

UNIVERSIDAD POLITÉCNICA DE VALENCIA

DEPARTAMENTO DE BIOTECNOLOGÍA



INCREMENTO DEL DIAGNÓSTICO MÁS ALLÁ DEL ANÁLISIS ESTÁNDAR DEL EXOMA EN UNA COHORTE DE PACIENTES CON LGMD

TRABAJO FIN DE MÁSTER EN BIOTECNOLOGÍA BIOMÉDICA

ALUMNA: Ana Belén Casasús Tomás

TUTOR/A: Dra. Ana Töpf

TUTOR ACADÉMICO: Prof. José Ramón Murguía Ibáñez

Curso Académico: 2018/2019

VALENCIA, 06/09/2018

RESUMEN

La distrofia muscular es un grupo heterogéneo de trastornos neuromusculares. Muchos pacientes presentan debilidad progresiva y desgaste muscular, y pueden tener una esperanza de vida reducida debido a comorbilidades como la insuficiencia respiratoria y cardíaca. Dentro de los diferentes tipos de distrofias se encuentra la distrofia de cinturas (LGMD, *Limb-girdle muscular dystrophy*) cuya prevalencia se estima en alrededor de 20-40/1.000.000, frecuencia probablemente subestimada por la falta de diagnóstico debido a la heterogeneidad genética y clínica. En el año 2014 se propuso realizar un proyecto colaborativo europeo, basado en utilizar la secuenciación del exoma completo en una cohorte de 1466 pacientes con distrofia de cinturas sin diagnóstico, para ayudar a comprender mejor la etiología de la enfermedad y como iniciativa de emplear esta técnica en el diagnóstico rutinario, proyecto denominado MYO-SEQ.

La tasa de casos diagnosticados después de cuatro años ha sido del 54%, estando dentro de los valores esperados y observados en otros estudios. El objetivo de nuestro proyecto se ha centrado en profundizar en el análisis de datos y comprender mejor la falta de diagnóstico en el resto de pacientes, enunciando diferentes hipótesis explicativas: herencia digénica, presencia de variantes intrónicas, papel modificador de variantes patogénicas y estudio funcional de nuevos genes candidatos utilizando el sistema modelo *Danio rerio*. Para ello hemos analizado los datos de la secuenciación del exoma desde la plataforma Seqr, integrando al mismo tiempo la información clínica de los médicos de cada paciente obtenida del software PhenoTips, permitiendo reevaluar los hallazgos clínicos, histopatológicos y de resonancia magnética.

Los resultados obtenidos indican la necesidad de profundizar el análisis en los pacientes todavía no diagnosticados utilizando otras técnicas de secuenciación para dilucidar los posibles mecanismos involucrados en la LGMD, para asegurar en un futuro un diagnóstico más preciso y, por lo tanto, un asesoramiento genético y tratamiento que permita mejorar la calidad de vida de estos pacientes.

PALABRAS CLAVE

Distrofia muscular, LGMD, WES, MYO-SEQ, heterogeneidad clínica, heterogeneidad genética, herencia digénica, variantes intrónicas, variantes modificadoras y *Danio rerio*.

ABSTRACT

Muscular dystrophy is a heterogeneous group of neuromuscular disorders. Many patients present with proximal muscle weakness and wasting, and may have a reduced life expectancy due to respiratory and cardiac failure. Among the different types of dystrophies is the limb-girdle muscular disease (LGMD), whose prevalence is estimated at 20-40/1.000.000, a frequency probably underestimated because of the lack of diagnosis due to genetic and clinical heterogeneity. In 2014, it was proposed to carry out a European collaborative project, based on applying whole exome sequencing to a cohort of 1466 patients with unexplained limb-girdle weakness, to help better understand disease aetiology and as an initiative to use this technique in routine diagnosis; a project called MYO-SEQ.

The diagnostic rate has been 54%, being within the values observed in other studies. The aim of our project has focused on deepening data analysis and better understanding the lack of diagnosis in the rest of the patient cohort, establishing different explanatory hypotheses: digenic inheritance, presence of intronic variants, genetic modifiers and functional work of new candidate genes using *Danio rerio* as a model organism. For this, we have analyzed the exome sequencing data from the Seqr platform, while integrating the clinical information of the reference physicians of each patient obtained from the PhenoTips software, which allows to reevaluate the clinical, histopathological and magnetic resonance imaging reports.

Our findings suggest the need to deepen the analysis in undiagnosed patients using other sequencing techniques to elucidate the possible mechanisms involved in LGMD, to ensure in the future a more accurate diagnosis and, therefore, a genetic counseling and treatment that improves the quality of life of these patients.

KEY WORDS

Muscular dystrophy, LGMD, WES, clinical heterogeneity, genetic heterogeneity, digenic inheritance, intronic variants, genetic modifiers, *Danio rerio*.

AGRADECIMIENTOS

Agradecer al grupo de MYO-SEQ por el apoyo y la ayuda durante este proyecto, en especial a Jennifer, Magdalena, Dan y Volker, quienes me han dado la oportunidad de formar parte del equipo y de integrarme desde el primer día.

A Lorena, Héctor y Fran, por tantos años de amistad y por ayudarme cuando más lo he necesitado.

Agradecer en especial a la persona que ha estado ahí durante todo este tiempo, que ha sido mi mentora y mi amiga, que ha compartido sus conocimientos conmigo y de la que he aprendido y sigo aprendiendo tanto. Gracias por todo Ana.

ÍNDICE

1. INTRODUCCIÓN	1
1.1 LGMD (<i>LIMB-GIRDLE MUSCULAR DYSTROPHY</i>)	3
1.1.1 NOMENCLATURA.....	4
1.1.2 PREVALENCIA	4
1.1.3. DIRECTRICES DE LAS PRUEBAS DIAGNOSTICAS.....	5
1.1.4 NGS (<i>NEXT-GENERATION SEQUENCING</i>).....	9
2. OBJETIVOS	13
3. MATERIAL Y MÉTODOS	14
3.1 PACIENTES.....	14
3.2 SECUENCIACIÓN DEL EXOMA COMPLETO.....	14
3.3 ANÁLISIS DE DATOS	15
3.4 ESTUDIO DE HERENCIA DIGÉNICA	15
3.4.1 DISTROGLICANOPATÍAS	15
3.4.2 SARCOGLICANOPATIAS	17
3.5 ESTUDIO DE VARIANTES INTRÓNICAS	18
3.5.1 COHORTE DE PACIENTES	20
3.5.2 ANÁLISIS BIOINFORMÁTICO	21
3.5.3 DISEÑO DE CEBADORES	21
3.5.4 ANÁLISIS DE LAS VARIANTES.....	21
3.6 ESTUDIO DEL PAPEL MODIFICADOR DE UNA VARIANTE PATOGENICA EN SARCOGLICANO	24
3.7 NUEVOS GENES ASOCIADOS.....	24
3.7.1 OLIGONUCLEÓTIDOS ANTISENTIDO.....	25
3.7.2 MICROINYECCIÓN	26
3.7.3 DESARROLLO EMBRIONARIO.....	26
3.7.4 FENOTIPO Y OBSERVACIÓN DE LA MOTILIDAD	27
3.7.5 INMUNOHISTOQUÍMICA Y ANÁLISIS DE NMJ	27
4. RESULTADOS	29
4.1 HERENCIA DIGÉNICA.....	29
4.1.1 HALLAZGOS GENÉTICOS.....	30
4.1.2 INTERACCIÓN PROTEÍNA-PROTEÍNA.....	31

4.1.3 HALLAZGOS CLÍNICOS	32
4.1.4 HALLAZGOS GENÉTICOS.....	34
4.1.5 INTERACCIÓN PROTEÍNA-PROTEÍNA.....	34
4.1.6 HALLAZGOS CLÍNICOS	34
4.2 ESTUDIO DE VARIANTES INTRÓNICAS	35
4.2.1 DISFERLINOPATÍA	35
4.2.2 CALPAINOPATÍA	36
4.3 ESTUDIO DEL PAPEL MODIFICADOR DE UNA VARIANTE PATOGENICA EN EL COMPLEJO SARCOGLICANO.....	44
4.3.1 HALLAZGOS CLÍNICOS	44
4.4 NUEVOS GENES CANDIDATOS	48
4.4.1 LIPRIN- α -1	48
5. DISCUSIÓN	55
5.1 HERENCIA DIGÉNICA.....	56
5.1.1 SARCOGLICANOPATÍAS	56
5.1.2 DISTROGLICANOPATÍAS	57
5.2 VARIANTES INTRÓNICAS.....	58
5.2.1 LGMDR2.....	58
5.2.2 LGMDR1	59
5.3 GENES MODIFICADORES.....	59
5.4 IDENTIFICACIÓN DE NUEVOS GENES CAUSATIVOS.....	60
5.5 LIMITACIONES DEL ESTUDIO.....	62
5. CONCLUSIONES.....	63
6. REFERENCIAS.....	64

ÍNDICE DE FIGURAS

Figura 1. Distribución de los patrones de debilidad muscular.....	1
Figura 2. Proteínas involucradas en las distrofias musculares.....	2
Figura 3. Superposición fenotípica entra las distrofias musculares.	5
Figura 4. Ejemplos de imágenes de resonancia magnética transversal ponderadas en T1	7
Figura 5. Enfoque conceptual de un paciente con una presunta distrofia muscular de las cinturas.....	8
Figura 6. Proyectos traslaciones de investigación asociados al diagnóstico y a la nueva búsqueda de genes relacionados con LGMD.....	10
Figura 7. Diagrama del procesado de muestras desde la extracción del material genético hasta el informe de los resultados.....	11
Figura 8. Porcentajes de casos resueltos y no resueltos en el proyecto de MYO-SEQ y las causas genéticas más prevalentes.	11
Figura 9. Desarrollo embrionario del Danio rerio durante las primeras 72 horas post-fecundación (hpf).....	26
Figura 10. Diagrama esquemático de la sinapsis neuromuscular en el pez cebra.....	28
Figura 11. Diagrama de flujo empleado para seleccionar las cohortes en ambos estudios de herencia digénica.	29
Figura 12. Interactoma obtenido por STRING entre proteínas asociadas a la O-glicosilación. ...	31
Figura 13. La evaluación inmunohistoquímica de algunos pacientes revela ausencia en la afección del α -dístroglicano.	33
Figura 14. Interactoma obtenido por STRING entre dos proteínas del complejo sarcoglicano. .	34
Figura 15. Electroferogramas del intrón 50 del gen <i>DYSF</i> obtenidos por el software SeqScape™ en dos pacientes de la cohorte..	36
Figura 16. Identificación de una deleción en individuos portadores de la variante.	38
Figura 17. Alineación de las secuencias de aminoácidos utilizando la herramienta BLAST 2 Sequences.....	39
Figura 18. Distribución de las variantes en <i>CAPN3</i> del paciente MESS004 a lo largo de los exones (1-24) y los dominios proteicos (I-IV).....	40
Figura 19. Expresión de calpaína-3 en el tejido muscular del paciente MESS004.....	40
Figura 20. Identificación de un sitio de empalme que crea una inclusión intrónica en dos hermanos con LGMD.....	42
Figura 21. El empalme de ARNm de <i>CAPN3</i> se altera en los pacientes MBEL028 y MBEL029, lo que lleva a la inclusión del pseudoexón patógeno PE 10.....	43
Figura 22. Examen histológico e inmunohistoquímico de la biopsia muscular en MBEL119 mostró cambios en los niveles de la subunidad γ	46

Figura 23. Examen inmunohistoquímico de la biopsia muscular en MBEL050 mostró cambios en los niveles de la subunidad δ	47
Figura 24. Análisis de segregación de la variante c.850 C<T (SGC- α) en la familia de MIST069..	48
Figura 25. Confirmación por el análisis de Western Blot en los tejidos embrionarios de Danio rerio de la expresión de Liprin- α -1..	49
Figura 26. Análisis macroscópico del pez cebra.....	50
Figura 27. Desorganización miofibrilar de los PPFIA-MO.....	53
Figura 28. Estudios de inmunofluorescencia en Danio rerio a 48 hpf..	54
Figura 29. Representación del número de variantes en el gen <i>DYSF</i>	58
Figura 30. Impacto del aumento del número de genes analizados en la tasa resuelta.....	61

ÍNDICE DE TABLAS

Tabla 1. Genes asociados con distroglicanopatías	16
Tabla 2. Variantes identificadas en genes LGMD2 en las bases de datos LOVD (Leiden Open Variation Database) y HGMD (Human Gene Mutation Database).....	19
Tabla 3. Programa utilizado para realizar la PCR con las temperaturas, el tiempo y los ciclos óptimos.....	23
Tabla 4. Variantes encontradas en los genes asociados con distroglicanopatías con un alto grado de patogenicidad en la cohorte de MYO-SEQ.....	30
Tabla 5. Procesos biológicos y funciones moleculares de las once proteínas analizadas con su código de referencia (ID).....	31
Tabla 6. Resumen de los datos clínicos recogidos en la plataforma Phenotips.....	32
Tabla 7. Variantes encontradas en los genes asociados con sarcoglicanopatías con un alto grado de patogenicidad en la cohorte de MYO-SEQ.....	34
Tabla 8. Variantes patogénicas encontradas en el gen <i>DYSF</i> en nuestra cohorte de estudio.....	35
Tabla 9. Resultados de los predictores que evalúan el sitio de empalme.....	37
Tabla 10. Valores predictivos del algoritmo HSF al analizar c.1354+129T>A.....	41
Tabla 11. Resumen de las características clínicas de los cuatro pacientes con una variante patogénica en una subunidad del complejo sarcoglicano.....	44

ABREVIATURAS

AChR: Receptores de acetilcolina
ADN: Ácido Desoxirribonucleico
ADNc: ADN complementario
ARN: Ácido Ribonucleico
CK: Creatina quinasa
CMD: Distrofias musculares congénitas
CMP: Miopatías congénitas
DGC: Distroglicano
EDMD: Distrofias musculares de Emery-Dreifuss
EMG: Electromiografía
ExAC: Consorcio de Agregación de Exomas
gnomAD: Base de datos de agregación del genoma
HE: Hematoxilina-Eosina
HSF: Buscador de empalme humano
hpf: hora post-fecundación
LGMD: Distrofia muscular de las cinturas
ENMC: Centro Europeo Neuromuscular
MAF: Frecuencia del alelo menos común
MO: Morfolino
NGS: Secuenciación de Nueva Generación
NMJ: Unión Neuromuscular
OMIM: Herencia Mendeliana en el hombre
PCR: Reacción en Cadena de la Polimerasa
pb: Pares de bases
RM: Resonancia magnética
SGC: Sarcoglicano
SV2: Proteína de la Vesícula Sináptica
TER: Respuesta de toque táctil
WES: Secuenciación del exoma completo
WGS: Secuenciación del genoma completo

1. INTRODUCCIÓN

Las distrofias musculares son trastornos genéticos, progresivos y degenerativos que causan debilidad del músculo esquelético y la consiguiente sustitución del músculo por tejido graso y fibroso (Wicklund, 2013). Esta pérdida y debilidad muscular se produce de manera variable en los distintos trastornos que conforman las distrofias musculares, siendo la severidad también diversa entre y dentro de estas patologías (Emery, 2002). Las distrofias musculares están constituidas por un grupo heterogéneo de trastornos que suelen ser, en mayor o menor medida, incapacitantes debido a la pérdida de movilidad y al aumento de comorbilidades médicas, como son la insuficiencia respiratoria (Ambrosi et al., 2015) las arritmias (Groh et al., 2012) o la cardiomiopatía (Finsterer et al., 2013) apareciendo principalmente en la segunda década de la vida (Palladino et al., 2016).

Estos trastornos se han clasificado de acuerdo con la distribución clínica, la gravedad de la debilidad muscular y por el patrón de herencia, pudiendo transmitirse como rasgos autosómicos dominantes, autosómicos recesivos, ligados al cromosoma X o aparecer de forma esporádica, como resultado de una mutación de *novo*. La presentación temprana durante la infancia se asocia generalmente con un fenotipo más grave, y conforme los síntomas van apareciendo en la edad adulta la esperanza de vida aumenta de forma significativa (Emery, 2002; Palladino et al., 2016; Wicklund, 2013).

La mayoría de las distrofias musculares se asocian con debilidad muscular progresiva, atrofia y concentraciones elevadas de creatina quinasa sérica (CK) (Roland, 2000). Dentro de estas características clínicas, se observan claras diferencias en el patrón de debilidad muscular que ayuda a los clínicos a poder establecer un posible diagnóstico (Fig.1).

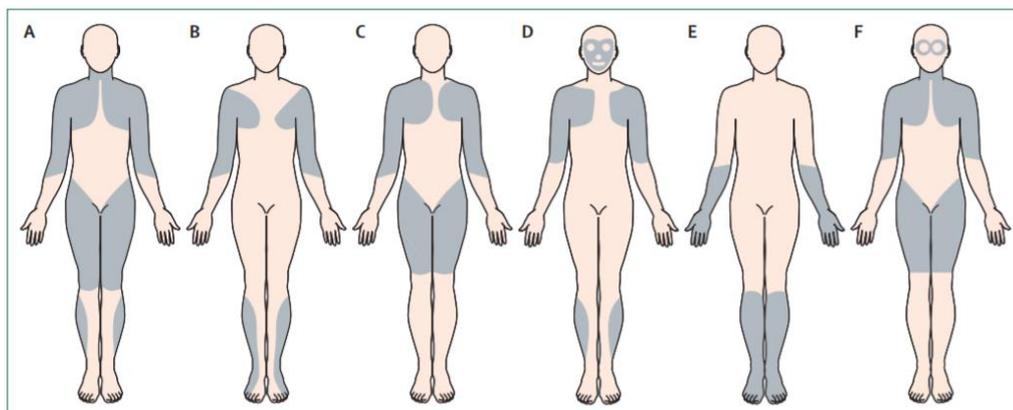


Figura 1. Distribución de los patrones de debilidad muscular

A) Distrofia muscular de Duchenne y Becker. B) Distrofia muscular de Emery-Dreifuss. C) Distrofia muscular de las cinturas. D) Distrofia muscular facioescapulohumeral. E) Distrofia muscular distal. F) Distrofia muscular oculofaríngea (Emery, 2002).

No obstante, los hallazgos de debilidad y concentración sérica elevada de CK son comunes a otros trastornos, como los asociados a la neurona motora inferior, lo que dificulta la exactitud del diagnóstico y su posterior tratamiento (Roland, 2000).

Hasta finales de los años 80 se clasificaban a los pacientes con distrofia muscular mediante el diagnóstico por exclusión, tomando únicamente la sintomatología del paciente como dato a considerar (Mahmood & Jiang, 2014). Sin embargo, el sistema de clasificación ha evolucionado durante las últimas décadas (Zundel & Tyler, 1965) y, actualmente, la mayoría de las distrofias musculares se clasifican en función de la confirmación molecular, es decir, de las proteínas involucradas en la distrofia muscular (Fig. 2), ya que la clasificación establecida por la sintomatología es una clasificación que dificulta el diagnóstico y pronóstico de los pacientes a causa de la gran heterogeneidad clínica (Mahmood & Jiang, 2014).

Entre los principales grupos de distrofias musculares se encuentran: la distrofia muscular de Duchenne y de Becker, facioescapulohumeral, congénita, distal, de Emery-Dreifuss, oculofaríngeal y de las cinturas (Mercuri & Muntoni, 2013). Dentro de los diferentes tipos de distrofias nos vamos a centrar en esta última.

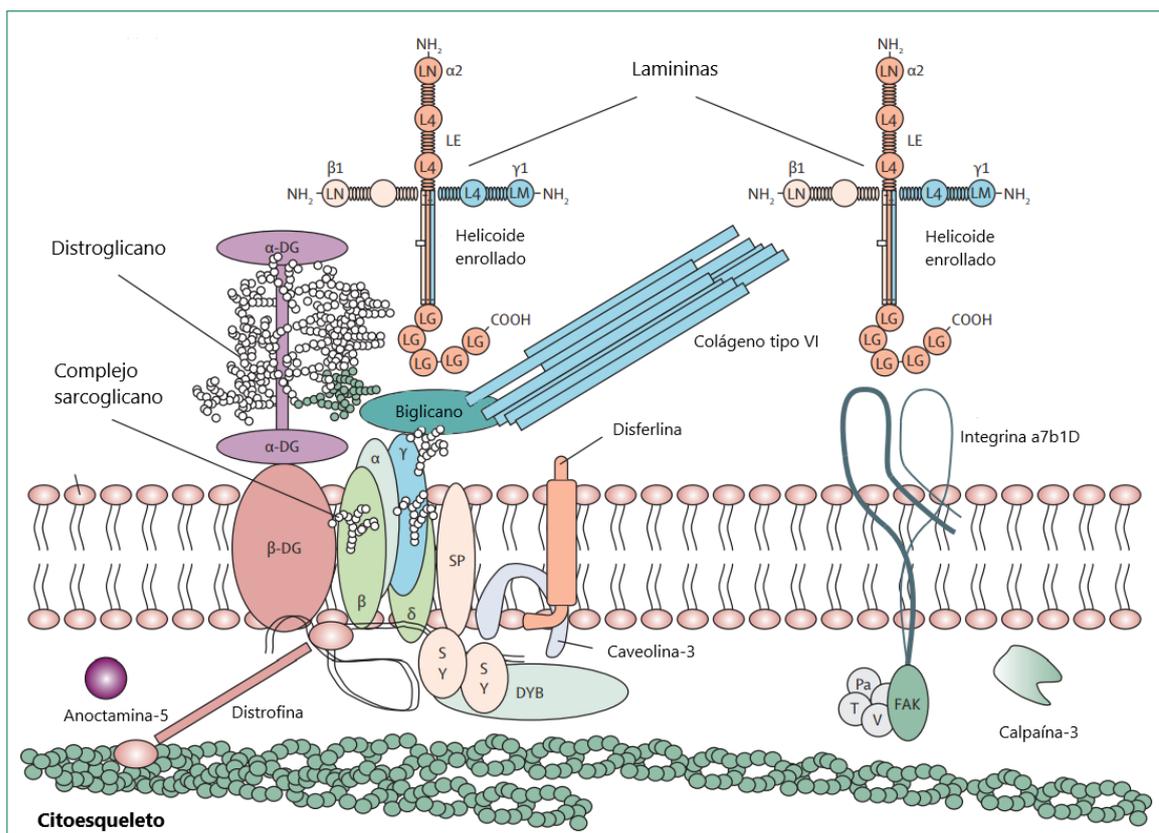


Figura 2. Proteínas involucradas en las distrofias musculares. Figura adaptada a Mercuri & Muntoni, 2013.

1.1 LGMD (*LIMB-GIRDLE MUSCULAR DYSTROPHY*)

Las distrofias musculares de las cinturas (del inglés, *Limb-girdle muscular dystrophy*, LGMD) son un grupo altamente heterogéneo de trastornos musculares, caracterizado por un deterioro progresivo que afecta primero a los músculos voluntarios de las áreas de la cadera y los hombros (Iyadurai & Kissel, 2016). La definición es altamente descriptiva y menos ambigua por exclusión: sin enlace X, sin FSH, no miotónica, no distal, no sindrómica y no congénita, diferenciándose a grandes rasgos de otras distrofias descritas. No obstante, el curso clínico y la expresividad pueden ser variables, desde formas graves con inicio y progresión rápidos hasta formas muy leves que permiten a las personas afectadas tener una vida y niveles de actividad bastante normales (Nigro & Savarese, 2014).

En el año 1954 Walton y Nattrass publicaron su artículo clásico (Walton & Nattrass, 1954) en el que describieron 105 casos de enfermedades musculares de Northumberland y Durham en el norte de Inglaterra. Propusieron una nueva clasificación de los trastornos musculares, basándose en sus observaciones clínicas detalladas y en las ideas ya propuestas previamente por otros investigadores, donde los autores identificaron la LGMD como una nueva entidad clínica.

Su definición ha ido ampliándose en estas últimas décadas, y actualmente recoge los datos estipulados por algunos especialistas en el congreso 229th ENMC (European Neuromuscular Centre), celebrado en el año 2017, donde un conjunto de clínicos y especialistas en enfermedades neuromusculares propusieron una definición más completa de la LGMD: “La distrofia muscular de las cinturas es una afección heredada genéticamente que afecta principalmente al músculo esquelético y conduce a una debilidad muscular predominantemente proximal causada por una pérdida de fibras musculares. Para ser considerada una forma de LGMD, la afección debe describirse en al menos dos familias no relacionadas con individuos afectados que logran caminar de forma independiente, deben tener una actividad elevada de CK, deben demostrar cambios degenerativos en las imágenes musculares durante el curso de la enfermedad, y cambios distróficos en la histología muscular, lo que finalmente lleva a una patología terminal para los músculos más afectados” (Straub et al., 2018).

La LGMD es causada por mutaciones autosómicas recesivas o dominantes en más de 30 genes diferentes que codifican para tipos distintivos de proteínas, con funciones completamente diferentes, ubicadas en todos los compartimentos de las células musculares: sarcolema, membrana nuclear, sarcómero, citosol o espacio extracelular (Mahmood & Jiang, 2014).

La evolución de las técnicas moleculares ha incrementado de forma exponencial la información que se tiene actualmente sobre las causas genéticas asociadas con este tipo de distrofia muscular, teniendo un enorme impacto, desde mejorar la comprensión de los antecedentes

teóricos de esta enfermedad, hasta implicaciones directas en su manejo clínico. Por otro lado, una mejor comprensión de sus bases moleculares ha llevado a definiciones más precisas de las características clínicas asociadas con los subtipos conocidos.

1.1.1 NOMENCLATURA

Hoy en día, existen más de 30 subtipos de LGMD distintos y este número se incrementa considerablemente cada poco tiempo. Al igual que en otras afecciones hereditarias que muestran una gran heterogeneidad genética, la nomenclatura se ha convertido en un problema importante debido a la gran velocidad con que se descubren nuevos genes y a las asignaciones incoherentes a causa de la no ocurrencia de la prueba de patogenicidad en poblaciones control y/o estudios de segregación en pequeñas familias (Di Fruscio et al., 2016).

En la clasificación tradicional, el número 'uno' o 'dos' se asignaba según el modo de herencia: dominante o recesivo, respectivamente, y se asignaba una letra en función del orden de descubrimiento del locus genético o gen. Aunque esta nomenclatura ha estado vigente en las últimas décadas, y es, dentro de lo que cabe, bastante precisa, el aumento de nuevos subtipos ha alertado a los clínicos por la falta de letras para nombrar a los próximos subtipos de LGMD.

En el congreso 229th ENMC se trató en profundidad este problema. Con el fin de facilitar la investigación y evitar los conceptos erróneos, se decidió incluir detalles de la proteína afectada, permitiendo agrupar las LGMD en cohortes con proteínas y fenotipos afectados similares para ensayos clínicos. También se incluyó un número según el orden de descubrimiento de la proteína afectada, reflejando la historia de la LGMD y permitiendo una mejor diferenciación entre los subtipos, y la letra 'D' o 'R' para indicar el modo de herencia.

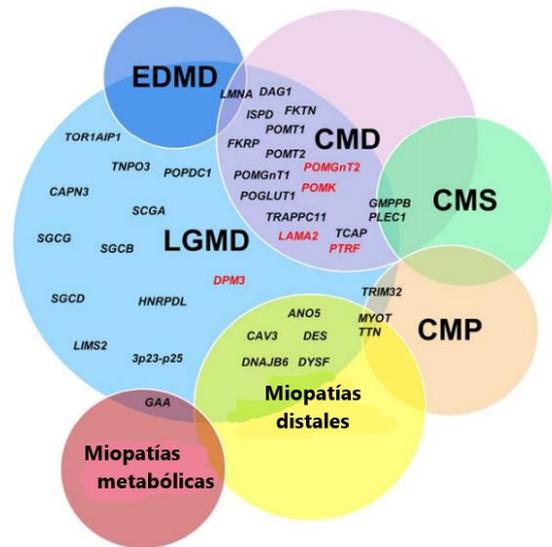
La clasificación de subtipo establecida después del *workshop* internacional fue la siguiente: “LGMD, herencia (R o D), orden de descubrimiento (número), proteína afectada” (Straub et al., 2018). Actualmente, la clasificación cuenta con 9 subtipos de LGMD con herencia autosómica dominante y 26 subtipos con herencia autosómica recesiva (Taghizadeh et al., 2018).

1.1.2 PREVALENCIA

La incidencia estimada de trastornos asociados con distrofia muscular es relativamente baja, sin embargo, el conjunto de todas estas incidencias posiciona a la distrofia muscular como una de las principales causas de morbilidad en la actualidad. Al mismo tiempo, la gran heterogeneidad de estos trastornos a nivel genético y clínico dificulta establecer unos valores estandarizados de las incidencias a nivel global, y es por ello que los estudios asociados con la prevalencia de estas enfermedades se suele realizar a nivel nacional o regional.

Calcular la prevalencia de cada forma de LGMD es actualmente difícil ya que todas las formas son individualmente raras y, para muchos subtipos, los diagnósticos genéticos se han establecido recientemente (Inashkina et al., 2016). Además, la prevalencia de LGMD varía en función del gen involucrado y la región analizada, lo que puede deberse a un efecto fundador de determinadas variantes en cada población.

Figura 3. Superposición fenotípica entre las distrofias musculares. Fenotipos superpuestos entre los genes causantes de LGMD y los responsables de las distrofias musculares congénitas (CMD), las miopatías congénitas (CMP), los síndromes miasténicos congénitos (CMS), las miopatías distales, las distrofias musculares de Emery-Dreifuss (EDMD) y las miopatías metabólicas. El rojo resalta los genes causantes de CMD que también se informaron con el fenotipo LGMD, pero aún no se han clasificado como LGMD (Liewluck & Milone, 2018).



La frecuencia general de todas las formas de LGMD, autosómica recesiva y autosómica dominante, se estima en alrededor de 20-40 / 1.000.000, siendo diferente entre las poblaciones, desde una frecuencia de 8 / 1.000.000 en los Países Bajos y 23 / 1.000.000 en el norte de Inglaterra (Norwood et al., 2009) a 69 / 1.000.000 en el País Vasco (España) (Inashkina et al., 2016). Sin embargo, esto es probablemente una subestimación, dado la variabilidad genética, fenotípica (Fig. 3) y la penetrancia variable e incompleta que presentan este grupo de distrofias (Iyadurai & Kissel, 2016).

1.1.3. DIRECTRICES DE LAS PRUEBAS DIAGNOSTICAS

En la mayoría de los casos se requiere de la revisión del fenotipo clínico, investigaciones de seguimiento de muestras de biopsia muscular, análisis de enzimas séricas y/o imágenes de resonancia magnética para respaldar el diagnóstico genético identificado mediante técnicas de secuenciación (Narayanaswami et al., 2014). Es necesario que los especialistas utilicen este enfoque clínico multidisciplinario para garantizar la precisión del diagnóstico.

La mayoría de los pacientes con LGMD presentan una debilidad simétrica lentamente progresiva principalmente en las extremidades proximales de los brazos y las piernas, sin embargo, son bastante comunes otros patrones de afectación muscular como la debilidad

escapulo-peroneal y la debilidad distal. Es por ello que las pruebas realizadas tienen que centrarse en las características patognomónicas.

Entre las pruebas rutinarias realizadas en el laboratorio se encuentran:

Medición CK en suero

La creatina quinasa (CK) es una enzima que actúa en la homeostasis energética de las células con altos requerimientos de energía, principalmente en el músculo esquelético y cardíaco (McLeish & Kenyon, 2005). Esta enzima es de gran interés clínico, siendo uno de los biomarcadores más utilizados en el diagnóstico de trastornos del músculo esquelético. Los niveles séricos de CK se correlacionan con el daño en las membranas celulares, que provocan la liberación de la enzima del interior celular al torrente sanguíneo. Los niveles normales de CK se encuentran entre: 50-110 $\mu\text{mol/l}$.

El grado de elevación de CK puede ser útil para diferenciar entre diferentes subtipos de LGMD: niveles de CK puede ser normal o ligeramente elevado en transportinopatía, moderadamente elevado en distroglicanopatía y muy elevada en disferlinopatía, calpainopatía y sarcoglicanopatía (Angelini et al., 2018). No obstante, no es una prueba concluyente que determine el tipo de patología, es meramente informativa.

Electromiografía (EMG)

La EMG es una prueba bastante común que se utiliza para evaluar la intensidad de la contracción muscular que, por un lado, mide la respuesta muscular ante una estimulación nerviosa y, por otro lado, mide la actividad eléctrica en el músculo mediante un electrodo de aguja (Schwartz et al., 2017)

La EMG generalmente muestra características miopáticas en todos los subtipos de LGMD y es útil para excluir causas neurogénicas de la debilidad muscular, incluidos los trastornos nerviosos periféricos. Se trata de un estudio consistente excluyente que permite asegurar que la debilidad muscular está ocasionada principalmente por una alteración en el músculo esquelético (Angelini et al., 2018).

Resonancia magnética (RM)

La RM se está empleando cada vez más en la evaluación de pacientes con trastornos neuromusculares al proporcionar un alto contraste de tejidos blandos que permite una excelente evaluación de los músculos estriados en relación con la forma, el volumen, la masa muscular, la arquitectura del tejido y los cambios inflamatorios (Chardon & Straub, 2017). En el caso de las

distrofias musculares, la RM se utiliza como una herramienta adicional para caracterizar la gravedad y el patrón de los músculos afectados (reemplazo del músculo por grasa y/o edema muscular) y preservados, ayudando al diagnóstico y al continuo seguimiento del paciente (Vissing, 2016).

Las imágenes musculares obtenidas muestran patrones característicos que ayudan a distinguir entre las distrofias musculares, y en concreto, entre los diferentes subtipos de LGMD (Fig.4). El grado de patología muscular en la RM se puede describir de forma semicuantitativa utilizando varias escalas, pero cada vez más se están utilizando escalas cuantitativas con una variedad de secuencias diferentes, encargadas de cuantificar la grasa intramuscular (Wattjes et al., 2010).

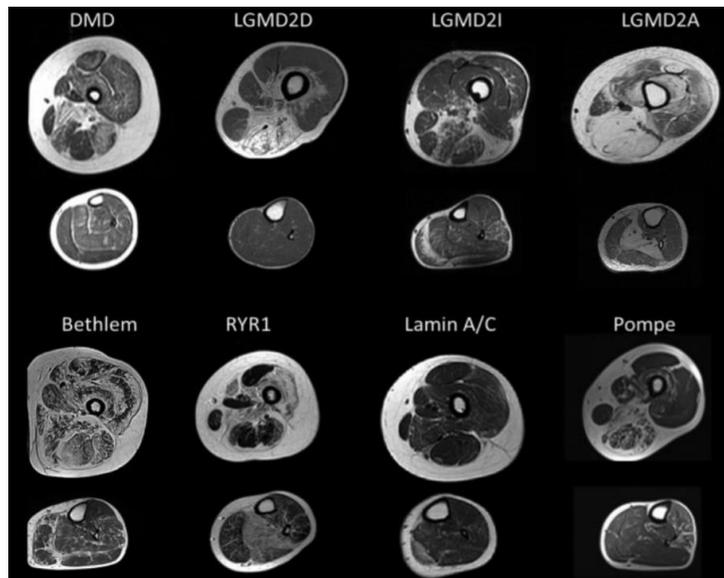


Figura 4. Ejemplos de imágenes de resonancia magnética transversal ponderadas en T1.

Para cada enfermedad, se representa una imagen de la mitad superior de la pierna y del tercio proximal de la parte inferior de la pierna. DMD (Distrofia muscular de Duchenne); LGMD2D (LGMD asociada a sarcoglicano); LGMD2I (LGMD asociada a *FKRP*); LGMD2A (LGMD asociada a calpaina-3); Bethlem (miopatía de Bethlem); RYR1 (miopatía asociada al receptor de rianodina 1); Lamin A / C (laminopatía); Pompe (enfermedad de Pompe).(Chardon & Straub, 2017).

No obstante, existen una serie de limitaciones en los escáneres de resonancia magnética entre las que destaca la falta de estandarización con respecto a la obtención de imágenes, debido a la falta de puntos de referencia a nivel

anatómico, impidiendo conseguir un alto nivel de reproducibilidad y una normalización de los patrones obtenidos (Fischer et al., 2016).

Reconocer patrones de patología mediante imágenes del músculo puede ayudar a confirmar las pruebas genéticas y evitar el procedimiento de técnicas más invasivas, como es la biopsia muscular. Al mismo tiempo, la RM puede optimizar la selección del músculo objetivo para la biopsia muscular (Straub & Bertoli, 2016).

Biopsia muscular

En LGMD las características distróficas musculares, como el aumento de la variabilidad del tamaño de la fibra, el aumento de los núcleos centrales, la fibrosis endomisial, la necrosis y la

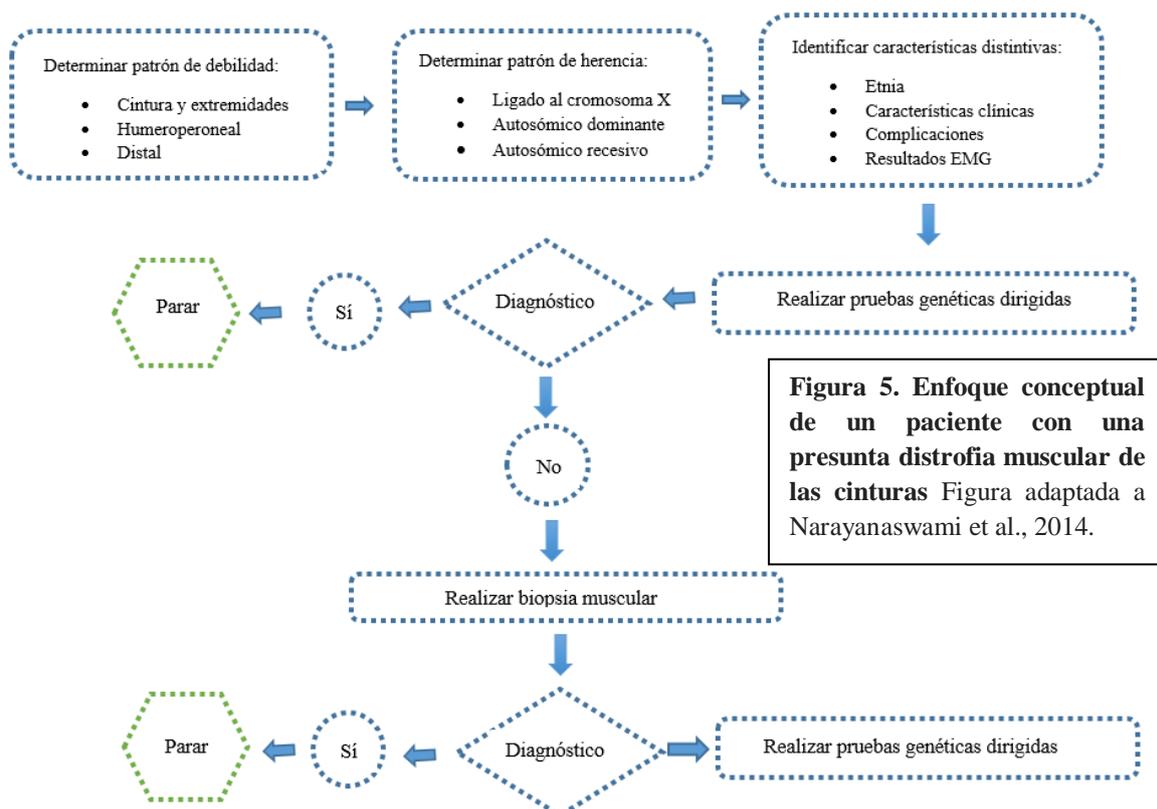
regeneración están muy extendidas, el grado de estos cambios es variable y no siempre se correlaciona con la gravedad de la enfermedad (Sewry & Wallgren-Pettersson, 2017).

Las biopsias musculares se obtienen de diferentes partes del muslo como es el músculo tibial anterior, el gastrocnemio, el deltoides, etc. A continuación, se procede a la tinción, generalmente con hematoxilina-eosina (HE), para la evaluación histopatológica general. Dependiendo de los valores serológicos de la CK, los resultados de la EMG y la RM, se focaliza la evaluación inmunohistoquímica a las regiones que se esperan observar dañadas, mediante la tinción con los anticuerpos correspondientes utilizando protocolos estandarizados (Østergaard et al., 2018).

Se han establecido patrones comunes entre los diferentes subtipos de LGMD que recogen alteraciones como, la miopatía miofibrilar con inclusiones citoplasmáticas o la acumulación de proteínas miofibrilares (miotilina, desmina y alfa-cristalina), característica patológica de algunas formas dominantes (Sewry & Wallgren-Pettersson, 2017). Los cambios estructurales observados en la biopsia muscular junto a la información recopilada en las pruebas anteriores, en algunas ocasiones, suelen determinar el diagnóstico del paciente.

Análisis genético

Entre las pruebas rutinarias de diagnóstico genético destaca el método de secuenciación tradicional, la secuenciación de Sanger. En esta técnica puede emplearse tanto ADN como ADN complementario (ADNc) para detectar las variantes de interés.



El empleo del ADN genómico debe usarse especialmente cuando las características clínicas o las pruebas de proteínas sugieren qué genes deben ser examinados: de esta manera, se puede diagnosticar molecularmente un gran número de casos de LGMD. También puede ser necesaria la utilización del ADNc para la identificación de variantes intrónicas profundas u otras variantes que interrumpen el empalme correcto, siendo estas últimas más del 10% del total de mutaciones patogénicas (Angelini et al., 2018). Por lo tanto, se destaca la dependencia continua en la evaluación clínica y las investigaciones de diagnóstico establecidas para aumentar y validar los hallazgos genéticos.

En el artículo publicado por Narayanaswami et al., 2014 revisan las pruebas diagnósticas y establecen un enfoque algorítmico (Fig.5), que enfatiza el patrón de debilidad, el patrón de herencia y la presencia de características clínicas y de biopsia muscular distintivas para aumentar de manera eficiente el diagnóstico en dichos trastornos.

No obstante, se trata de un proceso secuencial y largo, donde se analizan los genes de uno en uno, por lo que se requieren de otros enfoques para incrementar el diagnóstico en menor tiempo.

1.1.4 NGS (*NEXT-GENERATION SEQUENCING*)

La tasa de diagnóstico obtenida mediante el empleo de técnicas dirigidas, ya sea a través del enfoque de un solo gen o mediante el uso de paneles de genes asociados a LGMD, suele encontrarse entre el 20-60% (Thompson et al., 2016) considerando que el clínico está familiarizado con las características clínicas distintivas referidas a las pruebas.

Hoy en día, la lista de genes que se analizan es demasiado grande para el enfoque gen a gen y, en general, es muy adecuado el uso de paneles de secuenciación dirigida de próxima generación (Monies et al., 2016) que incluyen cualquier gen que hasta el momento está asociado con un cuadro clínico de LGMD. No obstante, la tasa de casos no resueltos sigue siendo altamente significativa y, es por ello, que en los últimos años las colaboraciones internacionales están obteniendo una gran importancia en el abordaje del diagnóstico mediante el uso de programas de secuenciación de LGMD a gran escala (Thompson et al., 2016).

La cohorte de pacientes con un fenotipo LGMD que permanece sin diagnosticar ha sido el foco de recientes investigaciones ómicas internacionales y proyectos de descubrimiento de genes (Fig. 6), basados en el uso de la secuenciación de todo el exoma (WES) y, cada vez más, la secuenciación de todo el genoma (WGS), así como otros enfoques moleculares de alto rendimiento, incluyendo la transcriptómica, para evaluar el efecto de la mutación en la transcripción, y la proteómica, para examinar la resultante expresión de proteínas (Nigro & Savarese, 2014).

Principalmente, WES ha revolucionado la genética clínica al proporcionar un método completo e independiente para la evaluación del paciente. Se trata de un método que tiene el potencial de capturar la región codificante completa de todos los genes conocidos, incluidas las regiones intrónicas flanqueantes, y proporcionar datos de secuencia de estas regiones genómicas enriquecidas con una profundidad de lectura suficiente.

Las tasas de diagnóstico varían de 25 a 50% y WES ha permitido la identificación de nuevos genes y la comprensión de la heterogeneidad clínica y genética de los trastornos mendelianos (Pena et al., 2018).

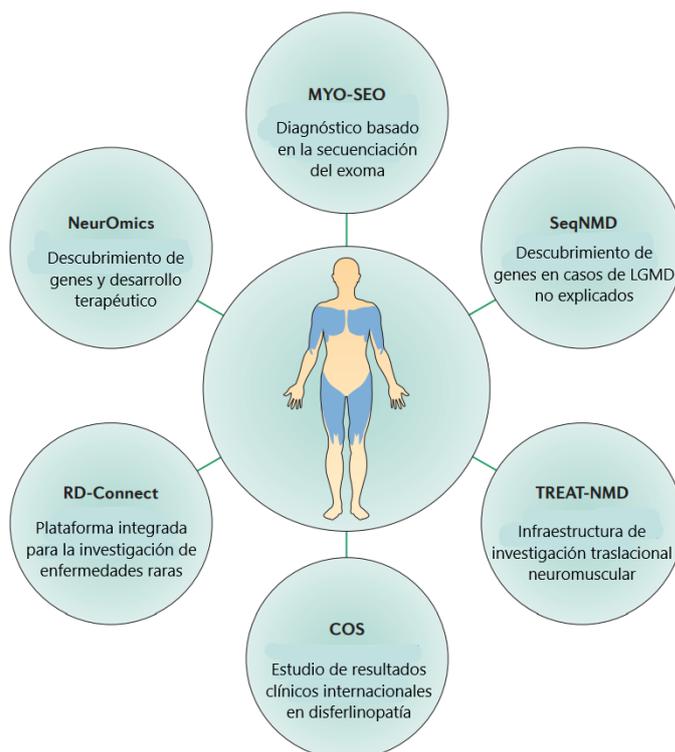


Figura 6. Proyectos traslaciones de investigación asociados al diagnóstico y a la nueva búsqueda de genes relacionados con LGMD (Thompson et al., 2016).

Entre los diferentes proyectos basados en la identificación de genes asociados con los casos no resueltos de LGMD, en este trabajo vamos a centrarnos en el proyecto MYO-SEQ, proyecto establecido en el año 2014 cuyo objetivo se basa en utilizar la secuenciación completa del exoma para (i) contribuir a la ruta diagnóstica de los pacientes afectados por debilidad muscular de las cinturas, (ii) mejorar la conciencia diagnóstica de las enfermedades neuromusculares genéticas raras y (iii) acelerar la integración de tecnologías de secuenciación de próxima generación en la asistencia sanitaria (<http://treatnmd.ncl.ac.uk>).

Se han examinado 1466 pacientes con debilidad inexplicable de las cinturas en busca de variantes potencialmente patógenas que se sabe que están asociadas con enfermedades neuromusculares (Johnson et al., 2016).

La secuenciación de todo el exoma es una nueva herramienta para los médicos neuromusculares, y los hallazgos recientes muestran que mejora el diagnóstico de la LGMD. La técnica tiene un doble papel como herramienta para el diagnóstico y descubrimiento de enfermedades neuromusculares genéticamente heterogéneas, consumiendo menos tiempo que el análisis de proteínas tradicional y la secuenciación de Sanger.



Figura 7. Diagrama del procesado de muestras desde la extracción del material genético hasta el informe de los resultados.

Al ser un proyecto colaborativo el procesado de las muestras pasa por diversos grupos de investigación (Fig.7) hasta conseguir un diagnóstico de los pacientes de MYO-SEQ. Nuestro grupo coordinó el proyecto, al mismo tiempo que participó en el procesado de las muestras previo a la secuenciación del exoma, en el análisis de los datos de WES y en la comunicación de los resultados a los respectivos clínicos.

El proyecto MYO-SEQ ha conseguido mejorar estadísticamente los casos difíciles de diagnosticar gracias al empleo de técnicas de nueva secuenciación, no obstante, aún existe un 44% de casos sin resolver (Fig.8). Este porcentaje no se aleja de los porcentajes obtenidos en otros estudios (Alfares et al., 2018, Ghaoui et al., 2015) demostrando las limitaciones que presenta este abordaje en el diagnóstico de ciertos pacientes.

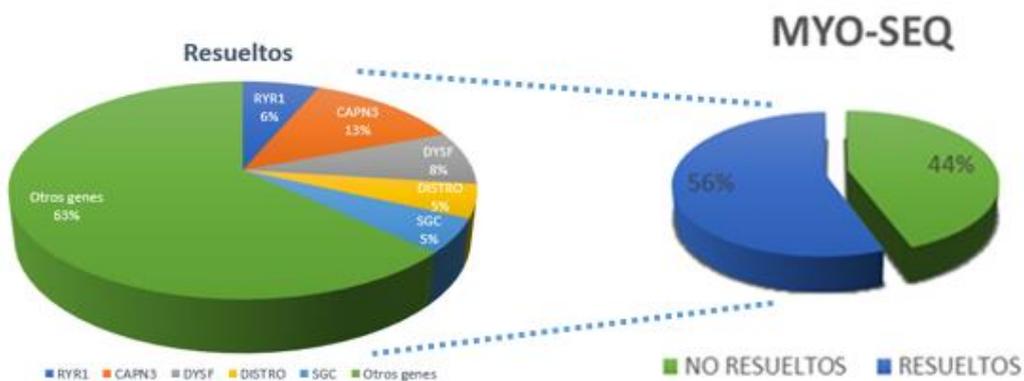


Figura 8. Porcentajes de casos resueltos y no resueltos en el proyecto de MYO-SEQ y las causas genéticas más prevalentes.

Entre las posibles razones que pueden explicar la falta de diagnóstico mediante WES se encuentran: (i) variantes en genes que aún no están asociados con la enfermedad, (ii) variantes que se encuentran dentro de regiones intrónicas profundas y, por lo tanto, no son detectadas por los métodos de captura del exoma, y (iii) limitaciones del método empleado que impiden una identificación eficiente de las variantes (Alfares et al., 2018; Ghaoui et al., 2015; Meienberg et al., 2015).

Otra idea explicativa es la presencia de una nueva condición hereditaria en algunos subtipos de LGMD. En los últimos años, la idea de considerar la LGMD como un conjunto de trastornos oligogénicos ha supuesto un punto de inflexión a la hora de diagnosticar a los pacientes (Fichna et al., 2018) cuestionando la definición clásica de causalidad genética y el concepto de trastornos estrictamente monogénicos, y subrayando la heterogeneidad y complejidad del genoma humano (Schäffer, 2013). Recientemente, se ha publicado un artículo de un caso con un fenotipo típico de LGMD cuyo diagnóstico se corresponde con la presencia de herencia digénica en los genes *DES* y *CAPN3* (Peddareddygari, et al., 2018). Por otro lado, la identificación de genes que tienen un papel modificador en ciertas patologías neuromusculares ha ido estableciéndose en los últimos años (Hogarth et al., 2017; Semplicini et al., 2015; Zhao et al., 2018) explicando, en parte, la gran heterogeneidad fenotípica que presentan algunos pacientes.

Nuestro trabajo se va a centrar en abordar el porcentaje de casos sin resolver en la cohorte de MYO-SEQ atendiendo a las hipótesis explicativas anteriores para esclarecer, en cierto modo, el futuro diagnóstico de diferentes subtipos de LGMD. Para poder abordar estos casos no resueltos se va a enfocar el estudio tanto en la complejidad genética como en la complejidad clínica a partir de un conjunto de genes asociados con LGMD.

Por un lado, la heterogeneidad genética se va a intentar explicar estableciendo la hipótesis de herencia digénica en las distroglicanopatías y sarcoglicanopatías, estudiando variantes intrónicas en el gen *DYSF* y *CAPN3*, y buscando nuevos genes no asociados previamente a distrofias musculares. Por otro lado, se va a analizar la heterogeneidad clínica estudiando variantes patogénicas en los genes sarcoglicanos como posibles modificadores del fenotipo.

2. OBJETIVOS

Previamente a este trabajo se han llevado a cabo otros estudios centrados en identificar nuevos genes candidatos para ampliar el listado de genes que actualmente utiliza MYO-SEQ, para determinar, tras la secuenciación del exoma, la causa genética de 1001 pacientes con distrofia inexplicable de las cinturas. Actualmente, el listado está conformado por 471 genes asociados a enfermedades neuromusculares y el número de pacientes ha incrementado a 1466.

En base al listado de genes asociados, a los casos con y sin diagnóstico y a la información de las bases de datos, se han planteado diversas hipótesis para explicar el 44% de los casos que no están resueltos tras la secuenciación del exoma en nuestra cohorte.

Por lo tanto, el objetivo principal del presente trabajo se centra en explicar este porcentaje tomando como punto de partida la heterogeneidad genética y clínica asociada con la LGMD.

Para ello, los objetivos específicos del trabajo se basarán en:

- 1) Considerar la herencia digénica como posible mecanismo molecular en las distroglicanopatías y sarcoglicanopatías.
- 2) Analizar variantes intrónicas descritas previamente y nuevas variantes intrónicas encontradas en nuestra cohorte que afecten a las proteínas disferlina y calpaína-3.
- 3) Evaluar el posible efecto modificador de una variante patogénica en genes del complejo sarcoglicano para explicar la variedad fenotípica en diferentes familias de estudio que presentan disferlinopatía (LGMDR2) o calpainopatía (LGMDR1).
- 4) Identificar nuevos genes candidatos y llevar a cabo estudios funcionales en nuestro sistema modelo, *Danio rerio*.

3. MATERIAL Y MÉTODOS

3.1 PACIENTES

La aprobación ética (número de referencia REC 08/H0906/28) fue otorgada por el Comité de Ética de Investigación de Newcastle upon Tyne. Se reclutaron 1466 pacientes de 43 centros diferentes en 21 países de Europa y Oriente Medio, y todos dieron su consentimiento por escrito. Los criterios de inclusión estipularon que los pacientes debían presentar debilidad muscular de las cinturas y/o actividad elevada de la creatina quinasa sérica.

A partir de estos pacientes se incluyeron en nuestro proyecto aquellos cuya causa genética no se hubiera confirmado hasta la fecha, siendo el número de estos casos 650.

Se recopiló información fenotípica utilizando el software en línea PhenoTips (<https://phenotips.org>), un software de código abierto que utiliza términos de ontología del fenotipo humano (HPO, del inglés *Human Phenotype Ontology*). Estos datos incluyen el historial médico y familiar, las mediciones físicas y de laboratorio, y los hallazgos físicos, que permiten estandarizar las notas clínicas y comparar a los pacientes.

La edad media de los pacientes fue de 37 años (rango 2-98), y el porcentaje de mujeres y hombres se situó en torno al 50%, siendo un poco superior el de los hombres (53%). Cincuenta y cinco pacientes tenían antecedentes familiares positivos que sugerían una herencia autosómica dominante, 356 pacientes presentaron una afección esporádica siendo en la mayoría de los casos signo de herencia autosómica recesiva, aunque se contempló también la idea de mutaciones de *novo* en el individuo afecto. El resto de casos en nuestra cohorte fueron casos con herencia ligada al sexo o sin información del historial familiar.

La nomenclatura de cada paciente estuvo conformada por cuatro letras que referenciaban el centro de origen y un número que indicaba el orden de llegada de estas muestras a nuestro centro.

3.2 SECUENCIACIÓN DEL EXOMA COMPLETO

La plataforma genómica del Broad Institute de Harvard y MIT (Broad Institute, Cambridge, MA, EE. UU) realizó la secuenciación completa del exoma y el procesamiento de datos de todos estos pacientes. WES se realizó con muestras de >250 ng de ADN a una concentración mayor de 2 ng/μl, utilizando la captura del exoma de Illumina (objetivo de 38 Mb). La alineación y el llamado de variantes se llevaron a cabo utilizando Burrows – Wheeler Aligner y Genome Analysis ToolKit, respectivamente.

Las variantes, finalmente, se cargaron en la plataforma de Seqr del laboratorio MacArthur para el análisis de los datos genómicos (<https://seqr.broadinstitute.org/>).

3.3 ANÁLISIS DE DATOS

La selección de variantes potenciales se restringió a 471 genes asociados con trastornos neuromusculares. Los criterios de filtrado de las variantes potenciales incluyeron una frecuencia de alelo menor (MAF, del inglés *minor allele frequency*) < 1%, un impacto de moderado a alto en la estructura de la proteína (es decir, con cambio de sentido, sin sentido, empalme alternativo, cambio de marco y con desplazamiento del marco de lectura), y un balance alélico superior a 0,10.

La relevancia biológica de las variantes identificadas se determinó considerando (1) la frecuencia de los alelos en la población general detallada por el *Exome Aggregation Consortium* (ExAC) del Broad Institute, (2) la frecuencia de los alelos en la población general detallada por la base de datos de agregación de genomas (gnomAD, *The Genome Aggregation Database*), (3) los valores obtenidos por los cuatro programas de predicción de patogenicidad *in silico*: PolyPhen-2, SIFT, MutationTaster2 y FATHMM, y (4) los informes de patogenicidad recogidos en la base de datos ClinVar y/o en OMIM, y en la literatura publicada hasta el momento.

3.4 ESTUDIO DE HERENCIA DIGÉNICA

El análisis de posibles casos asociados con herencia digénica se ha focalizado en dos subtipos de LGMD muy frecuentes en la cohorte de MYO-SEQ, las distroglicanopatías y las sarcoglicanopatías, debido al conjunto de proteínas asociadas.

3.4.1 DISTROGLICANOPATÍAS

El α -distroglicano (DGC) funciona como un receptor para los ligandos extracelulares como la laminina, mediando la adhesión de las células musculares a la lámina basal. El α -DGC requiere de su glicosilación en los residuos de serina, conocida como O-glicosilación, para su actividad funcional, y la pérdida de esta glicosilación puede llevar a una variedad de síntomas clínicos que incluyen distrofia muscular y una migración neuronal anormal (Michele & Campbell, 2003; Wicklund, 2013). Los trastornos asociados con la glicosilación anormal del α -DGC se conocen colectivamente como distroglicanopatías.

Actualmente, se han identificado mutaciones en 18 genes asociados con distroglicanopatías, todas ellas con herencia autosómica recesiva. Entre estos genes destaca el gen que codifica para α -DGC (*DAG1*), responsable de la distroglicanopatía primaria. El resto de genes, asociados a las distroglicanopatías secundarias, aparecen en la tabla 1 junto a su función proteica.

Tabla 1. Genes asociados con distroglicanopatías. Tabla adaptada a Johnson et al., 2018.

Gen	Función proteica	Distroglicanopatía asociada según OMIM
<i>B3GALNT2</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 11, con anomalías cerebrales y oculares
<i>B4GAT1</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 13, con anomalías cerebrales y oculares
<i>DOLK</i>	Quinasa	Trastorno congénito de glicosilación tipo Im
<i>DPM1</i>	Transferasa	Trastorno congénito de glicosilación tipo Ie
<i>DPM2</i>	Transferasa	Trastorno congénito de glicosilación tipo Iu
<i>DPM3</i>	Transferasa	Trastorno congénito de glicosilación tipo Io
<i>FKRP</i>	Glicosiltransferasa	LGMD2I
		Distroglicanopatía muscular congénita tipo A 5, con o sin retraso mental.
		Distroglicanopatía muscular congénita tipo B 5, con o sin retraso mental.
<i>FKTN</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 4, con anomalías cerebrales y oculares
		Distroglicanopatía muscular congénita tipo B 4, sin retraso mental.
		LGMD2M
<i>GMPPB</i>	Transferasa	Distroglicanopatía muscular congénita tipo A 14, con anomalías cerebrales y oculares
		Distroglicanopatía muscular congénita tipo B 14, con retraso mental.
		LGMD2T
<i>ISPD</i>	Sintasa	Distroglicanopatía muscular congénita tipo A 7, con anomalías cerebrales y oculares
		LGMD2U
<i>LARGE1</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo B 6, con retraso mental.
<i>POMGNT1</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 3, con anomalías cerebrales y oculares
		Distroglicanopatía muscular congénita tipo B 3, con retraso mental.
		LGMD2O
<i>POMGNT2</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 8, con anomalías cerebrales y oculares

<i>POMK</i>	Quinasa	Distroglicanopatía muscular congénita tipo A 12, con anomalías cerebrales y oculares
<i>POMT1</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 1, con anomalías cerebrales y oculares
		Distroglicanopatía muscular congénita tipo B 1, con retraso mental. LGMD2K
<i>POMT2</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 2, con anomalías cerebrales y oculares
		Distroglicanopatía muscular congénita tipo B 2, con retraso mental. LGMD2N
<i>TMEM5</i>	Glicosiltransferasa	Distroglicanopatía muscular congénita tipo A 10, con anomalías cerebrales y oculares

Hasta el momento, solo se han informado casos de distroglicanopatías en los pacientes que presentan mutaciones en el mismo gen, ya sea de forma homocigota recesiva o heterocigota compuesta. Teniendo en cuenta que las distroglicanopatías están causadas por 18 genes que actúan directa o indirectamente en la ruta de la O-glicosilación, enfocamos parte del trabajo a comprobar si dos variantes patogénicas en heterocigosis en dos genes diferentes podían desencadenar un fenotipo típico de distroglicanopatía.

3.4.2 SARCOGLICANOPATIAS

Los sarcoglicanos (SGCs) son un grupo de glicoproteínas transmembrana que actúan fortaleciendo la interacción entre la distrofina y el distroglicano (Moore & Hewitt, 2009). En el músculo esquelético y cardíaco el complejo de sarcoglicano consta de cuatro subunidades proteicas: α , β , γ y δ (Tarakci & Berger, 2016) las cuales están relacionadas entre sí estructural y funcionalmente, pero cada una tiene una ubicación cromosómica distinta (Pegoraro & Hoffman, 2012).

No obstante, se han descrito dos subunidades en los últimos años que, aunque no están directamente relacionados con subtipos de LGMD, se consideran parte del complejo sarcoglicano: subunidad ϵ , altamente relacionado con α -SGC (McNally et al., 1998) y subunidad ζ , altamente relacionado con γ -SGC y δ -SGC (Wheeler et al., 2002). Los genes que codifican estas dos subunidades del SGC tienen una expresión más ubicua, lo que podría explicar porque no se ha asociado ningún subtipo de LGMD a estas subunidades hasta el momento.

Los subtipos de LGMD asociados con el complejo sarcoglicano solo se han informado en los casos donde las dos variantes patogénicas se encuentran en el mismo gen, no obstante, teniendo en cuenta la estrecha asociación dentro del complejo vamos a considerar, como en el caso anterior, la herencia digénica como posible mecanismo desencadenante de LGMD.

A partir de las variantes obtenidas tras el análisis de datos, los casos que presentaron variantes posiblemente patogénicas en dos genes asociados con el distroglicano y su glicosilación (*DAG1*, *POMT1*, *POMT2*, *POMGNT1*, *POMGNT2*, *POMK*, *DPM1*, *DPM2*, *DPM3*, *DOLK*, *GMPPB*, *ISPD*, *FKTN*, *FKRP*, *LARGE1*, *B3GALNT2*, *B4GAT1*, *TMEM5*) o variantes en dos genes diferentes del complejo sarcoglicano (α , β , δ , γ , ϵ , ζ) fueron incluidos en nuestras cohortes de estudio. Al mismo tiempo, gracias a la información recogida en la plataforma PhenoTips se eliminaron los casos con antecedentes familiares que determinaban una herencia autosómica dominante o asociada al cromosoma X.

A continuación, se realizó un sesgo de los datos tomando diferentes factores como marcadores de patogenicidad:

- 1) Identificada en ClinVar como patogénica o probablemente patogénica.
- 2) Predicción perjudicial por parte de las herramientas *in silico* disponibles en la plataforma Seqr.
- 3) Variantes con un impacto alto en la estructura proteica (sin sentido, con desplazamiento del marco de lectura).
- 4) Variantes múltiples, es decir, variantes observadas en más de tres casos, con un valor $MAF < 1\%$.

Para estudiar si variantes en diferentes genes eran la causa del fenotipo del paciente, se pidió información clínica de cada paciente a sus centros correspondientes, junto a imágenes de la biopsia muscular y la RM.

3.5 ESTUDIO DE VARIANTES INTRÓNICAS

La identificación de variantes patogénicas en las regiones no codificantes suele ser una de las principales causas de falta de diagnóstico tras realizar la secuenciación completa del exoma. Para analizar la presencia de estas variantes en nuestra cohorte nos centramos en dos genes cuya proporción de variantes en las regiones no codificantes es superior al resto de genes: *DYSF* y *CAPN3* (Tabla 2).

DISFERLINOPATÍAS

El gen *DYSF* comprende 55 exones y abarca más de 150 kb, codificando una proteína transmembrana de 230 kDa que se expresa altamente en el músculo, denominada disferlina. Esta proteína se localiza en el sarcolema de las células musculares y se sabe que desempeña un papel en la remodelación del citoesqueleto, la translocación de vesículas, la fusión de la membrana y la reparación de la membrana muscular dependiente de calcio después de un daño (Cacciottolo et al., 2011; Lek et al., 2012). La distrofia muscular deficiente de disferlina se denomina disferlinopatía, siendo el LGMDR2 la más frecuente (Nguyen et al., 2005).

Se han determinado al menos 450 variantes patogénicas en *DYSF* debido a su gran tamaño y al avance de las técnicas moleculares. Frente a los avances en la determinación de variantes patogénicas, se han publicado en los últimos años casos típicos de LGMD provocados por variantes en las regiones exónicas e intrónicas. Actualmente, existen dos variantes intrónicas directamente relacionadas con esta enfermedad. Ambas variantes provocan un empalme aberrante en el ARNm de *DYSF* como un nuevo pseudoexón, tanto en el intrón 44 como en el intrón 50 (Dominov et al., 2014; Dominov et al., 2019).

Tabla 2. Variantes identificadas en genes LGMD2 en las bases de datos LOVD (*Leiden Open Variation Database*) y HGMD (*Human Gene Mutation Database*) (Di Fruscio et al., 2016).

LGMD2					
Gen	Sin sentido	Nonsense	Silente	Intrónica	Total
<i>ANOS</i>	19	5	0	5	29
<i>CAPN3</i>	214	29	4	42	289
<i>DAG</i>	1	1	0	0	2
<i>DYSF</i>	105	60	3	45	213
<i>FKRP</i>	82	12	1	0	95
<i>FKTN</i>	12	9	0	0	21
<i>POMGNT1</i>	19	7	0	9	35
<i>POMT1</i>	20	8	0	3	31
<i>POMT2</i>	22	2	0	6	30
<i>SGCA</i>	50	5	0	8	63
<i>SGCB</i>	21	5	0	5	31
<i>SGCD</i>	5	3	0	0	8
<i>SGCG</i>	10	4	0	3	17
<i>TCAP</i>	0	3	0	0	3
<i>TRIM32</i>	3	0	0	0	3
<i>TIN</i>	8	2	0	1	11
Total	591	155	8	127	881

La mutación c.4886+1249 G>T se produce en un sitio aceptor de empalme conservado en el extremo 5' del pseudoexón 44 (PE 40), y la mutación puntual c.5668-824 C> T crea un sitio consenso de donante de empalme en el extremo 3' de la secuencia del pseudoexón 50 (PE 50). Se predice que la inclusión de PE40 o de PE50 dará como resultado una inserción de 59 aminoácidos o 46 aminoácidos, respectivamente, y una terminación prematura de la traducción dentro del dominio C2G y la pérdida del dominio transmembrana C-terminal, dominios necesarios para la función de la proteína disferlina. Ambas variantes fueron secuenciadas en nuestro estudio.

CALPAINOPATÍAS

El gen *CAPN3* está formado por 24 exones que abarcan cerca de 53 kb de ADN, el cual codifica una proteasa de cisteína no lisosomal, denominada calpaína-3, involucrada en la regeneración muscular, la remodelación del sarcómero, la regulación del citoesqueleto y la homeostasis del calcio (Kramerova et al., 2007). El fenotipo característico causado por deficiencia en la calpaína-3 se denomina LGMDR1.

Hasta la fecha, se han identificado más de 300 mutaciones patógenas diferentes en el gen *CAPN3*, la mayoría de las cuales se encuentran en la región de codificación del gen y representan sustituciones de aminoácidos no sinónimas (Blázquez et al., 2008). Sin embargo, se estima que cerca del 25% de las mutaciones sean sustituciones sinónimas, así como deleciones e inserciones ubicadas en los límites intrón-exón que actúan modificando el empalme del ARNm precursor (pre-ARNm) (Blázquez et al., 2013).

Dentro de la cohorte de MYO-SEQ se detectaron varias mutaciones intrónicas presentes en los límites de exón-intrón cuyo efecto no estaba esclarecido. Las variantes seleccionadas presentaron, por un lado, valores probablemente patogénicos por parte de los programas predictivos y, por otro lado, una frecuencia elevada en los pacientes de nuestro estudio. Para la demostración experimental de la patogenidad de estas variantes fue necesario analizar el ARNm de *CAPN3* para determinar si había un empalme aberrante inducido por estas alteraciones: c.1354+129T>A y c.1746-20C>G.

3.5.1 COHORTE DE PACIENTES

Para establecer un grupo de muestra se llevaron a cabo varios filtrados para seleccionar los casos con una variante probablemente patogénica en la región codificante del gen *DYSF* o *CAPN3*, y cuyo fenotipo se relacionara con los casos típicos de disferlinopatías o calpainopatías. Tras esta serie de filtrados se tuvieron en cuenta los factores de patogenidad establecidos en las cohortes de sarcoglicanopatías y distroglicanopatías.

Se obtuvo un listado de 22 casos que presentaban una variante patogénica o probablemente patogénica en la región codificante del gen *DYSF* y a partir de esta cohorte se procedió al estudio de las variantes c.4886+1249 G>T y c.5668-824 C> T. Por otro lado, se consiguió un listado de 20 casos que presentaban una variante patogénica en la región codificante de *CAPN3* y que al mismo tiempo eran portadores de una de las variantes intrónicas cerca del límite exón-intrón que se iban a estudiar.

3.5.2 ANÁLISIS BIOINFORMÁTICO

La predicción del efecto potencial de las variantes intrónicas se determinó en primera instancia mediante los siguientes programas de predicción de sitios de empalme (PPSE) : *Human Splicing Finder* (<http://www.umd.be/HSF3/HSF>) y *SD-Score* (https://www.med.nagoya-u.ac.jp/neurogenetics/SD_Score/sd_score.html). La variación de la puntuación en el uso del sitio de empalme canónico y la activación potencial de sitios de empalme crípticos en presencia de mutaciones se evaluaron mediante el análisis de secuencias de exones e intrones utilizando la configuración predeterminada.

3.5.3 DISEÑO DE CEBADORES

Se utilizó el genoma de referencia hg19/GRCh37 y el diseño se realizó a partir de la base de datos genómicos, Ensembl (<http://www.ensembl.org/index.html>), donde se hizo la búsqueda de la región de interés, seleccionando un fragmento de secuencia flanqueante tanto *upstream* como *downstream* a la variante, superior a 200 pb. Determinada la secuencia se procedió al diseño de los cebadores empleando la herramienta bioinformática Primer3, y dos herramientas bioinformáticas que nos permitieron obtener información acerca de la calidad y especificidad de cada pareja de cebadores: UCSC In-Silico PCR y Nucleotide BLAST.

Los pares de cebadores que abarcan la unión del exón 44-45 en *DYSF* generan un amplicón de 603 pb a partir del ADN, mientras que el amplicón que abarca la unión del exón 50-51 tiene un tamaño de 332 pb. El mismo procedimiento se llevó a cabo para diseñar los cebadores de *CAPN3*, sin embargo, para estudiar si afectaban de alguna forma las variantes intrónicas al sitio de empalme se diseñaron varios cebadores por cada variante, un par que abarcara la región de interés y el otro par que estuviera en el límite del exón-intrón.

3.5.4 ANÁLISIS DE LAS VARIANTES

A partir del material genético de los pacientes se procedió al análisis de las variantes intrónicas, sin embargo, la metodología no fue la misma en los dos estudios. Para estudiar si los casos de nuestra cohorte presentaban las variantes en *DYSF* se realizó la secuenciación Sanger a partir

del ADN genómico de cada paciente. Por otro lado, para estudiar si las variantes intrónicas identificadas en la cohorte de *CAPN3* afectaban al empalme del transcrito, se pidieron a los centros correspondientes muestras de sangre y tejido muscular de los pacientes para extraer el ARN.

Las extracciones del ADN se llevaron a cabo utilizando el kit PureLink Genomic DNA Mini Kit (Ambion), mientras que las extracciones del ARN se realizaron empleando el kit mirVana™ miRNA Isolation (Ambion), mediante el uso de los protocolos estandarizados, utilizando sangre periférica y 30 µg del tejido muscular. La concentración, calidad e integridad del ADN y el ARN se analizó mediante espectrofotometría utilizando el NanoDrop® ND-1000 Spectrophotometer.

RT-PCR

Una vez cuantificado el ARN, se añadió un volumen 0,1 del buffer 10X ADNasa y 1 µl de rADNasa y se incubó a 37°C durante 30 minutos. Posteriormente, se resuspendió en un volumen de 0,1 con el reactivo que inactiva la ADNasa y se dejó a temperatura ambiente durante 2 minutos. Por último, se centrifugó durante 1 minuto a 10.000 rpm y se transfirió el ARN a un nuevo tubo.

A continuación, se realizó la reacción en cadena de la polimerasa con transcriptasa reversa (del inglés *Reverse transcription polymerase chain reaction*, RT-PCR), empleando el kit comercial High-Capacity cDNA Reverse Transcription Kit y la polimerasa MultiScribe™ Reverse Transcriptase (Applied Biosystems), siguiendo el protocolo estandarizado.

El programa para sintetizar el ADNc se dividió en un paso de preincubación para activar la enzima, la retrotranscripción y un paso final para desnaturalizar, diferenciándose en la temperatura y tiempo de incubación: una incubación de 25°C durante 10 minutos, una siguiente incubación a 37°C durante 2 horas, y a 85°C para finalizar durante 5 minutos. A partir de este punto el protocolo a seguir con el estudio de ambos genes fue el mismo.

PCR

La PCR se realizó utilizando el kit MyTaq HS Red Mix 2X (BioLine), empleando el tampón, los dNTPs, los cebadores y la MyTaq polimerasa a los volúmenes indicados. En la tabla 3 se encuentra el programa del termociclador utilizado para la correcta amplificación, siendo la temperatura de hibridación la única que variamos en función de los cebadores:

<p>Tabla 3. Programa utilizado para realizar la PCR con las temperaturas, el tiempo y los ciclos óptimos.</p>
--

Programa de PCR	Temperatura	Tiempo	Ciclos
Activación inicial	95°C	1:00	
Desnaturalización	95°C	0:15	35
Hibridación	- °C	0:15	35
Extensión	72°C	0:20	35
Extensión final	72°C	0:20	
Conservación	4°C	∞	

Electroforesis en gel de agarosa

Se empleó la técnica de electroforesis en gel de agarosa para analizar los fragmentos de ADN y ADNc obtenidos tras la amplificación por PCR, para separar los fragmentos amplificados por su tamaño y carga. Se preparó el gel con un porcentaje de agarosa del 2%: 2 g de Seakem LE Agarose y 100mL de tampón 1X TAE (Tris, acetato y EDTA).

Una vez calentada la agarosa en el tampón se añadió el agente intercalante GelRed®, y se corrió el gel durante 40 minutos a 80 voltios.

Amplificación bandas del gel de agarosa

Para estudiar la presencia de un empalme aberrante en las variantes de *CAPN3* se realizaron amplificaciones de las bandas obtenidas en la electroforesis, utilizando el kit de extracción de gel Qiaquick (Qiagen). Este kit se emplea para extraer y purificar ADN de 70 pb a 10 kb a partir de geles de agarosa de bajo punto de fusión en tampón TAE o TBE. Para ello, se cortaron las bandas de interés con un bisturí, se pesaron y se añadió un volumen x3 del tampón QG (tampón que actúa como indicador del pH óptimo). A continuación, lo dejamos incubar a 50°C durante 10 minutos para disolver el gel. Una vez disuelto se añadió un volumen x1 de isopropanol y posteriormente se pasó la muestra a las columnas QIAquick spin, donde se llevaron a cabo los lavados con los tampones respectivos, y finalmente el ADNc se eluyó en un tampón EB.

Los amplicones se purificaron utilizando el reactivo ExoSAP-IT™, reactivo que elimina los contaminantes gracias a la exonucleasa I, que degrada el ADN de cadena sencilla, y a la fosfatasa alcalina, encargada de desfosforilar los dNTPs. Se incubaron las muestras durante 15 minutos a 37°C para la activación de la enzima, y durante otros 15 minutos a 80°C para su inactivación.

Secuenciación Sanger

Cada muestra se traspasó a un nuevo tubo y se le añadieron 2 µl de un solo cebador. Los cebadores empleados en la secuenciación fueron los mismos que los empleados en la amplificación por PCR.

Las secuenciaciones se realizaron en Eurofins (<https://www.eurofinsgenomics.eu/>). Cuando los resultados de las secuenciaciones ya estuvieron listos, se seleccionaron los archivos adjuntos con la extensión. abi. para proceder con el alineamiento de las secuencias.

El software empleado para esta función fue el SeqScape. Dentro de su paquete de análisis están incluidos diferentes softwares: GeneMapper® (genotipado), Variant Reporter (comparaciones de secuencias) y el de análisis de secuenciación.

En este proyecto se empleó el software principalmente para identificar las variantes mediante la comparación de las secuencias de nuestros casos y las secuencias de la población control.

3.6 ESTUDIO DEL PAPEL MODIFICADOR DE UNA VARIANTE PATOGENICA EN SARCOGLICANO

Para estudiar si la presencia de una variante patogénica en alguno de los genes que codifican el complejo sarcoglicano agrava el fenotipo de los casos diagnosticados con LGMD, se recogieron algunos pacientes en los que se había identificado la causa genética del desorden neuromuscular que padecían, pero cuyo fenotipo era más severo de lo esperado, y que al mismo tiempo presentaban una variante patogénica en alguno de los genes sarcoglicano descrita en Clinvar. Descartamos los casos diagnosticados con genes asociados a subtipos de LGMD severos con un componente neurogénico, centrándonos en pacientes que presentaban calpainopatía o disferlinopatía.

Atendiendo a estos criterios nuestra cohorte se redujo significativamente, siendo el número muestral de cuatro casos, dos de estos casos fueron diagnosticados previamente como LGMDR2 y los otros dos casos como LGMDR1.

En esta parte del proyecto se llevaron a cabo dos abordajes diferentes en la cohorte. Por un lado, se realizaron en los casos con presencia de familiares afectados segregaciones de las variantes patogénicas, para comparar el fenotipo clínico en presencia o ausencia de la variante de sarcoglicano en los familiares. Por otro lado, se estudió la histopatología en secciones de tejido muscular para observar la reducción de la expresión de las subunidades de sarcoglicano.

3.7 NUEVOS GENES ASOCIADOS

Parte de los pacientes no presentaron variantes patogénicas en ninguno de los 471 genes asociados a enfermedades neuromusculares. Para analizar nuevos genes, se filtraron las variantes con una alta expresión en el músculo esquelético que hasta el momento no habían sido asociados a enfermedades neuromusculares. A continuación, se realizó una búsqueda bibliográfica para estudiar la funcionalidad de los nuevos genes junto a la información recopilada en OMIM. Una vez reunida la información publicada de los nuevos genes y su

función proteica, comparamos los niveles de expresión tisular de las proteínas recogidas en el proyecto Expresión Genotipo-Tejido (<https://gtexportal.org>) para determinar su expresión en el músculo esquelético.

Los candidatos con una expresión elevada en el tejido muscular y cuyos datos apoyaban su papel en el correcto funcionamiento de la unión neuromuscular, fueron listados, y se procedió a la recolección de muestras de los pacientes y sus familiares para realizar las segregaciones. Ante unos resultados positivos se llevó a cabo el estudio funcional, utilizando como organismo modelo uno de los sistemas más empleados en la investigación biomédica: el *Danio rerio*, comúnmente conocido como pez cebra, para finalmente determinar una posible causalidad.

Danio rerio

Su breve tiempo de generación lo hace susceptible de manipulación y análisis genéticos, y su pequeño tamaño y alta fecundidad lo hacen especialmente adecuado para ensayos a gran escala. Además, hay varios factores que hacen que el *Danio rerio* sea un organismo útil para estudiar la sinaptogénesis, lo que le convierte en un buen sistema modelo para estudiar nuevos genes candidatos asociados con las enfermedades neuromusculares.

En este estudio se trabajó a partir de la línea *Golden* (*slc24a5b1/+*), una cepa hipopigmentada que hace más visible los tintes vitales, marcadores fluorescentes, anticuerpos y las sondas de hibridación de ARN.

3.7.1 OLIGONUCLEÓTIDOS ANTISENTIDO

Para estudiar la funcionalidad de los genes en nuestro organismo modelo se modificaron genéticamente los embriones mediante el uso de morfolinós (MO) generando *knockdowns*. Los MO son oligonucleótidos sintetizados químicamente que se inyectan típicamente en la etapa celular del embrión, se unen a los ARNm diana complementarios y evitan su traducción o empalme. Son similares a los pequeños ARN de interferencia (ARNi) y los ARN de horquilla corta (ARNhc), ya que interfieren con la función de un gen sin alterar su secuencia.

Se diseñaron los MO para bloquear la traducción a partir de la plataforma *Gene Tools* (<https://www.gene-tools.com>, Pilomath, OR), y al mismo tiempo se adquirieron los controles estándar, dirigidos a un gen de β -hemoglobina humana (5'-CCTCTTACCTCAGTTACAATTTATA-3'). Este control se usó como control negativo de los efectos de la inyección de MO. Los cigotos se inyectaron con 5 ng, 10 ng y 15 ng del MO o el control-MO siguiendo los protocolos estándar.

El MO se disolvió en *Danieau buffer* para obtener una concentración final de 5mM. A continuación, se diluyó a una concentración final de 1mM con rojo de fenol al 0,05%. Una vez

preparado el morfolino se incubó a 65°C durante 5 minutos y rápidamente se colocó en hielo durante 1 minuto, almacenándose posteriormente a -20°C.

3.7.2 MICROINYECCIÓN

Para proceder con la microinyección se recolectaron los embriones del pez cebra, los cuales se enjuagaron con agua azul (agua del sistema que contiene solución de azul de metileno) y se trasladaron a una placa Petri que contenía medio E3.

Con una pipeta Pasteur se transfirieron alrededor de 20-30 embriones a una placa de microinyección, utilizando principalmente embriones en la etapa de una y dos células (Fig.9) para inyectar el morfolino. Una vez inyectados todos los embriones en la placa de microinyección, se transfirieron a una placa Petri limpia que contenía medio E3 y se incubaron a 28.5°C.

3.7.3 DESARROLLO EMBRIONARIO

Los embriones de pez cebra se criaron en medio E3 a 28,5 °C y se estadificaron horas después de la fertilización (hpf) según protocolos estándar (Kimmel et al., 1995), observando al microscopio los estadios a las 24,48 y 72 horas (Fig.9).

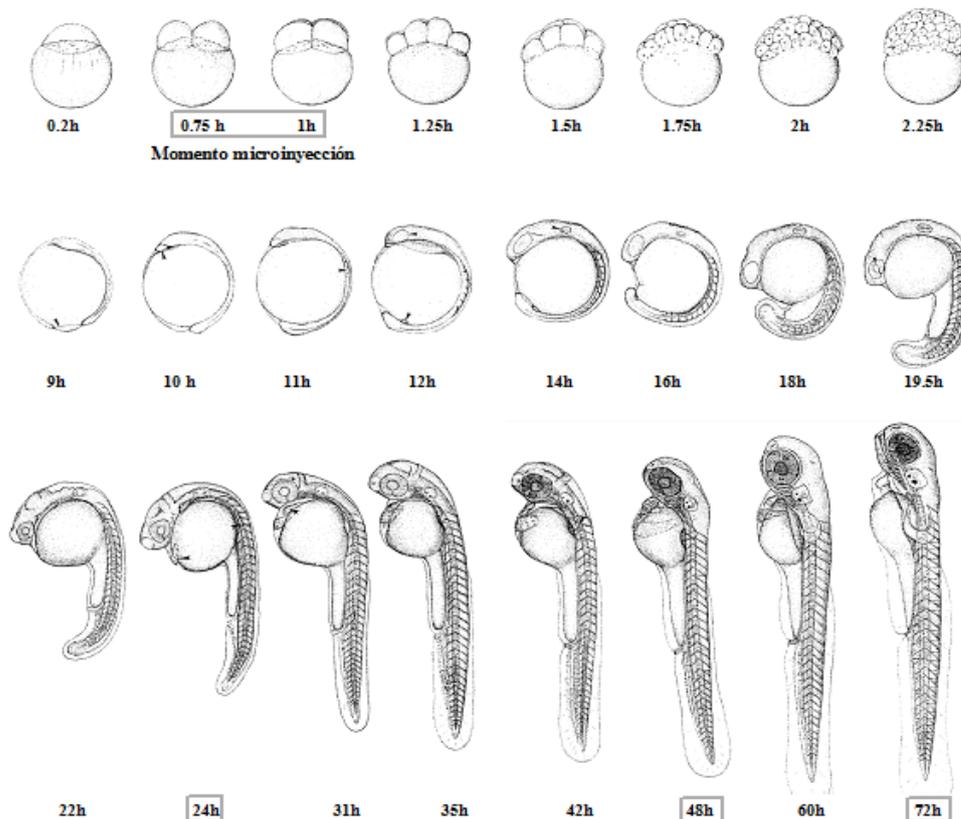


Figura 9. Desarrollo embrionario del *Danio rerio* durante las primeras 72 horas post-fecundación (hpf) (Kimmel, et al., 1995).

3.7.4 FENOTIPO Y OBSERVACIÓN DE LA MOTILIDAD

Debido a que la LGMD es una enfermedad neuromuscular, observamos el primer movimiento esquelético realizado durante el desarrollo embrionario: el enrollamiento espontáneo. La primera etapa de comportamiento locomotor observada en el pez cebra es la alternancia espontánea de movimientos de la cola dentro del corion a partir de las 17 hpf, y a partir de las 21 hpf comienzan a enrollarse en respuesta al tacto.

Los giros de la cola en el corion a 24 hpf se midieron observando cada embrión durante 45 segundos y contando el número total de giros completos. Se midieron un total de 15 embriones para cada categoría (WT, 5ng, 10ng, 15 ng). La respuesta a la natación provocada por el tacto (del inglés *touch-evoked response*, TER) se observó tocando la cabeza o la cola del pez cebra con una punta de pipeta fina. La TER se definió como anormal si los embriones exhibían movimientos en círculos, se contraían sin alejarse del estímulo o no tenían ningún movimiento como respuesta.

Los videos se importaron a Fiji ImageJ como películas .avi. Dentro de los complementos del programa se utilizó Trackmate para medir el movimiento del pez cebra, con edición manual de cada cuadro para garantizar que solo se detectara el pez cebra y que el movimiento identificado fuera preciso. Se exportaron los valores para analizar la distancia recorrida (mm).

El análisis estadístico se realizó utilizando el software RStudio (versión 3.6.1). Una vez seleccionados los embriones que fueron analizados, se realizaron las pruebas de Shapiro-Wilk y de Levene para comprobar la normalidad y homocedasticidad de los datos, respectivamente. Esto llevó a emplear la prueba paramétrica ANOVA para estudiar las medias de las distancias recorridas entre los 4 grupos de embriones (WT, 5ng, 10 ng y 15 ng). Los experimentos *in vivo* fueron cegados antes de la grabación *in vivo* y para la adquisición de imágenes.

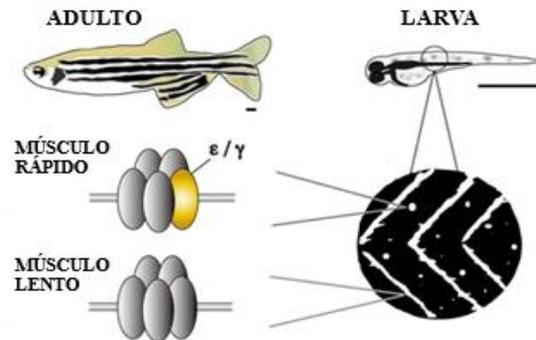
3.7.5 INMUNOHISTOQUÍMICA Y ANÁLISIS DE NMJ

Para observar la afección de la unión neuromuscular y el músculo estriado (Fig.10) se realizó una prueba inmunohistoquímica utilizando una serie de anticuerpos frente a las proteínas de interés en embriones de más de 48 horas.

La unión neuromuscular (*neuromuscular junction*, NMJ) es una sinapsis química entre las neuronas motoras y las miofibras de los músculos. La pre-sinapsis contiene vesículas sinápticas enriquecidas con moléculas como la acetilcolina (ACh) y la proteína de la vesícula sináptica (SV2). Después de la exocitosis y la transmisión a través de la sinapsis, la ACh se une a los receptores de acetilcolina (AChR) de la sinapsis y transmite las señales necesarias para la contracción muscular. La formación de la NMJ en el pez cebra comienza con la distribución de

grupos de AChR que se han modelado previamente (Daikoku et al., 2015). Las neuronas motoras primarias comienzan a crecer, lo que al entrar en contacto con los grupos de AChR aneurales, hace que los AChR se integren y se estabilicen, formando la NMJ.

Figura 10. Diagrama esquemático de la sinapsis neuromuscular en el pez cebra. En las larvas, los grupos de AChR se encuentran en el borde de las fibras musculares lentas y forman líneas en forma de chevrón en el límite de los segmentos del cuerpo, mientras que los de las fibras musculares rápidas se observan como puntos redondos y puntiagudos distantes de los bordes. Los AChR en fibras lentas y rápidas tienen diferentes subunidades. Escala 1 mm.



Los embriones se fijaron en paraformaldehído (PFA) al 4% en solución salina tamponada con fosfato (PBS) durante la noche. Al día siguiente se realizaron una serie de lavados utilizando PBS y dH₂O, y a continuación se permeabilizaron los embriones en acetona fría a -20°C y se procedió con otra serie de lavados utilizando 1XPBS-Tween20 (PBS-T) al 0.1%, traspasando finalmente los embriones a placas de 6 pocillos.

Una vez traspasados los embriones se añadió colagenasa en cada pocillo (Roche Diagnostics, 1 mg/ml) y se dejó incubar durante 60 min. Para retirar la colagenasa se llevaron a cabo 3 lavados con PBS-T y posteriormente los embriones se bloquearon en suero de caballo al 5% en PBS-T al 0,1% y se incubaron durante al menos 1 hora. Pasada la hora se retiró el suero de caballo y se añadieron 250 µl del anticuerpo anti-SV2 (vesícula sináptica 2) de ratón, diluido en suero de caballo al 5% a una dilución de 1:200, para visualizar las neuronas motoras, incubándose toda la noche a 4 °C. Posteriormente, se lavaron varias veces con PBS-T, para eliminar el anticuerpo primario, y se incubaron con el anticuerpo secundario Alexa Fluor 488, Alexa Fluor 594, Alexa Fluor 680 (1:500; Invitrogen). Los AChR post-sinápticos se visualizaron utilizando la α -bungarotoxina conjugada con Alexa Fluor 594 (1 µg/ml, Invitrogen). La tinción de inmunofluorescencia se fotografió usando un microscopio confocal Nikon (Nikon A1R Invert). Utilizamos el software ImageJ para analizar estas imágenes. Para la cuantificación de los cambios morfológicos de NMJ, se tomaron imágenes de 16 embriones para cada categoría. Debido al gradiente rostral a caudal del desarrollo embrionario en el pez cebra, utilizamos los mismos segmentos miotómicos para su comparación.

Al mismo tiempo, se realizó la tinción muscular utilizando faloidina, toxina que se une a la F-actina y que permite la visualización del músculo de contracción rápida.

4. RESULTADOS

4.1 HERENCIA DIGÉNICA

El filtrado realizado para seleccionar nuestros pacientes de estudio se resume en este esquema (Fig.11), mediante el cual obtuvimos dos pequeñas cohortes para profundizar en la herencia digénica: distroglicanopatías (n= 8) y sarcoglicanopatías (n=1).

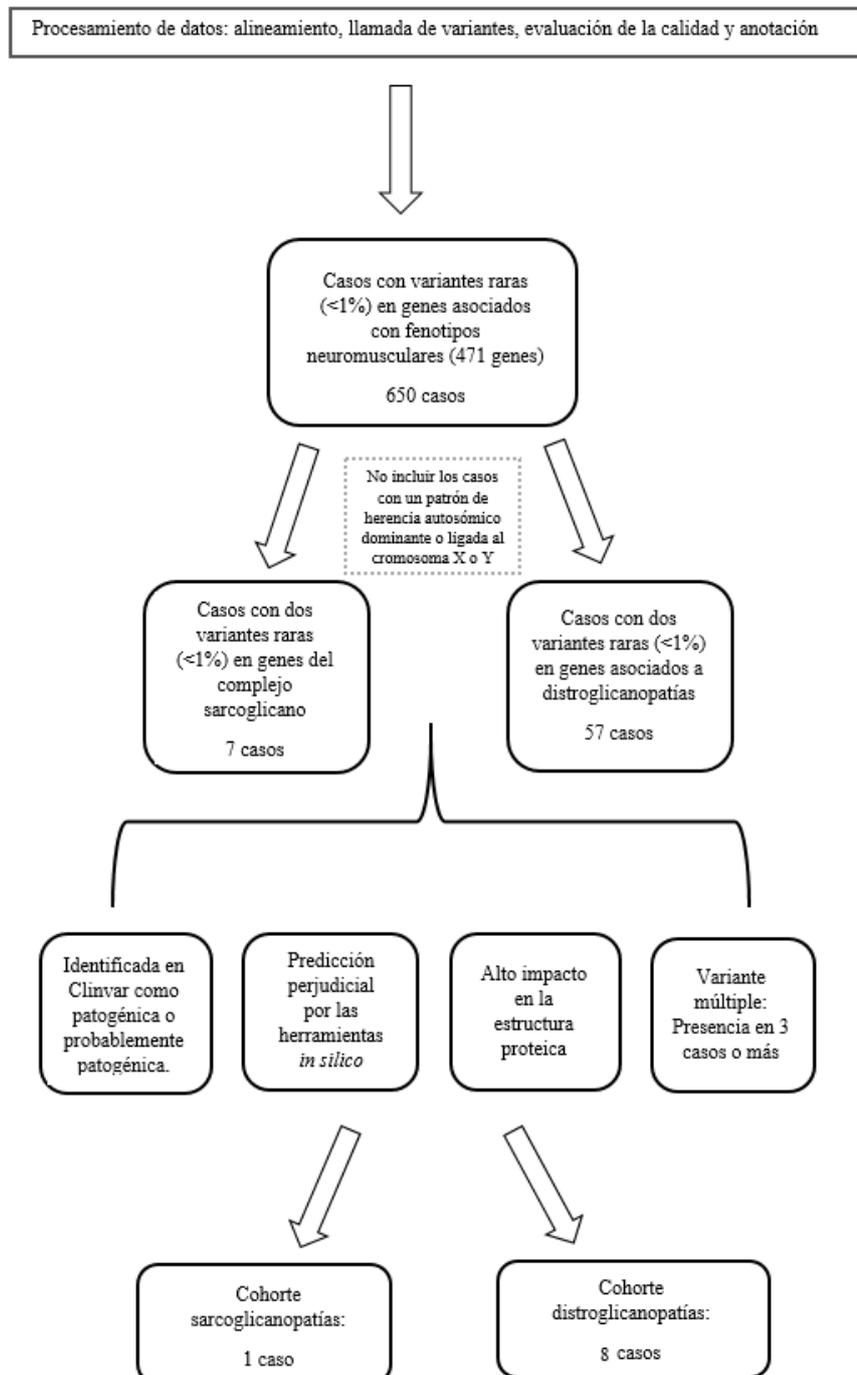


Figura 11. Diagrama de flujo empleado para seleccionar las cohortes en ambos estudios de herencia digénica.

DISTROGLICANOPATÍAS

4.1.1 HALLAZGOS GENÉTICOS

Tras realizar el análisis genético, se reconocieron dos variantes en dos de los dieciocho genes asociados con la ruta de la glicosilación del α -distroglucano. Identificamos variantes en 11/18 genes asociados a distroglicanopatías, de las cuales nueve variantes habían sido asociadas previamente a otros casos de distrofia muscular y siete no habían sido descritas hasta el momento. En la tabla 4 aparecen las dos variantes identificadas en cada paciente, siendo la mayoría mutaciones con cambio de sentido.

Existe una variante que no cumplía con los parámetros establecidos de patogenicidad, la variante en DPM1 del caso MANT048, cuya frecuencia era superior a 0,001, pero al no haber evidencias descritas ni observadas del efecto de la variante por los predictores, la incluimos en nuestro estudio.

Tabla 4. Variantes encontradas en los genes asociados con distroglicanopatías con un alto grado de patogenicidad en la cohorte de MYO-SEQ. Se describe el impacto de la variante, los valores de los predictores *in silico* y la frecuencia poblacional.

ID	Gen	rsID	Impacto	Sift	Polyphen	Muttaster	Fathmm	Clinvar	freq_exac_v3	freq_gnomad_exomes
MCOP141	GMPPB	rs142336618	Cambio de sentido	Perjudicial	Posiblemente perjudicial	Causante de enfermedad	Perjudicial	Patogénico	0,0008159	0,000697664
MCOP141	POMT2		Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial			4,29786E-06
MWAR016	FKTN		Desplazamiento marco lectura					Posiblemente patogénico		4,06286E-06
MWAR016	ISPD	rs374054216	Cambio de sentido	Tolerado	Posiblemente perjudicial	Causante de enfermedad	Perjudicial	Significado incierto	0,00007448	7,75681E-05
MANT069	B3GALNT2	rs185213208	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Tolerado		0,000313	0,00028317
MANT069	TMEM5		Variante de empalme							
MSAN043	FKTN		Cambio de sentido		Benigno	Causante de enfermedad	Perjudicial	Significado incierto		4,06415E-06
MSAN043	POMT2	rs200690151	Cambio de sentido	Perjudicial	Posiblemente perjudicial	Causante de enfermedad	Perjudicial		0,0000906	9,34374E-05
MANT048	DPM1		Variante de empalme						0,003	0,002338604
MANT048	POMK	rs149297443	Cambio de sentido	Perjudicial	Posiblemente perjudicial	Causante de enfermedad	Tolerado	Interpretaciones conflictivas de patogenicidad	0,0009112	0,0012631
MPAU045	DOLK	rs145310298	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Perjudicial	Interpretaciones conflictivas de patogenicidad	0,001296	0,001856426
MPAU045	ISPD		Desplazamiento marco lectura							0,000250784
MANT062	B3GALNT2		Cambio de sentido	Tolerado	Posiblemente perjudicial	Polimorfismo	Tolerado		0,00003295	1,6258E-05
MANT062	LARGE	rs150861748	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Tolerado	Significado incierto	0,0008648	0,000850008
MGIE011	POMT2	rs202192851	Cambio de sentido						0,000726	0,000718521
MGIE011	TMEM5	rs145516652	Cambio de sentido	Tolerado	Posiblemente perjudicial	Polimorfismo	Tolerado	Significado incierto	0,00014	0,000135677

4.1.2 INTERACCIÓN PROTEÍNA-PROTEÍNA

Para proporcionar soporte adicional a nuestra hipótesis, realizamos un análisis con el software STRING (<https://string-db.org>) para investigar posibles interacciones entre los productos de los genes identificados asociados con la O-glicosilación.

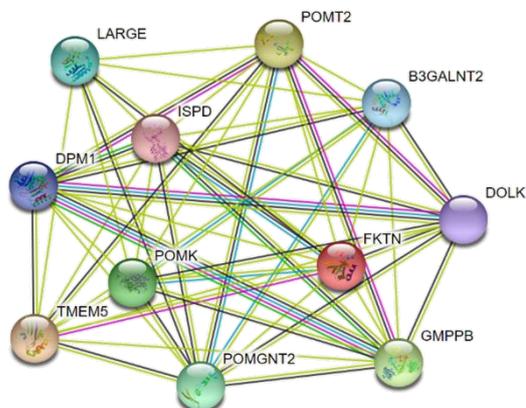


Figura 12. Interactoma obtenido por STRING entre proteínas asociadas a la O-glicosilación. El color de las líneas representa diferentes categorías de evidencia: vecindario en el genoma (línea verde oscura), fusión de genes (línea roja), coocurrencia génica (línea azul marino), coexpresión (línea negra), datos experimentales (línea morada), asociación en bases de datos biológicas (línea azul clara) y minería de textos (línea verde clara).

Obtuvimos este interactoma entre las once proteínas de nuestra cohorte, donde las redes establecen la información esperada y observada en otros estudios (Fig.12). Al mismo tiempo, obtuvimos el P-valor $<1.0e-16$, indicativo de la presencia de más interacciones entre las proteínas de lo que se esperaría para un conjunto aleatorio de proteínas de tamaño similar, extraídas del genoma.

Por otro lado, el análisis de la vía STRING predijo vías comunes recogidas en la tabla 5, proporcionando información sobre los procesos biológicos y la función molecular de las proteínas, respaldando la fuerte interacción en la ruta de la O-glicosilación, y la posible herencia digénica en las distroglicanopatías.

Tabla 5. Procesos biológicos y funciones moleculares de las once proteínas analizadas con su código de referencia (ID).

ID	Proceso biológico
IR: 0035269	Manosilación de proteína O-ligada
GO: 0006493	Glicosilación ligada a proteínas O
GO: 0006486	Glicosilación de proteínas
IR: 1901137	Proceso biosintético derivado de carbohidratos
GO: 0009100	Proceso metabólico de la glucoproteína
ID	Función molecular
IR: 0016740	Actividad transferasa
IR: 0016757	Actividad transferasa, transferencia de grupos glicosilo
IR: 0016758	Actividad transferasa, transferencia de grupos hexosilo
GO: 0008194	Actividad de UDP-glicosiltransferasa
GO: 0008375	Actividad acetilglucosaminiltransferasa

4.1.3 HALLAZGOS CLÍNICOS

Nuestra cohorte consistió en ocho pacientes de ocho familias no relacionadas. Tres de estos pacientes eran de origen belga, un paciente de origen polaco, otro paciente de origen danés, dos pacientes de origen español y otro de origen alemán.

Los datos recogidos en la tabla 6 mostraron la gran heterogeneidad clínica en los pacientes, no habiendo un claro patrón en ninguno de los síntomas analizados. La falta de datos publicados con respecto a variantes en dos genes involucrados en la O-glicosilación dificultaron el análisis clínico y no pudimos descartar ningún caso por presentar un fenotipo atípico.

Tabla 6. Resumen de los datos clínicos recogidos en la plataforma Phenotips.

PACIENTE	INICIO	PROGRESIÓN	PATRÓN DE DEBILIDAD Y DISTROFIA	SIGNOS NEUROLÓGICOS	CK	INFORMACIÓN ADICIONAL
MANT069*	Juvenil	Lenta	Debilidad progresiva proximal y distal. Debilidad muscular espinal. Atrofia	No disfunción cognitiva	Normal	
MANT048*	Juvenil	Rápida	Debilidad progresiva proximal y distal. Debilidad facial y espinal. Atrofia.	No disfunción cognitiva	Elevado menos de 10x	Disfunción respiratoria
MANT062*	Adulto	Lenta	Debilidad progresiva proximal y distal. Debilidad del dorsiflexor del pie. Debilidad muscular espinal	No disfunción cognitiva	Elevado menos de 10x	Camina con soporte
MCOP141	Juvenil	No progresiva	Debilidad estática. Atrofia escapular y del miembro superior proximal	No disfunción cognitiva	Normal	
MWAR016	Vejez	Variable	Debilidad estática proximal y progresiva distal	Hiperreflexia, neuropatía media constrictiva	Elevado más de 10x	Escoliosis, retinopatía hipertensiva
MIST117	Infancia	Lenta	Debilidad progresiva proximal y distal. Debilidad del dorsiflexor del pie	No disfunción cognitiva	Elevado más de 10x	Ha perdido la habilidad de caminar independientemente
MGIE011	Infancia	Lenta	Debilidad progresiva proximal y distal. Hipertrofia proximal del músculo del miembro inferior	No disfunción cognitiva	Elevado más de 10x	Fatiga
MPAU045	Vejez	Rápida	Debilidad progresiva proximal y distal. Debilidad facial y espinal. Parálisis facial bilateral	No disfunción cognitiva	No hay datos	Escoliosis torácica
MSAN043	Adulto	Rápida	Debilidad progresiva proximal. Debilidad facial y espinal	Problemas de habla	Elevado menos de 10x	

Atendiendo a la patogenicidad de las variantes identificadas, a las interacciones obtenidas por el software STRING y a los síntomas clínicos recopilados, consideramos como un buen candidato para estudiar las imágenes de la biopsia muscular al paciente MCOP141.

Glicosilación intacta en α -dystroglicano

Las técnicas inmunohistoquímicas de las biopsias musculares se realizaron en el centro de referencia del paciente. Estas biopsias musculares se obtuvieron de los músculos tibial anterior, gastrocnemio o deltoides. Para la evaluación inmunohistoquímica de la glicosilación α -DG, las secciones musculares se tiñeron con los anticuerpos VIA4-1 y I1H6C, ampliamente utilizados para estudiar la glicosilación funcional de α -DG (Merck-Millipore, Temecula, California, EE. UU). Usando la inmunohistoquímica con los dos anticuerpos específicos de la O-glicosilación observamos una tinción uniforme en ambas imágenes (Fig. 13), determinándose que la glicosilación se mantenía prácticamente intacta en este caso.

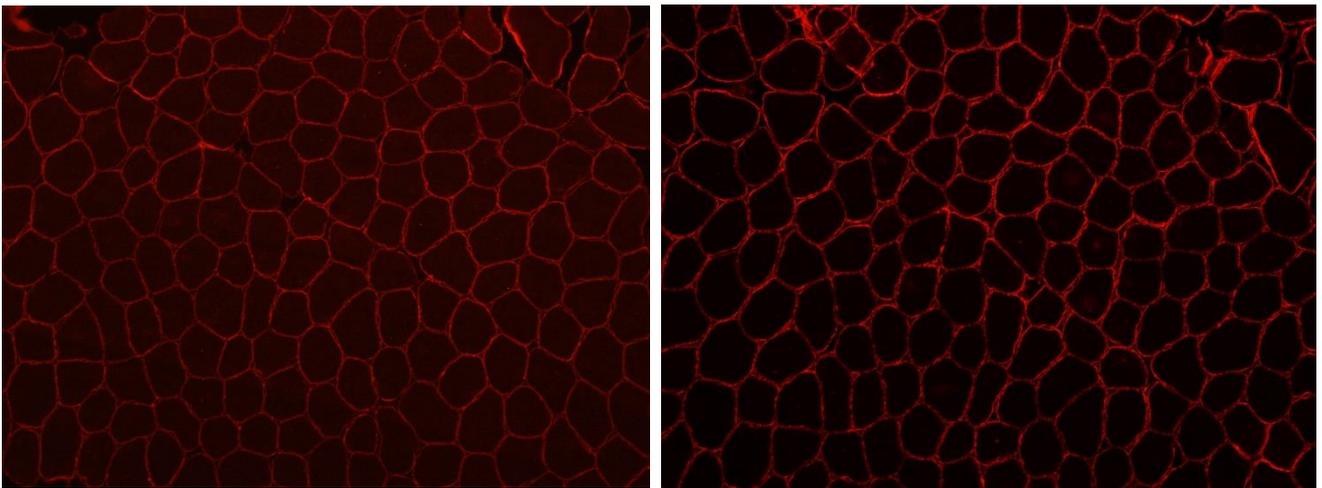


Figura 13. La evaluación inmunohistoquímica de algunos pacientes revela ausencia en la afección del α -dystroglicano. Se obtuvieron imágenes de inmunohistoquímica para la tinción del α -dystroglicano a partir de secciones musculares, utilizando los anticuerpos VIA4-1 (izquierda) y I1H6C (derecha).

Deficiencias en la O-glicosilación impiden que el dystroglicano pueda unirse a proteínas extracelulares y realizar su función de nexo de unión entre el citoesqueleto de la célula y la matriz extracelular. Al no haber deficiencia en la O-glicosilación la función del α -DG no se ve alterada, rechazando la dystroglicanopatía como causa del fenotipo del paciente.

Atendiendo a estas imágenes y a las valoraciones de los clínicos excluimos como mecanismo molecular la herencia digénica en los ocho pacientes.

SARCOGLICANOPATÍAS

4.1.4 HALLAZGOS GENÉTICOS

Tras realizar el análisis genético, se reconocieron en un paciente dos variantes posiblemente patogénicas en dos de las subunidades de SGC: γ y ϵ , no encontradas en la población control (Tabla 7).

Tabla 7. Variantes encontradas en los genes asociados con sarcoglicanopatías con un alto grado de patogénesis en la cohorte de MYO-SEQ. Se describe el impacto de la variante, los valores de los predictores *in silico* y la frecuencia poblacional.

ID	Gen	Variante	Impacto	Sift	Polyphen	Muttaster	Fathmm
MSAN016	SGCG	c.205G> A	Cambio de sentido	perjudicial	posiblemente perjudicial	causante de enfermedad	perjudicial
MSAN016	SGCE	c.446A> C	Cambio de sentido	perjudicial	probablemente perjudicial	causante de enfermedad	perjudicial

Los predictores *in silico* determinaron que ambas variantes podrían desencadenar un cuadro clínico específico del gen afectado. Para contrarrestar estos datos se analizó el interactoma obtenido por el software STRING.

4.1.5 INTERACCIÓN PROTEÍNA-PROTEÍNA

Las redes que establecieron las relaciones entre ambas proteínas apoyaron una asociación estrecha entre ambas subunidades, representadas en la figura 14.

Las funciones biológicas en este caso no se establecieron por pertenecer al mismo complejo proteico. Sin embargo, el p-valor obtenido ($0,00246 < 0,05$) indicó, como en el caso anterior, que ambas proteínas tenían más interacciones entre ellas de lo que se esperaría para un conjunto aleatorio de proteínas extraídas del genoma.



Figura 14. Interactoma obtenido por STRING entre dos proteínas del complejo sarcoglicano.

4.1.6 HALLAZGOS CLÍNICOS

La paciente, de origen español, comenzó a mostrar los síntomas en torno a los 30 años. Presentaba una debilidad progresiva de los músculos proximales inferiores y superiores, ptosis y disfonía con una clara fatiga. Entre los síntomas clínicos más relevantes se encontró la presencia de miocardiopatía dilatada y el bloqueo de la aurícula ventricular, lo que ha provocado que lleve marcapasos. Los clínicos consideraron que era poco probable que se debiera a una sarcoglicanopatía. No se pudieron realizar pruebas complementarias como la segregación de las variantes o la inmunotinción del sarcoglicano, por falta de material serológico y tisular de la paciente.

4.2 ESTUDIO DE VARIANTES INTRÓNICAS

4.2.1 DISFERLINOPATÍA

Nuestro grupo de muestra para estudiar mutaciones descritas en el gen *DYSF* se compuso de veintidós casos que presentaban una variante patogénica en dicho gen, siguiendo los criterios explicados anteriormente. Una vez establecido el número se analizaron los datos clínicos. Uno de los pacientes tuvo un inicio congénito, un fenotipo más grave, un patrón de EMG neurológico y un acusado deterioro cognitivo, por lo que decidimos descartarlo. El resto de casos presentaron un fenotipo bastante similar con la sintomatología esperada de LGMDR2, mostrando debilidad muscular progresiva de las extremidades a partir de la adolescencia o la edad adulta temprana, niveles elevados de creatina quinasa en sangre que indicaban daño muscular y un patrón de EMG miopático. En 5/21 se observó miocardiopatías y en 9/21 atrofia muscular en los miembros proximales inferiores y superiores.

Examinamos el ADN genómico de estos pacientes para determinar si eran o no portadores de una de las variante intrónicas descritas: c.4886+1249 G>T y c.5668-824C> T. Como se resume en la tabla 8, la mayoría de los casos presentaban una variante con cambio de sentido (19/21). Como indicación de prevalencia, la variante c.1351-A>G (rsID 141818764) se encontró en 5/21 pacientes, siendo dos pacientes daneses, dos españoles y un paciente serbio, y la variante c.4819-A>G (rsID 138357301) en 3/21 casos cuyas nacionalidades fueron ucraniana, serbia y polaca.

Tabla 8. Variantes patogénicas encontradas en el gen *DYSF* en nuestra cohorte de estudio. Se describe el impacto de la variante, los valores de los predictores *in silico* y la frecuencia poblacional.

ID	rsID	Impacto	sift	polyphen	muttaster	fathmm
MANT029_3		Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MANT069		Sitio aceptor empalme			Causante de enfermedad	
MAAC004	rs149732545	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MCOP140	rs144636654	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Perjudicial
MANT041	rs150139276	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MCAI005	rs115407852	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MCOP021	rs115407852	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MBEL002	rs139754493	Cambio de sentido	Perjudicial	Posiblemente perjudicial	Causante de enfermedad	Perjudicial
MINN015	rs149357145	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MRIG004		Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Tolerado
MBEL060	rs141818764	Cambio de sentido	Perjudicial	Benigno	Causante de enfermedad	Tolerado
MCOP145	rs141818764	Cambio de sentido	Perjudicial	Benigno	Causante de enfermedad	Tolerado
MCOP091	rs141818764	Cambio de sentido	Perjudicial	Benigno	Causante de enfermedad	Tolerado
MSAN043	rs141818764	Cambio de sentido	Perjudicial	Benigno	Causante de enfermedad	Tolerado
MVAL037	rs141818764	Cambio de sentido	Perjudicial	Benigno	Causante de enfermedad	Tolerado
MANT062	rs144598063	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial
MWAR042_2	rs138357301	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Tolerado
MKHA013	rs138357301	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Tolerado
MBEL001	rs138357301	Cambio de sentido	Tolerado	Benigno	Causante de enfermedad	Tolerado
MCOP120		Codón parada			Causante de enfermedad	
MPOR018	rs201049092	Cambio de sentido	Perjudicial	Probablemente perjudicial	Causante de enfermedad	Perjudicial

Diagnóstico de LGMDR1 en el paciente MRIG004

Al secuenciar con el método de Sanger ambas variantes se compararon nuestras secuencias con las secuencias de referencia, obteniendo unos electroferogramas como los observados en la figura 15.

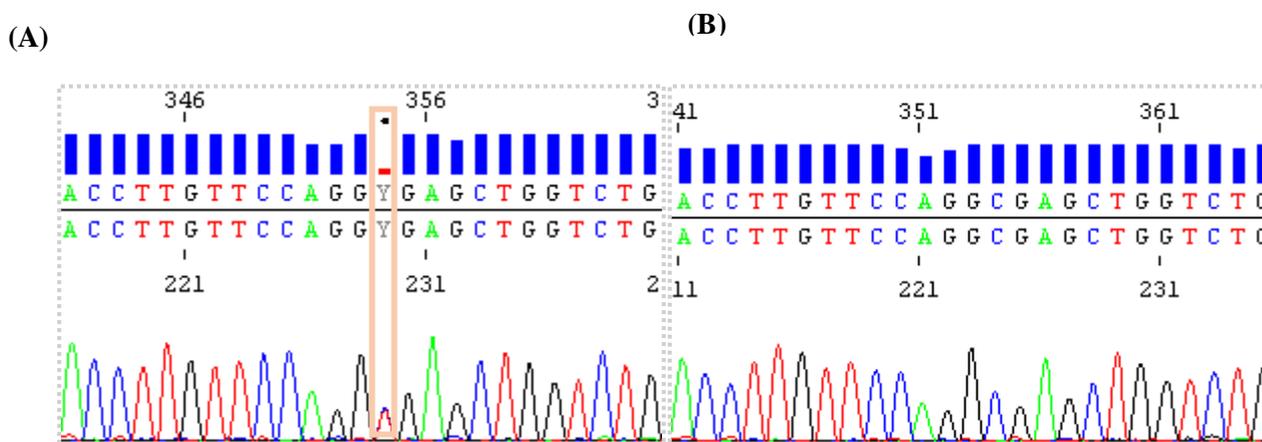


Figura 15. Electroferogramas del intrón 50 del gen *DYSF* obtenidos por el software SeqScape™ en dos pacientes de la cohorte. (A) El paciente MRIG004 es heterocigoto para la variante intrónica patogénica c.5668-824C> T. (B) El paciente MINN015 no presenta ningún cambio en la región de interés.

La tasa de diagnóstico fue del 4,76% (1/21) en nuestro grupo de estudio, siendo identificada la variante c.5668-824C> T en el caso MRIG004.

4.2.2 CALPAINOPATÍA

Los pacientes de nuestra cohorte, individuos con la variante c.1354+129T>A o c.1746-20C>G, mostraron debilidad muscular progresiva principalmente en las extremidades proximales, con un inicio desde de la infancia hasta la edad adulta temprana (n=20). Seis pacientes presentaron hipertrofia en la pantorrilla, cuatro pacientes presentaron una escápula alada, y cinco pacientes mostraron contracturas en las articulaciones de las extremidades inferiores. En 2/20 casos hubo complicaciones cardíacas, destacando un caso con defectos en el tabique ventricular. Trece de los pacientes tuvieron los niveles elevados de creatina quinasa en sangre, y la mayoría mostraron un patrón de EMG miopático, a excepción de dos casos que también presentaron un patrón neurogénico. En vista de estos datos, no se descartó ningún paciente del estudio por los hallazgos clínicos.

c.1746-20C>G

La variante c.1746-20C>G se identificó en dieciocho pacientes. Todos los casos tenían un patrón de herencia autosómica recesiva o esporádica, y ninguna familia era consanguínea.

Para realizar el estudio contamos con sangre periférica y biopsia muscular de un paciente de origen alemán (MESS004), paciente que presentaba, junto a la variante intrónica, una variante patogénica en la región codificante del gen, c.598_612del.

Predicciones bioinformáticas no esclarecedoras

Por un lado, la variante fue predicha por el SD-Score como benigna al ser el valor de la consecuencia del empalme prevista normal, y no habiendo diferencias en el valor SD-Score entre ambas secuencias (Tabla 9a). Por otro lado, la misma variante fue predicha por el algoritmo HSF para crear un nuevo sitio aceptor de empalme (A1). En las matrices HSF el valor umbral es de 65, por lo que al ser la puntuación del tipo salvaje (WT, del inglés *wild-type*) inferior y la variación del valor superior a más del 10% (53'23%) se considera que la mutación crea un nuevo sitio de empalme (Tabla 9b).

Tabla 9. Resultados de los predictores que evalúan el sitio de empalme. (A) Valores obtenidos por el algoritmo SD-Score comparando la secuencia WT de la secuencia mutada. El programa predice diferencias en el puntaje SD (Δ SD-Score), los contenidos de información (Δ Ri) y la matriz de peso de posición (Δ CV), así como la consecuencia de empalme prevista. **(B) Predicciones de posibles variaciones en el sitio de empalme utilizando el algoritmo HSF.** El porcentaje de variación indica que c.1746-20C>G puede crear un nuevo sitio de empalme en el transcrito de *CAPN3*.

(A)

Wt.Seq				Mt.Seq				Diferencia			
Wt.Seq (601 bases)	SD-Score	Rhode Island	CV	Mt.Seq (601 bases)	SD-Score	Rhode Island	CV	Δ SD-Score	Δ Ri	Δ CV	Predicción
CATGTCTGT	-5,879	2.270	0,674	GATGTCTGT	-5,879	1.307	0,643	0,000	-0,963	-0,031	Normal

(B)

Posición de ADNc	Tipo de sitio de empalme	Motivo	Nuevo sitio de empalme	Tipo salvaje	Mutante	Si se usa un sitio críptico, la variación de la longitud del exón	Variación (%)
290	Aceptor	CTGAGACCCACAT	ctgagaccccagAT	54.37	83.31	-290	Nuevo sitio +53.23
301	Donante	CATGTCTGT	GATgtctgt	75.3	73.54	-301	-2.34

c.1746-20C>G provoca la delección del exón 14

Se analizaron las muestras del paciente, sus padres y un individuo control que no presentaba ninguna afección muscular. El amplicón de la muestra control estaba formado por una banda cuyo tamaño oscilaba los 250 pb. La muestra del padre tenía una banda similar a la del individuo sano, sin embargo, el paciente y la madre presentaban una banda adicional en la reacción de la RT-PCR de un tamaño menor al amplicón salvaje, indicativo de una posible delección.

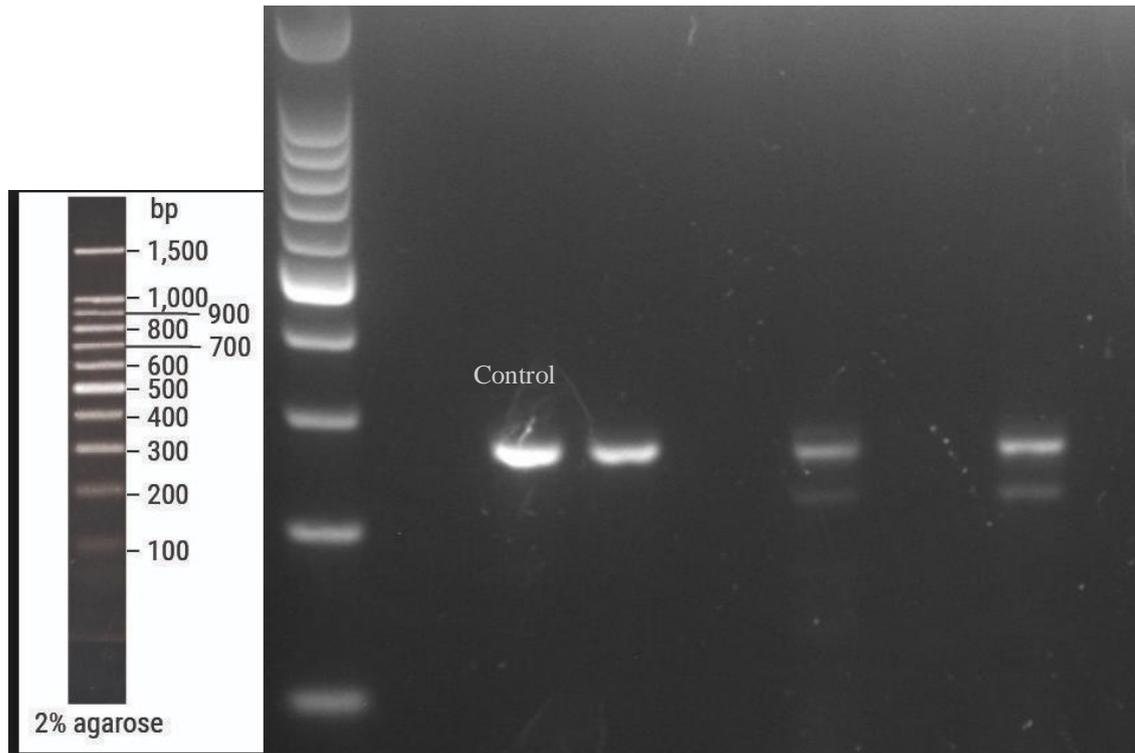


Figura 16. Identificación de una deleción en individuos portadores de la variante c.1746-20C>G. El evento de deleción es observable en los amplicones del paciente y de su madre, que muestran dos bandas: una del mismo tamaño que la de tipo salvaje (control) y la otra de un menor tamaño (en torno a los 220pb). El marcador de peso molecular se cargó en el primer carril.

A partir de las bandas de la electroforesis se secuenciaron los amplicones de cada muestra para identificar la afección del transcrito. La secuencia acotada por los cebadores empleados era de 277 pb, englobando parte del exón 13 hasta el exón 16 (Fig. 16). Los fragmentos tanto del individuo control como del paciente mostraron longitudes diferentes a la esperada, habiendo diferencias entre las secuencias. En el caso del control aparecía la deleción completa del exón 15 (AAAAAGAAAAAACCAAG) siendo el tamaño del amplicón de 259 pb, mientras que en la secuencia del paciente había una deleción de los exones 14 y 15, (GGAAGTTGAAAATACCATCTCCGTGGATCGGCCAGTGAAAAAGAAAAAACCAA) cuyo tamaño era de 222 pb.

La ausencia del exón 15 es una isoforma presente en la población control. Como se indican en otros estudios (Blázquez et al., 2008) a partir de muestras de biopsia muscular se obtiene una isoforma, en la que la proteína está íntegramente, pero a partir de muestras serológicas pueden diferenciarse hasta cuatro tipos de isoformas, todas ellas con ausencia del exón 15, derivadas de empalmes y promotores alternativos.

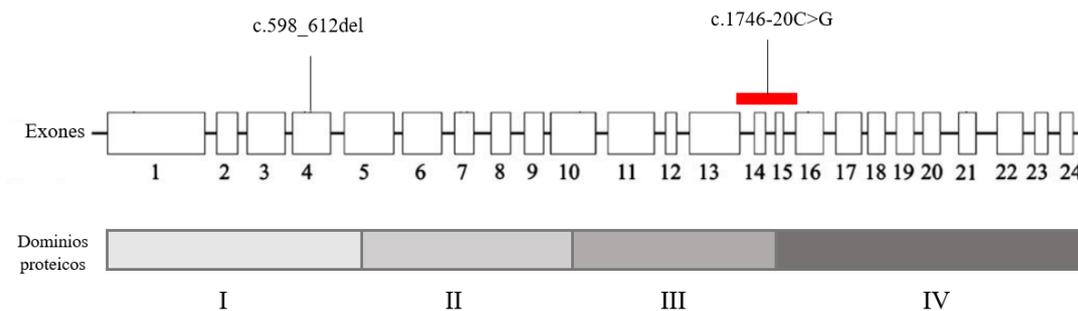


Figura 18. Distribución de las variantes en CAPN3 del paciente MESS004 a lo largo de los exones (1-24) y los dominios proteicos (I-IV).

A continuación, se analizó la expresión de la proteína por Western blot. El anticuerpo monoclonal empleado fue 3d/2C4 (NovoCastra, Newcastle), anticuerpo que reconoce el extremo N terminal y la región limitante del residuo catalítico Asn-358 de calpaína-3. El clon 2C4 reconoce el tamaño completo de la proteína y bandas de 30 kDa. Esta técnica fue realizada en el hospital del paciente de estudio.

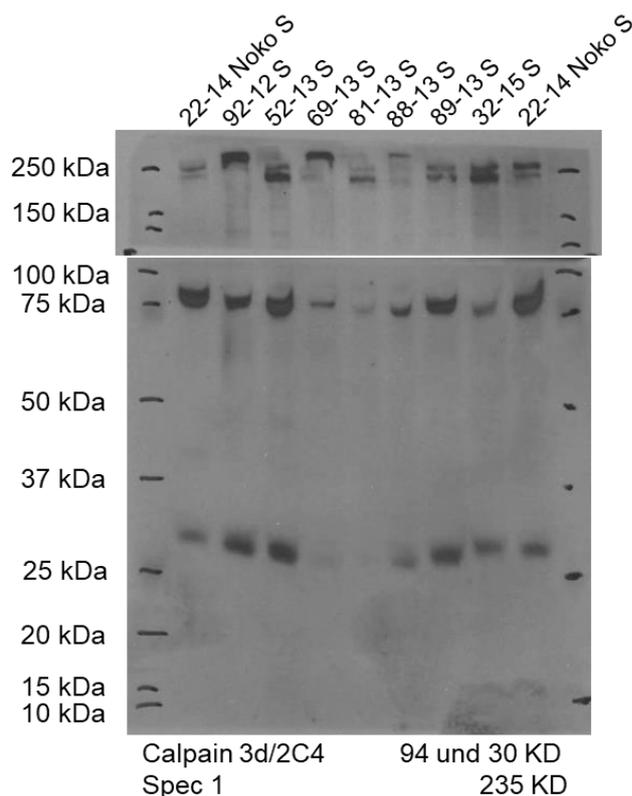


Figura 19. Expresión de calpaína-3 en el tejido muscular del paciente MESS004. La imagen muestra la membrana en la que se ha utilizado el anticuerpo 3d/2C4. La ID del paciente MESS004 se corresponde con el número 81-13S, la ID del control con 22-14 Noko S, y el resto son pacientes en los que se sospechaba una calpainopatía.

La reducción de la banda en MESS04 fue más destacada que en el resto de muestras, aunque había cierta variabilidad entre los otros pacientes. Se observó una clara reducción de la expresión de calpaína-3 en MESS004 en comparación con los pacientes 92-12S, 52-13S, 88-13S, 89-13S, 32-15S.

La variante intrónica c.1746-20C>G en *CAPN3* compromete a la maduración del ARNm actuando en el sitio de empalme y eliminando los exones 14 y 15, dando lugar a diferencias significativas en la transcripción del ARNm entre el paciente y el individuo sano.

c.1354+129T>A

La otra variante a analizar fue c.1354+129T>A, presente en dos hermanos (MBEL028 y MBEL029), con un fenotipo típico de calpainopatía, siendo ambos heterocigotos para una mutación con cambio del marco de lectura en el exón 4 (c. 550delA) que causa un codón de parada prematuro dentro del dominio peptidasa C2 (Dorobek et al., 2015).

Predicciones bioinformáticas apuntan a un nuevo sitio aceptor de empalme

Los predictores *in silico* (SD-Score y el algoritmo HSF) predijeron que el empalme ocasionado por la variante intrónica tenía una alta probabilidad de generar un sitio aceptor de empalme (Tabla 10), como en el caso anterior, o de generar un sitio silenciador del empalme exónico, encargado de impedir la inclusión del exón en el ARNm y, por lo tanto, afectando negativamente a la maquinaria del empalme.

Tabla 10. Valores predictivos del algoritmo HSF al analizar c.1354+129T>A. Hipótesis de los posibles cambios en el ADNc, por causa de un nuevo sitio aceptor de empalme o por la presencia de un sitio silenciador del empalme exónico (ESS, del inglés *Exonic splicing silencer*,).

Señal predicha	Algoritmo de predicción	Posición de ADNc	Interpretación
Nuevo sitio de aceptación	1 - Matrices HSF		Activación de un sitio aceptor críptico exónico, con presencia de uno o más puntos de ramificación crípticos. Posible alteración del empalme.
	2 - MaxEnt		
Nuevo sitio de ESS	1 - Sironi y col. - Patrón 1		Creación de un sitio exónico de ESS. Posible alteración del empalme.
	2 - Matrices HSF - hnRNP A1		
	3 - Secuencias de ESR de Goren et al.		

Tipo de sitio de empalme	Motivo	Nuevo sitio de empalme	Tipo salvaje	Mutante	Si se usa un sitio críptico, la variación de la longitud del exón	Variación (%)
Aceptor	CCTGTTTACTGAG	cctgttttacagAG	58.89	87.83	-279	Nuevo sitio +49.14

c.1354+129T>A provoca la inserción de un pseudoexón en la transcripción de CAPN3

El gel de la RT-PCR mostró las bandas correspondientes a la muestra control y a los dos pacientes en tres sets diferentes, en función de los cebadores empleados (Fig.20).

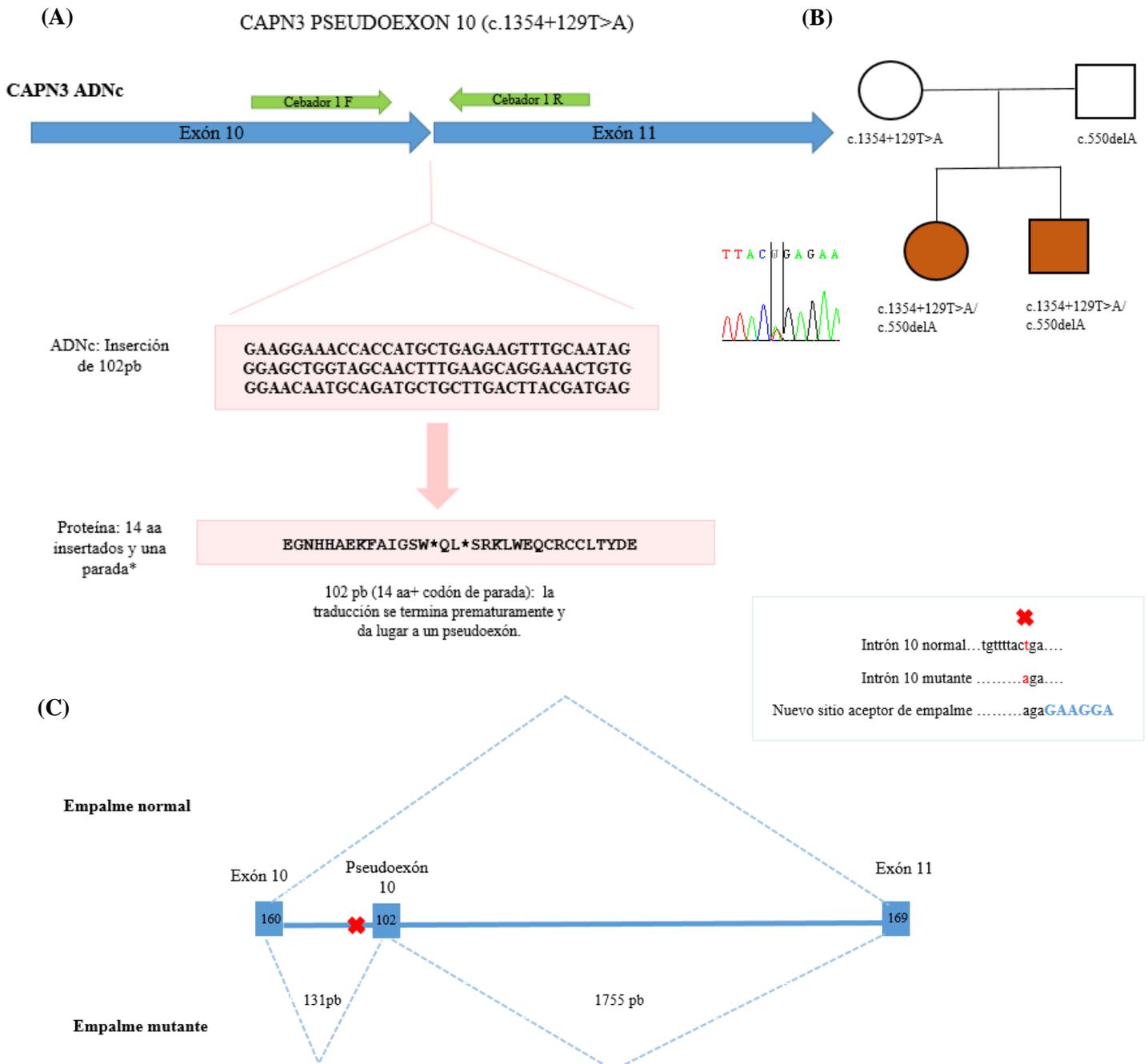


Figura 21. El empalme del ARNm de CAPN3 se alteró en los pacientes MBEL028 y MBEL029, lo llevando a la inclusión del pseudoexón patógeno PE 10. (A) La secuenciación de ADNc mostró que los pacientes tenían una inserción de 102 nt en la unión de los exones 10 y 11, revelando la inclusión de un nuevo pseudoexón, PE 10, empalmado en la secuencia de codificación. (B) El pedigrí de los familiares directos de ambos pacientes muestra los resultados de las segregaciones realizadas. (C) El ADN genómico del intrón 10i mostró que los pacientes eran heterocigotos para una mutación puntual (c.1354+129T>A) (aspa roja) que crea un nuevo sitio aceptor de empalme en el extremo 5' de la secuencia. En estas células heterocigotas, tanto el ARNm que contiene PE 10 normal como el mutante se expresan, lo que representa las formas alternativas de empalme de los transcritos CAPN3. Se muestra la estructura del ARNm dentro del gen, junto con los patrones de empalme normales y mutantes. Los números a lo largo del ARN indican el tamaño de cada elemento en pb.

Este sitio aceptor de empalme generó una transcripción aberrante que llevó la inserción de 102 pares de bases en el intrón 10 (Fig.21A), verificando los resultados mostrados en el gel de la RT-PCR. La segregación de los pacientes y sus progenitores confirmaron vía Sanger la presencia de las variantes en los hermanos afectados y mostró que cada variante fue heredada de un progenitor distinto, indicando que las variantes están en *trans* y, por lo tanto, afectan a los dos alelos del gen (Fig. 21B).

La secuenciación y los estudios de ADNc muscular confirmaron que la variante intrónica causa un empalme aberrante de un exón artificial/fragmento intrónico en transcripciones del ARNm de ambos pacientes (Fig. 21C), resultando en la codificación de un codón de terminación prematuro.

No se pudo realizar hasta el momento la transferencia Western para confirmar que el pseudoexón 10 afecta a la estabilidad proteica de calpaína-3.

4.3 ESTUDIO DEL PAPEL MODIFICADOR DE UNA VARIANTE PATOGENICA EN EL COMPLEJO SARCOGLICANO

4.3.1 HALLAZGOS CLÍNICOS

Los cuatro casos analizados y sus variantes patogénicas identificadas fueron MTEH057 (sarcoglicano- γ c.101 G>A), MIST069 (sarcoglicano- α c.850-C>T), MBEL050 y MBEL119, portadores de la misma variante (sarcoglicano- β c.87_89delAAG). MTEH057 y MBEL119 estaban diagnosticados como calpainopatías y MIST069 y MBEL050 como disferlinopatías.

Los detalles clínicos de los pacientes se enumeran en la tabla 11. Todos los pacientes mostraron debilidad en la cintura y extremidades y niveles elevados de creatina quinasa sérica (4–100 veces el valor normal).

Tabla 11. Resumen de las características clínicas de los cuatro pacientes con una variante patogénica en una subunidad del complejo sarcoglicano.

<i>Hallazgos principales</i>	<i>LGMDR1</i>	<i>LGMDR2</i>
Inicio en adultos (segunda a cuarta década)	0/2	2/2
Creatina quinasa elevada (superior a 5 veces)	2/2	2/2
Miopatía proximodistal	2/2	1/2
<i>Debilidad muscular</i>		
Músculos proximales más afectados	2/2	2/2
Debilidad distal inferior	1/2	1/2

Debilidad espinal	1/2	1/2
Debilidad distal superior	0/2	1/2
<i>Daño muscular</i>		
Atrofia muscular del cuello	1/2	1/2
Hipertrofia pantorrillas	1/2	0/2
Aleteo escapular (leve a moderado)	1/2	1/2
Contracturas en las extremidades inferiores distales	1/2	0/2
Función respiratoria normal	2/2	2/2
Cardiomiopatía ausente	2/2	2/2
Fatiga	0/2	1/2
Biopsia muscular compatible con proceso distrófico.	2/2	2/2
EMG: miopatía	2/2	2/2

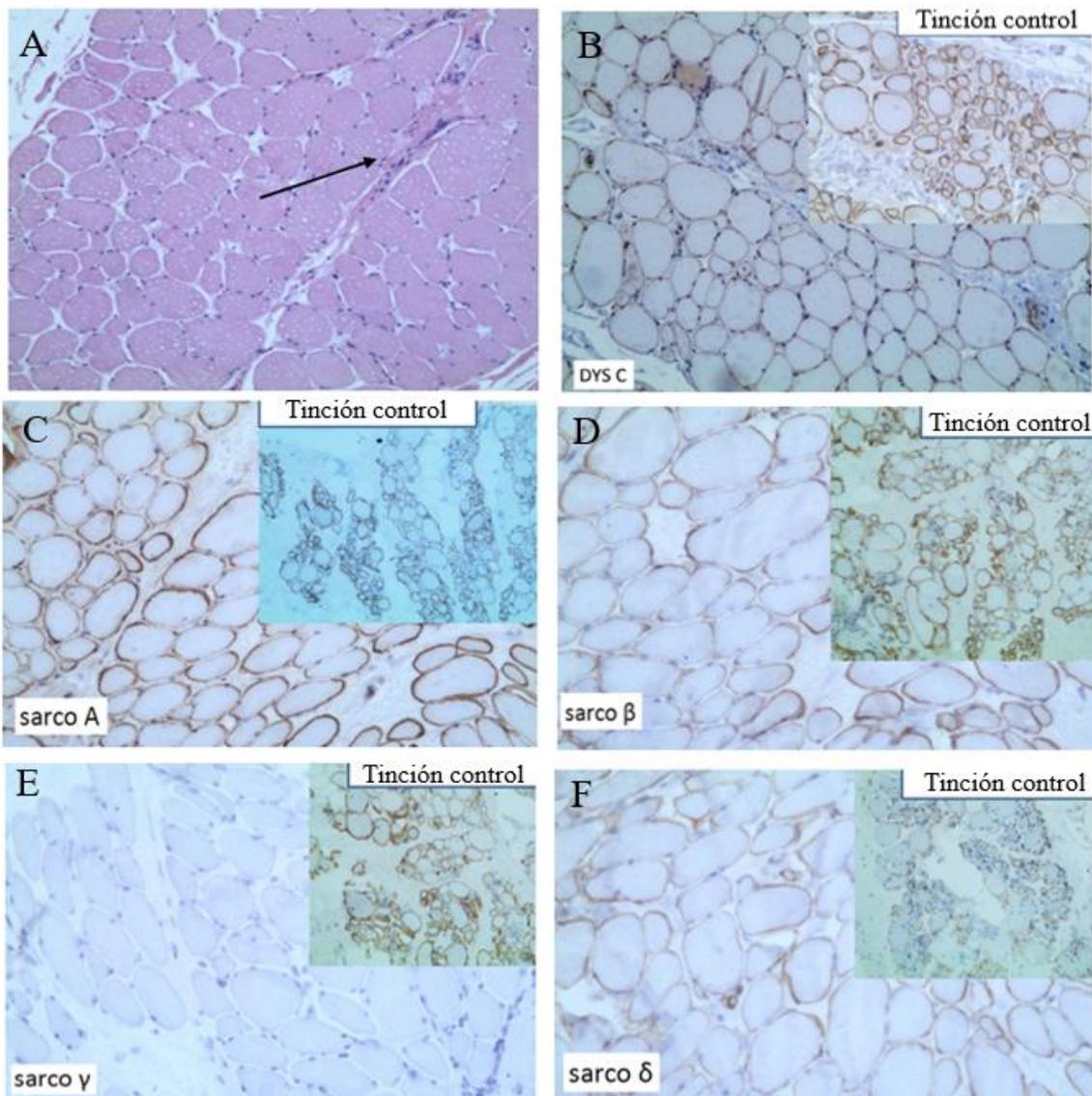
Ante la gran heterogeneidad clínica presente en los pacientes con calpainopatía, disferlinopatía y sarcoglicanopatías, estos datos no nos dieron evidencias de un efecto de la variante patogénica en el progreso de la enfermedad.

Ausencia de la expresión de gamma sarcoglicano en el paciente MBEL119

La tinción con HE en MBEL119 (Fig. 22A), realizada en el centro de referencia, reveló cambios distróficos como la variación en el tamaño de las fibras musculares y la presencia de fibrosis endomisial. El tejido conectivo y el tejido adiposo del músculo proliferaron moderadamente. No se observó infiltración celular inflamatoria.

La biopsia muscular del paciente mostró unos niveles de distrofina (Fig. 22.B) preservados al igual que en la muestra control, cuyas secciones fueron etiquetadas sin anticuerpos primarios, y cuyo fenotipo no estaba asociado a ningún trastorno neuromuscular. En la tinción inmunohistoquímica de las subunidades sarcoglicano se observó como el SGC- α (Fig. 22.C) estaba preservado y como, tanto SGC- β (Fig. 22.D) como SGC- δ (Fig. 22.F) estaban prácticamente intactos. Sin embargo, el patrón de expresión de la subunidad γ no se detectó en el paciente (Fig. 22.E), lo que podía ser indicativo de una afección por la variante “modificadora”. No obstante, el paciente presentaba una variante patogénica en SGC- β (c.87_89delAAG), lo que dificulta determinar que esta disminución en la expresión de SGC- γ se debiera a la variante y no a errores en la técnica.

Figura 22. Examen histológico e inmunohistoquímico de la biopsia muscular en MBEL119 mostró cambios en los niveles de la subunidad γ . En la tinción con hematoxilina-eosina (A) se observaron características distróficas como es la variación del tamaño de la fibra y la fibrosis endomisial (flecha negra). (B) La inmunotinción de la distrofina mostró patrones similares en el paciente y en la muestra control. (C-F) Secciones de la biopsia muscular inmunohistoquímicamente teñidas para detectar los niveles de las cuatro subunidades sarcoglicanas. Las inmunoreactividades para los sarcoglicanos α , β , y δ fueron detectables en las fibras musculares del control y del paciente (A, B y D), sin embargo, la subunidad γ apareció intacta en la muestra control, pero indetectable en las fibras musculares del



Ausencia de la expresión de delta sarcoglicano en el paciente MBEL050

La biopsia muscular del paciente MBEL050 mostró cambios aleatorios y difusos en la forma y en el tamaño de las fibras musculares, apoyando la presencia de miopatía y distrofia. Además, se observó un aumento en los núcleos internos y cambios distróficos con reemplazo

significativo de tejido fibroso y graso. Este paciente estaba diagnosticado con disferlinopatía, sin embargo, sus características clínicas eran más severas de las esperadas. Para ver si había afección en el complejo sarcoglicano se realizó la inmunotinción de las subunidades de SGC con los anticuerpos NCL- α (Fig.23A-B), NCL- β (Fig.23C-D), δ Sarc3/12C1 (Fig.23E-F) y 35DAG/21B5 (Fig. 23G-H) , en su respectivo centro.

En la tinción inmunohistoquímica se observó cómo tanto el SGC- α (Fig. 23A), SGC- β (Fig. 23C) y SGC- γ (Fig. 23G) estaban intactos en el paciente, aunque la expresión fue difícil de comparar con el control por el estado del tejido muscular. Sin embargo, el patrón de expresión de la subunidad δ se detectó a niveles muy bajos en MBEL050 (Fig. 22.E), lo que podía ser indicativo de una afección por la variante patogénica. No obstante, como en el caso anterior, el paciente presentaba una variante patogénica en SGC- β (c.87_89delAAG), lo que dificulta determinar que esta disminución se debiera a la variante y no a errores en la técnica.

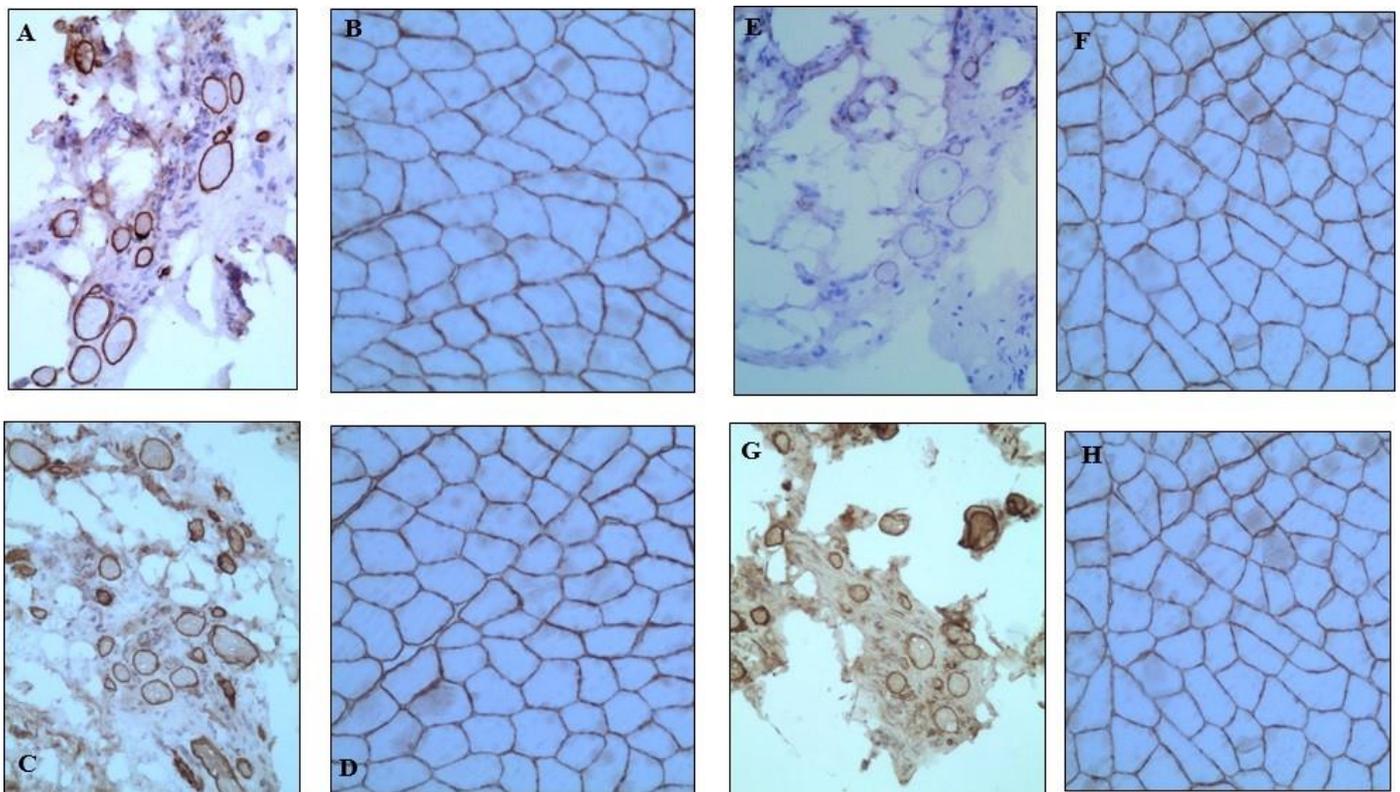


Figura 23. Examen inmunohistoquímico de la biopsia muscular en MBEL050 mostró cambios en los niveles de la subunidad δ . (A-H) Secciones de la biopsia muscular inmunohistoquímicamente teñidas para detectar los niveles de las cuatro subunidades sarcoglicanas. Las inmunoreactividades para los sarcoglicanos α , β , y δ fueron detectables en las fibras musculares del control (B, D, F y H) y del paciente (A, C, E, G), sin embargo, la subunidad δ apareció intacta en la muestra control, pero casi indetectable en las fibras musculares del paciente.

La variante en sarcoglicano no explica la severidad del fenotipo del paciente MIST069

La segregación de los familiares directos del paciente MIST069 se representan en el pedigrí de la figura 24, donde se muestra que el único individuo con disferlinopatía y portador de la variante c.850 C<T en la subunidad- α fue nuestro paciente de estudio, heredada del padre.

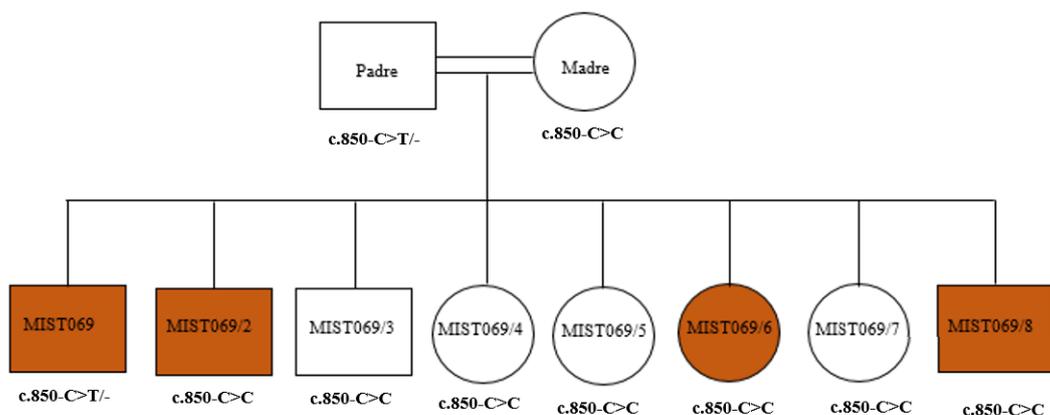


Figura 24. Análisis de segregación de la variante c.850 C<T (SGC- α) en la familia de MIST069. Familia consanguínea donde los padres son sanos (blancos) pero el 50% de sus hijos presentan disferlinopatía (marrón). Se indicant los portadores de la variante c.850 C<T en negrita.

Se analizaron los síntomas clínicos de los hermanos afectados y tanto el individuo MIST069/6 como MIST69/8 presentaron síntomas más leves que el paciente. Sin embargo, el hermano MIST069/2 presentaba valores de CK y una debilidad muscular proximal y distal más severa que nuestro paciente. No se pudo asociar la variante patogénica en sarcoglicano con un fenotipo más grave en esta familia.

4.4 NUEVOS GENES CANDIDATOS

Identificamos una serie de genes no asociados con trastornos neuromusculares en algunos pacientes. A continuación, se procedieron con sus respectivas segregaciones, sin embargo, el único gen que segregaba fue el gen PPFIA1, que codifica para la proteína Liprin- α -1.

4.4.1 LIPRIN- α -1

Liprin- α -1 pertenece a la familia liprin de proteínas de armazón citosólicas. Hasta el momento, se han realizado diferentes estudios asociados con el papel de Liprin- α -1 en los compartimientos pre-sinápticos y post-sinápticos, donde actúan como un ancla molecular para los receptores y otras proteínas durante el ensamblaje sináptico. Para conocer si Liprin- α -1 también está asociada con el correcto funcionamiento de la DGC y de la integridad funcional de las fibras musculares, al mismo tiempo que ejerce su papel en la unión neuromuscular, se realizó el estudio funcional.

*Expresión de Liprin- α -1 durante el crecimiento embrionario de *Danio rerio**

Se confirmó la validez de nuestro organismo modelo en el estudio funcional de PPFIA1 realizando la técnica RT-PCR, para ver la expresión de Liprin- α -1 en distintos estadios del crecimiento del pez cebra (Fig. 25), utilizando como proteína control el factor de elongación 1- α (EF1 α).

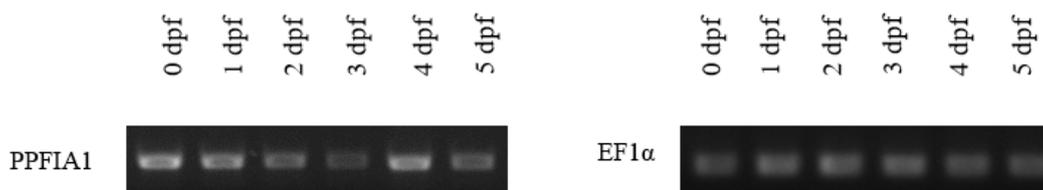


Figura 25. Confirmación por RT-PCR en los tejidos embrionarios de *Danio rerio* de la expresión de Liprin- α -1. Izquierda, se observa la expresión de la proteína de interés, y derecha la expresión de la proteína control.

Las imágenes de la figura 25 mostraron que Liprin- α -1 estaba presente desde la fecundación del embrión y se mantenía durante los primeros cinco días, lo que significa que esta proteína tiene un papel crucial en las primeras etapas del desarrollo embrionario.

Los embriones PPFIA1-MO 48hpf presentaron una respuesta anormal táctil y un fenotipo moderado/severo conforme aumentó la concentración del morfolino

Después de salir del corion, los embriones de pez cebra responden al TER rápidamente alejándose del estímulo. La respuesta normal a la TER es nadar lejos de los estímulos por al menos una longitud del propio cuerpo (Saint-Amant & Drapeau 1998), sin embargo, los embriones inyectados dieron un espectro de respuestas anormales que incluyeron nadar lejos de los estímulos, pero mostrando movimientos nerviosos, nadar en círculos o no responder al estímulo. Descubrimos que los PPFIA1-MO tenían un defecto en sus capacidades de enrollamiento después de la decorionización manual, que consideramos como un marcador fenotípico temprano de mutantes con posible desorden en la estructura muscular.

Para analizar la afección de la motilidad se midieron las distancias recorridas (mm) ante el estímulo en 15 embriones de cada grupo, observándose diferencias significativas entre los embriones WT y los embriones de todos los grupos PPFIA-MO, siendo el p-valor en todos los casos $< 0,005$ (Fig. 26A). No obstante, no se encontraron claras diferencias entre los grupos 5ng-10 ng, y 10ng-15 ng, lo que podría ser a causa de una superposición fenotípica entre los embriones a esas concentraciones del morfolino. En la figura 26B se representa mediante un diagrama de cajas las distancias recorridas de cada categoría.

Al mismo tiempo, los peces controles y PPFIA1-MO fueron sometidos a análisis de morfología macroscópica. El fenotipo general se clasificó como normal, leve/moderado o severo según el grado de curvatura de la cola. A las 48 hpf los embriones inyectados en cada una de las concentraciones del morfolino, mostraron una curvatura de la cola (Fig.26C), natación anormal y edema leve, generalmente en la region cardíaca. La gravedad de la curvatura de la cola aumentó con el incremento de la concentración del morfolino, sin embargo, hubo cierta variabilidad en la gravedad entre los embriones que también recibieron la misma concentración.

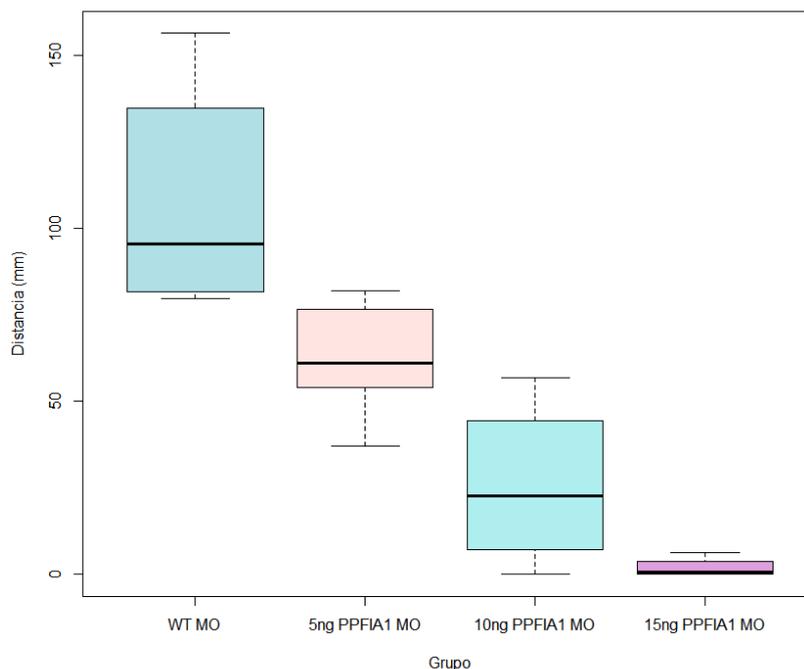
(A)

```
> summary(anova)
              Df Sum Sq Mean Sq F value    Pr(>F)
Grupo          3  37968   12656   29.26 1.64e-07 ***
Residuals     20   8652     433
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> posthoc
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = distancia ~ Grupo, data = data)

$Grupo
              diff          lwr          upr      p adj
15ng PPFIA1 MO-10ng PPFIA1 MO -23.67567 -57.286470  9.935137 0.2316217
5ng PPFIA1 MO-10ng PPFIA1 MO  36.46200  2.851196  70.072804 0.0304953
WT MO-10ng PPFIA1 MO          81.65767  48.046863 115.268470 0.0000073
5ng PPFIA1 MO-15ng PPFIA1 MO  60.13767  26.526863  93.748470 0.0003640
WT MO-15ng PPFIA1 MO         105.33333  71.722530 138.944137 0.0000002
WT MO-5ng PPFIA1 MO           45.19567  11.584863  78.806470 0.0061722
```

(B)



(C)

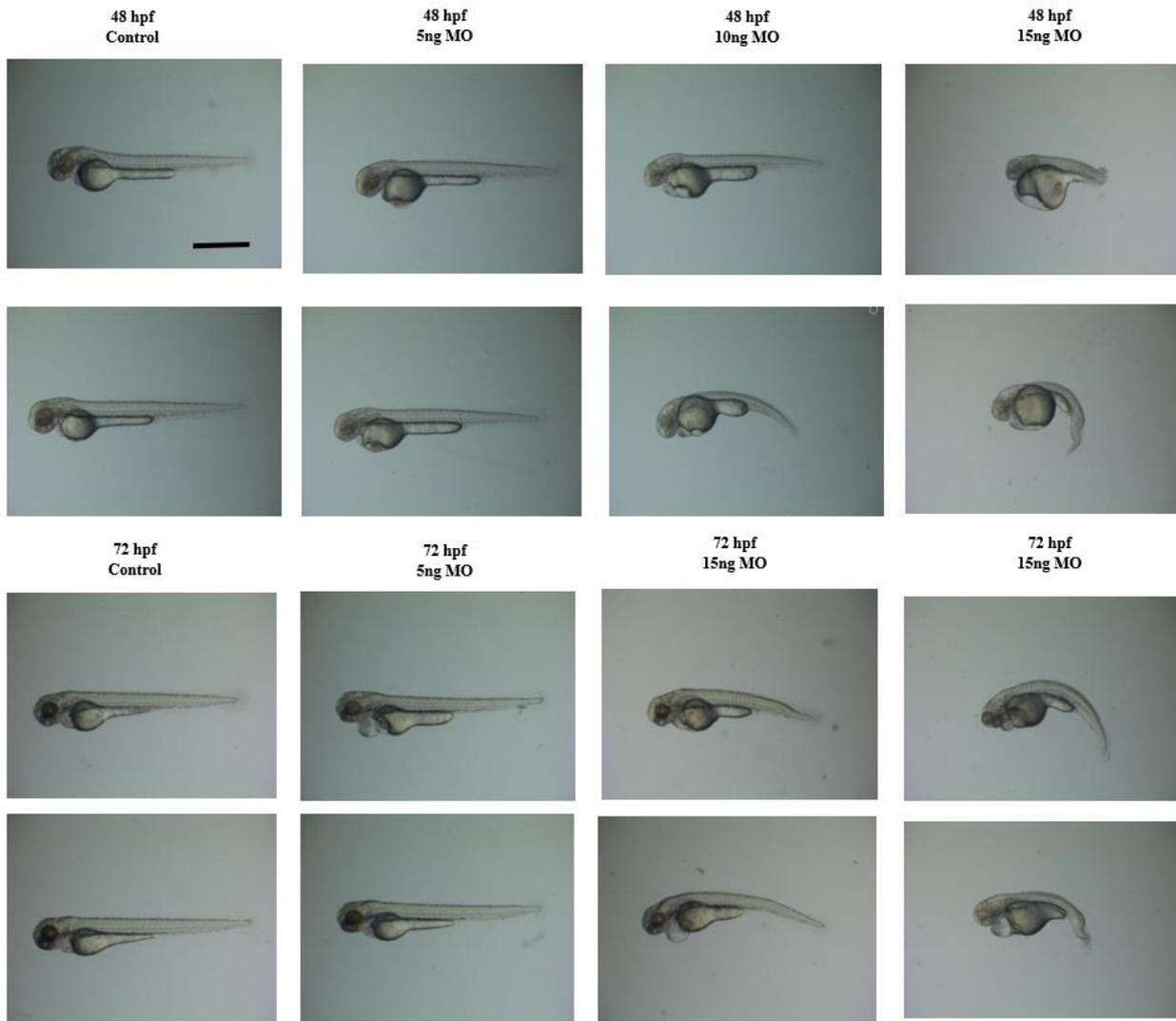


Figura 26. Estudio macroscópico del pez cebra. (A) Defecto en la morfología del pez cebra a 48 hpf a diferentes dosis de PPFIA1-MO. Embriones vivos inyectados con MO control o PPFIA1 MO. Más del 95% de los embriones inyectados con MO control no mostraron anomalías macroscópicas. Las imágenes de los PPFIA1-MO representan defectos leves (izquierda, 15% de los embriones vivos), moderados (medio, 32% de los embriones vivos) y severos (derecha, 68% de los embriones vivos) en términos de longitud y curvatura de las colas. Se contaron al menos 15 embriones vivos por condición. La barra de escala representa 1 mm. (B) Resultados obtenidos por la prueba paramétrica ANOVA. (C) Diagrama de cajas de las distancias recorridas, medidas en mm, por cada grupo de estudio (WT, 5 ng, 10 ng y 15 ng).

La desestructuración de la actina filamentosa apoya un defecto muscular temprano en los embriones PPFIA-MO

Nuestra observación de embriones PPFIA1-MO decorionizados a 48 hpf sugirió que el músculo esquelético mutante podía ser anormal en esta etapa temprana del desarrollo muscular. Para investigar esto más a fondo, utilizamos la tinción de faloidina como un medio para detectar cambios en la organización de las miofibrillas. Los embriones control a 48 hpf mostraron el patrón esperado de tinción con la faloidina. Los embriones de PPFIA1-MO, en contraste, mostraron una estructura alterada en la organización de las miofibrillas, en concreto en los filamentos de actina, siendo más destacada en la región donde se produce la unión miotendinosa (Fig. 27.B).

En la representación 27.A se observa de forma esquematizada los cambios morfológicos que las fibras musculares sufrieron frente a una disminución en los niveles de Liprin- α -1. Los peces cebrá con una concentración de 5ng comenzaron a presentar fibras divididas y cortas, perdiendo parte de su integridad. A unas concentraciones superiores (10-15 ng), las fibras musculares de los embriones se retorcieron y ondularon, y perdieron parte de la alineación con la F-actina.

Si bien el significado completo de esta observación merece una mayor investigación, estos datos respaldan un fenotipo temprano del músculo esquelético en mutantes PPFIA1.

Algunos embriones PPFIA-MO mostraron un retraso en el desarrollo de la NMJ

La morfología de la NMJ de los embriones se visualizó usando los anticuerpos anti-SV2 y α -bungarotoxina (Fig. 28). En los embriones más gravemente afectados en cada concentración del morfolino había un fenotipo NMJ claro.

En los controles y en algunos embriones inyectados, los axones motores se ramificaron en todas las partes del somita donde formaron contactos con los grupos AChR (Fig.28 C). En los embriones inyectados con PPFIA1-MO más severos, los grupos AChR se ubicaron predominantemente a lo largo de la superficie medial, al mismo tiempo que los axones motores se acortaron a largo del miosepto, disminuyendo su ramificación. Las membranas de las neuronas motoras y las membranas de las fibras musculares deben estar relativamente colocadas, sin embargo, en algunos de los embriones hubo desviaciones en esta organización, mostrando separaciones entre ambas membranas (Fig.28 D). Aunque se observa cierta afección en el desarrollo del NMJ, en la mayoría de los embriones se mantuvo la estructura de los somitas y un desarrollo de la NMJ bastante similar a la observada en los embriones controles (Fig.28 A y B).

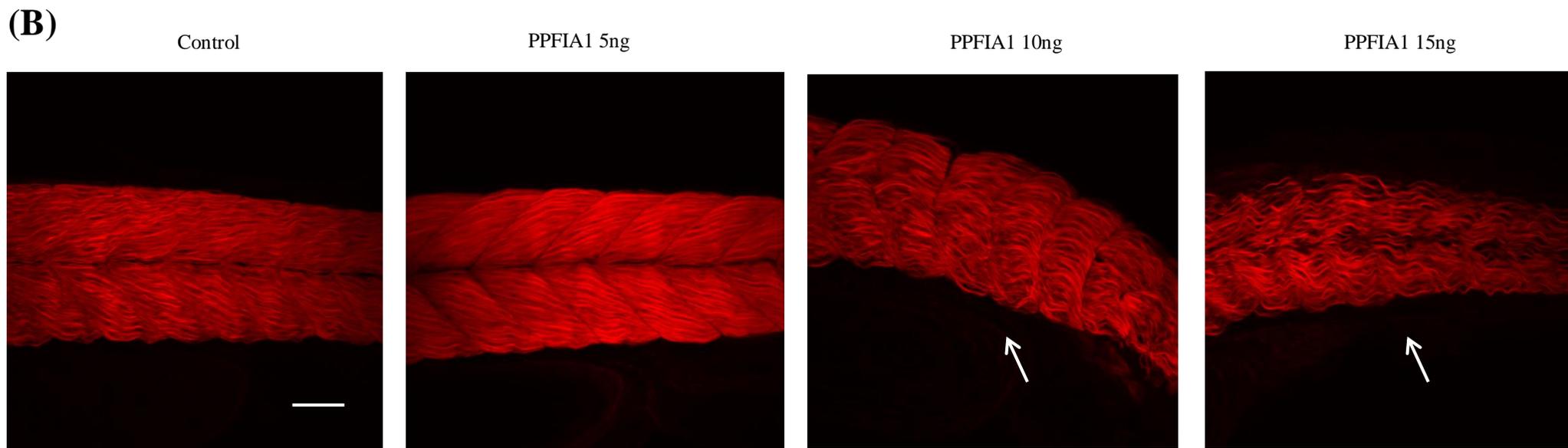
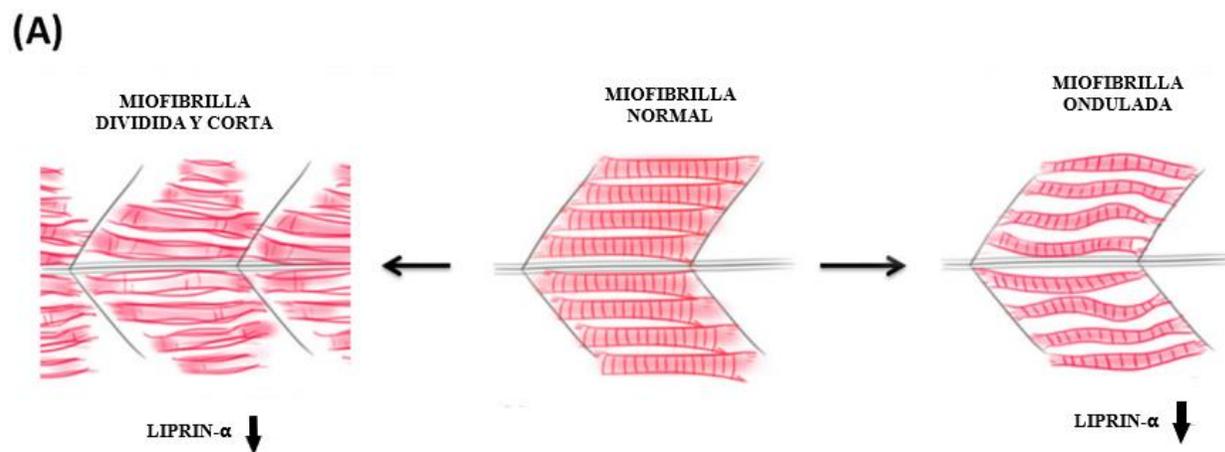


Figura 27. Desorganización miofibrilar de los PPFIA-MO. (A) Representación esquemática de los cambios morfológicos que sufren las miofibrillas ante una reducción en los niveles de la proteína Liprin- α . (B) Los peces control y PPFIA1-MO antisentido se tiñeron con faloidina para marcar la actina filamentosa y luego se visualizó el montaje completo por microscopía confocal. Los controles muestran el patrón esperado de tinción a 48 hpf, mostrando la naturaleza organizada de las fibras rápidas de los peces. En el músculo de mutantes PPFIA1, hay una desorganización intensa en el alineamiento miofibrilar de la actina (flecha). Barra de escala = 100 μ m.

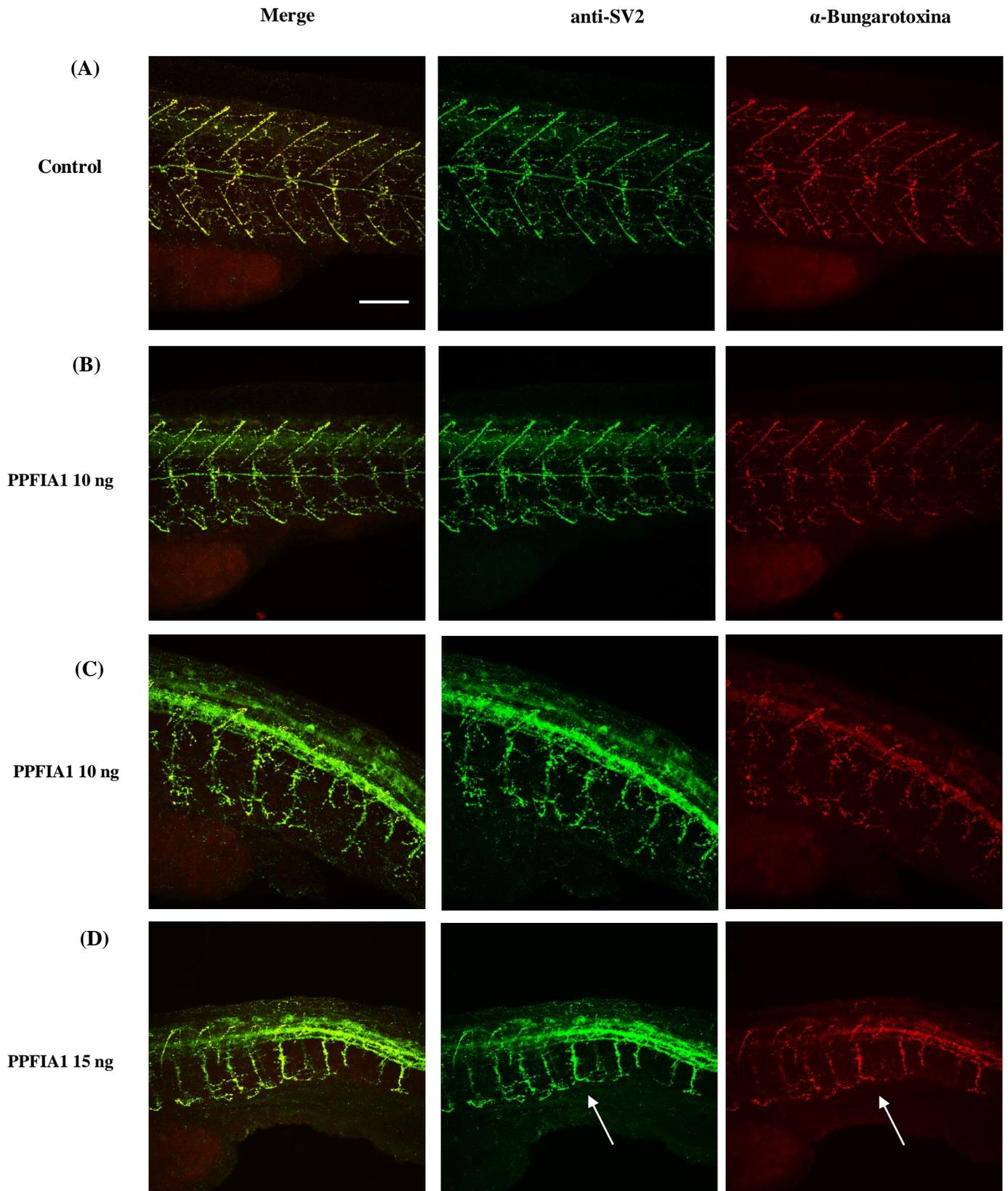


Figura 28. Estudios de inmunofluorescencia en Danio rerio a 48 hpf. Vistas laterales de de las sinapsis neuromusculares marcadas con anticuerpos contra SV2 y α -Bungarotoxina en embriones de 48 hpf (A-D). Las señales amarillas (Merged) indican la co-localización de las señales anti-SV2 y α -BTX. Los embriones inyectados con PPFIA1-MO a 15 ng mostraron acortamiento de las neuronas motoras, y presencia de algunas neuronas que no se extendieron a lo largo de la miosepta, entre los somitas (flecha blanca). Barra de escala = 100 μ m.

5. DISCUSIÓN

Los avances recientes en los procedimientos de diagnóstico genético y las tecnologías de alto rendimiento, como la secuenciación de todo el gen, han permitido el diagnóstico de casos comunes y atípicos. La secuenciación completa del exoma se está convirtiendo en una tecnología asequible para investigar enfermedades raras. Las regiones funcionales y de codificación representan solo el 1% de todo el genoma humano, pero albergan el 85% de las variantes causantes de enfermedades conocidas (Majewski, et al., 2011). Sobre esta base, se secuenciaron los exomas de una cohorte de 1466 pacientes con debilidad de las cinturas sin diagnóstico.

Dada la superposición fenotípica y la rareza de algunas de estas afecciones musculares, es casi imposible para los médicos, por expertos que sean, diagnosticar individualmente cada enfermedad. En este proyecto intentamos mejorar la vía diagnóstica de los pacientes con LGMD mediante la implementación de WES dirigida como estrategia de diagnóstico. Al tratarse de un proyecto colaborativo a nivel mundial ha sido esencial la comunicación con los respectivos médicos para proporcionar un fenotipado inverso; esto es, una reevaluación de los hallazgos clínicos, histopatológicos y de resonancia magnética.

Los primeros resultados del proyecto MYO-SEQ enfatizan la gama de fenotipos asociados con mutaciones en genes identificados originalmente en pacientes con CMD y miopatía de Bethlem, y su superposición en el espectro de LGMD. Estos datos recomendaron una consideración abierta de todos los genes conocidos relacionados con miopatía para realizar la WES (listado de 471 genes) (manuscrito en proceso de publicación). La heterogeneidad genética junto a la creciente heterogeneidad fenotípica de los genes asociados con las enfermedades neuromusculares, han cambiado la visión clásica de entender las enfermedades Mendelianas, considerando la posibilidad de observar un fenotipo más leve para un gen implicado en la miopatía de aparición congénita (Wallgren-Pettersson, et al., 1999), manifestaciones clínicas muy diversas provocadas por el mismo gen (McCarthy, et al., 2000), variabilidad en los síntomas y en la edad de inicio (Sung, et al., 2018), así como una posible novedosa presentación recesiva vs herencia típica dominante (Findlay, et al., 2018) o viceversa (Martinez-Thompson, et al., 2018).

El objetivo principal de este trabajo ha sido ampliar el porcentaje de casos resueltos en la cohorte de MYO-SEQ, teniendo en cuenta diversas hipótesis anteriormente no estudiadas para explicar la falta de diagnóstico tras la secuenciación del exoma completo. Las hipótesis propuestas han sido: herencia digénica, variantes intrónicas y genes modificadores. Al mismo tiempo, hemos realizado un estudio funcional que provee evidencias para un gen recientemente identificado como posiblemente causante de una enfermedad neuromuscular.

5.1 HERENCIA DIGÉNICA

Estudiar si existe o no herencia digénica requiere de diversas pruebas que complementen la información genética obtenida para sopesar un nuevo mecanismo genético causante del fenotipo estudiado. El enfoque de esta hipótesis tanto en sarcoglicanopatías como en distroglicanopatías se ha debido al conjunto de genes involucrados en ambas distrofias musculares, y en el porcentaje de pacientes sin diagnosticar con una variante patogénica en dos de las subunidades de SGC (3,34%) o en una variante de los dieciocho genes asociados con distroglicanopatías (10,6%) sin diagnosticar.

La existencia de una fuerte interacción proteica entre las subunidades del complejo sarcoglicano y las proteínas asociadas con la ruta de la O-glicosilación se ha basado en ensayos publicados o en bases de datos biológicas como STRING.

5.1.1 SARCOGLICANOPATÍAS

Los trastornos musculares asociados con mutaciones de pérdida de función en alguna de las subunidades del SGC tienen inicio en la infancia e involucran funciones tanto cardíacas como respiratorias. Estos subtipos representan el 20-25% de todos los LGMD y el 40-65% de los casos de LGMD con inicio infantil (Semplicini, et al., 2015). Por otro lado, ensayos *in vitro* establecieron la asociación entre las subunidades del complejo sarcoglicano, mediante el uso de detergentes desnaturizantes como octil glucósido (Yoshida, et al., 1994) y SDS, sugiriendo que los SGC están estrechamente asociados entre sí, principalmente los SGC β , γ y δ gracias a la presencia de enlaces de disulfuro intramoleculares (Chan, et al., 1998).

La gran tasa de portadores de una variante patogénica en SGC, la fuerte interacción proteína-proteína entre las subunidades del complejo (Schäffer, 2013), la alta frecuencia de diagnóstico en pacientes con LGMD y el hecho de que el diagnóstico molecular de las sarcoglicanopatías siga siendo una tarea difícil (Trabelsi, et al., 2008), apuntan a un posible patrón de herencia digénica en estos subtipos de LGMD.

Los casos obtenidos tras la serie de filtrados realizados (n=1) dificultan la verificación de este mecanismo molecular y patológico como causa de la enfermedad. Además, la poca información clínica obtenida, las variantes identificadas c.205-G> A (p.Gly69Ser) en SGC- γ y c.446A> C (p.Asn149Thr) en SGC- ϵ , que no están descritas en la literatura, y la falta de pruebas complementarias impiden aceptar o rechazar la hipótesis de una posible herencia digénica.

Una de las principales incógnitas es la baja frecuencia de pacientes portadores de dos variantes patogénicas en dos subunidades de SGC, hecho que puede darse porque aquellos pacientes con dos variantes están diagnosticados con otro subtipo de LGMD, debido a la falta de causalidad, o porque la afección de dos subunidades puede desencadenar en un fenotipo más grave que el esperado, dando lugar incluso a una muerte prematura.

Hasta el momento, no se ha publicado ningún ensayo basado en el estudio de herencia digénica en estos subtipos de LGMD, sin embargo, en un artículo publicado en el año 2008 aparece en la cohorte de estudio un paciente que presenta mutaciones heterocigotas dobles en β y δ , concluyendo que su debilidad muscular y miocardiopatía podría ser el resultado de un posible efecto sinérgico perjudicial de ambas variantes sobre la función muscular (Trabelsi, et al., 2008).

Se requiere, por tanto, de una recopilación mayor de casos sin diagnóstico y portadores de variantes en diferentes subunidades del complejo para esclarecer este posible mecanismo molecular.

5.1.2 DISTROGLICANOPATÍAS

La herencia digénica se ha informado en algunos tipos de distrofia muscular, principalmente en la distrofia facioescapulohumeral tipo 2 (Lemmers, et al., 2012), y en un caso de LGMD causado por variantes en *DES* y *CAPN3* (Peddareddygar, et al., 2018), donde genes que no pertenecen al mismo complejo proteico actúan de manera sinérgica en el desarrollo de la enfermedad.

La idea de enfocar este tipo de herencia en las distroglicanopatías se debe al gran número de genes asociados con estos subtipos de distrofia muscular y al porcentaje de portadores de una variante patogénica en nuestra cohorte (10,6%). Además, el análisis de la vía STRING predice vías comunes que implican a estos genes y proporciona una verosimilitud biológica acerca de cómo las proteínas codificadas por estas mutaciones podrían interactuar y provocar una miopatía.

Sin embargo, los datos clínicos recogidos por los médicos de cada paciente de nuestro estudio y las pruebas inmunohistoquímicas realizadas indican que la causa genética de estos pacientes no se ajusta a una herencia digénica, porque el fenotipo no se asocia al típico de distroglicanopatías o bien porque la glicosilación de α -distroglicano se mantiene intacta.

A pesar de ello, no podemos descartar este tipo de herencia en todos los genes asociados a la O-glicosilación, porque existe una mayor interacción entre diferentes genes de la ruta como, por ejemplo: *DPM1-DOLK*, *DPM1-GMPPB* y *DPM1-DPM2-DPM3*, que no aparecen en ninguno de nuestros casos. Esto, junto a la falta de pruebas complementarias en algunos pacientes, no excluye del todo la herencia digénica en las distroglicanopatías. Sugerimos que para profundizar más en este abordaje se requiere de una estrategia metodológica y analítica más completa apoyada en algoritmos aplicados en la identificación de múltiples variables candidatas que contribuyen al fenotipo, como es el caso de OligoPVP (Boudellioua et al., 2018).

5.2 VARIANTES INTRÓNICAS

La sobrerrepresentación en nuestra cohorte de portadores de variantes patogénicas previamente informadas en genes recesivos probablemente implica que una segunda variante críptica aún no se ha identificado. De hecho, cuando se interrogaron estos casos aún más, se identificaron algunas variantes sospechosas recurrentes intrónicas.

5.2.1 LGMDR2

Este subtipo de LGMD es la segunda causa de distrofia muscular más frecuente en la cohorte de MYO-SEQ (8%) al igual que en la población europea. La alta tasa de casos sin diagnosticar junto al número de pacientes con una variante patogénica (Fig. 29) son indicadores de posibles casos de LGMDR2 todavía no resueltos, que junto a los valores observados en otros estudios (Dastur et al., 2017) apoyan la teoría de la presencia de nuevas variantes intrónicas patogénicas en este gen.

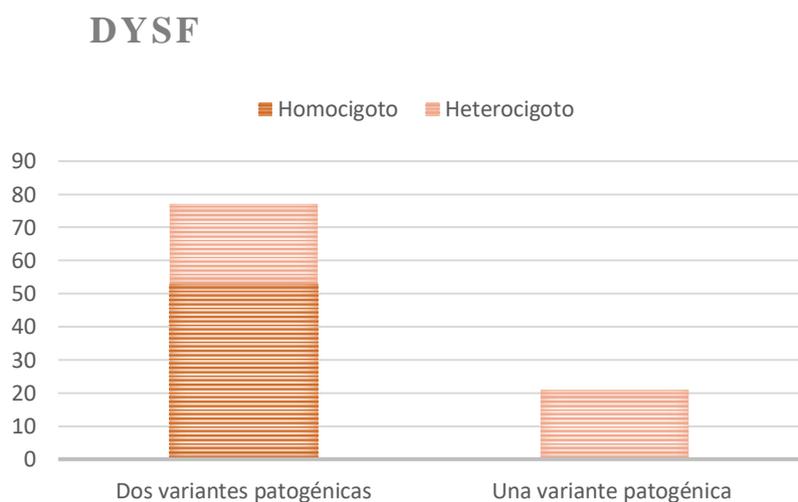


Figura 29. Representación del número de variantes en el gen *DYSF*. En la primera columna se observa el porcentaje de casos con dos variantes patogénicas (53 homocigotas y 24 heterocigotas); y en la segunda columna aquellos pacientes que tienen una sola variante patogénica (22 individuos).

Para reducir el análisis obtuvimos, mediante una serie de filtrados, un listado de pacientes (n=21) con un fenotipo asociado a LGMDR2 que presentaban una única variante patogénica. En lugar de estudiar nuevas variantes intrónicas y corroborar su patogenicidad, se analizaron primariamente dos variantes ya publicadas (c.4886+1249 G>T y c.5668-824C> T) y obtuvimos un caso positivo (MRIG004). Los veinte pacientes restantes no consiguieron un diagnóstico, considerando que presentan una variante intrónica todavía no descrita en *disferlina* o, por otro lado, que están asociados a otros subtipos de LGMD cuyas variantes no se han detectado mediante WES.

En vista de los resultados obtenidos, dado que la tasa de diagnóstico ha sido del 4,76% en pacientes con una variante patogénica en *DYSF*, se recomienda la confirmación mediante la secuenciación de Sanger de ambas variantes intrónicas en futuros ensayos de análisis genético de las enfermedades neuromusculares.

5.2.2 LGMDR1

La distrofia muscular de la cintura y las extremidades tipo R1 es la más frecuente en la población europea y en la cohorte de MYO-SEQ (13%). La alta frecuencia junto al gran tamaño del gen lo hacen un buen candidato para profundizar el estudio en las regiones no codificantes.

A pesar de la frecuente identificación de cambios intrónicos en el gen *CAPN3*, la demostración de su efecto patogénico requiere de laboriosos análisis y, sobre todo, de la disponibilidad de tejido muscular de los pacientes. Para facilitar este abordaje se utilizaron predicciones *in-silico* combinadas con el análisis de RT-PCR en el ARNm extraído del músculo y también de sangre periférica. Esto demostró la importancia patológica de las variantes c.1746-20C>G y c.1354+129T>A, mostrando sus efectos en los sitios aceptores de empalme y en la traducción de la proteína.

Aunque la demostración experimental de los efectos patogénicos de las mutaciones intrónicas es laborioso, descubrimos que a menudo tiene éxito cuando este esfuerzo se basa en un análisis bioinformático preliminar empleando predictores *in silico*. Este enfoque debe llevarse a cabo con mayor frecuencia debido a sus importantes consecuencias para el asesoramiento clínico y genético, para establecer correlaciones genotipo-fenotipo, proporcionar nuevas ideas sobre el complejo mecanismo de empalme y para facilitar un futuro tratamiento.

5.3 GENES MODIFICADORES

Estudios recientes destacaron cómo las interacciones entre genes en los trastornos neuromusculares pueden ser más frecuentes de lo estimado contribuyendo a la variabilidad del fenotipo. Descubrir los mecanismos moleculares involucrados podría tener una amplia aplicabilidad en el diagnóstico y como estrategia terapéutica para muchas enfermedades hereditarias y adquiridas.

Somos conscientes de que este estudio presenta algunas limitaciones: no se pudo obtener material genético del paciente MTEH057 y de sus familiares para segregar la variante, el tamaño muestral es pequeño (n=3), y las imágenes obtenidas de los centros no tienen una gran calidad, por lo que no podemos descartar errores en la técnica de tinción. Sin embargo, que los pacientes MBEL050 y MBEL119 diagnosticados con LGMDR2 y LGMDR1, respectivamente, presenten la misma variante y una reducción en una subunidad del complejo sarcoglicano puede indicar una contribución de la variante patogénica en el fenotipo.

Algunos estudios han determinado que el patrón de tinción puede ser variable e inespecífico en las diferentes formas de sarcoglicanopatía (Klinge et al., 2008), datos que respaldan nuestros resultados al haber reducción de una subunidad (SGC- γ o SGC- δ) en presencia de una variante patogénica en SGC- β . Aunque está más allá del alcance del presente estudio, en el futuro será importante replicar estos hallazgos en un mayor número de pacientes con LGMD para validar estos resultados.

Por otro lado, la segregación del paciente MIST069 no pudo explicar la variabilidad fenotípica entre los familiares afectados. Aun así, no podemos descartar el papel de la variante c.850-C>T en la progresión de la LGMD, debido a la posibilidad de que el hermano más afecto sea portador de otra variante patogénica que agrave su sintomatología.

En conjunto, nuestros datos demuestran que estas variantes heterocigotas patogénicas podrían estar actuando como modificadores de la enfermedad, contribuyendo a la gravedad y/o progresión del fenotipo.

5.4 IDENTIFICACIÓN DE NUEVOS GENES CAUSATIVOS

Dentro de nuestro estudio identificamos y caracterizamos funcionalmente un gen nuevo, como posible candidato para explicar la causa genética de un paciente español que presentaba una distrofia progresiva proximal, fatigabilidad y disfagia neuromuscular. Este gen no ha sido asociado con ninguna enfermedad o fenotipo previamente. Sin embargo, estudios funcionales apoyan su papel en la actividad sináptica (Hoogenraad, et al.,2007), en la unión neuromuscular y en la organización de la maquinaria post-sináptica (Bernadzki, et al., 2017).

Los embriones del pez cebra inyectados con el MO mostraron curvaturas anormales de la cola, respuestas motoras deterioradas y axones aberrantes, lo que respalda la causalidad del gen PPFIA1 en la distrofia muscular y en la afección del NMJ, siendo el fenotipo muscular mayor que el neurológico, acorde al fenotipo del paciente. La presencia del edema cardíaco en la mayoría de los embriones MO no está claro si se debe a la caída de la expresión de Liprin- α -1 o a un efecto del proceso de inyección, por lo que debería ser discernido con el uso de un oligonucleótido morfolino p53 para ver si se alivia este síntoma en próximos estudios (Robu et al., 2007). Estos datos preliminares obtenidos tras el estudio funcional indican que variantes en el gen PPFIA1 promueven un fenotipo con afección neuromuscular.

Ante el porcentaje de casos sin resolver, el proyecto MYO-SEQ amplió la lista de 169 genes (Johnson, et al., 2016) a un total de 471 genes causantes de enfermedades neuromusculares, incluidos los asociados con neuropatías, cardiomiopatías y otras afecciones que presentan debilidad muscular. El proyecto MYO-SEQ apoya para reducir el tiempo de análisis el empleo de paneles de genes antes de considerar WES y, en los casos no resueltos, analizar nuevas variantes predichas como patogénicas en busca de nuevos genes asociados a enfermedades neuromusculares, como en el caso de PPFIA1.

Al centrarnos en un panel de 471 genes, se resolvieron el 56% de nuestra cohorte. El aumento relativamente pequeño que observamos al expandir la lista original de 169 a 471 genes (Fig. 30) está en línea con un panel NGS de >700 genes neuromusculares para los cuales el WES de los casos no diagnosticados restantes no mejoró la tasa resuelta (Monies, et al.,2016).

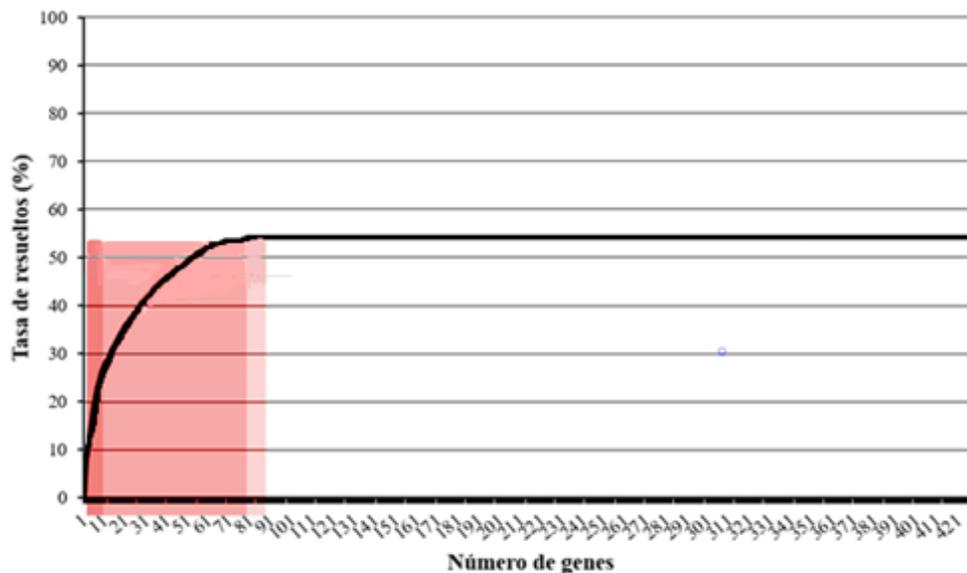


Figura 30. Impacto del aumento del número de genes analizados en la tasa resuelta.

En la figura 30 se aprecia como los ocho genes de enfermedades más comunes representan una cuarta parte de la cohorte completa. Los siguientes diez genes representan menos del 10% adicional, seguido de un aumento muy lento en la tasa de resolución ya que solo se detectan genes de enfermedades muy raras en el análisis. La constancia en la tasa de casos resueltos conforme aumenta el número de genes asociados a enfermedades neuromusculares se confirma con el nuevo gen caracterizado en nuestro estudio (PPFIA1) que aumentaría tan solo un 0,07% la tasa del diagnóstico.

El aumento de la tasa de resolución tras haber finalizado este proyecto ha sido de un 1,5%, siendo un 1,43% por la identificación de variantes en regiones no codificantes. Estos datos respaldan la idea de una falta de diagnóstico a causa de las limitaciones de WES en lugar de la falta de identificación de nuevos genes candidatos.

El gran porcentaje de portadores de una variante patogénica, la relación entre el número de genes y la tasa de resolución, las variantes intrónicas analizadas en el estudio y la comprobación de variantes publicadas en nuestra cohorte, apoyan la idea de cambiar el enfoque de diagnóstico en la clínica integrando información genómica funcional (RNA-seq), que proporciona información directa sobre las perturbaciones transcripcionales causadas por cambios genéticos (Cumming, et al., 2017). Esta estrategia podría utilizarse no solo para validar variantes con significado incierto identificadas en la secuenciación del exoma, sino también para conocer nuevos mecanismos causantes de estas enfermedades (Gonorazky, et al., 2016), (Kremer, et al. 2017).

Una proporción adicional de los casos restantes sin resolver podría ser causada por nuevos mecanismos moleculares no descubiertos hasta ahora, como la impronta genómica, o incluso la herencia digénica, mecanismo que no ha sido demostrado en nuestro estudio pero que debido a las

limitaciones del proyecto no podemos descartar. También es posible que algunos de los pacientes no resueltos puedan haber adquirido formas de enfermedad muscular, tales como miopatías inmunitarias, inflamatorias o incluso inducidas por estatinas. Todas estas opciones deben de considerarse en los próximos estudios a realizar, manteniendo la base de nuestro proyecto: integrar la información clínica del paciente con los datos analizados, realizando un trabajo multidisciplinar y colaborativo.

5.5 LIMITACIONES DEL ESTUDIO

Al tratarse de un proyecto de colaboración europeo, la necesidad de estar en un continuo contacto con los clínicos de cada paciente ha dificultado, en algunos casos, la obtención de las muestras serológicas o de biopsia muscular para proceder con los ensayos, al igual que las imágenes de las tinciones inmunohistoquímicas o de resonancia magnética para confirmar o rechazar los hallazgos genéticos.

Otra de las limitaciones la hemos encontrado al realizar la técnica Western blot en los embriones del pez cebra PPFIA-MO para analizar la expresión de la proteína Liprin- α -1. El marcaje con los anticuerpos empleados no ha permitido diferenciar si existe una reducción significativa en los embriones inyectados con el MO frente a los controles.

Por último, la dificultad de extender la estancia en el centro de investigación John Walton Muscular Dystrophy, ha impedido realizar ciertos ensayos durante el verano para optimizar las condiciones de algunas técnicas y mejorar los resultados obtenidos.

Por tanto, debido a la falta de datos, al gran abordaje durante este estudio y al tiempo limitado, el presente trabajo debe considerarse como un estudio inicial en el que se necesita seguir trabajando para poder aumentar la tasa de diagnóstico en la cohorte de MYO-SEQ.

5. CONCLUSIONES

1. La herencia digénica como posible mecanismo molecular para las sarcoglicanopatías y distroglicanopatías no pudo explicar el fenotipo clínico de los pacientes interrogados. Tanto el pequeño tamaño muestral como las variantes identificadas en los pacientes de estudio dificultaron la comprobación de esta hipótesis, sin embargo, en vista de los datos publicados y de la alta tasa de casos sin diagnosticar con una variante patogénica, no se puede descartar este mecanismo por el momento. Se requieren de más estudios y de un abordaje más completo para confirmar la presencia o ausencia de herencia digénica en estos subtipos de LGMD.
2. La tasa de diagnóstico de la cohorte de MYO-SEQ tras el análisis de variantes intrónicas publicadas en *DYSF* e identificadas en *CAPN3* aumentó en la cohorte de MYO-SEQ en un 1,43%, lo que confirma las limitaciones de WES y la alta tasa de variantes en regiones no codificantes que todavía no han sido descritas como causantes de enfermedad.
3. La variabilidad clínica de los pacientes con formas comunes de LGMD puede estar asociada a la presencia de variantes modificadoras que contribuyen a la gravedad y/o progresión del fenotipo. El papel modificador de una variante patogénica en *SGC* no ha sido demostrada en este estudio, aunque los datos obtenidos pueden suponer un punto de partida para próximos ensayos.
4. El estudio funcional de *PPFIA1* en *Danio rerio* confirmó el papel de este gen en la estructura de la actina filamentosa y en la unión neuromuscular, principalmente en el área post-sináptica, datos que sostienen este gen como causante del fenotipo de un paciente con dos variantes patogénicas en este gen. Se prevé realizar un estudio más completo para analizar con detalle las características del pez cebra extrapolando los resultados a la clínica.
5. El aumento de genes en la lista analizada en este estudio (471 genes) no ha conseguido aumentar el porcentaje de diagnóstico, datos que demuestran la necesidad de otro enfoque para dilucidar la causa genética de estos pacientes. Proponemos para futuros estudios el empleo de WGS junto con RNA-seq, manteniendo la continua integración de los datos clínicos de cada centro.

6. REFERENCIAS

- 1) Alfares, A., Aloraini, T., Alissa, A., Al Qudsi, A., Alahmad, A., Al Mutairi, F., ... & Alturki, S. (2018). Whole-genome sequencing offers additional but limited clinical utility compared with reanalysis of whole-exome sequencing. *Genetics in Medicine*, 20(11), 1328.
- 2) Ambrosi, X., Lamothe, L., Heming, N., & Orlikowski, D. (2015). Respiratory and intensive care aspects of muscular dystrophies. *Archives de pediatrie: organe officiel de la Societe francaise de pediatrie*, 22(12 Suppl 1), 12S51-7.
- 3) Angelini, C., Giaretta, L., & Marozzo, R. (2018). An update on diagnostic options and considerations in limb-girdle dystrophies. *Expert review of neurotherapeutics*, 18(9), 693-703.
- 4) Bernadzki, K. M., Gawor, M., Peźniński, M., Mazurek, P., Niewiadomski, P., Rędownicz, M. J., & Prószyński, T. J. (2017). Liprin- α -1 is a novel component of the murine neuromuscular junction and is involved in the organization of the postsynaptic machinery. *Scientific reports*, 7(1), 9116.
- 5) Blázquez, L., Azpitarte, M., Sáenz, A., Goicoechea, M., Otaegui, D., Ferrer, X., ... & De Munain, A. L. (2008). Characterization of novel *CAPN3* isoforms in white blood cells: an alternative approach for limb-girdle muscular dystrophy 2A diagnosis. *Neurogenetics*, 9(3), 173-182.
- 6) Blázquez, L., Aiastui, A., Goicoechea, M., Martins de Araujo, M., Avril, A., Beley, C., ... & Lopez de Munain, A. (2013). In Vitro Correction of a Pseudoexon-Generating Deep Intronic Mutation in *LGMDR1* by Antisense Oligonucleotides and Modified Small Nuclear RNA s. *Human mutation*, 34(10), 1387-1395.
- 7) Boudellioua, I., Kulmanov, M., Schofield, P. N., Gkoutos, G. V., & Hoehndorf, R. (2018). OligoPVP: Phenotype-driven analysis of individual genomic information to prioritize oligogenic disease variants. *Scientific reports*, 8(1), 14681.
- 8) Cacciottolo, M., Numitone, G., Aurino, S., Caserta, I. R., Fanin, M., Politano, L., ... & Nigro, V. (2011). Muscular dystrophy with marked *DYSF*erlin deficiency is consistently caused by primary *DYSF*erlin gene mutations. *European journal of human genetics*, 19(9), 974.
- 9) Chan, Yiu-mo, et al. "Molecular organization of sarcoglycan complex in mouse myotubes in culture." *J Cell Biol* 143.7 (1998): 2033-2044.
- 10) Chardon, J. W., & Straub, V. (2017). The role of muscle imaging in the diagnosis and assessment of children with genetic muscle disease. *Neuropediatrics*, 48(04), 233-241.
- 11) Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., ... & Estrella, E. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Science translational medicine*, 9(386), eaal5209.
- 12) Daikoku, E., Saito, M., & Ono, F. (2015). Zebrafish mutants of the neuromuscular junction: swimming in the gene pool. *The Journal of Physiological Sciences*, 65(3), 217-221.

- 13) Dastur, R. S., Gaitonde, P. S., Kachwala, M., Nallamilli, B. R., Ankala, A., Khadilkar, S. V., ... & Shira, S. (2017). Detection of *DYSF*erlin gene pathogenic variants in the indian population in patients predicted to have a *DYSF*erlinopathy using a blood-based monocyte assay and clinical algorithm: a model for accurate and cost-effective diagnosis. *Annals of Indian Academy of Neurology*, 20(3), 302.
- 14) Di Fruscio, G., Garofalo, A., Mutarelli, M., Savarese, M., & Nigro, V. (2016). Are all the previously reported genetic variants in limb girdle muscular dystrophy genes pathogenic?. *European Journal of Human Genetics*, 24(1), 73.
- 15) Dominov, Janice A., et al. "A novel *DYSF*erlin mutant pseudoexon bypassed with antisense oligonucleotides." *Annals of clinical and translational neurology* 1.9 (2014): 703-720.
- 16) Dominov, J. A., Uyan, Ö., McKenna-Yasek, D., Nallamilli, B. R. R., Kergourlay, V., Bartoli, M., ... & Krahn, M. (2019). Correction of pseudoexon splicing caused by a novel intronic *DYSF*erlin mutation. *Annals of Clinical and Translational Neurology*.
- 17) Emery, A. E. (2002). The muscular dystrophies. *The Lancet*, 359(9307), 687-695.
- 18) Fichna, J. P., Macias, A., Piechota, M., Korostyński, M., Potulska-Chromik, A., Redowicz, M. J., & Zekanowski, C. (2018). Whole-exome sequencing identifies novel pathogenic mutations and putative phenotype-influencing variants in Polish limb-girdle muscular dystrophy patients. *Human genomics*, 12(1), 34.
- 19) Findlay, A. R., Harms, M. B., Pestronk, A., & Weihl, C. C. (2018). Homozygous recessive *MYH2* mutation mimicking dominant *MYH2* associated myopathy. *Neuromuscular Disorders*, 28(8), 675-679.
- 20) Finsterer, J., Stöllberger, C., & Wahbi, K. (2013). Cardiomyopathy in neurological disorders. *Cardiovascular Pathology*, 22(5), 389-400.
- 21) Fischer, D., Bonati, U., & Wattjes, M. P. (2016). Recent developments in muscle imaging of neuromuscular disorders. *Current opinion in neurology*, 29(5), 614-620.
- 22) Ghaoui, R., Cooper, S. T., Lek, M., Jones, K., Corbett, A., Reddel, S. W., ... & O'grady, G. (2015). Use of whole-exome sequencing for diagnosis of limb-girdle muscular dystrophy: outcomes and lessons learned. *JAMA neurology*, 72(12), 1424-1432.
- 23) Gonorazky, H., Liang, M., Cummings, B., Lek, M., Micallef, J., Hawkins, C., ... & Marshall, C. R. (2016). RNA seq analysis for the diagnosis of muscular dystrophy. *Annals of clinical and translational neurology*, 3(1), 55-60.
- 24) Groh, W. J. (2012). Arrhythmias in the muscular dystrophies. *Heart rhythm*, 9(11), 1890-1895.
- 25) Hogarth, M. W., Houweling, P. J., Thomas, K. C., Gordish-Dressman, H., Bello, L., Vishwanathan, V., ... & Tulinius, M. (2017). Evidence for *ACTN3* as a genetic modifier of Duchenne muscular dystrophy. *Nature communications*, 8, 14143.

- 26) Hoogenraad, C. C., Feliu-Mojer, M. I., Spangler, S. A., Milstein, A. D., Dunah, A. W., Hung, A. Y., & Sheng, M. (2007). Liprin α 1 degradation by calcium/calmodulin-dependent protein kinase II regulates LAR receptor tyrosine phosphatase distribution and dendrite development. *Developmental cell*, 12(4), 587-602.
- 27) Inashkina, I., Jankevics, E., Stavusis, J., Vasiljeva, I., Viksne, K., Micule, I., ... & Krumina, A. (2016). Robust genotyping tool for autosomal recessive type of limb-girdle muscular dystrophies. *BMC musculoskeletal disorders*, 17(1), 200.
- 28) Iyadurai, S. J. P., & Kissel, J. T. (2016). The limb-girdle muscular dystrophies and the dystrophinopathies. *CONTINUUM: Lifelong Learning in Neurology*, 22(6), 1954-1977.
- 29) Johnson, K., et al. "The MYO-SEQ project: Application of exome sequencing technologies of 1000 patients affected by limb-girdle weakness of unknown origin." *Neuromuscular Disorders* 27 (2017): S194.
- 30) Johnson, K., Bertoli, M., Phillips, L., Töpf, A., Van den Bergh, P., Vissing, J., ... & Kostera-Pruszczyk, A. (2018). Detection of variants in dystroglycanopathy-associated genes through the application of targeted whole-exome sequencing analysis to a large cohort of patients with unexplained limb-girdle muscle weakness. *Skeletal muscle*, 8(1), 23.
- 31) Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental dynamics*, 203(3), 253-310.
- 32) Klinge, L., Dekomien, G., Aboumoussa, A., Charlton, R., Epplen, J. T., Barresi, R., ... & Straub, V. (2008). Sarcoglycanopathies: can muscle immunoanalysis predict the genotype?. *Neuromuscular Disorders*, 18(12), 934-941.
- 33) Kramerova, I., Beckmann, J. S., & Spencer, M. J. (2007). Molecular and cellular basis of calpainopathy (limb girdle muscular dystrophy type 2A). *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1772(2), 128-144.
- 34) Kremer, L. S., Bader, D. M., Mertes, C., Kopajtich, R., Pichler, G., Iuso, A., ... & Koňářiková, E. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nature communications*, 8, 15824.
- 35) Lek, A., Evesson, F. J., Sutton, R. B., North, K. N., & Cooper, S. T. (2012). Ferlins: regulators of vesicle fusion for auditory neurotransmission, receptor trafficking and membrane repair. *Traffic*, 13(2), 185-194.
- 36) Lemmers, RJ, Tawil, R., Petek, LM, Balog, J., Block, GJ, Santen, GW, ... & Krom, YD (2012). La herencia digénica de una mutación SMCHD1 y un alelo D4Z4 permisivo para FSHD causa distrofia muscular facioescapulohumeral tipo 2. *Nature genetics* , 44 (12), 1370.
- 37) Liewluck, T., & Milone, M. (2018). Untangling the complexity of limb-girdle muscular dystrophies. *Muscle & nerve*, 58(2), 167-177.
- 38) Mahmood, O. A., & Jiang, X. M. (2014). Limb-girdle muscular dystrophies: where next after six decades from the first proposal. *Molecular medicine reports*, 9(5), 1515-1532.

- 39) Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A., & Jabado, N. (2011). What can exome sequencing do for you?. *Journal of medical genetics*, 48(9), 580-589.
- 40) Martinez-Thompson, J. M., Niu, Z., Tracy, J. A., Moore, S. A., Swenson, A., Wieben, E. D., & Milone, M. (2018). Autosomal dominant calpainopathy due to heterozygous *CAPN3* c. 643_663del21. *Muscle & nerve*, 57(4), 679-683.
- 41) McCarthy, T. V., Quane, K. A., & Lynch, P. J. (2000). Ryanodine receptor mutations in malignant hyperthermia and central core disease. *Human mutation*, 15(5), 410-417.
- 42) McLeish, M. J., & Kenyon, G. L. (2005). Relating structure to mechanism in creatine kinase. *Critical reviews in biochemistry and molecular biology*, 40(1), 1-20.
- 43) McNally, E. M., Ly, C. T., & Kunkel, L. M. (1998). Human ϵ -sarcoglycan is highly related to α -sarcoglycan (adhalin), the limb girdle muscular dystrophy 2D gene 1. *FEBS letters*, 422(1), 27-32.
- 44) Meienberg, J., Zerjavic, K., Keller, I., Okoniewski, M., Patrignani, A., Ludin, K., ... & Schlapbach, R. (2015). New insights into the performance of human whole-exome capture platforms. *Nucleic acids research*, 43(11), e76-e76.
- 45) Mercuri, E., & Muntoni, F. (2013). Muscular dystrophies. *The Lancet*, 381(9869), 845-860.
- 46) Michele, D. E., & Campbell, K. P. (2003). Dystrophin-glycoprotein complex: post-translational processing and dystroglycan function. *Journal of Biological Chemistry*, 278(18), 15457-15460.
- 47) Monies, D., Alhindi, H. N., Almuhaizea, M. A., Abouelhoda, M., Alazami, A. M., Goljan, E., ... & Subhani, S. (2016). A first-line diagnostic assay for limb-girdle muscular dystrophy and other myopathies. *Human genomics*, 10(1), 32.
- 48) Moore, C. J., & Hewitt, J. E. (2009). Dystroglycan glycosylation and muscular dystrophy. *Glycoconjugate journal*, 26(3), 349.
- 49) Narayanaswami, P., Weiss, M., Selen, D., David, W., Raynor, E., Carter, G., ... & Gronseth, G. (2014). Evidence-based guideline summary: diagnosis and treatment of limb-girdle and distal dystrophies: report of the guideline development subcommittee of the American Academy of Neurology and the practice issues review panel of the American Association of Neuromuscular & Electrodiagnostic Medicine. *Neurology*, 83(16), 1453-1463.
- 50) Nguyen, K., Bassez, G., Bernard, R., Krahn, M., Labelle, V., Figarella-Branger, D., ... & Eymard, B. (2005). *DYSF*erlin mutations in LGMDR2, Miyoshi myopathy, and atypical *DYSF*erlinopathies. *Human mutation*, 26(2), 165-165.
- 51) Nigro, V., & Savarese, M. (2014). Genetic basis of limb-girdle muscular dystrophies: the 2014 update. *Acta Myologica*, 33(1), 1.
- 52) Nilsson, M. I., Laureano, M. L., Saeed, M., & Tarnopolsky, M. A. (2013). *DYSF*erlin aggregation in limb-girdle muscular dystrophy type 2B/myoshi myopathy necessitates mutational screen for diagnosis. *Muscle & nerve*, 47(5), 740-747.

- 53) Norwood, F. L., Harling, C., Chinnery, P. F., Eagle, M., Bushby, K., & Straub, V. (2009). Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population. *Brain*, 132(11), 3175-3186.
- 54) Østergaard, S. T., Johnson, K., Stojkovic, T., Krag, T., De Ridder, W., De Jonghe, P., ... & Topf, A. (2018). Limb girdle muscular dystrophy due to mutations in POMT2. *J Neurol Neurosurg Psychiatry*, 89(5), 506-512.
- 55) Palladino, A., D'Ambrosio, P., Papa, A. A., Petillo, R., Orsini, C., Scutifero, M., ... & Politano, L. (2016). Management of cardiac involvement in muscular dystrophies: paediatric versus adult forms. *Acta Myologica*, 35(3), 128.
- 56) Peddareddygari, L. R., Oberoi, K., & Grewal, R. P. (2018). Limb Girdle Muscular Dystrophy due to Digenic Inheritance of DES and CAPN3 Mutations. *Case reports in neurology*, 10(3), 272-278.
- 57) Pegoraro, E., & Hoffman, E. P. (2012). Limb-girdle muscular dystrophy overview. In *GeneReviews*@[Internet]. University of Washington, Seattle.
- 58) Pena, Loren DM, et al. "Looking beyond the exome: a phenotype-first approach to molecular diagnostic resolution in rare and undiagnosed diseases." *Genetics in Medicine* 20.4 (2018): 464.
- 59) Roland, E. H. (2000). Muscular dystrophy. *Pediatrics in review*, 21(7), 233.
- 60) Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A., & Ekker, S. C. (2007). p53 activation by knockdown technologies. *PLoS genetics*, 3(5), e78.
- 61) Schäffer, A. A. (2013). Digenic inheritance in medical genetics. *Journal of medical genetics*, 50(10), 641-652.
- 62) Schwartz, C., Tubez, F., Wang, F. C., Croisier, J. L., Brûls, O., Denoël, V., & Forthomme, B. (2017). Normalizing shoulder EMG: An optimal set of maximum isometric voluntary contraction tests considering reproducibility. *Journal of Electromyography and Kinesiology*, 37, 1-8.
- 63) Semplicini, C., Vissing, J., Dahlqvist, J. R., Stojkovic, T., Bello, L., Witting, N., ... & Eymard, B. (2015). Clinical and genetic spectrum in limb-girdle muscular dystrophy type 2E. *Neurology*, 84(17), 1772-1781.
- 64) Sewry, C. A., & Wallgren-Pettersson, C. (2017). Myopathology in congenital myopathies. *Neuropathology and applied neurobiology*, 43(1), 5-23.
- 65) Straub, Volker, and Marta Bertoli. "Where do we stand in trial readiness for autosomal recessive limb girdle muscular dystrophies?." *Neuromuscular Disorders* 26.2 (2016): 111-125.
- 66) Straub, V., Murphy, A., & Udd, B. (2018). 229th ENMC international workshop: Limb girdle muscular dystrophies – Nomenclature and reformed classification Naarden, the Netherlands, 17–19 March 2017. *Neuromuscular Disorders*.

- 67) Sung, A. R., Moretti, P., & Shaibani, A. (2018). Case of late-onset Sandhoff disease due to a novel mutation in the HEXB gene. *Neurology Genetics*, 4(4), e260.
- 68) Taghizadeh, E., Rezaee, M., Barreto, G. E., & Sahebkar, A. (2018). Prevalence, pathological mechanisms, and genetic basis of limb-girdle muscular dystrophies: A review. *Journal of Cellular Physiology*.
- 69) Tarakci, H., & Berger, J. (2016). The sarcoglycan complex in skeletal muscle. *Front Biosci (Landmark Ed)*, 21, 744-56.
- 70) Thompson, R., & Straub, V. (2016). Limb-girdle muscular dystrophies—international collaborations for translational research. *Nature Reviews Neurology*, 12(5), 294.
- 71) Trabelsi, M., Kaviani, N., Daoud, F., Commere, V., Deburgrave, N., Beugnet, C., ... & Leturcq, F. (2008). Revised spectrum of mutations in sarcoglycanopathies. *European journal of human genetics*, 16(7), 793.
- 72) Vissing, J. (2016). Limb girdle muscular dystrophies: classification, clinical spectrum and emerging therapies. *Current opinion in neurology*, 29(5), 635-641.
- 73) Wallgren-Pettersson, C., Pelin, K., Hilpelä, P., Donner, K., Porfirio, B., Graziano, C., ... & Sewry, C. (1999). Clinical and genetic heterogeneity in autosomal recessive nemaline myopathy. *Neuromuscular Disorders*, 9(8), 564-572.
- 74) Walton, J. N., & Nattrass, F. J. (1954). On the classification, natural history and treatment of the myopathies. *Brain*, 77, 169-231.
- 75) Wattjes, M. P., Kley, R. A., & Fischer, D. (2010). Neuromuscular imaging in inherited muscle diseases. *European radiology*, 20(10), 2447-2460.
- 76) Wheeler, M. T., Zarnegar, S., & McNally, E. M. (2002). ζ -sarcoglycan, a novel component of the sarcoglycan complex, is reduced in muscular dystrophy. *Human molecular genetics*, 11(18), 2147-2154.
- 77) Wicklund, M. P. (2013). The muscular dystrophies. *CONTINUUM: Lifelong Learning in Neurology*, 19(6, Muscle Disease), 1535-1570.
- 78) Yoshida, Mikiharu, et al. "Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl β -D-glucoside." *European journal of biochemistry* 222.3 (1994): 1055-1061.
- 79) Zhao, M., Maani, N., & Dowling, J. J. (2018). Dynamin 2 (DNM2) as cause of, and modifier for, human neuromuscular disease. *Neurotherapeutics*, 15(4), 966-975.
- 80) Zundel, W. S., & Tyler, F. H. (1965). The Muscular Dystrophies. *New England Journal of Medicine*, 273(10), 537–543.
- 81) <http://treatnmd.ncl.ac.uk/about/related-projects/MYO-SEQ/>