. Coli*, cultivadas a 37°C toda la noche en placas con medio LB y ampicilina (concentración final de 50 µg/mL). Las colonias aisladas obtenidas fueron testadas mediante PCR para comprobar la presencia de las distintas construcciones, confirmándose mediante secuenciación Sanger los respectivos cambios además de la correcta direccionalidad del inserto en el vector, tal como se describe en el apartado 3.2.2.2.

Los clones correspondientes a las construcciones *wild-type* y mutante se cultivaron en medio LB líquido con ampicilina (concentración final de 50 µg/mL) durante 16 horas para proceder a la extracción de los plásmidos con el kit *NucleoSpin® Plasmid* (Macherey-Nagel). Se sembraron células HeLa en placas de 6 pocillos (80.000-100.000 células/pocillo) en medio *Dulbecco's Modified Eagle's Medium* (DMEM) *High Glucose* (Sigma-Aldrich) suplementado con 10 % (v/v) de FBS inactivado, 2 mM de glutamina, penicilina/estreptomicina al 1 % (100 U.I./mL de penicilina, y 100 µg/mL de estreptomicina), e incubadas en un incubador humidificado a 37 °C y 5 % CO₂. Las células se transfectaron al día siguiente con las construcciones de interés, así como vector pSPL3 vacío, con *FuGENE HD* (Promega) y ratio de transfección 5:2 (µL FuGENE: µg DNA), siguiendo el protocolo del fabricante.

A 24 horas post-transfección se recogieron las células para la extracción de RNA total con *TRizol® Reagent* (Life Technologies). Mediante transcripción reversa se sintetizó el cDNA correspondiente, empleando 300 ng de RNA total como molde, con el kit *qScript cDNA SuperMix* (Quantabio). Se amplificaron por PCR los transcritos de interés con los cebadores SD6 y SA2, específicos del vector pSPL3 (Tabla 3 del Anexo II). Los productos obtenidos se testaron mediante electroforesis en gel de agarosa al 1,5% para posteriormente comprobar la secuencia, tal como se describe en el apartado 3.2.2.2.

3.3.2. Caracterización de variantes de *splicing* en mRNA de sangre periférica

Tanto la variante c.704+2T>G en *INSIG1* como c.5003-23T>C en *NOTCH2*, se analizaron en mRNA a partir de muestra de sangre periférica recuperada en tubos *PAXgene* de EW27 y MD489, pacientes portadores de los cambios. Para ello, en primer lugar, se aisló RNA total de sangre periférica con el kit *PAXgene Blood RNA* (Qiagen). Mediante transcripción reversa se sintetizó cDNA, empleando 300 ng de RNA total como molde con el kit *qScript cDNA SuperMix* (Quantabio). Posteriormente, se analizó el patrón de *splicing* generado por ambas variantes mediante amplificación por PCR con cebadores localizados en exones contiguos y flanqueantes a la variante de interés en el mRNA. Los cebadores, diseñados siguiendo los criterios indicados en el apartado 3.2.2.1, se recogen en la Tabla 4 del Anexo II. Los productos de PCR se testaron mediante electroforesis en gel de agarosa al 2%, y se analizó su secuencia según se ha descrito anteriormente en el apartado 3.2.2.2. Para determinar el patrón de *splicing* obtenido en el análisis del cambio

c.704+2T>G en *INSIG1*, fue necesario subclonar los productos de PCR correspondientes a control y paciente en el vector pcDNA3 (Figura 1 del Anexo III). Para ello, se introdujeron los sitios de corte *HindIII* y *BamHI* en los cebadores utilizados anteriormente para la PCR desde cDNA (Tabla 5 del Anexo II) siguiendo a continuación la misma metodología ya descrita para la obtención de construcciones de minigenes en el apartado 3.3.1. Se realizó un cribado por PCR y posterior secuenciación Sanger de colonias aisladas, tal como se describe en el apartado 3.2.2.2, con el fin de determinar las secuencias correspondientes a las distintas variantes de *splicing* alternativo.

3.4. Estudio de miRNA circulantes en plasma de ratón

Con la finalidad de estudiar miRNA circulantes como posibles biomarcadores en la EW, se llevó a cabo un estudio en 10 muestras de plasma de ratones: cinco de ellas proceden de ratones *knockout* para *Atp7b* de ocho semanas de edad (*Atp7b*^{-/-}), y los otros cinco de ratones *wild-type* de 21 semanas de edad, todos ellos con fondo genético puro de C57BL/6. Dichas muestras fueron proporcionadas por el laboratorio de la Dra. Gloria González-Aseguinolaza, del Centro de Investigación Médica Aplicada (CIMA, Pamplona).

3.4.1. Extracción de RNA total circulante

Se procedió con la extracción de RNA total circulante a partir del plasma de los ratones con el *miRNeasy Serum/Plasma Kit* (QIAGEN). La extracción se realizó a partir de 50 µL de plasma de los ratones. El procedimiento se llevó a cabo siguiendo las instrucciones del fabricante del kit, a excepción de que no se añadió el *miRNeasy Serum/Plasma Spike-In control* (cel-miR-39 de *C.elegans*). Finalmente, el RNA aislado se eluyó en 12 µL de agua libre de RNasas y se almacenó a -80°C hasta el momento de su análisis.

3.4.2. Transcripción reversa a cDNA

Para obtener el cDNA a partir del RNA aislado anteriormente se utilizó el *TaqMan Advanced miRNA Assay kit* (Applied Biosystems). Este kit permite, a partir una pequeña cantidad de RNA circulante total (2 µL) aislado de muestras de suero o plasma, detectar un gran número de miRNAs maduros. Este procedimiento se realizó siguiendo los pasos descritos por el fabricante, obteniendo un volumen final de 50 µL de cDNA pre-amplificado que se almacenó a -20°C hasta su posterior uso.

3.4.3. PCR cuantitativa (qPCR) de los miRNA de interés

El estudio de los diferentes miRNAs de interés junto con los de referencia se llevó a cabo mediante amplificación por qPCR. Para ello se utilizaron sondas específicas para cada miRNA *TaqMan Advanced miRNA Assay* (Applied Biosystems) y *TaqMan Fast Advanced MasterMix* (Applied Biosystems). Los experimentos se realizaron en placas de 384 pocillos, y para cada reacción se utilizaron 7,5 µL de mezcla de reacción y 2,5 µL de cDNA en una dilución 1/10 con tampón TE 0,1X (como indica el protocolo del fabricante). Por cada muestra y miRNA a estudiar se realizaron 3 réplicas técnicas.

El programa llevado a cabo para el instrumento *LightCycler 480 II* (Roche) consistía en un paso de pre-incubación a 95°C durante 10 minutos, seguido por 50 ciclos de amplificación que podemos desglosar en 10 segundos de desnaturalización a 95°C y un paso de 30 segundos a 37°C en el que se produce la hibridación de la sonda y los cebadores, así como la elongación, durante la que se mide la emisión de fluorescencia que va a permitir detectar el producto de reacción. Finalmente, hay un paso de enfriamiento a 37°C durante 30 segundos.

Con el procedimiento descrito se analizaron 3 miRNAs en plasma de ratón: miR146a-5p, miR340-3p y miR485-3p, utilizando miR-191-5p y miR-16-5p como miRNAs de referencia endógenos para la normalización. La información de las sondas empleadas en el estudio se detalla en la Tabla 3.2.

Tabla 3.2. Sondas *TaqMan Advanced* utilizadas para la amplificación de miRNAs circulantes mediante qPCR.

miRNA	Referencia	Secuencia
hsa-miR146a-5p	478399_mir	UGAGAACUGAAUCCAUGGGUU
hsa-miR340-3p	478041_mir	UCCGUCUCAGUUACUUUAUAGC
hsa-miR485-3p	478125_mir	GUCAUACACGGCUCUCCUCUCU
hsa-miR191-5p	477952_mir	CAACGGAAUCCCAAAGCAGCUG
hsa-miR16-5p	477860_mir	UAGCAGCACGUAAAUAUUGGCG

3.4.4. Análisis estadístico de los datos de qPCR

Los resultados se obtuvieron en un fichero con formato ixi y fueron analizados con el programa *LightCycler® 480 Software, Version 1.5* (Roche) con el método de cuantificación basado en el primer máximo de la segunda derivada de la curva. En nuestro caso se seleccionó el miR16-5p de referencia para la normalización del Ct (umbral de ciclo) obtenido en cada reacción, ya que presentó un comportamiento más estable entre ambos grupos en comparación con el miR-191-5p de referencia.

Para cada miRNA de interés se calculó el nivel de expresión relativo en muestra de ratones *Atp7b*^{-/-} con respecto a *wild-type* como $2^{-(\Delta Ct \text{ Atp7b}^{-/-} - \text{Media } \Delta Ct \text{ wild-type})}$, siendo

$\Delta Ct = Ct_{miRNA\ de\ interés} - Ct_{miRNA\ referencia}$, representándose gráficamente como logaritmo en base 2 (Log_2). Los tres conjuntos de datos pasaron las pruebas de normalidad y homogeneidad, por lo que se les realizó una prueba t de Student para comparar los grupos *wild-type* con los *Atp7b^{-/-}*. Los análisis estadísticos y los histogramas se realizaron usando GraphPad Prism 9.3.1

4. RESULTADOS

4.1. Variantes candidatas identificadas en la familia fMD-218

Con la finalidad de encontrar la posible causa genética del fenotipo clínico observado en el paciente MD489, se estudiaron mediante WES las muestras de tres individuos de la familia fMD-218 (MD489, MD491 y MD492). Aplicando los distintos criterios de filtrado y priorización en el análisis de los datos de secuenciación de exoma mostrados en el apartado 3.1, se lograron identificar 7 cambios en 4 genes (Tabla 4.1). Estos cambios se validaron mediante secuenciación Sanger como se ha descrito en Material y Métodos y se realizó la segregación familiar, cuyos resultados podemos ver en la Figura 1 del Anexo V.

Tabla 4.1. Variantes candidatas priorizadas en el WES de la familia fMD-218.

Gen	Individuo	Genotipo	HGVS cDNA	HGVS proteína	Frecuencia en gnomAD (rs)	Predictores <i>in silico</i>		
						SIFT	Polyphen2	CADD
NOTCH2	MD489	0/1	c.5003-23T>C	-	1/230.582 (rs926669050)	-	-	5,745
	MD491	0/0						
	MD492	0/1						
CARD10	MD489	0/1	c.478C>T	p.Arg160Trp	3/241.322 (rs919323160)	D	PD	25
	MD491	0/1						
	MD492	0/0						
	MD489	0/1	c.2262T>A	p.Thr754=	-	T	-	2,634
	MD491	0/0						
	MD492	0/1						
CBFA2T3	MD489	0/1	c.232C>G	p.Pro78Ala	2/191.878 (rs141249194)	T	PD	23,3
	MD491	0/0						
	MD492	0/1						
	MD489	0/1	c.1381G>C	p.Gly461Arg	3/217.960 (rs772795593)	D	B	13,55
	MD491	0/0						
	MD492	0/1						
COG5	MD489	0/1	c.*2089A>T	-	29/280.922 (rs768129883)	-	-	23,9
	MD491	0/1						
	MD492	0/1						
	MD489	0/1	c.*2090A>G	-	30/280.888 (rs776238759)	-	-	24,6
	MD491	0/1						
	MD492	0/1						

HGVS: Human Genome Variation Society; cDNA: DNA complementario; rs: identificador dbSNP; UTR: untranslated region; D: Deletéreo; T: Tolerado; PD: Probablemente deletéreo; B: Benigno

- **NOTCH2 (c.5003-23T>C)**: NOTCH2 es una proteína transmembrana que forma parte de la red de señalización Notch, una vía de señalización que regula la expresión génica, pudiendo desempeñar un papel fundamental en el desarrollo vascular, renal y hepático durante la embriogénesis (*NOTCH2 notch receptor 2 [Homo sapiens (human)] - Gene - NCBI, s. f.*). Mutaciones en el gen *NOTCH2* pueden dar lugar a enfermedades como el síndrome de Alagille (ALGS; MIM:118450) o el síndrome de Hajdu-Cheney (HJCYS; MIM:102500). El cambio

identificado en el paciente es una variante intrónica, a 23 pb del sitio aceptor de *splicing* del exón 27 y tiene un MAF <0,01. Los predictores *in silico* utilizados no mostraron pérdidas ni ganancias de sitios de *splicing* 3' (*splicing acceptor site*) o 5' (*splicing donor site*) (Anexo VI.B). No obstante, al tratarse de una variante de muy baja frecuencia en la población en un gen asociado a enfermedad *Wilson-like* y observar en la segregación familiar que solo son portadores de la variante el paciente (MD489) y el hijo posiblemente afecto (MD492), se procedió a realizar un análisis funcional (descrito en el apartado 3.2.1).

- **CARD10 (c.478C>T y c.2262T>A):** CARD10 es una proteína que pertenece a la familia MAGUK (guanilato quinasa asociada a membrana). Esta forma el complejo CBM junto con linfoma de células B 10 (*BCL10*) y *MALT1* (D. Yang *et al.*, 2020). En el análisis de *CARD10* se han identificado dos variantes, una sinónima novel (c.2262T>A; p.Thr754=) y otra no sinónima con una MAF <0,01 (c.478C>T; p.Arg160Trp). La variante p.Thr754= se clasificó como tolerada por el predictor *in silico* de patogenicidad SIFT y con un valor para el predictor CADD de 2,634 (considerado no patológico). En cambio, la variante de cambio de aminoácido p.Arg160Trp se clasificó como deletérea por SIFT, probablemente deletérea por Polyphen2 y con un valor en CADD de 25 (considerado patológico). Estudios de validación y segregación familiar confirmaron la presencia de ambas mutaciones en muestra del caso índice (MD489), así como la presencia de la variante sinónima en muestra de ambos hijos (MD492 y MD493) y la variante de cambio de aminoácido únicamente en muestra de la hija (MD491).
- **CBFA2T3 (c.232C>G y c.1381G>C):** CBFA2T3 (también conocido como MTG16) es un correpresor transcripcional que actúa mediante la asociación con factores de unión al DNA y el reclutamiento de correpresores adicionales y enzimas modificadoras de histonas (Rossetti *et al.*, 2004). En el estudio de este gen se han identificado dos variantes con un MAF <0,01 y ambas de cambio de aminoácido. La variante p.Pro78Ala (c.232C>G) se clasificó como tolerable por el predictor SIFT, pero como probablemente deletérea por Polyphen2 y con un valor de CADD de 23,3 (considerado patológico). Por otra parte, la variante p.Gly461Arg (c.1381G>C) se clasificó como deletérea por SIFT, en cambio tuvo un resultado de benigna por el predictor *in silico* Polyphen2 y una puntuación de 13,55 en CADD (considerado patológico con baja probabilidad). Estudios de validación y segregación familiar de ambos cambios confirmó la presencia de las variantes en muestra del paciente (MD489) así como en muestra de los hijos (MD492 y MD493), no detectándose ninguna de las mutaciones en muestra de su cónyuge e hija (MD490 y MD491). Ello indicaría que ambas variantes se encuentran en el mismo alelo (en *cis*).
- **COG5 (c.*2089A>T y c.*2090A>G):** COG5 es una de las ocho proteínas que forman parte del complejo oligomérico de Golgi (COG), indispensable para una correcta morfología y funcionamiento del aparato de Golgi. En este gen se han validado dos cambios, c.*2089A>T y c.*2090A>G, que se localizan en la región 3' UTR, ambos con una puntuación de CADD elevada

(23,9 y 24,6 respectivamente). Estudios de validación y segregación han confirmado la presencia de ambas variantes tanto en muestra del caso índice (MD489) como en muestra de dos de sus hijos (MD491 y MD492), estando estos ausentes en muestra de su cónyuge y del otro hijo (MD490 y MD493), lo que confirmaría que ambos cambios se encuentran en *cis*.

4.2. Análisis funcional de variantes de *splicing*

4.2.1. Variante de *splicing* c.704+2T>G en *INSIG1*

Se estudió la variante c.704+2T>G en el gen *INSIG1* que podría afectar al proceso de *splicing*. Previo a los estudios funcionales de la variante, se realizaron análisis computacionales para evaluar las posibles consecuencias de ésta. Los resultados de los distintos predictores *in silico* de *splicing* consultados (Anexo VI.A) determinaron una pérdida del sitio de *splicing* 5' (*splicing donor site*) del exón 4 en presencia de la variante. Para corroborar estas predicciones, se llevó a cabo el análisis del mRNA en sangre periférica del paciente (EW27) portador en heterocigosis de la mutación. Además, se realizó un estudio mediante ensayo de minigenes.

El análisis del mRNA mostró un patrón de *splicing* alternativo en presencia de la variante. Al testar la PCR en el gel de agarosa, solo se observó la presencia de una banda de 405 pb. Tras secuenciar el producto de PCR se determinó que había más de un producto (Figura 4.1.B); estos eran de tamaño similar, por lo que no pudieron resolverse fácilmente en la electroforesis. Con la finalidad de distinguir los diversos productos obtenidos en la amplificación por PCR, estos fueron clonados en un vector. En el cribado de las colonias obtenidas mediante amplificación por PCR de colonias y secuenciación Sanger, se identificaron colonias con la secuencia de mRNA de referencia, que contienen los exones 4 y 5 seguidos (Figura 4.1.C), y colonias con el producto de *splicing* alternativo como consecuencia de la mutación c.704+2T>G, correspondiéndose con una secuencia de mRNA que contiene 4 pares de bases (pb) más que la secuencia de referencia (Figura 4.1.D). En esta se ha perdido el sitio de *splicing* 5' canónico como anticipaban los predictores, pero el espliceosoma sería capaz de detectar un nuevo sitio a 4 pb de distancia en dirección 3' de la secuencia consenso canónica, dando lugar a un producto que difiere de la secuencia de mRNA de referencia en la adición de 4 bases intrónicas entre los exones 4 y 5. Para comprobar el efecto que causaría a nivel de proteína, se analizó cómo se modificaría la traducción del mRNA en presencia de esta variante. Para ello, se tradujo la secuencia de mRNA correspondiente al producto de *splicing* alternativo a secuencia de aminoácidos. La secuencia de la proteína mutante diferiría en la región C-terminal, ya que esta inserción de 4 pb conduciría a un desplazamiento de la pauta de lectura y a la generación de un codón de parada prematuro, como se indica en la Figura 4.1.E.

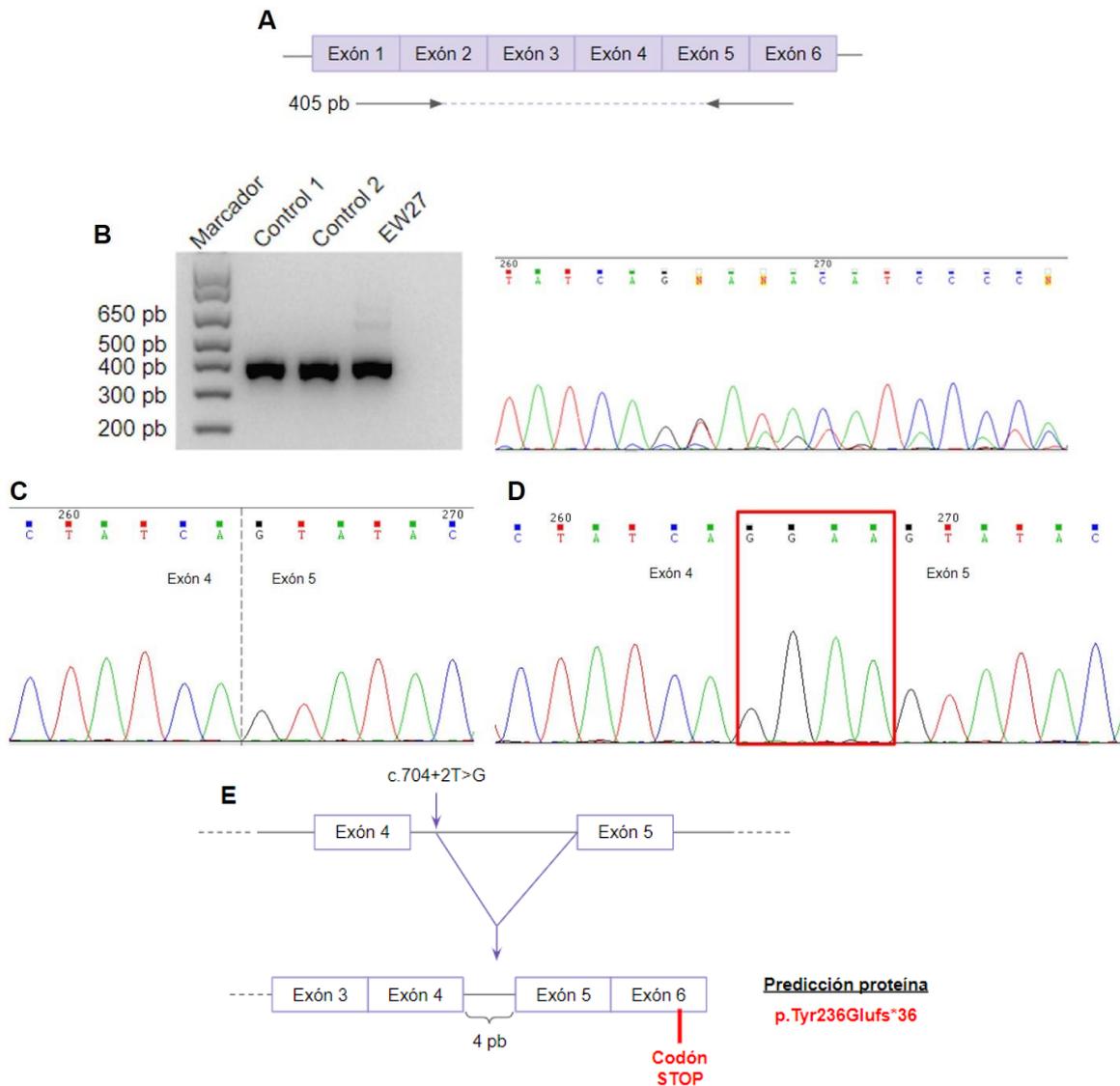


Figura 4.1. Resultados del análisis del mRNA en sangre periférica de la variante de *splicing* c.704+2T>G en *INSIG1*. **A)** Diseño de cebadores para la amplificación del mRNA de *INSIG1* **B)** Productos de la amplificación por RT-PCR del mRNA aislado desde muestra de sangre periférica del paciente EW27, portador del cambio en heterocigosis, y dos controles junto con el electroferograma de la secuenciación Sanger del producto de PCR obtenido desde muestra de EW27. **C)** Electroferograma de la secuenciación Sanger del producto de PCR de colonias con la secuencia del mRNA de referencia. **D)** Electroferograma de la secuenciación Sanger del producto de PCR de colonias con la secuencia del mRNA correspondiente al *splicing* alternativo en presencia de la mutación. **E)** Predicción del efecto que causaría la mutación a nivel del procesado del mRNA y de la proteína.

Por otra parte, se estudió el efecto de la variante c.704+2T>G en *INSIG1* mediante ensayo *in vitro* de minigenes. En este estudio se utilizaron células Hela como huésped para la sobreexpresión de las construcciones de minigenes, *wild-type* y mutante, basadas en la utilización del vector reportero pSPL3. Los resultados de este ensayo, recogidos en la Figura 4.2, mostraron que tanto la

construcción *wild-type* como la mutante eran procesadas de igual forma, dando lugar a transcritos con una pérdida del exón 4 en ambos casos.

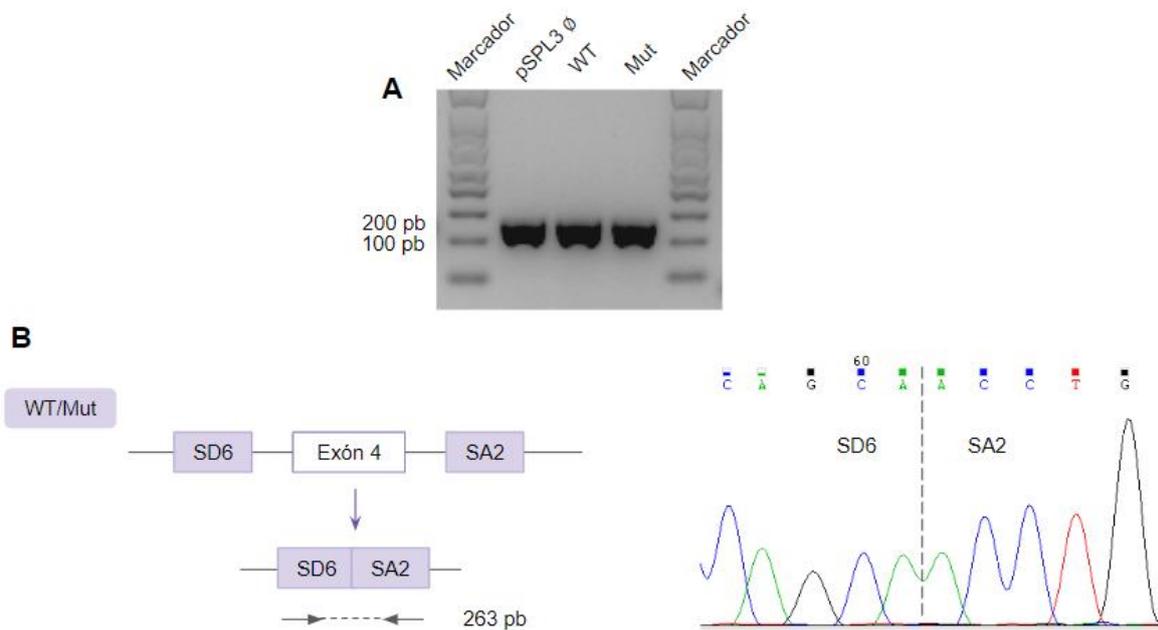


Figura 4.2. Resultados del ensayo de minigenes de la variante de splicing c.704+2T>G en *INSIG1*. A) Productos de la amplificación por PCR con los cebadores SD6 y SA2 de los transcritos generados tras la sobreexpresión de las construcciones de minigenes pSPL3_Vacío, pSPL3_INSIG1_WT y pSPL3_INSIG1_Mut a las 24 horas post-transfección en células HeLa. B) Esquema del procesamiento del mRNA de las construcciones de minigenes y electroferogramas de la secuenciación mediante Sanger de las bandas detectadas en A). WT: *wild-type*; Mut: *mutante*; Ø: *vacío*

4.2.2. Variante de *splicing* c.5003-23T>C en *NOTCH2*

Se estudió la variante c.5003-23T>C en el gen *NOTCH2* que podría afectar al proceso de *splicing*. Previo a los estudios funcionales de la variante, al igual que para el análisis de la variante en *INSIG1*, se realizaron análisis computacionales para evaluar las posibles consecuencias de esta. Los resultados de los distintos predictores *in silico* de *splicing* (Anexo VI.B) no determinaron ninguna pérdida ni ganancia de sitios de *splicing* 5' o 3' en presencia de la variante. Independientemente de ello, se llevó a cabo el análisis del mRNA a partir de muestra de sangre periférica del paciente MD489, portador en heterocigosis de la mutación, para investigar una posible alteración del procesamiento del mRNA en presencia de la variante que no hubiera podido ser predicha por los programas consultados. En el análisis del mRNA no se detectaron alteraciones del patrón de *splicing* en presencia del cambio c.5003-23T>C (Figura 4.3). En la secuenciación de los productos de la amplificación por PCR identificados en la electroforesis en gel de agarosa (Figura 4.3.A), se confirmó que todos ellos se corresponden con la secuencia de mRNA de referencia para los exones 27 y 28 de *NOTCH2* (Figura 4.3.B).

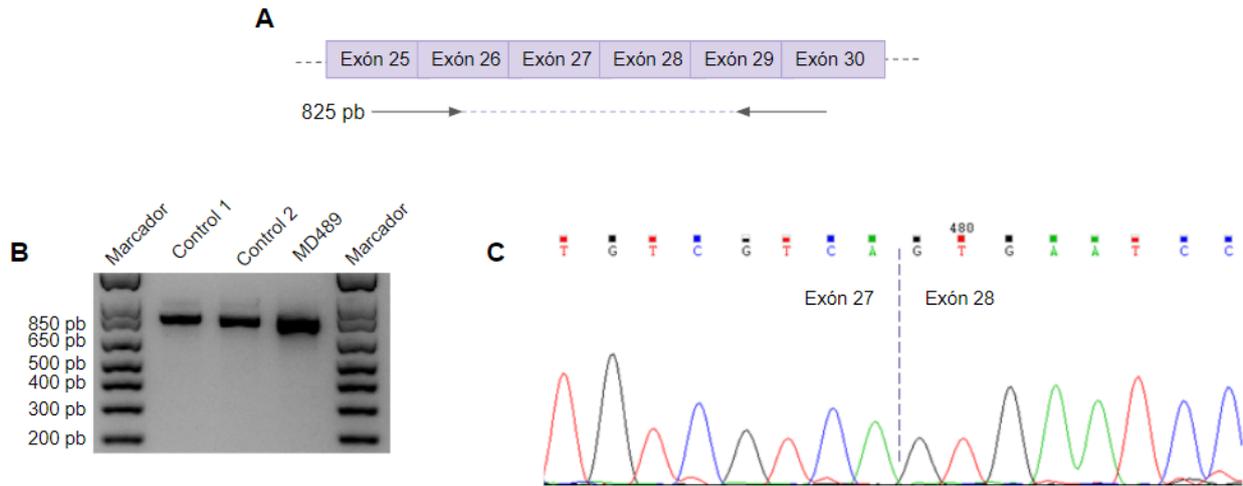


Figura 4.3. Resultados del análisis del mRNA en sangre periférica de la variante de *splicing* c.5003-23T>C en *NOTCH2* A) Diseño de cebadores para la amplificación del mRNA de *NOTCH2*. B) Productos de la amplificación por RT-PCR del mRNA aislado de muestra de sangre periférica del paciente MD489, portador del cambio en heterocigosis, y dos individuos control. C) Electroferograma de la secuenciación Sanger del producto de RT-PCR del paciente MD489. En la secuenciación de los productos de PCR obtenidos desde muestra de individuos control se obtuvo el mismo resultado (no mostrado). WT: *wild-type*; Mut: *mutante*; Ø: *vacío*

Tras el resultado obtenido en el análisis del mRNA de *NOTCH2* en sangre periférica y teniendo en cuenta que mutaciones patológicas en este gen se asocian al síndrome de Alagille, se decidió ahondar en el estudio de la variante c.5003-23T>C mediante ensayo de minigenes. Este no reveló diferencias en el patrón de *splicing* correspondiente a los transcritos generados tras la sobreexpresión y procesamiento de las construcciones *wild-type* y mutante, derivadas de ambos alelos para los que el paciente MD489 es portador, en células HeLa (Figura 4.4.A). Tanto el transcrito que concierne al procesamiento de la construcción *wild-type* como el transcrito producto del procesamiento de la construcción contienen los exones 27 y 28 consecutivos, como se aprecia en la secuenciación de los productos de PCR obtenidos (Figura 4.4B). Consecuentemente, no se pudo demostrar que la variante pudiera estar implicada en la alteración del proceso de *splicing*.

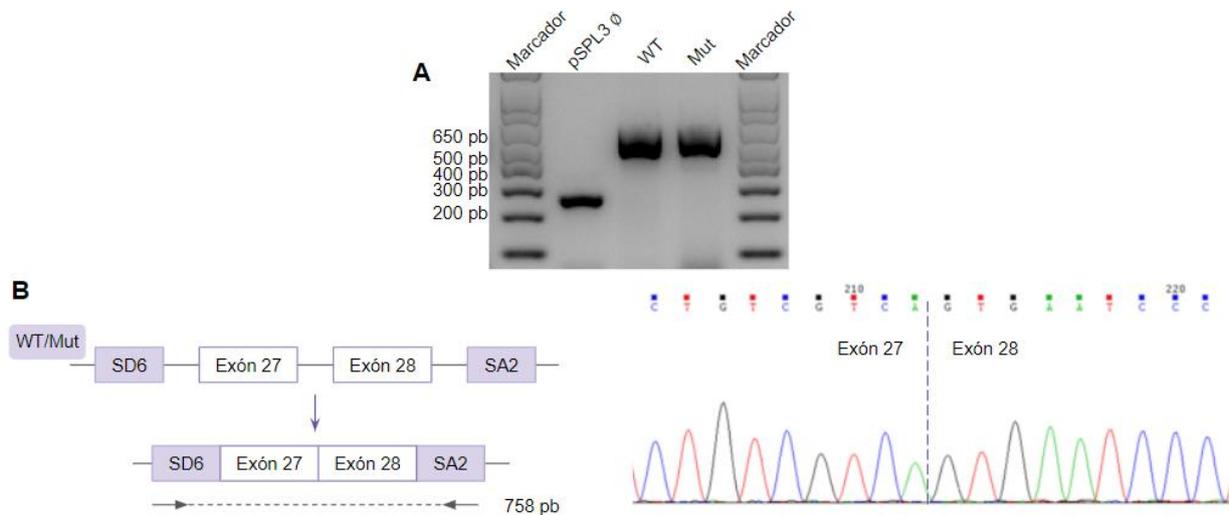


Figura 4.4. Resultados del ensayo de minigenes de la variante de *splicing* c.5003-23T>C en *NOTCH2*. A) Productos de la amplificación por PCR con los cebadores SD6 y SA2 de los transcritos generados tras la expresión de las construcciones de minigenes pSPL3_Vacío, pSPL3_NOTCH2_WT y pSPL3_NOTCH2_Mut a las 24 horas post-transfección en células HeLa. B) Esquema del procesamiento del mRNA de las construcciones de minigenes y electroferogramas de la secuenciación mediante Sanger de las bandas detectadas en A) WT: *Wild-type*; Mut: *mutante*; ∅: *vacío*.

4.3. Estudio de miRNAs circulantes en plasma como biomarcadores de la EW

Se realizó el análisis mediante PCR cuantitativa (qPCR) de tres miRNAs, miR-146a-5p, miR-340-3p y miR485-3p, en muestra de plasma de cinco ratones *knockout* (KO) para *Atp7b* (*Atp7b*^{-/-}) y otros cinco ratones *wild-type*.

Como se muestra en las gráficas de la Figura 4.5, en dos de los tres miRNAs de interés, miR-146a-5p y miR-340-3p, no se identificaron diferencias significativas entre los grupos KO y *wild-type*. Los p-valores obtenidos en el test estadístico aplicado fueron 0,0738 para miR-146a-5p y 0,0672 para miR340-3p. En cambio, para miR-485-3p se observó un incremento significativo en el grupo KO con respecto al *wild-type*, obteniéndose un p-valor de 0,0014.

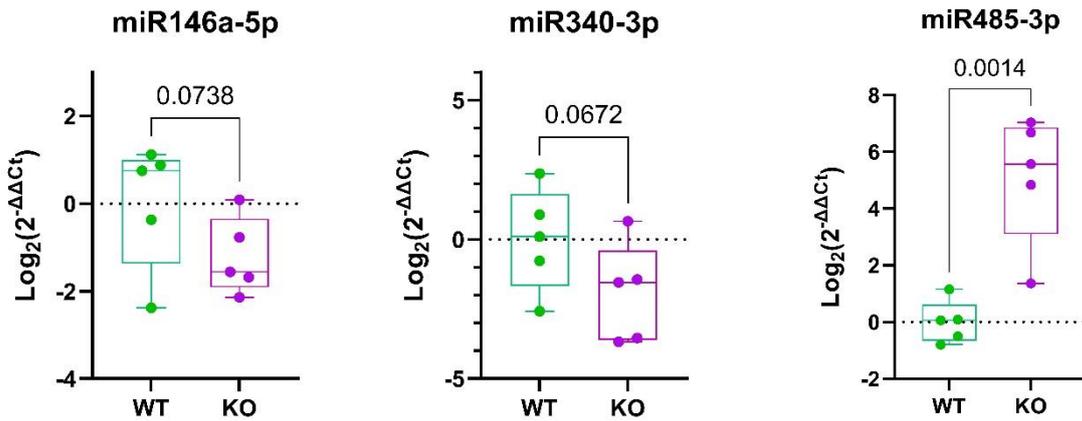


Figura 4.5. Análisis comparativo entre grupos de los niveles de miR146a-5p, miR340-3p y miR485-3p mediante qPCR. Se representa el rango intercuartílico del log₂ de la expresión relativa (2^{-ΔΔCt}) de las muestras de cada grupo (cinco ratones *wild-type*, en verde, y cinco ratones KO, en morado). Para cada miRNA, se indica el p-valor obtenido en el test estadístico aplicado (prueba t de Student con dos colas para muestras de distinta varianza).

5. DISCUSIÓN

5.1. Diagnóstico genético

La EW es una enfermedad genética rara caracterizada por la acumulación de cobre en varios tejidos del cuerpo, principalmente en el hígado y el cerebro. El diagnóstico temprano junto con el tratamiento adecuado son capaces de prevenir complicaciones potencialmente mortales. El diagnóstico genético es definitivo e implica el análisis de *ATP7B*, ya que se trata de un trastorno monogénico causado por mutaciones en dicho gen. No obstante, otras enfermedades hereditarias presentan un fenotipo solapante con la EW y pueden dar lugar a un diagnóstico clínico inicial erróneo. Las enfermedades *Wilson-like* cursan con una sintomatología similar a la de los pacientes con EW. Un ejemplo de esto ocurre con los pacientes que padecen algunos trastornos congénitos de glicosilación y que cursan, entre otros, con signos neurológicos, niveles séricos elevados de alanina aminotransferasa y alteraciones del metabolismo del cobre que resultan en su acumulación en el hígado, dando lugar a un diagnóstico equivoco de la EW (Sánchez-Monteagudo *et al.*, 2021). Estas confusiones en el diagnóstico que reciben inicialmente los pacientes en base a las características clínicas de la enfermedad, también las podemos ver en otras enfermedades como en el caso de los síndromes de neurodegeneración con acumulación cerebral de hierro (NBIA). Hay diez formas NBIA aceptadas causadas por mutaciones en diez genes distintos, pero se han descrito casos que cursan con fenotipos anormales de hierro cerebral compatibles con NBIA que realmente corresponden a otras enfermedades mendelianas (Hinarejos *et al.*, 2020).

Por tanto, si no se identifican mutaciones en *ATP7B* que confirmen el diagnóstico de la EW, es necesario llevar a cabo pruebas genéticas más exhaustivas utilizando herramientas NGS como la secuenciación de exoma o WES (*Whole-exome sequencing*) con el fin de identificar las mutaciones implicadas en el cuadro clínico observado.

Actualmente, más del 50% de los pacientes con una enfermedad rara genética permanecen sin diagnosticar (Weiss *et al.*, 2018). Una limitación es la anotación del genoma que continuamente cambia. El reanálisis de casos de WES no concluyentes (análisis dinámico), pasado un tiempo (al menos un año), aumenta el ratio diagnóstico en un 10% (Boycott *et al.*, 2019). Esto se debe a que en el momento en que se analizaron, no había evidencia suficiente para confirmar la patogenicidad de la variante candidata o la causalidad del gen, pero con la mejora en la anotación del genoma, al volver a analizar los datos se podría establecer un diagnóstico genético certero.

Entre las dificultades que impiden mejorar el éxito diagnóstico está el posterior análisis de los datos crudos de herramientas NGS, principalmente la interpretación de la gran cantidad de variantes genéticas cuyo significado patológico no es obvio. En este sentido, hay iniciativas tanto nacionales como internacionales, que fomentan la creación de plataformas especializadas en estudios funcionales que permitan investigar el carácter deletéreo de las variantes novedades. Así, la convocatoria de EJP (*European Joint Programme*) en enfermedades raras de este mismo año 2022, tenía entre las líneas prioritarias: “*Functional strategies to globally stratify variants of unknown*

significance (VUS) for clinical use; setting up of (in vitro) systems to distinguish between VUS and pathogenic variants” (Joint Transnational Call 2022 – EJP RD – European Joint Programme on Rare Diseases (ejprarediseases.org)). Adicionalmente, las técnicas NGS empleadas tienen limitaciones metodológicas, tales como la incapacidad para detectar grandes cambios estructurales o repeticiones en tándem características de las mutaciones dinámicas. De hecho, el desarrollo de la secuenciación de tracto largo (*long read sequencing*) ha comenzado a solventar, al menos parcialmente, este desafío, y se han caracterizado las bases moleculares de muy diversas patologías (Mitsuhashi and Matsumoto, 2020). Por ejemplo, la mutación implicada en una demencia autosómica dominante ligada a 7q36 se resolvió mediante secuenciación de tracto largo al descubrir que la alteración responsable es una inversión cromosómica de 4-Mb que interrumpe la secuencia de *DPP6* (Cacace *et al.*, 2019).

Sin embargo, el diagnóstico tiene otros retos a los que enfrentarse. Una enfermedad mendeliana es el resultado de la mutación causativa, y también, del fondo genético que acompaña a esta, de la epigenética y de los hábitos. Dos pacientes con las mismas mutaciones deletéreas que causan una enfermedad mendeliana dada pueden tener clínica muy dispar porque esta está influenciada por los factores mencionados. Así, un paciente con la enfermedad neuromuscular de Charcot-Marie-Tooth tipo 2K (CMT2K), causada por mutaciones en *GDAP1* heredadas de forma autosómica dominante, presentará un fenotipo más grave si, además, es portador de variantes en *JPH1*, cuando *JPH1* por sí solo no está implicado en el desarrollo de ninguna condición genética; la coexistencia de cambios en ambos genes conduce a un cuadro clínico más grave (Pla-Martín *et al.*, 2015). Igualmente, otros mecanismos moleculares pueden conducir a fenotipos distintos. La activación sesgada del cromosoma X puede hacer que hermanas heterocigotas para la misma mutación en el gen *FIX* implicado en la hemofilia B con herencia ligada al X recesiva, estén enfermas o no dependiendo de si se expresa el gen sano o no según el patrón de inactivación (Espinós *et al.*, 2000). Son muchos los eventos mutacionales que no son de fácil detección con las herramientas NGS que se suelen utilizar, tales como disomías uniparentales, o mutaciones intrónicas profundas. En su conjunto conducen a una deficiencia en el diagnóstico de prácticamente el 50% de los casos (Bis-Brewer and Züchner, 2018).

- **Familia fMD-218**

Tras obtener resultados negativos en las distintas pruebas genéticas para el diagnóstico de la EW, se concluyó que los pacientes de la familia fMD-218 debían cursar con un fenotipo *Wilson-like*, por lo que dichos casos han sido investigados mediante WES. La utilidad diagnóstica de la secuenciación de exoma se ha implementado en la clínica, con un rendimiento de diagnóstico en un rango entre 25-30% en grandes cohortes de enfermedades raras y heterogéneas (Boycott *et al.*, 2019).

En el análisis de WES se pueden detectar alrededor de unas 20.000 variantes en regiones codificantes de proteínas por individuo. Siguiendo los pasos de filtrado estándar, podemos reducir el número de variantes a un orden de 10 a 100 variantes candidatas (Graham *et al.*, 2018). El resultado del análisis de WES en nuestra familia nos proporcionó un gran listado de variantes identificadas, el cual, tras el proceso de filtrado y priorización, se redujo a siete variantes en un total de cuatro genes. Dos de estos genes, *COG5* y *NOTCH2*, se priorizaron en el filtrado de variantes mediante términos HPO al estar asociados a una clínica similar a la observada en el paciente. Los otros dos, *CBFA2T3* y *CARD10*, se priorizaron en el estudio del resto del exoma en base a la baja frecuencia poblacional registrada en la base de datos gnomAD para las variantes identificadas (inferior al 0,1%), evidencias de expresión de ambos tanto a nivel de mRNA como de proteína en hígado según se recoge en la base de datos *Human Protein Atlas* y participación en rutas de señalización e interacción con otras proteínas relacionadas con fenotipos *Wilson-like*.

Las variantes identificadas y validadas en *COG5* y *CBFA2T3* se descartaron como posibles causantes de la enfermedad en base a la segregación familiar observada. En el caso del gen *COG5*, deficiencias en la proteína se han asociado con una forma leve del trastorno congénito de glicosilación tipo III (CDG2I), una enfermedad autosómica recesiva caracterizada por alteraciones neurológicas leves y, en ciertos casos, por alteraciones hepáticas como la presencia de transaminasas elevadas (Fung *et al.*, 2011). En el estudio de la segregación familiar se observó que ambas variantes descritas se encontraban en el mismo alelo, puesto que ambas estaban presentes tanto en el paciente (MD489) como en dos de sus hijos (MD491 y MD492) y ausentes en la cónyuge (MD490). Basándonos en el tipo de herencia de la enfermedad (autosómica recesiva), en este caso podemos descartarlas como posibles causantes de la enfermedad. Por otra parte, *CBFA2T3/MTG16* es un correpresor transcripcional que se ha asociado a la leucemia megacarioblástica aguda sin síndrome de Down y a leucemia mieloide, pero en este caso el interés reside en su relación con la ruta de señalización Notch. Se ha comprobado que la proteína codificada por *CBFA2T3* interactúa con los dominios intracelulares de todos los receptores Notch de mamíferos, insinuando una relación importante entre las proteínas de la familia MTG y los fenotipos dependientes de Notch (Engel *et al.*, 2010). No obstante, al realizar estudios de segregación familiar, se determinó la presencia de ambas variantes de *CBFA2T3* en muestra de los dos hijos, MD492 y MD493, posiblemente afecto y sano respectivamente, así como ausentes en muestra de la hija y la cónyuge (MD491 y MD490), por tanto, ambos cambios se encuentran en *cis*. En este caso, si alguna de las variantes fuese patogénica, debería cursar como una enfermedad autosómica dominante, la presencia de las variantes en el hijo sano nos permite descartar dichos cambios como posibles causantes de la enfermedad.

Otras dos variantes fueron priorizadas en el gen *CARD10*. El probando (MD489) es heterocigoto compuesto y con la segregación familiar se confirmó que ambas variantes se encontraban en alelos independientes siendo los hijos portadores de una de las dos variantes. Mutaciones en este gen son responsables de un tipo de inmunodeficiencia y autoinmunidad (IMD89; MIM 607209) con

herencia autosómica recesiva, caracterizadas por la aparición en adultos de infecciones recurrentes, alergias, anemia microcítica y la enfermedad de Crohn (Yang *et al.*, 2020). Los síntomas de esta enfermedad no son similares a los observados en el paciente, además, las variantes priorizadas no se han relacionado con el desarrollo de la enfermedad. No obstante, se ha observado que el complejo CBM, formado por CARD10 y otras proteínas mencionadas anteriormente, es capaz de activar la proteína TAK1. Así mismo, recientemente, se ha correlacionado la ausencia de TAK1 en hepatocitos con el desarrollo de fibrosis (Wang *et al.*, 2021), pudiendo sugerir que la ausencia o no funcionalidad de la proteína CARD10 cause en última instancia un efecto similar al observado con la ausencia de TAK1. En muchas ocasiones, las variantes identificadas como plausibles a menudo tienen un bajo nivel de evidencia como respaldo y no son adecuadas para establecer un diagnóstico basado en la genética (Graham *et al.*, 2018), siendo este el caso de las variantes identificadas en *CARD10*.

Finalmente nos centramos en la variante c.5003-23T>C del gen *NOTCH2*. Esta variante se analizó por la implicación de dicho gen en el síndrome de Alagille. Este es un trastorno multisistémico autosómico dominante causado por mutaciones en los genes *JAG1* o *NOTCH2* (los casos debidos a mutaciones en *NOTCH2* son mucho menos frecuentes que en *JAG1*). Clínicamente se caracteriza por colestasis y escasez de conductos biliares en la biopsia hepática, además de presentar defectos cardíacos, esqueléticos y anomalías oftalmológicas (Ayoub and Kamath, 2020; Sánchez-Monteagudo *et al.*, 2021), pudiendo dar lugar a un fenotipo similar al observado en pacientes con EW. MD489 y su hijo posiblemente afecto MD491 son los únicos portadores de la variante en la familia, por lo que fue un candidato interesante a estudiar mediante ensayos funcionales, dirigidos a investigar su posible implicación en la alteración del proceso de *splicing*. Para ello se llevaron a cabo el estudio de mRNA en sangre periférica y el ensayo de minigenes. En ambos estudios se obtuvo el mismo resultado, el procesamiento del transcrito en presencia de la variante se corresponde con el patrón de *splicing* obtenido para el transcrito generado a partir del alelo de referencia, es decir en ausencia de la mutación, confirmando lo previamente revelado por los predictores *in silico*, que señalaban que la variante no alteraría los sitios de *splicing* 5' y 3' de la secuencia en la que se localiza. Por tanto, al menos con las dos técnicas realizadas para el análisis, se podría descartar como variante patológica a c.5003-23T>C, puesto que no se observaron diferencias en el procesamiento del transcrito en presencia de dicha variante.

- **Familia fEW-18**

En este caso se estudió la variante c.704+2T>G en el gen *INSIG1* identificada en la WGS en heterocigosis en muestra del probando EW27. Esta es interesante por la implicación de *INSIG1* en la homeostasis del colesterol. Estudios previos han confirmado que la proteína INSIG1 es responsable de la retención del complejo SCAP/SREBP en el retículo endoplásmico, lo que disminuye la síntesis de colesterol (Yang *et al.*, 2002). Además, se ha observado que una

sobreexpresión hepática de esta proteína en ratas Zucker diabéticas grasas reduce significativamente la esteatosis hepática y la hipertrigliceridemia (Takaishi *et al.*, 2004). Estos hallazgos hacen pensar que un déficit de esta proteína podría causar el efecto contrario, dando lugar a un aumento de la activación de SREBP y, por ende, de la síntesis de triglicéridos con los efectos perjudiciales que ello conlleva.

Para el análisis de la variante se estudió cómo podría afectar esta al proceso de *splicing* mediante el estudio de mRNA en sangre periférica y el estudio de minigenes. En este caso se obtuvieron resultados distintos entre ambos ensayos. Estas discrepancias entre ambas aproximaciones también las podemos ver en otros estudios donde se han observado diferencias ocasionales en los patrones de *splicing* entre minigenes y el análisis del mRNA del paciente (Van Der Klift *et al.*, 2015). Las diferencias en el *splicing* pueden deberse a diversos factores como la fuerza del sitio de *splicing*, la longitud del exón/intrón resultante (Sánchez-Monteagudo *et al.*, 2020), la cercanía a elementos de *splicing* y el orden temporal de eliminación de los intrones en el pre-mRNA (Ramalho *et al.*, 2003). Por tanto, aunque los ensayos funcionales de *splicing* con minigenes son una estrategia importante para la caracterización inicial de las variantes de *splicing*, no son siempre efectivos y estos deben compararse con el RNA del paciente utilizando tejidos fisiopatológicamente relevantes para la enfermedad. Estos proporcionan una información más precisa y confiable en el estudio de este tipo de variantes, sin embargo, no siempre se tiene un acceso sencillo al tejido de interés. En nuestro caso la obtención de tejido hepático mediante biopsia para su estudio consiste en una prueba invasiva, por lo que en estos casos el análisis utilizando células sanguíneas representa una opción alternativa siempre y cuando el gen de interés se exprese normalmente en ellas. Por tanto, estas discrepancias entre ambas aproximaciones pueden deberse en parte a que ninguna de ellas reflejaría necesariamente el procesamiento real en los tejidos afectados.

Además, durante el ensayo de minigenes observamos que tanto la secuencia de referencia como la mutada producían un *skipping* del exón 4. Esto se ha visto previamente en otros estudios que han hecho uso del vector reportero pSPL3, donde han observado que el minigen de la secuencia de referencia producía un *skipping* del exón a estudiar (Lin *et al.*, 2021). Esto se debe en parte a que actualmente no existe un vector reportero de minigenes que sea adecuado para el análisis de todas las variantes de *splicing*. También se ha observado que hay diferencias en el ensayo con diferentes tipos celulares, ya que no en todos los tipos celulares se lleva a cabo el mismo procesamiento de los transcritos. Por tanto, sería interesante plantear nuevos enfoques para mejorar el ensayo de esta variante con minigenes. Se pueden abordar distintas estrategias para ello como realizar el ensayo con una línea celular hepática en lugar de las células HeLa con el fin de proporcionar el entorno adecuado para el procesamiento del transcrito de forma similar a como lo hace en el entorno fisiopatológico. También, puesto que se ha visto en otros casos que el vector reportero pSPL3 puede causar estos fallos, se puede plantear el ensayo con un vector reportero de minigenes diferente, como por ejemplo el vector pET01 (Chen *et al.*, 2018). Finalmente, podría plantearse un ensayo de minigenes, basado en la utilización de construcciones multiexon, en el que se puede

contextualizar mejor el entorno de la secuencia por si hay otros factores en esta que participan en el proceso de *splicing* (Sangermano *et al.*, 2018).

5.2. Biomarcadores en la EW

La EW es una enfermedad crónica, que precisa de un diagnóstico temprano y preciso para proporcionar un mejor pronóstico y una mayor respuesta al tratamiento. Sin embargo, el diagnóstico de esta enfermedad, basado en la escala de Leipzig, es un proceso complicado que involucra una combinación de resultados de pruebas clínicas, bioquímicas y genéticas, por lo que es común un retraso en este. Los mayores retrasos en el diagnóstico se observan en pacientes que presentan síntomas psiquiátricos, debido a que las características clínicas asociadas a estos son variables e inicialmente se suelen atribuir a trastornos más comunes de forma errónea (Shribman *et al.*, 2021).

Los marcadores bioquímicos usados tradicionalmente, el análisis de la ceruloplasmina, el cobre intrahepático y el cobre urinario, son insuficientes de forma aislada para establecer un diagnóstico; en cambio de forma conjunta son bastante útiles y, de hecho, son pruebas incluidas en la escala de Leipzig. No obstante, un 15% de los heterocigotos son positivos a estos marcadores y un 3% de los pacientes confirmados mediante diagnóstico genético presentan unos niveles de cobre en orina normales (Woimant *et al.*, 2019). Por tanto, es necesario buscar nuevos marcadores de la enfermedad más efectivos y que permitan evaluar la respuesta al tratamiento. En esta línea, en los últimos años se le ha dado un enfoque alternativo a la forma de medir el cobre sérico, pasando a estudiar el cobre intercambiable, que se corresponde con la fracción lábil unida a la albúmina y otros péptidos. Se ha visto que esta nueva prueba es capaz de diferenciar la EW de otras enfermedades hepáticas, pero aún no se tiene constancia de que sea capaz de discernir la EW de otros trastornos neurológicos y psiquiátricos (Shribman *et al.*, 2021).

Actualmente también se están estudiando los miRNAs como marcadores valiosos de la enfermedad, el estadio en el que se encuentra y la respuesta a terapia. Estos son pequeños RNA no codificantes capaces de regular la expresión génica que tienen el potencial de convertirse en nuevos biomarcadores no invasivos debido a su gran estabilidad y su relativa facilidad de detección en circulación. Se sabe que la expresión de miRNAs y, consecuentemente, los niveles detectables de estos pueden variar dependiendo de las condiciones fisiológicas del hígado, destacando su utilidad de pronóstico tanto a nivel de enfermedad hepática como predictor temprano de la respuesta a un tratamiento. Por otro lado, los miRNAs circulantes en sangre son fáciles de aislar y el diseño de test para el análisis de estos es relativamente sencillo. Además, se ha descrito que existe una fuerte correlación entre miRNAs circulantes y de tejido en cáncer, enfermedades neurodegenerativas, cardiovasculares y hepáticas, y en infecciones virales (Hayes and Chayama,

2016; Assmann *et al.*, 2019; Szelenberger *et al.*, 2019). Un ejemplo de su utilidad en la EW es el miR-122; se ha visto en el modelo animal de rata LEC que elevados niveles de este miRNA prolongados en el tiempo pueden predecir de forma más precisa que mediante el estudio de las enzimas hepáticas, el agravamiento de la enfermedad hepática hasta derivar en fallo hepático fulminante (Mahgoub and Steer, 2016).

En nuestro estudio se evaluaron tres miRNAs como posibles biomarcadores de la EW: miR-146a-5p, miR-340-3p y miR-485-3p. miR-146a-5p se asocia con la regulación de la respuesta inflamatoria (Chen *et al.*, 2016). Además, es sabido que este miRNA se encuentra disminuido durante la diferenciación a hepatocitos de las células progenitoras (Mahgoub and Steer, 2016). En el modelo murino *knockout* de la EW (*Atp7b*^{-/-}) no obtuvimos diferencias estadísticamente significativas entre el grupo control y el enfermo; no obstante, sí se observa una tendencia de este miRNA a presentar menores niveles en circulación en ratones con fenotipo Wilson respecto a los sanos. Sería necesaria una cohorte de muestras mayor para poder comprobar si esta tendencia a la baja es significativa realmente, ya que nuestra cohorte no es suficientemente grande como para poder discernirlo.

También se estudió miR-340-3p, este es un miRNA multifuncional involucrado en la supervivencia celular, la apoptosis y la diferenciación celular. miR-340-3p se encuentra elevado en las vesículas extracelulares sanguíneas de ratones con esteatohepatitis alcohólica (Eguchi *et al.*, 2017), y también, en receptores de trasplante de hígado estables a largo plazo respecto con receptores con rechazo agudo (Zhang *et al.*, 2015). En su análisis mediante qPCR no observamos diferencias estadísticamente significativas entre ambos grupos de estudio. Sin embargo, al igual que en el análisis de miR-146a-5p, se observa una tendencia a la baja en el grupo de ratones enfermos frente a los sanos. Por tanto, sería interesante ampliar el tamaño muestral del estudio para poder corroborar si esta tendencia que observamos es realmente una diferencia significativa entre ambos grupos.

Finalmente, estudiamos miR-485-3p. Este miRNA es capaz de reprimir la expresión de ferroportina, un exportador celular de hierro, por lo que su sobreexpresión se ha asociado con un aumento de los niveles de hierro celular. Se expresa de forma anormal en enfermedades relacionadas con inflamación, afectando al grado de respuesta inflamatoria (Sangokoya *et al.*, 2013). En nuestro estudio observamos una clara diferencia entre el grupo control y el KO. En el caso del grupo de ratones sanos no se logró amplificar dicho miRNA en ninguna de las cinco muestras; sin embargo, este sí amplificaba a partir de muestra de ratones *Atp7b*^{-/-}, observando una diferencia significativa entre ambos grupos. La liberación incrementada de miR-485-3p podría indicar el daño hepático en cuanto a inflamación de dicho tejido, pudiendo ser un futuro biomarcador de la EW que, junto con otros, pudiese usarse como una firma de miRNAs que permitiese mejorar el diagnóstico y el seguimiento de la enfermedad.

Estos tres miRNA se estudiaron previamente en el laboratorio en cohortes de pacientes con resultados similares a los observados en ratones para miR-146a-5p y miR-340-3p (ambos disminuidos en el grupo de pacientes comparado con el control), pero en el caso de miR-485-3p, se encuentra disminuido en los pacientes, a diferencia del resultado obtenido en ratones. Además, en el estudio de la serie clínica se analizó una posible relación de los miRNAs con diferentes parámetros bioquímicos asociados a la EW como la ceruloplasmina, las enzimas hepáticas o la bilirrubina entre otros. Ninguno de los tres miRNAs se correlacionó con los parámetros bioquímicos estudiados, lo que indicaría que no se encuentran vinculados a la progresión de la enfermedad. Puesto que los pacientes estudiados en la serie clínica reciben tratamiento, se procedió a su análisis en el modelo animal de la enfermedad. En este sentido, con el estudio en el modelo murino de la EW, podemos reproducir la historia natural de la enfermedad y determinar si estos miRNAs realmente están relacionados con el desarrollo de esta. El estudio llevado a cabo en este trabajo es una prueba de concepto para determinar si dichos miRNAs pueden detectarse en ratón y estudiarse correctamente mediante qPCR. El siguiente paso será ampliar la cohorte de ratones a estudiar y realizar un seguimiento de estos a distintos tiempos para determinar cómo varían en circulación, si lo hacen, los diferentes miRNAs con la progresión de la enfermedad y con qué aspectos de la misma se correlacionan.

6. CONCLUSIONES

1. El análisis de WES permite el diagnóstico de pacientes con enfermedades *Wilson-like*. No obstante, se trata de un proceso complejo en el que se deben tener en cuenta numerosos criterios para poder confirmar la asociación de una variante con el fenotipo observado.
2. En la familia fMD-218 investigada mediante WES, se identificaron variantes candidatas en *NOTCH2*, *CARD10*, *CBFA2T3* y *COG5*. Mediante análisis de segregación familiar se descartan las variantes en *CBFA2T3* y *COG5* como posibles causantes de la enfermedad.
3. En el paciente MD489, probando de la familia fMD-218, se identificaron en heterocigosis compuesta las variantes c.478C>T y c.2262T>A en *CARD10*, sin embargo, no hay evidencias suficientes que avalen la implicación de mutaciones en este gen en la patología.
4. Tanto el paciente MD489 como su hijo posiblemente afecto MD492 son portadores en heterocigosis del cambio c.5003-23T>C en *NOTCH2*, gen asociado a síndrome de Allagile. El análisis del transcrito y el estudio de minigenes han permitido concluir que la variante c.5003-23T>C en *NOTCH2* no modifica el proceso de *splicing*, descartándose como variante patológica.
5. El análisis del transcrito y el estudio de minigenes de la variante c.704+2T>G en *INSIG1* no fueron concluyentes puesto que se obtuvieron resultados dispares. Se precisa de un análisis más exhaustivo de la variante para poder determinar su patogenicidad.
6. Los miRNAs miR-146a-5p, miR-340-3p y miR-485-5p, previamente detectados en plasma de pacientes con EW como significativamente desregulados, pudieron detectarse en muestras de plasma del modelo murino *knockout* para la EW (*Atp7b*^{-/-}) permitiendo así su análisis.
7. Es necesaria una cohorte de muestras mayor con tal de poder verificar si las tendencias a la baja detectadas en plasma de ratones *Atp7b*^{-/-} para miR-146a-5p y miR-340-3p y el incremento de miR-485-3p son significativos.

7. BIBLIOGRAFÍA

- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature methods*, 7(4), 248-249. <https://doi.org/10.1038/nmeth0410-248>
- Assmann, T. S., Milagro, F. I., & Martínez, J. A. (2019). Crosstalk between microRNAs, the putative target genes and the lncRNA network in metabolic diseases. *Molecular Medicine Reports*, 20(4), 3543-3554. <https://doi.org/10.3892/mmr.2019.10595>
- Ayoub, M. D., & Kamath, B. M. (2020). Alagille Syndrome: Diagnostic Challenges and Advances in Management. *Diagnostics*, 10(11), 907. <https://doi.org/10.3390/diagnostics10110907>
- Backes, C., Meese, E., & Keller, A. (2016). Specific miRNA Disease Biomarkers in Blood, Serum and Plasma: Challenges and Prospects. *Molecular Diagnosis & Therapy*, 20(6), 509-518. <https://doi.org/10.1007/s40291-016-0221-4>
- Bearn, A. G. (1953). Genetic and biochemical aspects of Wilson's disease. *The American Journal of Medicine*, 15(4), 442-449. [https://doi.org/10.1016/0002-9343\(53\)90134-X](https://doi.org/10.1016/0002-9343(53)90134-X)
- Bis-Brewer, D. M., & Züchner, S. (2018). Perspectives on the Genomics of HSP Beyond Mendelian Inheritance. *Frontiers in Neurology*, 9. <https://www.frontiersin.org/article/10.3389/fneur.2018.00958>
- Bitter, R. M., Oh, S., Deng, Z., Rahman, S., Hite, R. K., & Yuan, P. (2022). Structure of the Wilson disease copper transporter ATP7B. *Science Advances*, 8(9), eabl5508. <https://doi.org/10.1126/sciadv.abl5508>
- Boycott, K. M., Hartley, T., Biesecker, L. G., Gibbs, R. A., Innes, A. M., Riess, O., Belmont, J., Dunwoodie, S. L., Jojic, N., Lassmann, T., Mackay, D., Temple, I. K., Visel, A., & Baynam, G. (2019). A Diagnosis for All Rare Genetic Diseases: The Horizon and the Next Frontiers. *Cell*, 177(1), 32-37. <https://doi.org/10.1016/j.cell.2019.02.040>
- Brunak, S., Engelbrecht, J., & Knudsen, S. (1991). Prediction of human mRNA donor and acceptor sites from the DNA sequence. *Journal of Molecular Biology*, 220(1), 49-65. [https://doi.org/10.1016/0022-2836\(91\)90380-o](https://doi.org/10.1016/0022-2836(91)90380-o)
- Cacace, R., Heeman, B., Van Mossevelde, S., De Roeck, A., Hoogmartens, J., De Rijk, P., Gossye, H., De vos, K., De Coster, W., Strazisar, M., De Baets, G., Schymkowitz, J., Rousseau, F., Geerts, N., De Pooter, T., Peeters, K., Sieben, A., Martin, J.-J., Engelborchs, S., ... Van Broeckhoven, C. (2019). Loss of DPP6 in neurodegenerative dementia: A genetic player in the dysfunction of neuronal excitability. *Acta Neuropathol*, 137, 901-918. <https://doi.org/10.1007/s00401-019-01976-3>
- Chen, Y., Huang, L., Jiao, X., Riazuddin, S., Riazuddin, S. A., & Fielding Hetmancik, J. (2018). A novel LRAT mutation affecting splicing in a family with early onset retinitis pigmentosa. *Human Genomics*, 12(1), 35. <https://doi.org/10.1186/s40246-018-0165-3>
- Chen, Y., Zeng, Z., Shen, X., Wu, Z., Dong, Y., & Cheng, J. C.-H. (2016). MicroRNA-146a-5p Negatively Regulates Pro-Inflammatory Cytokine Secretion and Cell Activation in Lipopolysaccharide Stimulated Human Hepatic Stellate Cells through Inhibition of Toll-Like Receptor 4 Signaling Pathways. *International Journal of Molecular Sciences*, 17(7), 1076. <https://doi.org/10.3390/ijms17071076>
- Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLOS ONE*, 7(10), e46688. <https://doi.org/10.1371/journal.pone.0046688>
- Czlonkowska, A., Litwin, T., Dusek, P., Ferenci, P., Lutsenko, S., Medici, V., Rybakowski, J. K., Weiss, K. H., & Schilsky, M. L. (2018). Nature Reviews Disease Primers article: Wilson disease. *Nature reviews. Disease primers*, 4(1), 21. <https://doi.org/10.1038/s41572-018-0018-3>

- Eguchi, A., Lazaro, R. G., Wang, J., Kim, J., Povero, D., Williams, B., Ho, S. B., Stärkel, P., Schnabl, B., Ohno-Machado, L., Tsukamoto, H., & Feldstein, A. E. (2017). Extracellular vesicles released by hepatocytes from gastric infusion model of ALD contain a miRNA barcode that can be detected in blood. *Hepatology (Baltimore, Md.)*, *65*(2), 475-490. <https://doi.org/10.1002/hep.28838>
- Engel, M. E., Nguyen, H. N., Mariotti, J., Hunt, A., & Hiebert, S. W. (2010). Myeloid Translocation Gene 16 (MTG16) Interacts with Notch Transcription Complex Components To Integrate Notch Signaling in Hematopoietic Cell Fate Specification . *Molecular and Cellular Biology*, *30*(7), 1852-1863. <https://doi.org/10.1128/mcb.01342-09>
- Espinós, C., & Ferenci, P. (2020). Are the new genetic tools for diagnosis of Wilson disease helpful in clinical practice? *JHEP Reports*, *2*(4), 100114. <https://doi.org/10.1016/j.jhepr.2020.100114>
- Espinós, C., Lorenzo, J. I., Casaña, P., Martínez, F., & Aznar, J. A. (2000). Haemophilia B in a female caused by skewed inactivation of the normal X-chromosome. *Haematologica*, *85*(10), 1092-1095. <https://doi.org/10.3324/%25x>
- Ferenci, P., Stremmel, W., Członkowska, A., Szalay, F., Viveiros, A., Stättermayer, A. F., Bruha, R., Houwen, R., Pop, T. L., Stauber, R., Gschwantler, M., Pfeiffenberger, J., Yurdaydin, C., Aigner, E., Steindl-Munda, P., Dienes, H.-P., Zoller, H., & Weiss, K. H. (2019). Age and Sex but Not ATP7B Genotype Effectively Influence the Clinical Phenotype of Wilson Disease. *Hepatology*, *69*(4), 1464-1476. <https://doi.org/10.1002/hep.30280>
- Fung, C. W., Matthijs, G., Sturiale, L., Garozzo, D., Wong, K. Y., Wong, R., Wong, V., & Jaeken, J. (2011). COG5-CDG with a Mild Neurohepatic Presentation. *JIMD Reports*, *3*, 67-70. https://doi.org/10.1007/8904_2011_61
- Graham, E., Lee, J., Price, M., Tarailo-Graovac, M., Matthews, A., Engelke, U., Tang, J., Kluijtmans, L. A. J., Wevers, R. A., Wasserman, W. W., van Karnebeek, C. D. M., & Mostafavi, S. (2018). Integration of genomics and metabolomics for prioritization of rare disease variants: A 2018 literature review. *Journal of Inherited Metabolic Disease*, *41*(3), 435-445. <https://doi.org/10.1007/s10545-018-0139-6>
- Guías de Práctica Clínica de la EASL: Enfermedad de Wilson - PDF Descargar libre.* (s. f.). Recuperado 26 de mayo de 2022, de <https://docplayer.es/85147479-Guias-de-practica-clinica-de-la-easl-enfermedad-de-wilson.html>
- Hahn, S. H. (2014). Population screening for Wilson's disease. *Annals of the New York Academy of Sciences*, *1315*(1), 64-69. <https://doi.org/10.1111/nyas.12423>
- Hayes, C., & Chayama, K. (2016). MicroRNAs as Biomarkers for Liver Disease and Hepatocellular Carcinoma. *International Journal of Molecular Sciences*, *17*(3), 280. <https://doi.org/10.3390/ijms17030280>
- HGMD ® Professional 2020.1.* (s. f.). Recuperado 21 de abril de 2022, de <http://www.hgmd.cf.ac.uk/ac/index.php>
- HMDD v3.2.* (s. f.). Recuperado 8 de abril de 2022, de <http://www.cuilab.cn/hmdd>
- Huang, Z., Shi, J., Gao, Y., Cui, C., Zhang, S., Li, J., Zhou, Y., & Cui, Q. (2019). HMDD v3.0: A database for experimentally supported human microRNA–disease associations. *Nucleic Acids Research*, *47*(D1), D1013-D1017. <https://doi.org/10.1093/nar/gky1010>
- Kasztelan-Szczerbinska, B., & Cichoz-Lach, H. (2021). Wilson's Disease: An Update on the Diagnostic Workup and Management. *Journal of Clinical Medicine*, *10*(21), 5097. <https://doi.org/10.3390/jcm10215097>
- Kerkar, N., & Roberts, E. A. (2018). *Clinical and Translational Perspectives on WILSON DISEASE.* Academic Press.

- Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, 4(7), 1073-1081. <https://doi.org/10.1038/nprot.2009.86>
- Laurie, S., Piscia, D., Matalonga, L., Corvó, A., Fernández-Callejo, M., Garcia-Linares, C., Hernandez-Ferrer, C., Luengo, C., Martínez, I., Papakonstantinou, A., Picó-Amador, D., Protasio, J., Thompson, R., Tonda, R., Bayés, M., Bullich, G., Camps-Puchadas, J., Paramonov, I., Trotta, J.-R., ... Beltran, S. (2022). The RD-Connect Genome-Phenome Analysis Platform: Accelerating diagnosis, research, and gene discovery for rare diseases. *Human Mutation*, 43(6), 717-733. <https://doi.org/10.1002/humu.24353>
- Lin, J.-H., Wu, H., Zou, W.-B., Masson, E., Fichou, Y., Le Gac, G., Cooper, D. N., Férec, C., Liao, Z., & Chen, J.-M. (2021). Splicing Outcomes of 5' Splice Site GT>GC Variants That Generate Wild-Type Transcripts Differ Significantly Between Full-Length and Minigene Splicing Assays. *Frontiers in Genetics*, 12(August), 1-12. <https://doi.org/10.3389/fgene.2021.701652>
- Lucena-Valera, A., Perez-Palacios, D., Muñoz-Hernandez, R., Romero-Gómez, M., & Ampuero, J. (2021). Wilson's disease: Revisiting an old friend. *World Journal of Hepatology*, 13(6), 634-649. <https://doi.org/10.4254/wjh.v13.i6.634>
- Mahgoub, A., & Steer, C. (2016). MicroRNAs in the Evaluation and Potential Treatment of Liver Diseases. *Journal of Clinical Medicine*, 5(5), 52. <https://doi.org/10.3390/jcm5050052>
- Mitsushashi, S., & Matsumoto, N. (2020). Long-read sequencing for rare human genetic diseases. *Journal of Human Genetics*, 65(1), 11-19. <https://doi.org/10.1038/s10038-019-0671-8>
- Mukherjee, S., Dutta, S., Majumdar, S., Biswas, T., Jaiswal, P., Sengupta, M., Bhattacharya, A., Gangopadhyay, P. K., Bavdekar, A., Das, S. K., & Ray, K. (2014). Genetic defects in Indian Wilson disease patients and genotype-phenotype correlation. *Parkinsonism & Related Disorders*, 20(1), 75-81. <https://doi.org/10.1016/j.parkreldis.2013.09.021>
- Murillo, O., Luqui, D. M., Gazquez, C., Martinez-Espartosa, D., Navarro-Blasco, I., Monreal, J. I., Guembe, L., Moreno-Cermeño, A., Corrales, F. J., Prieto, J., Hernandez-Alcoceba, R., & Gonzalez-Aseguinolaza, G. (2016). Long-term metabolic correction of Wilson's disease in a murine model by gene therapy. *Journal of Hepatology*, 64(2), 419-426. <https://doi.org/10.1016/j.jhep.2015.09.014>
- NOTCH2 notch receptor 2 [Homo sapiens (human)]—Gene—NCBI.* (s. f.). Recuperado 24 de febrero de 2022, de <https://www.ncbi.nlm.nih.gov/gene/4853>
- O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Frontiers in Endocrinology*, 9, 402. <https://doi.org/10.3389/fendo.2018.00402>
- Pla-Martín, D., Calpena, E., Lupo, V., Márquez, C., Rivas, E., Sivera, R., Sevilla, T., Palau, F., & Espinós, C. (2015). Junctophilin-1 is a modifier gene of GDAP1-related Charcot-Marie-Tooth disease. *Human Molecular Genetics*, 24(1), 213-229. <https://doi.org/10.1093/hmg/ddu440>
- Ramalho, A. S., Beck, S., Penque, D., Gonska, T., Seydewitz, H. H., & Mall, M. (2003). Transcript analysis of the cystic fibrosis splicing mutation 1525-1G>A shows use of multiple alternative splicing sites and suggests a putative role of exonic splicing enhancers. *J Med Genet*, 40, 88. <https://doi.org/10.1136/jmg.40.7.e88>
- RD-Connect Dashboard.* (s. f.). Recuperado 13 de mayo de 2022, de <https://platform.rd-connect.eu/>
- Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved Splice Site Detection in Genie. *Journal of Computational Biology*, 4(3), 311-323. <https://doi.org/10.1089/cmb.1997.4.311>
- Rentzsch, P., Schubach, M., Shendure, J., & Kircher, M. (2021). CADD-Splice—Improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Medicine*, 13(1), 31. <https://doi.org/10.1186/s13073-021-00835-9>

- Rossetti, S., Hoogeveen, A. T., & Sacchi, N. (2004). The MTG proteins: Chromatin repression players with a passion for networking. *Genomics*, *84*(1), 1-9. <https://doi.org/10.1016/j.ygeno.2004.02.011>
- Ryan, A., Nevitt, S. J., Tuohy, O., & Cook, P. (2019). Biomarkers for diagnosis of Wilson's disease. *The Cochrane Database of Systematic Reviews*, *2019*(11), CD012267. <https://doi.org/10.1002/14651858.CD012267.pub2>
- Sánchez-Monteagudo, A., Álvarez-Sauco, M., Sastre, I., Martínez-Torres, I., Lupo, V., Berenguer, M., & Espinós, C. (2020). Genetics of Wilson disease and Wilson-like phenotype in a clinical series from eastern Spain. *Clinical Genetics*, *97*(5), 758-763. <https://doi.org/10.1111/cge.13719>
- Sánchez-Monteagudo, A., Álvarez-Sauco, M., Sastre, I., Martínez-Torres, I., Lupo, V., Berenguer, M., & Espinós, C. (2020). Genetics of Wilson disease and Wilson-like phenotype in a clinical series from eastern Spain. *Clinical Genetics*, *97*(5), 758-763. <https://doi.org/10.1111/cge.13719>
- Sánchez-Monteagudo, A., Ripollés, E., Berenguer, M., & Espinós, C. (2021a). Wilson's Disease: Facing the Challenge of Diagnosing a Rare Disease. *Biomedicines*, *9*(9), 1100. <https://doi.org/10.3390/biomedicines9091100>
- Sánchez-Monteagudo, A., Ripollés, E., Berenguer, M., & Espinós, C. (2021b). Wilson's Disease: Facing the Challenge of Diagnosing a Rare Disease. *Biomedicines*, *9*(9), 1100. <https://doi.org/10.3390/biomedicines9091100>
- Sandahl, T. D., Laursen, T. L., Munk, D. E., Vilstrup, H., Weiss, K. H., & Ott, P. (2020). The Prevalence of Wilson's Disease: An Update. *Hepatology*, *71*(2), 722-732. <https://doi.org/10.1002/hep.30911>
- Sangermano, R., Khan, M., Cornelis, S. S., Richelle, V., Albert, S., Garanto, A., Elmelik, D., Qamar, R., Lugtenberg, D., van den Born, L. I., Collin, R. W. J., & Cremers, F. P. M. (2018). ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. *Genome Research*, *28*(1), 100-110. <https://doi.org/10.1101/gr.226621.117>
- Sangokoya, C., Doss, J. F., & Chi, J.-T. (2013). Iron-Responsive miR-485-3p Regulates Cellular Iron Homeostasis by Targeting Ferroportin. *PLoS Genetics*, *9*(4), e1003408. <https://doi.org/10.1371/journal.pgen.1003408>
- Schwarz, J. M., Rödelberger, C., Schuelke, M., & Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*, *7*(8), 575-576. <https://doi.org/10.1038/nmeth0810-575>
- Shribman, S., Poujois, A., Bandmann, O., Czlonkowska, A., & Warner, T. T. (2021). Wilson's disease: Update on pathogenesis, biomarkers and treatments. *Journal of Neurology, Neurosurgery & Psychiatry*, *92*(10), 1053-1061. <https://doi.org/10.1136/jnnp-2021-326123>
- Smedley, D., Smith, K. R., Martin, A., Thomas, E. A., McDonagh, E. M., & Cipriani, V. (2021). 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care—Preliminary Report. *New England Journal of Medicine*, *385*(20), 1868-1880. <https://doi.org/10.1056/NEJMoa2035790>
- Szelenberger, R., Kacprzak, M., Saluk-Bijak, J., Zielinska, M., & Bijak, M. (2019). Plasma MicroRNA as a novel diagnostic. *Clinica Chimica Acta*, *499*(September), 98-107. <https://doi.org/10.1016/j.cca.2019.09.005>
- Takaishi, K., Duplomb, L., Wang, M.-Y., Li, J., & Unger, R. H. (2004). Hepatic insig-1 or -2 overexpression reduces lipogenesis in obese Zucker diabetic fatty rats and in fasted/refed normal rats. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(18), 7106-7111. <https://doi.org/10.1073/pnas.0401715101>
- Tang, N., Sandahl, T. D., Ott, P., & Kepp, K. P. (2019). Computing the Pathogenicity of Wilson's Disease ATP7B Mutations: Implications for Disease Prevalence. *Journal of Chemical Information and Modeling*, *59*(12), 5230-5243. <https://doi.org/10.1021/acs.jcim.9b00852>

- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigartyo, C. A.-K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., ... Pontén, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. <https://doi.org/10.1126/science.1260419>
- Van Der Klift, H. M., Jansen, A. M. L., Van Der Steenstraten, N., Bik, E. C., Tops, C. M. J., Devilee, P., & Wijnen, J. T. (2015). Splicing analysis for exonic and intronic mismatch repair gene variants associated with Lynch syndrome confirms high concordance between minigene assays and patient RNA analyses. *Molecular Genetics & Genomic Medicine*, *3*(4), 327-345. <https://doi.org/10.1002/mgg3.145>
- Wang, W., Gao, W., Zhu, Q., Alasbahi, A., Seki, E., & Yang, L. (2021). TAK1: A Molecular Link Between Liver Inflammation, Fibrosis, Steatosis, and Carcinogenesis. *Frontiers in Cell and Developmental Biology*, *9*. <https://www.frontiersin.org/article/10.3389/fcell.2021.734749>
- Weiss, K., Kurolap, A., Paperna, T., Mory, A., Steinberg, M., Hershkovitz, T., Ekhilevitch, N., & Baris, H. N. (2018). Rare Disease Diagnostics: A Single-center Experience and Lessons Learnt. *Rambam Maimonides Medical Journal*, *9*(3), e0018. <https://doi.org/10.5041/RMMJ.10341>
- Woimant, F., Djebrani-Oussedik, N., & Poujois, A. (2019). New tools for Wilson's disease diagnosis: Exchangeable copper fraction. *Annals of Translational Medicine*, *0*(0), 16-16. <https://doi.org/10.21037/atm.2019.03.02>
- Wright, C. F., FitzPatrick, D. R., & Firth, H. V. (2018). Paediatric genomics: Diagnosing rare disease in children. *Nature Reviews Genetics*, *19*(5), 253-268. <https://doi.org/10.1038/nrg.2017.116>
- Yang, D., Guo, T., Yuan, Z., Lei, C., Ding, S., Yang, Y., Tan, Z., & Luo, H. (2020). Mutant CARD10 in a family with progressive immunodeficiency and autoimmunity. *Cellular & Molecular Immunology*, *17*(7), 782-784. <https://doi.org/10.1038/s41423-020-0423-x>
- Yang, T., Espenshade, P. J., Wright, M. E., Yabe, D., Gong, Y., Aebersold, R., Goldstein, J. L., & Brown, M. S. (2002). Crucial Step in Cholesterol Homeostasis: Sterols Promote Binding of SCAP to INSIG-1, a Membrane Protein that Facilitates Retention of SREBPs in ER. *Cell*, *110*(4), 489-500. [https://doi.org/10.1016/S0092-8674\(02\)00872-3](https://doi.org/10.1016/S0092-8674(02)00872-3)
- Zhang, P., Guo, Z., Zhong, K., Li, Q., Ouyang, J., Chen, M., Hu, A., Jiao, X., Zhu, X., & He, X. (2015). Evaluation of Immune Profiles and MicroRNA Expression Profiles in Peripheral Blood Mononuclear Cells of Long-Term Stable Liver Transplant Recipients and Recipients With Acute Rejection Episodes. *Transplantation Proceedings*, *47*(10), 2907-2915. <https://doi.org/10.1016/j.transproceed.2015.10.048>

ANEXOS

ANEXO I

Listado de genes seleccionados por términos HPO:

ABCB11, ABCB4, ABCC2, ABCD3, ABCG8, ABHD5, ACAD9, ACADM, ACADVL, ACOX1, ACSF3, ADA2, ADAR, ADK, AGK, AGL, AGPAT2, AHCY, AKR1D1, ALAD, ALAS2, ALDOB, ALG11, ALG12, ALG13, ALG6, ALG8, ALMS1, AMACR, AP1S1, ASAH1, ASL, ASS1, ATP6AP1, ATP6AP2, ATP7A, ATP7B, ATP8B1, ATRX, AXIN1, BAAT, BCAT2, BCS1L, BICRA, BLVRA, BMP2, BSCL2, C10orf2, C1QBP, C8orf37, CAVIN1, CBS, CC2D2A, CCDC115, CD19, CD46, CD81, CDKN2A, CEP63, CFH, CFI, CFTR, CLDN1, CLPB, COG2, COG4, COG5, COG6, COG7, COG8, CORIN, COX16, COX5A, CP, CPT1A, CPT2, CR2, CSPP1, CYC1, CYP7B1, DCDC2, DDOST, DGUOK, DHDDS, DLD, DMPK, DNAJC19, DNAJC21, DOLK, DPAGT1, DPM1, DPM2, DPM3, DPYS, EARS2, EFL1, EIF2AK3, EPHX1, ERCC4, ERCC6, ERCC8, EVC2, F5, FADD, FAH, FAN1, FARSA, FARSB, FBP1, FBXL4, FOXA2, FOXP3, G6PC, GAA, GALT, GBA, GBE1, GDF2, GFM1, GLRX5, GNE, GNMT, GPD1, GSTZ1, GYS2, HADH, HADHA, HADHB, HAMP, HFE, HJV, HMGCL, HNF1B, HNF4A, HPD, HSD17B4, HSD3B7, IARS1, ICOS, IER3IP1, IFIH1, IFNAR2, IFT140, IFT172, IFT43, IKZF1, IL18BP, INPP5E, INVS, JAG1, JAK2, KIAA0586, KRAS, KRT18, KRT8, LARS1, LBR, LDHA, LIPA, LIPT1, LRP2, LYST, MARS1, MICOS13, MPV17, MRPL3, MRPL44, MRPS16, MRPS28, MS4A1, MTPP, MTX2, MVK, MYH9, MYO5B, NBAS, NFKB1, NFKB2, NFS1, NGLY1, NOS3, NOTCH2, NPC1, NPC2, NPHP3, NPHP4, NR1H4, NSMCE2, OCLN, OCRL, OFD1, OTUD5, PALB2, PALLD, PEPD, PEX1, PEX10, PEX12, PEX13, PEX14, PEX16, PEX19, PEX2, PEX26, PEX3, PEX5, PEX6, PGM1, PHKA2, PHKB, PHKG2, PKD1, PKD2, PKHD1, PMM2, PNPLA2, PNPLA3, POLD1, POLG, POLG2, POMC, PPOX, PRF1, PRKCD, PRKCSH, PRSS1, PSMB4, PSMB8, PSMB9, PYGL, RABL3, RAG1, RAG2, RBCK1, RINT1, RNASEH2A, RNASEH2B, RNASEH2C, RPGRIP1L, RRM2B, SAMHD1, SAR1B, SBDS, SC5D, SCO1, SDHD, SEC63, SERPINA1, SERPINA3, SLC10A1, SLC10A2, SLC17A3, SLC22A5, SLC25A13, SLC25A15, SLC25A20, SLC25A4, SLC29A3, SLC2A2, SLC30A10, SLC33A1, SLC35A2, SLC37A4, SLC40A1, SLC51B, SLC7A7, SMAD4, SMPD1, SP110, SPINK1, SRD5A3, SRP54, STEAP3, STOX1, STX11, STXBP2, SUCLA2, SUCLG1, TANGO2, TAT, TCF4, TFAM, TFR2, TJP2, TKFC, TMEM165, TMEM199, TMEM67, TMEM70, TNFRSF13B, TNFRSF13C, TP53, TRAPPC11, TREX1, TRMT10C, TRMU, TWNK, TYMP, UBR1, UGT1A1, UGT1A4, UNC13D, UQCRC2, UROD, USP18, UTP4, VIPAS39, VPS13A, VPS33B, WDR35, YARS2

ANEXO II

Tabla 1. Lista de cebadores y detalles de la reacción de PCR para la amplificación y validación de las distintas variantes encontradas en el análisis de WES del paciente MD489.

Gen	Cebadores	Secuencias 5'- 3'	Longitud del amplicón (pb)	Temperatura de <i>annealing</i>
<i>NOTCH2</i> (NM_024408)	NOTCH2_E27_D	GAGGCTCCTGTCACCCTTTATG	628	62°C
	NOTCH2_E27_R	CATCCTGTCCCCTGGCTC		
<i>CARD10</i> (NM_014550)	CARD10_E4_D	CTTGCCTGAGCTGCCTATTC	461	62°C
	CARD10_E4_R	CAATGGTGGCTCTGGGAC		
	CARD10_E15_D	GACCTGGGTTTGAGTCCTGG	798	64°C
	CARD10_E15_R	GTGTCTTCTAGGTTGCCAGAG		
<i>CBFA2T3</i> (NM_005187)	CBFA2T3_E2_D	GACACTCAGTGGGGCCTTC	539	62°C
	CBFA2T3_E2_R	CTGTGTGGACTCACCTGGC		
	CBFA2T3_E9_D	CAAACCTCTGCGTGCCTC	531	62°C
	CBFA2T3_E9_R	GTTGCATTGCCTGCCTGA		
<i>COG5</i> (NM_006348)	COG5_3UTR_D	CTAGAGGAAAGGCCGACTCC	642	64°C
	COG5_3UTR_R	GCTTGAGCCATCAGATATCAGC		

Tabla 2. Cebadores y detalles de la PCR para amplificar los fragmentos de *NOTCH2* e *INSIG1* a clonar en pSPL3 durante el estudio de minigenes. En verde la secuencia de corte para las enzimas *XhoI* (en el cebador directo) y *NheI* (en el reverso).

Cebadores	Secuencias 5'- 3'	Longitud del amplicón	Temperatura de <i>annealing</i>
NOTCH2_E27E28_ <i>XhoI</i> /D	ggaaCTCGAGGTCATGCATTGCCTCAGC	983 pb	64°C
NOTCH2_E27E28_ <i>NheI</i> /R	gcgcGCTAGCGGTTTCAGTCACTACAAGGC		
INSIG1_E4_ <i>XhoI</i> /D	ggaaCTCGAGGACTCAGTGATGGCAGGGAG	625 pb	64°C
INSIG1_E4_ <i>NheI</i> /R	gcgcGCTAGCGCAGAACGCATATCAACCATC		

Tabla 3. Cebadores y detalles de la PCR para amplificar los fragmentos clonados en pSPL3 durante el estudio de minigenes.

Cebadores	Secuencias 5'- 3'	Temperatura de <i>annealing</i>
SD6_F	TCTGAGTCACCTGGACAACC	55°C
SA2_R	ATCTCAGTGGTATTTGTGAGC	

Tabla 4. Lista de cebadores y detalles de la reacción de PCR para la amplificación del cDNA a partir de sangre periférica de los pacientes EW27 y MD489.

Cebadores	Secuencias 5'- 3'	Longitud del amplicón	Temperatura de <i>annealing</i>
INSIG1_cDNA_E1E2/D	CTGCTGTTGTTGGCCTACTG	405 pb	60°C
INSIG1_cDNA_E4E5/R	CAGGAACACCCATAGCTAACTG		
NOTCH2_E25E26_cDNA/D	AGCAAGACATGCAAGTATGAC	825 pb	60°C
NOTCH2_E29E30_cDNA/R	CTCATCTTCAGCCTTTACTTTC		

Tabla 5. Cebadores y detalles de la PCR para la amplificación del fragmento a clonar de *INSIG1*. En verde la secuencia de corte para las enzimas *HindIII* (en el cebador directo) y *BamHI* (en el reverso).

Cebadores	Secuencias 5'- 3'	Longitud del amplicón	Temperatura de <i>annealing</i>
INSIG1_cDNA_E1E2_ <i>HindIII</i> /D	ggaaAAGCTTCTGCTGTTGTTGGCCTACTG	415pb	60°C
INSIG1_cDNA_E4E5_ <i>BamHI</i> /R	gcgcGGATCCAGGAACACCCATAGCTAACTG		

ANEXO III

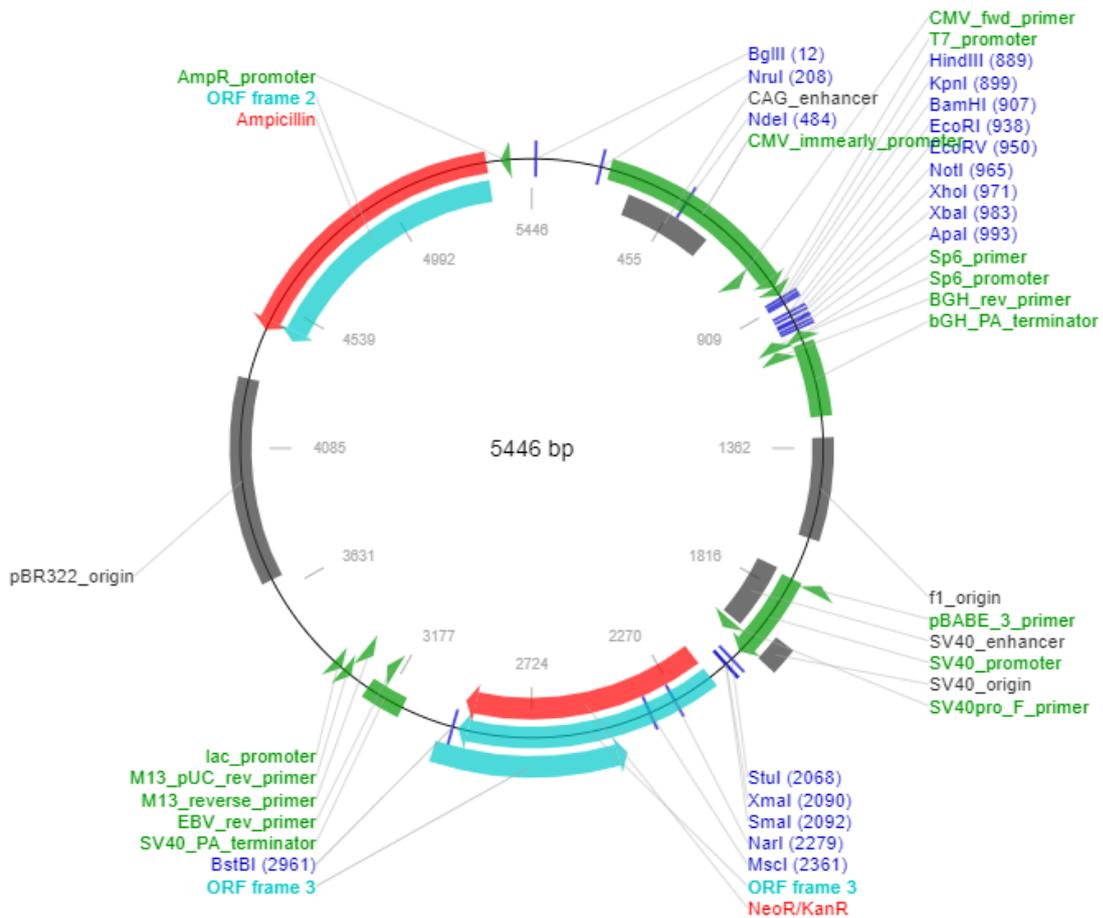


Figura 1. Mapa del vector pcDNA3. El vector contiene secuencias que permiten su replicación en *E. coli* y células huésped. También posee resistencia a ampicilina que permite la selección de subclones y un sitio de clonación múltiple (MCS) que proporciona varios sitios de restricción para la subclonación del DNA genómico.

ANEXO IV

Páginas WEB empleadas en el cribado de las variantes:

Blast2seq del NCBI: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROGRAM=blastn&BLAST_PROGRAM_DEF=megaBlast&BLAST_SPEC=blast2seq

CADD: <https://cadd.gs.washington.edu/snv>

Ensembl GRCh37: <https://grch37.ensembl.org/index.html>

FSplice: <http://www.softberry.com/berry.phtml?topic=fssplice&group=programs&subgroup=gfind>

Human Protein Atlas: <https://www.proteinatlas.org/>

MutationTaster: <https://www.mutationtaster.org/index.html>

NCBI Gene: <https://www.ncbi.nlm.nih.gov/gene>

NetGene2: <https://services.healthtech.dtu.dk/service.php?NetGene2-2.42>

NNSplice: https://www.fruitfly.org/seq_tools/splice.html

OMIM: <https://www.omim.org/>

PolyPhen2: <http://genetics.bwh.harvard.edu/pph2/index.shtml>

PubMed: <https://pubmed.ncbi.nlm.nih.gov/>

SIFT-PROVEAN: http://provean.jcvi.org/genome_submit_2.php?species=human

SpliceAI: <https://spliceailookup.broadinstitute.org/>

ANEXO V

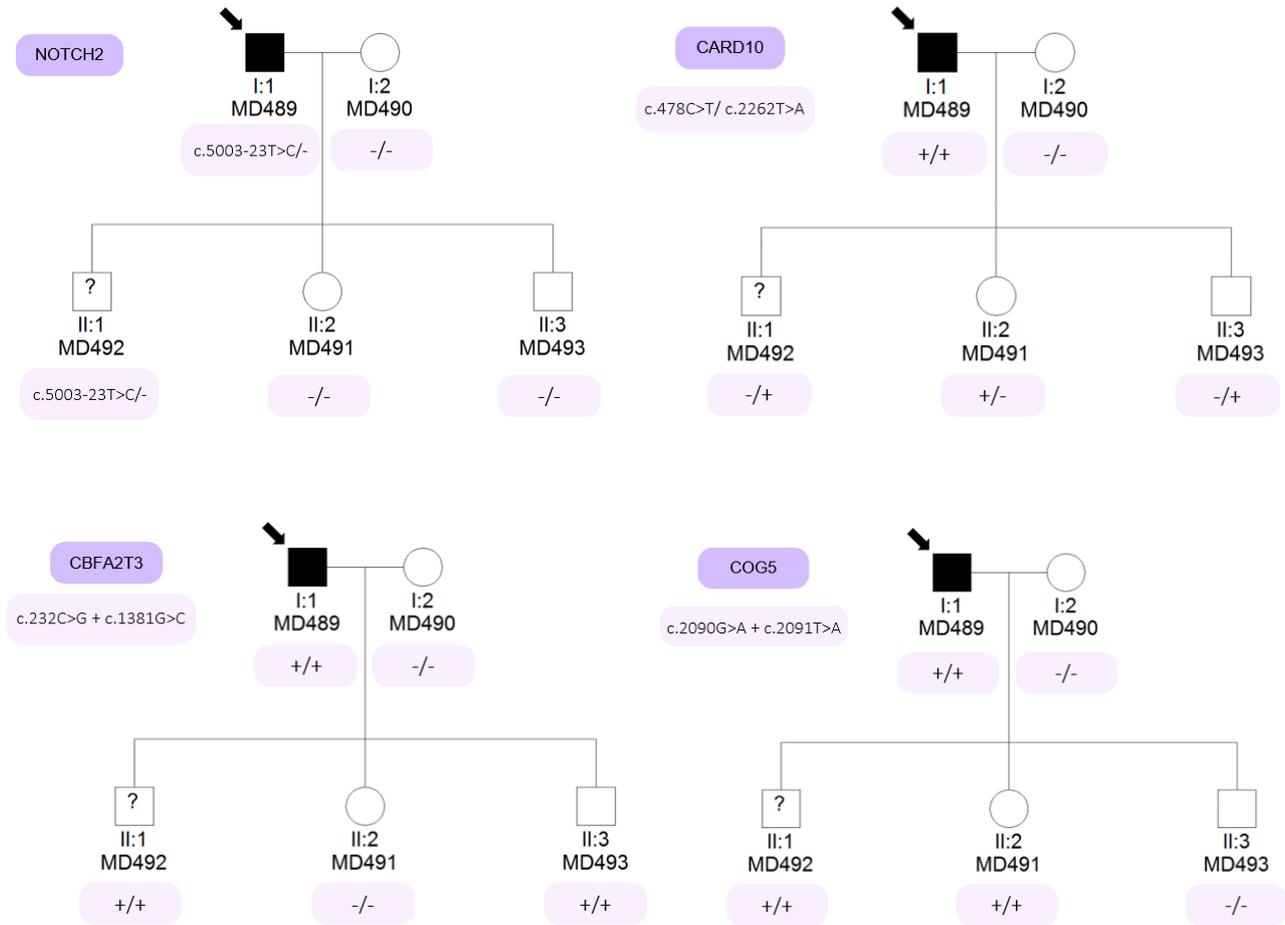


Figura 1. Segregación de las variantes validadas en la familia fMD-218

ANEXO VI

A) Predictores *in silico* de *splicing* para la variante c.704+2T>G en *INSIG1*

FSPLICE:

Secuencia de referencia:

Direct chain.

Acceptor(AG) sites. Threshold 4.175 (90%).

- 1 P: 123 W: 5.72 Seq: tcctttctttAGcttattacac
- 2 P: 260 W: 4.58 Seq: ttatatcaccAGaaattggatt Acceptor canónico
- 3 P: 346 W: 4.30 Seq: tgatcgttccAGaagtgcctt
- 4 P: 531 W: 6.72 Seq: tttttatatAGaatgatacag
- 5 P: 687 W: 6.60 Seq: tccccactacAGagaatctgtg

Donor(GT) sites. Threshold 6.099 (90%).

- 1 P: 75 W: 6.38 Seq: gagatttcagGTgatcttaatg
- 2 P: 428 W: 11.00 Seq: gtgtctatcaGTAagtgtgtg Donador canónico
- 3 P: 715 W: 7.78 Seq: attgagccagGTgaaattgact

Donor(GC) sites. Threshold 7.677 (90%).

Secuencia mutada:

Direct chain.

Acceptor(AG) sites. Threshold 4.175 (90%).

- 1 P: 123 W: 5.72 Seq: tcctttctttAGcttattacac
- 2 P: 260 W: 4.58 Seq: ttatatcaccAGaaattggatt Acceptor canónico
- 3 P: 346 W: 4.30 Seq: tgatcgttccAGaagtgcctt
- 4 P: 531 W: 6.72 Seq: tttttatatAGaatgatacag
- 5 P: 687 W: 6.60 Seq: tccccactacAGagaatctgtg

Donor(GT) sites. Threshold 6.099 (90%).

- 1 P: 75 W: 6.38 Seq: gagatttcagGTgatcttaatg
- 2 P: 715 W: 7.78 Seq: attgagccagGTgaaattgact

Donor(GC) sites. Threshold 7.677 (90%).

NetGene2:

Secuencia de referencia:

Donor splice sites, direct strand

pos 5'→3'	phase	strand	confidence	5' exon intron 3'	
428	2	+	0.76	GTGTCTATCA^GTAAGTGTGT	Donador canónico

Donor splice sites, complement strand

No donor site predictions above threshold.

Secuencia mutada:

Donor splice sites, direct strand

No donor site predictions above threshold.

Donor splice sites, complement strand

No donor site predictions above threshold.

NNSPLICE:

Donor site predictions for secuencia de referencia:

Start	End	Score	Exon	Intron
421	435	0.98	tctatcagtaagtgt	Donador canónico

Donor site predictions for secuencia mutada:

Start	End	Score	Exon	Intron
Sitio donador desaparece/no es reconocido				

B) Predictores *in silico* de *splicing* para la variante c.5003-23T>C en *NOTCH2*

FSPLICE:

Secuencia de referencia:

Direct chain.

Acceptor(AG) sites. Threshold 0.000 (100%).

- 1 P: 178 W: 10.15 Seq: catccacAGctctaaag Aceptor Exón27 canónico
- 2 P: 287 W: 3.70 Seq: cgccatacAGgggaccct
- 3 P: 322 W: 6.33 Seq: ctgtcgtcAGtgagtggc
- 4 P: 396 W: 0.80 Seq: cattatgtAGccctctat
- 5 P: 412 W: 9.93 Seq: atttttcAGgactgttg
- 6 P: 482 W: 3.65 Seq: ttttattcAGcactttcc
- 7 P: 510 W: 8.28 Seq: ggactttcAGgtgaatcc Aceptor Exón28 canónico
- 8 P: 528 W: 2.67 Seq: ctgactccAGAacgcact

Donor(GT) sites. Threshold 0.079 (100%).

- 1 P: 51 W: 7.08 Seq: tcccagagGTggggaccct
- 2 P: 212 W: 3.44 Seq: aaccgccaGTgtgttcaa
- 3 P: 318 W: 1.20 Seq: tgtctgtcGTcagtgagt
- 4 P: 322 W: 8.06 Seq: tgtcgtcaGTgagtggca Donador Exón27 canónico
- 5 P: 350 W: 3.02 Seq: ggggaaagGTggcggggg
- 6 P: 423 W: 1.90 Seq: ctgttgagGTcaaagtgtg
- 7 P: 466 W: 1.20 Seq: attgagagGTtaatgttt
- 8 P: 511 W: 7.78 Seq: actttcagGTgaatccct
- 9 P: 597 W: 3.16 Seq: tgctggggGTAatcatgg
- 10 P: 619 W: 7.08 Seq: acgaaagcGTAagcatgg
- 11 P: 688 W: 2.88 Seq: caagcgtcGTgagccagt
- 12 P: 722 W: 8.48 Seq: gggctgaaGTAagaggca Donador Exón28 canónico

Donor(GC) sites. Threshold 5.158 (100%).

Secuencia mutada:

Direct chain.

Acceptor(AG) sites. Threshold 0.000 (100%).

- 1 P: 178 W: 10.15 Seq: catccacAGctctaaag Aceptor Exón27 canónico
- 2 P: 287 W: 3.70 Seq: cgccatacAGgggaccct
- 3 P: 322 W: 6.33 Seq: ctgtcgtcAGtgagtggc
- 4 P: 396 W: 0.80 Seq: cattatgtAGccctctat
- 5 P: 412 W: 9.93 Seq: atttttcAGgactgttg

6 P: 482 W: 3.65 Seq: tttattcAGcacttccc
 7 P: 510 W: 8.28 Seq: ggactttcAGgtgaatcc Aceptor Exón28 canónico
 8 P: 528 W: 2.88 Seq: ctgactccAGaacgcact

Donor(GT) sites. Treshold 0.079 (100%).

1 P: 51 W: 7.08 Seq: tcccagagGTggggacct
 2 P: 212 W: 3.44 Seq: aaccgccaGTgtgttcaa
 3 P: 318 W: 1.20 Seq: tgtctgtcGTcagtgagt
 4 P: 322 W: 8.06 Seq: tgtcgtcaGTgagtggca Donador Exón27 canónico
 5 P: 350 W: 3.02 Seq: ggggaaagGTggcggggg
 6 P: 423 W: 1.90 Seq: ctgttgagGTcaaagtgtg
 7 P: 466 W: 1.20 Seq: attgagagGTaatgttt
 8 P: 511 W: 7.78 Seq: actttcagGTgaatccct
 9 P: 597 W: 3.16 Seq: tgctggggGTAatcatgg
 10 P: 619 W: 7.08 Seq: acgaaagcGTAagcatgg
 11 P: 688 W: 2.88 Seq: caagcgtcGTgagccagt
 12 P: 722 W: 8.48 Seq: gggctgaaGTAagaggca Donador Exón28 canónico

Donor(GC) sites. Treshold 5.158 (100%).

NetGene2:

Secuencia referencia:

Donor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' exon intron 3'
322	1	+	0.93	TCTGTCGTCAGTGAGTGGCA Donador exón 27
619	1	+	0.60	AAACGAAAGC^GTAAGCATGG
722	2	+	0.91	TGGGGCTGAA^GTAAGAGGCA Donador exón 28

Donor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' exon intron 3'
831	19	-	-	0.00	CATGGTCAAT^GTGAGTAAAT

Acceptor splice sites, direct strand

pos 5'->3'	phase	strand	confidence	5' intron exon 3'
178	2	+	0.34	CATCCCACAG^CTCTAAAGTC Aceptor exon 27
186	1	+	0.34	AGCTCTAAAG^TCTTTCTGGA
212	0	+	0.18	CAACCGCCAG^TGTGTTCAAG
412	2	+	0.43	ATTTTTTCAG^GACTGTTGAG
510	1	+	0.38	GGACTTTCAG^GTGAATCCCT Aceptor exon 28
528	1	+	0.19	CTGACTCCAG^AACGCACTCA
539	0	+	0.20	ACGCACTCAG^CTCCTCTATC

Acceptor splice sites, complement strand

pos 3'->5'	pos 5'->3'	phase	strand	confidence	5' intron exon 3'
717	133	0	-	0.26	CTTACTTCAG^CCCCACAGCA
709	141	2	-	0.23	AGCCCCACAG^CATCCTGTCC
591	259	0	-	0.19	TTACCCCCAG^CAGAATAATA
588	262	0	-	0.19	CCCCCAGCAG^AATAATAAAC

NNSPLICE:

Secuencia referencia:

Donor site predictions for 10.42.1.192.3377656.0:

Start	End	Score	Exon	Intron
44	58	0.64	cccagaggtggggac	
315	329	0.48	gtcgtcagtgagtgg	Donador canónico exón 27
612	626	0.96	cgaaagcgtaagcat	
715	729	0.80	ggctgaagtaagagg	Donador canónico exón 28

Acceptor site predictions for 10.42.1.192.3377656.0:

Start	End	Score	Intron	Exon
158	198	0.54	ttctcccttggcatcccacagctctaaagtctttctggaaa	Aceptor canónico exón 27
267	307	0.54	ctggcctctcacgccatacaggggaccctgtcataccctct	
392	432	0.99	tgtagccctctatTTTTcaggactgttgaggtaaagtg	
490	530	0.87	cctgtctctgtggactttcaggtgaatccctgactccagaa	Aceptor canónico exón 28

Secuencia mutada:

Donor site predictions for 10.42.1.192.3382012.0:

Start	End	Score	Exon	Intron
44	58	0.64	cccagaggtggggac	
315	329	0.48	gtcgtcagtgagtgg	Donador canónico exón 27
612	626	0.96	cgaaagcgtaagcat	
715	729	0.80	ggctgaagtaagagg	Donador canónico exón 28

Acceptor site predictions for 10.42.1.192.3382012.0:

Start	End	Score	Intron	Exon
158	198	0.54	ttctcccttggcatcccacagctctaaagtctttctggaaa	Aceptor canónico exón 27
267	307	0.54	ctggcctctcacgccatacaggggaccctgtcataccctct	
392	432	0.99	tgtagccctctatTTTTcaggactgttgaggtaaagtg	
490	530	0.87	cctgtctctgtggactttcaggtgaatccctgactccagaa	Aceptor canónico exón 28