

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN  
FACULTAD DE MEDICINA**



**IDENTIFICACIÓN DE VARIANTES GENÉTICAS PATOGENICAS EN  
PACIENTES CON SÍNDROMES RAROS MEDIANTE SECUENCIACIÓN MASIVA**

**PRESENTA**

**EDWIN ANTONIO REYES OLIVA**

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE  
MAESTRÍA EN CIENCIAS CON ORIENTACIÓN EN BIOLOGÍA  
MOLECULAR E INGENIERÍA GENÉTICA**

**AGOSTO, 2019**

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN  
FACULTAD DE MEDICINA**

**DEPARTAMENTO DE BIOQUÍMICA Y MEDICINA MOLECULAR**



**IDENTIFICACIÓN DE VARIANTES GENÉTICAS PATOGENICAS EN  
PACIENTES CON SÍNDROMES RAROS MEDIANTE SECUENCIACIÓN MASIVA**

**PRESENTA**

**EDWIN ANTONIO REYES OLIVA**

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE  
MAESTRÍA EN CIENCIAS CON ORIENTACIÓN EN BIOLOGÍA  
MOLECULAR E INGENIERÍA GENÉTICA**

**ASESOR  
DR. C. CARLOS CÓRDOVA FLETES**

**MONTERREY, NUEVO LEÓN, MÉXICO**

**AGOSTO, 2019**

**El presente trabajo se realizó bajo la dirección del**

**Dr. C. Carlos Córdova Fletes en el:**

**Laboratorio de Citogenómica y Microarreglos del Departamento de  
Bioquímica y Medicina Molecular de la Facultad de Medicina de la  
Universidad Autónoma de Nuevo León.**

**IDENTIFICACIÓN DE VARIANTES GENÉTICAS PATOGENICAS EN  
PACIENTES CON SÍNDROMES RAROS MEDIANTE SECUENCIACIÓN MASIVA**

**Aprobación de la tesis:**



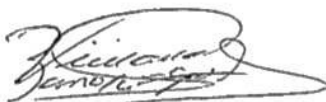
---

**DR. C. CARLOS CÓRDOVA FLETES**  
Director de Tesis



---

**DR. JOSÉ ELÍAS GARCÍA ORTIZ**  
Co-Director de Tesis



---

**DRA. C. VIVIANA CHANTAL ZOMOSA SIGNORET**  
Miembro de la Comisión de Tesis



---

**DR. med. FELIPE ARTURO MORALES MARTÍNEZ**  
Subdirector de Estudios de Posgrado

## **COMISIÓN DE TESIS**

### **DIRECTOR**

Dr. C. Carlos Córdova Fletes  
Laboratorio de Citogenómica y Microarreglos

### **CO-DIRECTOR**

Dr. José Elías García Ortiz  
Centro de Investigación Biomédica de Occidente del IMSS

### **MIEMBRO DE LA COMISIÓN**

Dra. C. Viviana Chantal Zomosa Signoret  
Biología estructural

### **COLABORACIONES**

Dra. C. Ana María Guadalupe Rivas Estilla  
Dr. C. Kame Alberto Galán Huerta  
M. C. Elda Ariadna Flores Contreras  
M.C.P. Miguel Zambrano Lucio.

## **FINANCIAMIENTO**

Convocatoria Apoyo Infraestructura Científica 2016, proyecto No. 268123.  
Responsable técnico: Dr. Carlos Córdova Fletes

*Nature is nowhere accustomed more openly to display her secret mysteries than in cases where she shows tracings of her workings apart from the beaten paths; nor is there any better way to advance the proper practice of medicine than to give our minds to the discovery of the usual law of nature, by careful investigation of cases of rarer forms of disease.*

**William Harvey, 1657**

## **DEDICATORIA**

**A mis padres Lourdes Oliva y Salvador Reyes  
por su infinito amor. Gracias.**

## AGRADECIMIENTOS

A mis padres, Lourdes Oliva y Salvador Reyes por siempre amarme tanto, por confiar en mi y por darme todo lo que está en sus manos para verme feliz, los amo. A mi hermano Salvador, por siempre apoyarme, guiarme y ser un ejemplo de bondad. A mi hermana Yoana, por siempre transmitirme su fortaleza, perseverancia y esas ansias de querer más. A mis increíbles sobrinos: Milagros, Ixchel y Valentino a quienes extrañé en todo momento y siempre me recibieron con una sonrisa.

A Brenda, que confió en mí, me dejó vivir en su casa, comer sus galletas y me guió a creer que yo podía hacer más, que merecía más y que la vida no espera.

A Catalán, por estar siempre cuando las cosas no salían bien, por escucharme, tranquilizarme y darme la mano en forma de buenos consejos. Gracias chiqui, te quiero.

A Alejo, por llevarme a reflexionar y elegir un futuro diferente, por enseñarme que los momentos “nunca van a ser perfectos, así que hazlo” por que no importa si creen en ti, lo haces porque puedes, y puedes porque en realidad quieres. Te extraño mucho amigo.

A Gris, quien siempre me recibió en su casa e hizo mi estancia menos complicada, por quererme tanto y por estar siempre para mí, por recibirme siempre con el corazón en la mano. Mi más mejor.

A los amigos de Aidé: Florencia, Juan, Alejandro, Martha y S. Extraños que conocí un día sin querer, y con el tiempo comenzamos a compartir alimento, bebida, compras, celebraciones, momentos divertidos, viajes, bromas, secretos y aventuras. Personas que admiro y respeto, en quien puedo confiar y a quien puedo llamar desde hace tiempo mis amigos. Los quiero a todos.

A mis amigos y compañeros de generación Gabs, Javi y M, con quien compartí momentos buenos, malos y otros raros. Sin ustedes no me hubiera divertido tanto.

A María Fernanda, a quien aprecio por quedarse siempre conmigo, por ayudarme y procurarme en todo momento, por su corazón bondadoso.

A Elda, mi amiga mediterránea a quien siempre recordaré con una bella sonrisa y con quien estaré siempre agradecido por estar en todos mis momentos: buenos y felices, difíciles y frustrantes, raros, bizarros y desesperados, por su dedicación para transmitirme sus conocimientos, por su complicidad, por todas las charlas que parecían no tener fin y por siempre confiar en mí. Te llevo en el corazón chica.

A Mike Wazowski, mi amigo médico con quien compartí laboratorio y a quien admiro por su dedicación y entrega.



Agradezco al Dr. José Elías por las valiosas contribuciones y colaboración realizada, por sus atenciones y el tiempo brindado en el proyecto.

A las Dra. C. Viviana Chantal y Ana María, por estar al tanto de mi desarrollo como profesional y por las contribuciones realizadas en el proyecto.

Al Dr. C. Kame, por sus atenciones y valiosa colaboración en el proyecto.

Agradezco al Dr. en C. Carlos Córdova por brindarme la oportunidad de trabajar a su lado, por su dedicación para transmitirme sus conocimientos y guiarme dentro de mi formación profesional.

Por ultimo, agradezco a la Universidad Autónoma de Nuevo León por haberme permitido formarme en ella, por ser mi casa durante mi estancia y brindarme un mundo de posibilidades.

## ÍNDICE

### CAPÍTULO I

<b>INTRODUCCIÓN</b> .....	<b>1</b>
Definición de enfermedades raras .....	1
Epidemiología.....	2
Etiología y clasificación de las ER .....	4
Diagnóstico genético .....	5
Métodos de diagnóstico convencionales.....	7
Exploración física detallada o fenotipificación.....	7
Historia clínica o antecedentes familiares detallados.....	8
Pruebas clínicas y de laboratorio. ....	8
Pruebas moleculares.....	13
Debilidades y fortalezas de WES.....	14

### CAPÍTULO II

<b>ANTECEDENTES</b> .....	<b>17</b>
Antecedentes generales .....	17
Antecedentes directos.....	20

### CAPÍTULO III

<b>HIPÓTESIS</b> .....	<b>22</b>
<b>JUSTIFICACIÓN</b> .....	<b>22</b>

### CAPÍTULO IV

<b>OBJETIVOS</b> .....	<b>23</b>
Objetivo general.....	23
Objetivos específicos .....	23

### CAPÍTULO V

<b>DISEÑO EXPERIMENTAL</b> .....	<b>24</b>
Diseño del experimento.....	24
Estándares éticos.....	24
Universo de estudio.....	24
Reclutamiento de pacientes .....	25
Evaluación clínica .....	25
Muestra de estudio .....	25

<b>Criterios de selección</b> .....	<b>26</b>
Criterios de inclusión .....	26
Criterios de no inclusión .....	26
Criterios de eliminación .....	26

## **CAPÍTULO VI**

<b>MATERIALES</b> .....	<b>27</b>
<b>Lugar de realización</b> .....	<b>27</b>
<b>Reactivos</b> .....	<b>27</b>
Soluciones de uso general .....	27
<b>WES</b> .....	<b>27</b>
Verificación de la cantidad y calidad de muestra de partida .....	27
Preparación de la biblioteca .....	27
Comprobación de biblioteca:.....	28
Desnaturalización y dilución de la biblioteca .....	28
Generación de clusters y secuenciación .....	29
<b>Secuenciación Sanger</b> .....	<b>29</b>
Estandarización de cebadores y obtención de amplicones mediante reacción en cadena de la polimerasa (PCR) punto final.....	29
Purificación de amplicones .....	29
Secuenciación Sanger .....	29
<b>Materiales</b> .....	<b>30</b>
Material de uso general .....	30
<b>Equipos</b> .....	<b>30</b>
Equipo de uso general .....	30
Equipos para secuenciación masiva.....	31
<b>Secuenciación Sanger</b> .....	<b>31</b>
<b>Programas computacionales</b> .....	<b>32</b>

## **CAPÍTULO VII**

<b>MÉTODOS</b> .....	<b>33</b>
<b>Naturaleza de las muestras</b> .....	<b>34</b>
<b>WES</b> .....	<b>34</b>
Comprobación de la calidad, cantidad e integridad de las muestras .....	34
Preparación de la biblioteca .....	35
Generación de grupos y secuenciación.....	38
Análisis de secuencia .....	39

## CAPÍTULO VIII

<b>RESULTADOS</b> .....	<b>41</b>
<b>Descripción de casos</b> .....	<b>41</b>
Caso 1 - Probando LCM121.....	41
Caso 2- Probando LCM124.....	43
Caso 3 - Probando LCM129.....	44
Caso 4 - Probando LCM119.....	45
Caso 5 - Probando LCM13.....	45
Caso 6 - Paciente LCM108.....	46
<b>Objetivo 1: Secuenciación de exoma de pacientes con diagnóstico de ER.</b> .....	<b>48</b>
Comprobación de cantidad, calidad e integridad de las muestras de partida.....	48
Generación de bibliotecas .....	50
Secuenciación y análisis de Secuencia .....	52
Resumen de enriquecimiento.....	53
Variantes por contexto de secuencia.....	56
Variantes por consecuencia .....	57
<b>Objetivo 2: Análisis integral de los datos e identificación de variantes que puedan explicar el fenotipo clínico.</b> .....	<b>61</b>
Caso 1.....	61
Caso 2 y 3 .....	63
Caso 4.....	66
Caso 5.....	67
Caso 6.....	68
<b>Objetivo 3. Validación de variantes encontradas potencialmente dañinas mediante secuenciación Sanger</b> .....	<b>70</b>
Diseño y estandarización de cebadores.....	70
Obtención y extracción de amplicones.....	73

## CAPÍTULO IX

<b>DISCUSIÓN</b> .....	<b>78</b>
<b>CASO 1</b> .....	<b>78</b>
<b>CASO 2 Y 3</b> .....	<b>82</b>
<b>CASO 4</b> .....	<b>86</b>
<b>CASO 5</b> .....	<b>88</b>
<b>CASO 6</b> .....	<b>90</b>

## CAPÍTULO X

<b>CONCLUSIONES</b> .....	<b>94</b>
---------------------------	-----------

<b>CAPÍTULO XI</b>	
<i>PERSPECTIVAS</i> .....	<b>96</b>
<b>CAPÍTULO XII</b>	
<i>BIBLIOGRAFÍA</i> .....	<b>97</b>
<b>CAPÍTULO XIII</b>	
<i>ANEXOS</i> .....	<b>107</b>

## ABREVIATURAS, FÓRMULAS Y SÍMBOLOS

AC	Anomalías congénitas	mar	Marcador cromosómico
aCGH	Microarreglo de hibridación genómica comparativa	Mb	Megabases
AGT	Alanina glicoxilato aminotransferasa	mg	Miligramo
CGH	Hibridación genómica comparativa	MINGE	Encefalomiopatía neurogastrointestinal mitocondrial
chr	Cromosoma	mL	Mililitro
cm	Centímetro	mM	Milimolar
CNV	Variación en el número de copias	mos	Mosaico
ddNTP	Didesoxinucleótidos trifosfato	N	Normalidad
dL	Decilitro	ng	Nanogramos
DNA	Ácido desoxirribonucleico	NGS	Secuenciación de nueva generación
DNAg	DNA genómico	nm	Nanómetro
dNTP	Desoxirribonucleótidos trifosfato	nM	Nanomolar
dsDNA	DNA de doble cadena	PCR	Reacción en cadena de la polimerasa
EDTA	Ácido etilendiaminotetra acético	PH1	Hiperoxaluria primaria tipo 1
ER	Enfermedades raras	pM	Picomolar
ERG	Enfermedades raras genéticas	RNA	Ácido ribonucleico
ExAC	Consortio de Agregación de Exome	SDS	Dodecil sulfato de sodio
FATCO	Aplasia fibular, campomelia de tibia y oligosindactilia	SMF	Síndrome de Marfan
FISH	Hibridación fluorescente in situ	SNP	Polimorfismos de nucleótido único
g	Gramo	SNV	Variantes de nucleótido único
Gbp	Pares de giga bases	TAC	Tomografía axial computarizada
GWAS	Estudio de asociación de genoma completo	TEA	Trastorno del espectro autista
hg38	Genoma de referencia 38	TGO	Transaminasa Glutámico Oxalacética
HTA	Hipertensión arterial	UPD	Disomía uniparental
IGV	Integrative Genome Viewer	VUS	Variante de significado incierto
INDELS	Contracción de "inserción o delección"	WES	Secuenciación de exoma completo
Kb	Kilobases	WGS	Secuenciación de genoma completo
Kg	kilogramo	µL	Microlitro
M	Masa molar		

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Estrategia general de trabajo.....	33
<b>Figura 2.</b> Signos pivote.....	47
<b>Figura 3.</b> Alineamientos de lecturas gen <i>AGXT</i> .....	62
<b>Figura 4.</b> Visualización de alineamientos gen <i>CNTN6</i> .....	68
<b>Figura 5.</b> Validación de variantes gen <i>PLXNB1</i> y <i>FGFR4</i> .....	75
<b>Figura 6.</b> Validación de variantes encontradas en gen <i>TBX6</i> , <i>BMP15</i> y <i>WNT3A</i> .....	75
<b>Figura 7.</b> Validación de variante encontrada en el gen <i>CNTN6</i> .....	76

## ÍNDICE DE IMÁGENES

<b>Imagen 1.</b> Calidad e integridad del DNAg.....	49
<b>Imagen 2.</b> Bibliotecas.....	50
<b>Imagen 3.</b> Estandarización de cebadores.....	72
<b>Imagen 4.</b> Obtención de amplicones.....	74

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Materiales de uso general.....	30
<b>Tabla 2.</b> Programas computacionales y plataformas bioinformáticas.....	32
<b>Tabla 3.</b> Cuantificación de muestras por NanoDrop.....	48
<b>Tabla 4.</b> Comprobación de bibliotecas enriquecidas.....	51
<b>Tabla 5.</b> Versiones de software.....	52
<b>Tabla 6.</b> Información de muestra.....	53
<b>Tabla 7.</b> Enriquecimiento a nivel de lecturas.....	54
<b>Tabla 8.</b> Enriquecimiento a nivel de base.....	55
<b>Tabla 9.</b> Enriquecimiento a nivel de objetivo.....	55
<b>Tabla 10.</b> Resumen de variantes.....	56
<b>Tabla 11.</b> Variantes por contexto de secuencia.....	58
<b>Tabla 12.</b> Variante por consecuencia.....	59
<b>Tabla 13.</b> Resumen de cobertura.....	60
<b>Tabla 14.</b> Aplicación de filtros.....	61
<b>Tabla 15.</b> Identificación de genes y variantes, caso 2 y 3.....	64
<b>Tabla 16.</b> Frecuencia poblacional y predicción de la patogenicidad in silico, caso 2 y 3...	65
<b>Tabla 17.</b> Identificación de genes y variantes, caso 4.....	66
<b>Tabla 18.</b> Frecuencia poblacional y predicción de la patogenicidad in silico, caso 4.....	67
<b>Tabla 19.</b> Genes involucrados en la biosíntesis organización y degradación de tejido conectivo.....	67
<b>Tabla 20.</b> Diseño de cebadores.....	70
<b>Tabla 21.</b> Mezcla de reacción para el ensayo de amplificación.....	71
<b>Tabla 22.</b> Condiciones experimentales de PCR cualitativa.....	73
<b>Tabla 23.</b> Identificación de genes caso 6.....	77



## RESUMEN

### IDENTIFICACIÓN DE VARIANTES GENÉTICAS PATOGENICAS EN PACIENTES CON SÍNDROMES RAROS MEDIANTE SECUENCIACIÓN MASIVA

**Introducción:** Las enfermedades raras (ER) son un diverso grupo de entidades clínicas que tienen una baja prevalencia en la población mundial, afectando a un número limitado de personas. Sin embargo, acumulativamente tienen un gran impacto en la salud pública. Se estima que existen cerca de 7.000 ER que afectan al 7% de la población mundial. Sin embargo, para muchas de estas ER la causa exacta sigue siendo desconocida. En contraste, para una parte importante de ellas, el problema puede ser rastreado a mutaciones en la porción codificante del genoma. Ante este panorama la secuenciación de nueva generación aplicada al exoma completo representa una oportunidad importante para la identificación y diagnóstico de variantes que puedan responder por el fenotipo clínico de los pacientes.

**Objetivo:** Identificar y caracterizar las alteraciones genómicas en pacientes con diagnóstico de ER mediante WES para establecer una relación genotipo-fenotipo integral.

**Materiales y métodos:** Se prepararon 6 bibliotecas de exoma a partir de muestras de DNAg de pacientes diagnosticados con alguna ER utilizando el kit TruSeq Rapid Exome Library Prep. Se realizó secuenciación paired-end de 75 ciclos en el equipo MiSeq (Illumina). Los datos de secuenciación se alinearon con el genoma humano de referencia (GRCh37/hg19) utilizando el software BWA Enrichment v.2.1.0.0, y se generó la llamada de variantes usando GATK v1.6. Los datos se importaron al software Variant Interpreter (VI, beta) v1.0.15,015 para su análisis.

**Resultados:** En el caso 1, se identificó en el probando una mutación puntual (c.33dupC) en el gen **AGXT** presente de manera heterocigota que produce el desplazamiento del marco de lectura, generando un codón de terminación prematura en la posición 156 (p.Lys12GlnfsTer156) asociado directamente a la enfermedad de hiperoxaluria primaria tipo 1. Para el caso 2 y 3, los probandos presentaban signos y síntomas consistentes con síndrome de FATCO. En el caso 2 se identificó un conjunto de variantes ubicadas en genes asociados directamente al desarrollo de miembros óseos pertenecientes a las familias FGF (**FGFR4**, c.1287C>A) y genes Hox (**HOXC9**, c.596A>G), (**HOXD12**, c.236G>A) y (**HOXD3**, c.257T>C). Para el caso 3, se identificaron variantes ubicadas en miembros de la superfamilia del factor de crecimiento transformante beta (TGF- $\beta$ ) (**GDF5**, c.536C>T) y (**BMP15**, c.433C>T), genes Hox (**HOXC12**, c.139C>T), factores de transcripción de la familia T-box (**TBX6**, c.484G>A) y morfógenos pertenecientes a la familia Wnt (**WNT3A**, c.860C>G). Estas variantes fueron predichas *in silico* como probablemente dañinas o deletéreas por diversos programas bioinformáticos, además presentan una frecuencia poblacional <0.001, aunque no todas podrían estar relacionadas con el fenotipo. En caso 4 se identificó una variante de cambio de sentido ubicada en el gen **NDUFS7** (c.434G>A, p.Arg145His) en un paciente con signos y síntomas consistentes con el síndrome MNGIE que podría ocasionar la enfermedad. En el caso 5 se identificó un grupo de variantes ubicadas en los genes **MAML3** y **NOTCH4** involucrados en la vía de señalización canónica de TGF- $\beta$ , además, se identificaron variantes patogénicas en un segundo grupo de genes, involucrados en la biosíntesis, organización y degradación de tejido conectivo. Todas las variantes fueron predichas por diferentes plataformas bioinformáticas como patogénicas con una frecuencia poblacional <0.001. En el probando 6, se identificó una variante de pérdida de función en el gen **CNTN6** (c.566C>A, p.Ser189Ter) relacionada con un fenotipo del espectro autista. La variante implica la ganancia de un codón de paro. Debido a que no poseía registro de identificación la CNV fue registrada en la base de datos ClinVar con el número de entrada SCV000786639.1.



---

DR. C. CARLOS CÓRDOVA FLETES

Director de Tesis

# CAPÍTULO I

## INTRODUCCIÓN

### **Definición de enfermedades raras**

Las enfermedades raras (ER) son un grupo diverso de condiciones, generalmente definidas por el pequeño número de personas que se ven afectadas por ellas en comparación con afecciones más comunes. El concepto de ER se acuñó por primera vez en 1983, con la Ley de Medicamentos Huérfanos de Estados Unidos de Norteamérica, cuyo objetivo era estimular la investigación y comercialización de medicamentos destinados a la prevención y tratamiento de enfermedades con baja prevalencia.<sup>1</sup>

Actualmente no existe un consenso a nivel mundial que defina el concepto de ER, y cada país establece el umbral de prevalencia que designa para confirmar una enfermedad como rara. Así por ejemplo la Ley de Medicamentos Huérfanos de Estados Unidos de Norteamérica del 2002 define a las ER como aquella que afecta a menos de 200,000 habitantes, es decir la definición se refiere a un número absoluto.<sup>2</sup> Mientras que la normativa europea utiliza una medida relativa, es decir menos de 5 casos por cada 10,000 habitantes, además requiere que los trastornos atenten contra la vida y/o sea crónicamente debilitante o inadecuadamente tratada o bien, la existencia o no de tratamientos específicos para dicha condición.<sup>3</sup>

En México la definición utilizada es similar a la propuesta por la Agencia Europea de medicamentos, que definen a las ER, como aquellas cuya frecuencia es de 5

casos por cada 10,000 habitantes, establecida a través del Art. 224, 224 BIS y 224 BIS I de la Ley General de Salud, en donde establecen las bases y las modalidades para impulsar y fomentar la disponibilidad de los medicamentos huérfanos.<sup>4</sup>

Diferentes portales y organizaciones reportan que existen cerca de 6,000 – 7,000 ER reconocidas y cada una plantea problemas específicos en relación con su rareza. Además, se señala que el concepto de ER puede referirse a diferentes entidades patológicas tales como enfermedades, síndromes y trastornos. Es decir la definición incluye una amplia diversidad de desórdenes y síntomas así como diferentes pacientes.<sup>5,6</sup>

Las ER tienden a ser graves, crónicas y, en muchos casos, son progresivas, debilitantes y ponen en peligro la vida de quienes las padecen. Frecuentemente son llamadas "enfermedades huérfanas", debido a la desventaja que presentan en estudios de investigación clínica y experimental vs enfermedades comunes, estando así "huérfanas" del interés del mercado y de las políticas de salud pública, se estima que solo el 5 por ciento de las ER tienen un tratamiento aprobado por la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA).

### **Epidemiología**

Actualmente, existen muchas dificultades que hacen complicado definir y contar las ER, tales como obtener diagnósticos definitivos y certeros, pocos sistemas que concentren e informen los diagnósticos, diferencias en los parámetros que utilizan en cada sistema, códigos de clasificación inespecíficos de la enfermedad, así como

un deficiente proceso para rastrear los diagnósticos, sumado a eso el hecho de que en cada país se han adoptado diferentes definiciones de ER, así como la continua identificación de nuevas enfermedades o variantes de la enfermedad. Por lo tanto, la epidemiología de las ER, incluida la determinación de la prevalencia (número de personas afectadas al mismo tiempo), la incidencia (número de casos nuevos en un año determinado) y los patrones de enfermedad (por ejemplo, distribución por edad) en la población, es inexacta.

Sin embargo, cuando se habla de epidemiología de las ER, la prevalencia en lugar de la incidencia es el parámetro que sirve como medida de impacto de las ER. A pesar de que es imposible obtener datos de prevalencia precisos y actualizados de cada ER y que no existe una lista mundial autorizada de ER, se estima que el 75% de las 6,000-7,000 condiciones raras reconocidas tienen una prevalencia de 0.1 - 10 personas por cada 100,000 habitantes<sup>7,8</sup>, afectando aproximadamente del 6 al 8% de la población mundial.<sup>9</sup>

Dentro del aspecto de la prevalencia, se debe aclarar que el concepto de ER para un padecimiento específico puede variar entre los países. Si bien para un país tener pocos casos representa una ER, para otro país su presencia puede ser muy común y entonces no es considerada como rara. Lo anterior puede ser reflejo de que un buen porcentaje de estas enfermedades tiene un origen genético, haciendo posible que en ciertas regiones del mundo exista un alto grado de endogamia y, por lo tanto, sea muy frecuente su presencia. Mientras que, en otros, su aparición es muy esporádica y entonces sí se apega al concepto de ER. Además, en otros casos

algunas condiciones inicialmente pueden ser clasificadas como raras, pero eventualmente pueden superar esa categorización. Por ejemplo, cuando el SIDA surgió en Estados Unidos de Norteamérica, se ajustó a la definición legislativa de una ER, que afecta a menos de 200,000 personas. A medida que se propagó la infección, hubo una mejora en los métodos de diagnóstico y los sistemas de recopilación de datos, logrando desarrollar tratamientos efectivos que redujeron la mortalidad sin curar la enfermedad; sin embargo, el número total de personas con SIDA aumentó a casi 470,000 en 2007 y la cantidad de personas con infección por VIH superó 1.1 millones dejando de ser catalogada como una ER según la normativa vigente en ese país.<sup>7</sup>

### **Etiología y clasificación de las ER**

En las últimas dos décadas, la investigación epidemiológica, molecular y de otro tipo han aprovechado los avances científicos y tecnológicos en las ciencias biológicas a favor de la identificación de las causas que dan origen a las ER. Sin embargo, para muchas ER, la causa exacta sigue siendo desconocida. Aun así, para una parte importante, el problema se puede rastrear a cambios o mutaciones en un sólo gen conocidos también como enfermedades monogénicas (o mendelianas).<sup>10</sup> Acumulativamente, estas enfermedades son conocidas también como enfermedades genéticas raras (ERG)<sup>6</sup>. Muchas de estas mutaciones genéticas a menudo ocurren como enfermedades hereditarias con un patrón de herencia mendeliano simple, lo que explica por qué ciertas ER se presentan en familias.

Sin embargo, hay que tener en cuenta que las causas genéticas son sólo una pieza del complejo rompecabezas que engloba al conjunto de ER. Diferentes fuentes de información reportan que existen ER originadas a partir de causas no genéticas tales como agentes infecciosos, agentes tóxicos, causas inmunológicas, causas degenerativas o proliferativas, entre otras causas, que pueden desempeñar un papel en el desarrollo de una ER. Dichos factores pueden causar enfermedad directamente o interactuar con factores genéticos para causar o aumentar la gravedad de la enfermedad. Algunos autores señalan que el conocer la causa de una enfermedad no significa necesariamente que se entienda el mecanismo de la enfermedad, además muchas de estas condiciones raras tienen múltiples tipos de causas posibles. Por otro lado, están las enfermedades que tienen causas conocidas y mecanismos razonablemente bien entendidos, pero carecen de curas, tratamientos satisfactorios o estrategias preventivas. No obstante, identificar la causa de una afección suele ser un paso importante en la creación de una base de conocimientos para la prevención o el tratamiento efectivo. <sup>11</sup>

### **Diagnóstico genético**

En un contexto genético clínico general, la tasa de diagnóstico actual es aproximadamente del 50%, pero para aquellos que no reciben un diagnóstico molecular después de la evaluación genética inicial, esa tasa es mucho más baja. El éxito diagnóstico de estos individuos afectados más desafiantes depende en gran medida del progreso en el descubrimiento de genes asociados y de los mecanismos subyacentes de las enfermedades raras. <sup>12</sup>

Una encuesta de la Organización Europea de Enfermedades Raras realizada en el año 2012 dio a conocer la odisea diagnóstica que enfrentan los pacientes con alguna ER para llegar a esclarecer su situación. La encuesta informó que de ocho ER que representan a 12,000 pacientes de 17 países encontró que el 25% de los pacientes tuvo que esperar entre 5 y 30 años desde los primeros síntomas de la enfermedad hasta el diagnóstico confirmatorio, el 40% recibió un diagnóstico erróneo inicial (lo que provocó intervenciones médicas incorrectas incluida la cirugía y los tratamientos psiquiátricos) y 25 % tuvo que viajar a una región diferente para obtener un diagnóstico.<sup>13</sup>

De manera similar, una encuesta realizada por la Organización Canadiense de Desórdenes Raros encontró que aproximadamente 20% de los pacientes esperaron entre 6 y 14 años para obtener un diagnóstico, el 60% consultaron 3 o más de 20 especialistas en el camino al diagnóstico. Sin duda estas encuestas reflejan lo complicado que es establecer el diagnóstico etiológico de una ER.

Algunos autores señalan que establecer un diagnóstico etiológico de una ER a partir de la identificación de la base molecular causal, es notablemente difícil<sup>14</sup> por varias razones: una de estas razones es la heterogeneidad genética que existen entre las ER, diferentes bases de datos enlistan cerca de 7,000 trastornos genéticos pero para la mayoría de ellos aún no ha sido posible establecerse una base molecular.<sup>10</sup> También se habla de la heterogeneidad clínica, las presentaciones genéticas de la enfermedad en lactantes son a menudo formas atenuadas de descripciones clásicas en niños mayores.<sup>15</sup> Por otra parte la comorbilidad frecuente en los

lactantes con enfermedades genéticas, como la prematuridad, encefalopatías traumáticas en el parto y sepsis, pueden llegar a confundir las presentaciones clínicas.<sup>16</sup> A estas razones, se les suma el constante solapamiento entre los síntomas derivando en la promulgación de más de un diagnóstico, estimando que el 4% de los niños tiene más de un diagnóstico genético.<sup>16</sup> Por último, la progresión de la enfermedad, se ha estimado que esta es más rápida en los niños, convirtiendo la odisea diagnóstica en una carrera contra el tiempo.<sup>17,18</sup>

### **Métodos de diagnóstico convencionales**

Tradicionalmente, el establecimiento de diagnósticos moleculares se realiza mediante pruebas seriales guiadas por diagnóstico diferencial<sup>14</sup>, este implica un examen clínico integral que consta de tres elementos principales:

#### **Exploración física detallada o fenotipificación.**

Corresponde a una exploración física completa del individuo, realizada por parte de un médico genetista tratante con el objetivo de evaluar de forma general órganos y sistemas comprometidos obteniendo información sobre el estado de salud. Este es un paso vital en el proceso de identificación del gen causante de la enfermedad, debido a que la información obtenida es crítica para identificar la heterogeneidad fenotípica y del locus no reconocida, así como para aumentar la sospecha de expansión fenotípica y fenotipos combinados.<sup>19</sup>



### **Historia clínica o antecedentes familiares detallados.**

La función esencial de este elemento es recopilar detalles relacionados con estilo de vida, detección de factores de riesgo, así como antecedentes familiares y/o detección de enfermedades hereditarias dentro de la familia.

En conjunto, estos dos elementos nos llevan a la sospecha diagnóstica. Además, sirven como guía para ordenar pruebas clínicas y de laboratorio necesarias.

### **Pruebas clínicas y de laboratorio.**

En cuanto a las pruebas clínicas y de laboratorio, actualmente se usan diferentes métodos para detectar algún tipo de anomalía, el tipo de prueba depende del tipo de anomalía que se esté evaluando. Estas pueden ser agrupadas en tres tipos: pruebas bioquímicas, citogenéticas y moleculares.

**Pruebas bioquímicas-** Este conjunto de pruebas tienen como objetivo detectar una enfermedad bioquímica usando técnicas que se centran en analizar los productos génicos dejando de lado los genes. Las pruebas pueden usarse para medir directamente la actividad de una proteína (enzima), el nivel de metabolitos (medición indirecta de la actividad de una proteína) y el tamaño o la cantidad de proteínas (proteínas estructurales).

**Pruebas citogenéticas-** La citogenética implica la evaluación de todos los cromosomas para detectar anomalías. Los cromosomas de las células humanas en

división pueden analizarse bajo un microscopio utilizando glóbulos blancos, particularmente linfocitos T, debido a que son células altamente disponibles y accesibles para análisis citogenéticos ya que pueden obtenerse fácilmente del tejido sanguíneo y se dividen rápidamente en un cultivo celular, los cuales son sometidos a un procedimiento de bandeo. El bandeo G (o bandeo GTG) es un procedimiento mediante el cual los cromosomas son expuestos a la acción enzimática controlada. Al ser posteriormente teñidos, presentan un patrón de bandas oscuras y claras características que permite su identificación.<sup>20</sup>

Una de las técnicas citogenéticas más utilizada para el diagnóstico de enfermedades genéticas es el cariotipo, denominado también como cariotipo de rutina, esta técnica se enfoca en analizar anomalías cromosómicas que pueden ser numéricas (aneuploidías) que involucran ganancias y/o pérdidas cromosómicas o estructurales que incluyen pérdidas, duplicaciones, inserciones, inversiones o translocaciones de un segmento cromosómico. Sin embargo, esta técnica presenta limitaciones a nivel de resolución, permitiendo detectar únicamente reacomodos de un tamaño de 6,000-8,000 Kb (6-8 Mb) de DNA correspondientes a la banda oscura o clara desarrolladas por el bandeo G.<sup>20</sup> Si se requiere obtener un mayor grado de resolución de las bandas se debe recurrir al cariotipo de alta resolución. Esta técnica consiste en obtener cromosomas con un menor grado de compactación. Al aplicar bandeo G, se observa un mayor número de bandas (>600), permitiendo evaluar microarreglos cromosómicos (>4-5 Mb).<sup>21</sup>

Sin embargo, dado que la resolución de las bandas no siempre es suficiente para la identificación de anomalías cromosómicas, es necesario recurrir a técnicas adicionales a las de citogenética clásica para identificar aquellas alteraciones cromosómicas que escapan al análisis de estas técnicas, esto al implicar una alteración de menores proporciones en la secuencia de DNA. Debido a esto se han desarrollado otras opciones complementarias que involucran un nuevo campo de la citogenética: la citogenética molecular, producto de una combinación de citogenética y biología molecular: que incluye técnicas como hibridación in situ fluorescente (FISH, por sus siglas en inglés, Fluorescence in situ hybridization), hibridación in situ fluorescente multicolor (M-FISH o SKY) e hibridación genómica comparada (CGH, por sus siglas en inglés, Comparative Genomic Hybridization).<sup>22</sup>

FISH es una técnica visualmente atractiva que proporciona un grado intermedio de resolución entre el análisis de DNA y las investigaciones cromosómicas. La técnica consiste en la utilización de una sonda de DNA de hibridación, que reconoce partes del cromosoma con un alto grado de complementariedad de secuencia. Estas sondas pueden ser marcadas de forma directa, utilizando nucleótidos fluorescentes o indirectamente utilizando la incorporación de moléculas informadoras que posteriormente se puedan detectar con anticuerpos fluorescentes u otras moléculas de afinidad. En esta técnica se utiliza una preparación de cromosomas en interfase o metafase. Mismos que están firmemente unidos a un sustrato, generalmente vidrio. Luego, la sonda marcada y el DNA son desnaturalizados e incubados durante determinado tiempo mientras se hibridan. Los resultados se visualizan y cuantifican utilizando un microscopio que es capaz de excitar el fluoróforo y grabar las

imágenes.<sup>23</sup> Esta técnica es aplicada en el diagnóstico clínico de diversas anomalías cromosómicas, incluidas fusiones genéticas, identificación de un número anormal de cromosomas en una célula o la pérdida de uno, también es útil para detectar microdeleciones, microduplicaciones y translocaciones en regiones cromosómicas. A menudo también se utiliza en el mapeo genético y la identificación de oncogenes novedosos.<sup>24</sup>

La técnica de CGH es un enfoque eficaz para la detección de anomalías cromosómicas no balanceadas, es decir variaciones cromosómicas en el número de copias (CNV) ya sea ganancias o duplicaciones y/o pérdidas o deleciones que involucran regiones subcromosómicas o un cromosoma completo. Esta técnica compara el nivel de ploidía del DNA de una muestra en comparación con una muestra de referencia, esto sin la necesidad de realizar un cultivo celular y realizado en un solo experimento. La técnica se desarrollo en 1992<sup>25</sup> y se introdujo inicialmente para estudiar anomalías cromosómicas en tumores sólidos respecto a un tejido normal.<sup>26</sup>

La CGH cromosómica se basa en FISH de dos colores cuantitativos. La técnica implica el aislamiento de DNA de dos fuentes a comparar en donde las muestras son marcadas diferencialmente utilizando un tinte verde fluorescente (isotiocianato de fluoresceína, FITC) y un tinte rojo (rojo Texas), después las muestras son desnaturalizadas volviéndolas monocatenarias y precipitadas en presencia de DNA bloqueante para suprimir secuencias repetitivas, posteriormente son co-hibridadas con cromosomas normales en metafase para su propagación. Debido a que el

proceso de hibridación es simultáneo en los cromosomas normales desnaturalizados, existe una competencia por la hibridación del DNA en los sitios homólogos. Después de la hibridación y lavados, se pueden observar las propagaciones de los cromosomas en metafase bajo un microscopio fluorescente. El análisis de imágenes se realiza utilizando un software especial que captura y cuantifica las intensidades de fluorescencia resultantes de las hibridaciones de prueba y de referencia.<sup>27</sup>

Sin embargo, aunque la CGH se utilizó ampliamente en investigaciones de cáncer, no resultó ser particularmente valiosa como método estándar en los laboratorios de diagnóstico, debido a la pobre resolución espacial que ofrece la CGH en metafase, limitándose en la detección de 5-10 Mb por la adquisición de imágenes basadas en señales producidas por las sondas en las propagaciones en metafase utilizando microscopía de fluorescencia. Además, los cromosomas en metafase utilizados en CGH es un reto técnico que requiere experiencia para obtener cromosomas adecuados, así como para la adquisición y análisis de imágenes.<sup>28</sup>

En cambio, la CGH basada en micromatrices (array-CGH o aCGH) es una opción diagnóstica adicional a las técnicas anteriormente mencionadas. A diferencia de su predecesora implica la sustitución de cromosomas en metafase con secuencias de DNA adheridas a una matriz de vidrio, aumentando de esta forma la resolución a nivel de nucleótidos, siendo las matrices de polimorfismo de un solo nucleótido (matrices SNP) las que tienen la resolución más alta (5–10 kb), además no necesariamente (depende de la casa comercial/diseño) requiere la co-hibridación del DNA de prueba

y del DNA de referencia, debido a que el DNA de prueba puede hibridar directamente con la matriz de SNP. Esta técnica además de detectar las CNV, permite la detección de tramos de homocigosidad y, por lo tanto, la identificación de genes alterados causantes de enfermedades recesivas, aneuploidía de mosaico o disomía uniparental (UPD). Sin embargo, una desventaja que presentan estos arreglos es la capacidad limitada para detectar CNV más pequeños debido a la distribución que presentan los SNP en todo el genoma.<sup>29</sup>

## **Pruebas moleculares**

### **Secuenciación masiva de exoma**

La secuenciación de exoma completo (WES, por sus siglas en inglés whole exome sequencing) es una tecnología de alto rendimiento que consiste en la captura, secuenciación y análisis de la fracción codificante del genoma humano, que en conjunto se denomina exoma. Esta tecnología tiene la capacidad de proporcionar datos de secuencia de todos los genes conocidos que codifican a proteínas<sup>30</sup> además de regiones intrónicas cortas (150 pb) flanqueantes.<sup>31</sup> En breve, el proceso involucra la creación de una biblioteca a partir de DNA genómico (DNAg) el cual es procesado por un enzima que corta el DNAg en fragmentos, posteriormente los fragmentos son unidos a secuencias adaptadoras y son enriquecidos para llevar a cabo la secuenciación a partir de secuencias de exónicas a través de la captura de hibridación. Después de la captura, los fragmentos hibridados se amplifican y luego se someten masivamente a secuenciación paralela. Por último, se realiza una comparación de secuencias de un probando y una secuencia de referencia, en

donde permite la identificación de variaciones en la secuencia de DNA que potencialmente pueden ser asociadas a una enfermedad o rasgo en estudio.<sup>32</sup>

Debido a que el exoma representa solo ~ 1.5% del genoma humano, pero alberga ~ 85% de todas las mutaciones conocidas que causan enfermedades, WES es una alternativa cada vez más utilizada como una herramienta de diagnóstico exitoso en el estudio de enfermedades genéticas y la identificación de nuevos genes asociados a enfermedades.<sup>33-36</sup> En este sentido, esta tecnología ha comenzado a introducirse como una prueba de laboratorio clínico modificando el rumbo de los ensayos clínicos y diagnósticos. Esto con la promesa de reducir la “odisea diagnóstica” que enfrentan pacientes afectados por ER clínicas y genéticamente heterogéneas con presentaciones poco claras o atípicas que a menudo son difíciles de diagnosticar. Actualmente, WES puede considerarse como el método más eficiente y menos costoso para identificar todas las posibles mutaciones en los genes, esto en comparación con otros métodos, como los estudios de asociación de todo el genoma (GWAS, por sus siglas en inglés, Genome-wide association study) o la secuenciación del genoma completo (WGS, por sus siglas en inglés, Whole genome sequencing).<sup>37</sup>

### **Debilidades y fortalezas de WES**

WES es una tecnología robusta que es extremadamente práctica para investigar la variación a nivel genómico. A pesar del costo de la secuencia en los últimos años, WES a una profundidad de cobertura media de 100 × todavía cuesta cinco veces menos que la secuenciación del genoma completo (WGS) con una profundidad de

cobertura media de 30 ×. Además, el tamaño de los datos de WES por paciente es aproximadamente una sexta parte de los datos de WGS, lo que reduce el tiempo de procesamiento e impone una carga menor en términos de almacenamiento de datos. Sin embargo, es importante comprender las limitaciones de la tecnología WES. Ciertas regiones que codifican proteínas podrían no estar cubiertas debido a la anotación incompleta del genoma humano. Además, WES no cubre elementos no codificantes potencialmente funcionales, incluidas las regiones no traducidas, los potenciadores y los ARN largos que no codifican, aunque estos, en sí mismos, no están claramente definidos. Otro inconveniente de WES es la capacidad limitada para detectar variaciones estructurales, como variaciones de número de copias, translocaciones e inversiones.

A pesar de estas limitaciones, WES sigue siendo la herramienta de elección para muchos investigadores: sus ventajas prácticas permiten que un gran número de pacientes se proyecten de manera robusta, un aspecto crucial del descubrimiento de la mutación en la investigación genética humana.<sup>38</sup>

Gran parte del éxito de WES se debe a que la técnica se enfoca en secuenciar solo la porción codificante del genoma, es decir cerca de 30 millones de pares de bases y no los 3 mil millones de pares de bases que comprende el genoma completo de un individuo. Las secuencias exónicas tienen la capacidad de ser traducidas en proteínas funcionales, no es de esperar que estas secuencias sean propensas a sufrir mutaciones las cuales al ser traducidas a proteínas podrían tener una consecuencia fenotípica grave sobre el portador.



Además de estas ventajas, WES a contribuido en la identificación de la etiología genética de algunas enfermedades raras: el diagnóstico específico puede alterar el manejo del caso, las recomendaciones para el tratamiento y el seguimiento; proporciona un final a la odisea de diagnóstico que consume tiempo y dinero; facilita las decisiones reproductivas y el asesoramiento genético, también facilita la investigación sobre estos trastornos.<sup>39</sup>

## CAPÍTULO II

### ANTECEDENTES

#### **Antecedentes generales**

Los primeros esfuerzos exitosos para identificar genes subyacentes a los fenotipos mendelianos a menudo requerían un amplio conocimiento previo de la biología de la enfermedad, incluida la identidad de la proteína afectada.<sup>40</sup> Sin embargo, a mediados de la década de 1980, y durante las siguientes dos décadas, el enfoque principal para el descubrimiento de genes fue una combinación de análisis de ligamiento, clonación posicional y secuenciación de genes candidatos o seleccionados regionalmente, la mayoría de los cuales se basaron en hipótesis.<sup>41</sup>

Sin embargo, con la introducción de estrategias de secuenciación de próxima generación (NGS, por sus siglas en inglés, Next-Generation Sequencing) para la identificación de genes asociados con enfermedades y el incremento en las capacidades analíticas de los resultados, la genética médica revolucionó, al conducir un cambio de paradigma en el diseño de las investigaciones basadas en hipótesis a enfoques sin hipótesis en el diagnóstico de enfermedades genéticas raras. Estas tecnologías prometen una identificación robusta y amplia de los genomas en estudio con una resolución a nivel de nucleótidos.<sup>42</sup> En el 2009, este cambio se hizo evidente al incrementar el ritmo en el descubrimiento de genes basándose principalmente en la técnica de WES.<sup>43</sup> En este sentido, WES posee una clara ventaja como herramienta de diagnóstico debido a que no requiere preselección de genes y permite la detección de genes que no se sospecha que estén asociados con el trastorno, esto en comparación con métodos tradicionales,

como la secuenciación de Sanger, en donde es necesario saber de antemano que genes deben secuenciarse.<sup>44</sup>

Hoy en día, la WES se usa habitualmente como el enfoque tecnológico primario para descubrir asociaciones de genes y enfermedades. Resultando en una mejora significativa en la tasa de diagnóstico de enfermedades raras.<sup>45</sup> Así mismo, también condujo a la identificación de un espectro fenotípico más amplio en enfermedades conocidas, al descubrimiento de nuevos fenotipos asociados con genes conocidos y al descubrimiento de nuevas ER asociadas con nuevos genes que no habían sido investigados previamente y al establecimiento de fenotipos combinados debido a más de un diagnóstico.<sup>19</sup>

Su principal ventaja sobre la secuenciación del genoma completo (WGS, por sus siglas en inglés, Whole Genome Sequencing) se debe principalmente a su costo significativamente más bajo y que la mayoría de las variantes patógenas continúan dentro de la porción codificadora de proteínas del genoma.<sup>46</sup>

El primer reporte de la utilización de WES en el diagnóstico de un paciente se informó en año 2009. En donde según las características clínicas, el paciente se diagnosticó originalmente con el síndrome de Bartter, enfermedad tubular renal en donde hay presencia de alcalosis metabólica hipocalémica. Sin embargo, al realizar el diagnóstico molecular por WES, se descubrió una mutación homocigótica de cambio de sentido en el gen **SLC26A3** conocido como el locus congénito de diarrea por cloruro. Así mismo se reveló que este locus está completamente conservado en

ortólogos y parálogos de invertebrados a humanos. Después de la reevaluación, los médicos concluyeron que la presentación clínica estaba relacionada con la diarrea por cloruros y aplicaron un plan de tratamiento adecuado. Posterior a esto se informó que el seguimiento clínico confirmó el diagnóstico.<sup>47</sup>

Sin embargo, fue a partir del año 2011, donde WGS y WES han ganado un uso relativamente amplio para el diagnóstico etiológico de enfermedades genéticas. Al permitir el análisis de todos o la mayoría de los genes, WGS y WES tienen el potencial de permitir una detección completa y oportuna de enfermedades genéticas. El diagnóstico molecular oportuno, a su vez, tiene el potencial de instituir una nueva era en la medicina de precisión para las enfermedades genéticas. Una revisión sistemática identificó 37 publicaciones en el período de enero de 2011 a agosto de 2017, que incluyen 20.068 niños afectados, que informaron la utilidad de diagnóstico de WGS y WES.<sup>14</sup>

Uno de los primeros antecedentes que evidenciaron la utilidad de WES, fue el establecimiento del diagnóstico de un paciente que presentó enfermedad inflamatoria intestinal intratable. En este caso se identificó una novedosa mutación homocigota sin sentido en el gen inhibidor de la apoptosis ligado a X, sustituyendo a una tirosina por una cisteína altamente conservada y funcionalmente importante. Este gen cumple un papel central en la respuesta proinflamatoria y la detección bacteriana a través de la vía de señalización NOD. Tomando en cuenta el historial médico, los datos genéticos y funcionales, el paciente se diagnosticó con una deficiencia en el gen inhibidor de la apoptosis ligado a X. Con base en este

hallazgo, se realizó un trasplante alogénico de células progenitoras hematopoyéticas para prevenir el desarrollo de linfohistiocitosis hemofagocítica potencialmente mortal, en concordancia con el tratamiento recomendado, a más de 42 días después del trasplante, el paciente pudo comer y beber, y no se reportó recurrencia de enfermedad gastrointestinal.<sup>48</sup>

### **Antecedentes directos**

En nuestro grupo de investigación contamos con antecedentes directos en cuanto a metodología, reportados en la tesis de maestría de Rangel-Sosa en el 2017 en donde se evaluó a un grupo de pacientes que presentaban un fenotipo con múltiples anomalías congénitas (AC), incluidas la discapacidad intelectual (DI) y el retraso en el desarrollo (RD), mismos que tenían un diagnóstico inconcluso por técnicas de rutina como cariotipo y técnicas moleculares como aCGH. Los casos estudiados en dicha tesis son únicos en el mundo y no hay otro igual.

En el primer caso se identificaron 3 mutaciones puntuales (c.A97G, c.T98C y c.T100A) en el gen **SNRPA** que estuvieron presentes de manera homocigota en ambas pacientes y de manera heterocigota en sus padres no afectados, sugiriendo entonces una patología recesiva asociada a dicho gen. Estas tres mutaciones ocasionaron el cambio de dos aminoácidos contiguos (p.Ile33Ala y p.Phe34Ile) y ambas variantes fueron clasificadas como de significancia clínica incierta por la plataforma Variant Interpreter, probablemente debido a que no habían sido reportadas previamente. Sin embargo, los diferentes análisis bioinformáticos determinaron una predicción alta de variantes patogénicas o dañinas.

El producto del gen *SNRPA* forma parte de la ribonucleoproteína U1, la cual se une al sitio de corte y empalme 5' de los mRNAs precursores durante el proceso de corte y empalme. Esta proteína es un componente del “empalmosoma”, el cual es un complejo multi-componente formado por ribonucleoproteínas nucleares pequeñas, las cuales se denominan snRNP U1, U2, U4, U5 y U6, el cual está encargado de llevar el proceso de empalme.

La proteína ***SNRPA*** tiene dos dominios de unión y reconocimiento de RNAm que abarcan de los aminoácidos 10-89 y 208-282; las mutaciones presentes en las pacientes 1 y 2 caen dentro del primer dominio (10-89 aa), lo cual indica que la capacidad de reconocimiento y unión de la proteína puede estar afectada como resultado de las mutaciones, ocasionando un mal reconocimiento de RNAs mensajeros blanco o el fracaso de la interacción entre estos. Este análisis derivó en un diagnóstico de un trastorno del empalmosoma.<sup>49</sup>

En el segundo caso se encontraron dos mutaciones patogénicas en dos genes, ***SPAG17*** (p.Asp357His) y ***WDR35*** (p.Arg472Gln), siendo estas las probables responsables de producir el fenotipo del paciente, afectando nucleótidos altamente conservados que involucran alteraciones en la estructura y el transporte intraflagelar de cilios. Tomando en cuenta el historial médico, los datos genéticos y funcionales, el paciente fue diagnosticado con una enfermedad de ciliopatía.<sup>50</sup>

## **CAPÍTULO III**

### **HIPÓTESIS**

El enfoque genómico personalizado usando WES nos permitirá revelar las variantes genéticas asociadas al fenotipo clínico de los pacientes con ER.

### **JUSTIFICACIÓN**

Las ER son condiciones de morbilidad caracterizadas por su baja prevalencia, por lo cual son condiciones poco conocidas y estudiadas, y en general muchas de ellas sin la identificación de las causas moleculares, lo que produce diagnósticos más tardíos y poco específicos; esto se asocia a mayores costos y perjudica aún más la calidad de vida de los pacientes y sus familias, quienes habitualmente deben consultar varias opiniones de especialistas antes de llegar a un diagnóstico preciso.

Establecer un diagnóstico etiológico, certero y precoz, es fundamental para guiar decisiones clínicas importantes tales como prevención de complicaciones, cirugías, selección terapéutica, cuidados paliativos y asesoramiento pre-concepcional de las familias en riesgo. Así, la aplicación de tecnologías de nueva generación para la caracterización molecular de este tipo de enfermedades permitirá establecer una relación genotipo-fenotipo más precisa, identificando genes asociados a la etiología de cada una de las enfermedades en concreto, beneficiando por ende a los pacientes y sus familias.

## **CAPÍTULO IV**

### **OBJETIVOS**

#### **Objetivo general**

Identificar y caracterizar las alteraciones genómicas en pacientes con diagnóstico de ER mediante WES para establecer una relación genotipo-fenotipo integral.

#### **Objetivos específicos**

1. Realizar la secuenciación del exoma completo de los pacientes con diagnóstico de ER.
2. Realizar el análisis integral de los datos para identificar las variantes que puedan explicar el fenotipo clínico de los pacientes.
3. Validar las variantes encontradas mediante secuenciación Sanger.



## **CAPÍTULO V**

### **DISEÑO EXPERIMENTAL**

#### **Diseño del experimento**

El estudio por realizar es un estudio descriptivo, transversal, observacional y retrospectivo con seguimiento.

#### **Estándares éticos**

El protocolo fue sometido al Comité de Ética de la Subdirección de Investigación de la Facultad de Medicina, UANL el cual aprobó el protocolo básico con número de registro BI15-003.

Todos los participantes, así como sus tutores legales recibieron una explicación completa sobre los objetivos y procedimientos del estudio, además firmaron un formulario de consentimiento informado por escrito.

#### **Universo de estudio**

El universo de estudio estará constituido por individuos seleccionados con base en una alta sospecha clínica de algún trastorno mendeliano basado en la presentación clínica de la enfermedad y los antecedentes familiares, la cual tendrá la característica de ser de etiología desconocida. Así mismo serán incluidos los padres biológicos de los individuos participantes. En este estudio no nos concentraremos en una presentación clínica específica, sino que incluiremos una variedad de casos en colaboración con diferentes sub-especialistas.

## **Reclutamiento de pacientes**

El reclutamiento de los pacientes fue realizado en el Centro de Investigación Biomédica de Occidente del IMSS en Guadalajara, Jalisco. Todos los pacientes que aceptaron participar firmaron carta de consentimiento informado. En donde los procedimientos se realizaron acordes a la Declaración de Helsinki.

## **Evaluación clínica**

Los genetistas médicos y sus médicos tratantes evaluaron a los participantes sanos (padres biológicos) y realizaron la caracterización fenotípica de los pacientes afectados del estudio, incluidos los fenotipos profundos, así como los antecedentes familiares y árbol genealógico. Los datos clínicos de las historias clínicas se utilizarán para un mayor acercamiento al fenotipo estudiado.

## **Muestra de estudio**

Se analizaron muestras de DNAg de pacientes con diagnóstico de ER, así como de sus padres biológicos, que previamente han sido referidos con un medico genetista y continúan con un diagnóstico genético inconcluso después de haber realizado técnicas de rutina como el cariotipo o técnicas moleculares como aCGH.

Dado que los casos fueron totalmente independientes, y no se buscaba relación o asociación entre ellos, no se requirió un gran número de muestras ni establecer una "n". De esta manera, se captaron los pacientes que acudieron a consulta de genética en un periodo de 1 año, por lo cual, esta muestra se considero censal (número de muestras x unidad de tiempo).

## **Criterios de selección**

### **Criterios de inclusión**

- Que aceptaran participar y firmen la autorización de consentimiento informado, en caso de ser menores de edad que los tutores legales firmen dicha autorización.
- Pacientes diagnosticados con ER referidos por instituciones participantes.
- Pacientes con cariotipo normal o cariotipo alterado sin delineación de la alteración.
- Que se les haya realizado aCGH y el diagnóstico molecular haya sido inconcluso.
- Que contaran con evaluación clínica integral por diferentes especialidades utilizando diferentes abordajes.
- Que se hayan descartado factores extrínsecos, como agentes teratógenos, infecciosos y exposición a tóxicos como causantes del fenotipo.

### **Criterios de no inclusión**

- Pacientes con alguna ER asociados a una alteración estructural genética conocida.
- Pacientes con alguna alteración genética que no calificaran como ER.

### **Criterios de eliminación**

- Muestras de mala calidad o cantidad insuficiente para realizar el análisis.

## **CAPÍTULO VI**

### **MATERIALES**

#### **Lugar de realización**

Este trabajo se realizó en el Laboratorio de Citogenómica y Microarreglos del Departamento de Bioquímica y Medicina Molecular de la Facultad de Medicina de la Universidad Autónoma de Nuevo León, en el periodo comprendido de agosto del 2018 a julio del 2019.

#### **Reactivos**

##### **Soluciones de uso general**

La limpieza de la zona de trabajo se realizó utilizando dos soluciones diferentes, una solución al 1% de dodecil sulfato de sodio (SDS, marca Sigma) y etanol al 70% (Etanol desnaturalizado, marca CTR Scientific), utilizando como diluyente agua grado biología molecular, marca Corning.

#### **WES**

##### **Verificación de la cantidad y calidad de muestra de partida**

- Agua grado biología molecular, marca Corning.
- Kit Qubit dsDNA HS Assay, marca Invitrogen.

##### **Preparación de la biblioteca**

Para la preparación de la biblioteca se utilizó el Kit TruSeq Rapid Expme Library Prep de la marca Illumina. Este kit utiliza la fragmentación del DNAG basado en la actividad de la enzima transposasa, el enriquecimiento del exoma mediante transposones, así como la captura de las regiones mediante sondas de captura

biotiniladas. El kit contiene sondas de captura exónicas que incluye >340,000 sondas 95 bases, construidas a partir del genoma de referencia humano hg19/NCBI37. El conjunto de sondas fue diseñado para enriquecer > 200,000 exones, abarcando 20,794 genes de interés. Si bien la longitud total de estas sondas es de 32 Mb, el kit se enfoca en realidad 62 Mb del genoma humano (117.5 Mb si se capturan las regiones de 150 bp corriente arriba y corriente abajo del objetivo).

### **Comprobación de biblioteca:**

#### **Gel de agarosa al 4%**

- Agarosa Ultrapura, marca Invitrogen.
- Tris-Base, marca Sigma-Aldrich.
- Ácido bórico, marca Ambresco.
- EDTA, marca J. T. Baker.
- Marcador de peso molecular 100pb, marca Invitrogen.

#### **Qubit**

- Kit Qubit dsDNA HS Assay, marca Invitrogen.

#### **Desnaturalización y dilución de la biblioteca.**

- Solución 1.0N de NaOH, grado biología molecular, marca Thermo Fisher Scientific.
- Buffer de hibridación (HT1), marca Illumina.
- PhiX Control, marca Illumina.
- Agua grado biología molecular, marca Corning.

### **Generación de clusters y secuenciación**

Para llevar a cabo la generación de clusters y secuenciación se utilizó el Kit MiSeq Reagent v3.0, marca Illumina.

### **Secuenciación Sanger**

#### **Estandarización de cebadores y obtención de amplicones mediante reacción en cadena de la polimerasa (PCR) punto final.**

Para la estandarización de los cebadores se utilizó la enzima Taq DNA polymerase recombinante, buffer 10X para PCR, solución de  $MgCl_2$  50mM y dNTP 50mM marca Invitrogen™. El volumen de reacción se ajustó a 50 $\mu$ L con agua grado biología molecular, marca Corning.

Los cebadores utilizados fueron construidos en la compañía Integrated DNA Technologies (Coralville, IA USA).

### **Purificación de amplicones**

Para la purificación de los amplicones se utilizó el Kit de extracción en gel QIAquick®.

### **Secuenciación Sanger**

- Applied Biosistem 3500 (Thermo Fisher Scientific)

## Materiales

### Material de uso general

Tabla 1. Materiales de uso general		
Material	Volumen	Marca
Micropipeta	100 a 1000 $\mu$ L	Eppendorf Research Plus
	20 a 2000 $\mu$ L	
	100 a 1000 $\mu$ L	
	2 a 2000 $\mu$ L	
	0.5 a 10 $\mu$ L	
	0.1 a 2.5 $\mu$ L	
Puntillas desechables para micropipeta	1000 $\mu$ L	Axygen
	200 $\mu$ L	
	10 $\mu$ L	
Puntillas desechables con filtro	1000 $\mu$ L	Axygen
	200 $\mu$ L	
	10 $\mu$ L	
Tubos	15 mL	Corning
	50 mL	
Microtubos	2 mL	Axygen
	1.5 mL	
	0.5 mL	
	0.2 mL	
Microtubos Qubit assay	0.5 mL	Axygen
Guantes de nitrilo	-	Microflex
Toallas desechables de papel	-	Kimberly Clark

## Equipos

### Equipo de uso general

- Microcentrifuga centrifugue 5415 C, marca Eppendorf
- Microcentrifuga, marca Labnet
- Plancha de calentamiento, marca Thermolyne.
- Vórtex, marca Thermolyne
- Balanza analítica BM-500, marca fraLib

- Refrigerador 4°/-20°C, marca Thermo Scientific.
- Cámara de electroforesis Mini-Sub cell GT, Marca Bio-Rad.
- Fuente de poder PowerPac Basic, marca Bio-Rad
- Fotodocumentador ChemiDoc XRS+, marca Bio-Rad
- Horno de microondas, marca LG.
- NanoDrop 2000, marca Thermo Scientific.

#### **Equipos para secuenciación masiva.**

- Fluorómetro Qubit, marca invitrogene.
- Secuenciador MiSeq, marca Illumina.
- Horno de Hibridación Hybex, marca SciGene
- Centrifuga para placas Legend XTR, marca Thermo Scientific.
- Termociclador SureCycler 8800, marca Agilent Technologies.

#### **Secuenciación Sanger**

- Termociclador SureCycler 8800, marca Agilent Technologies.
- Applied Biosistem 3500 (Thermo Fisher Scientific)



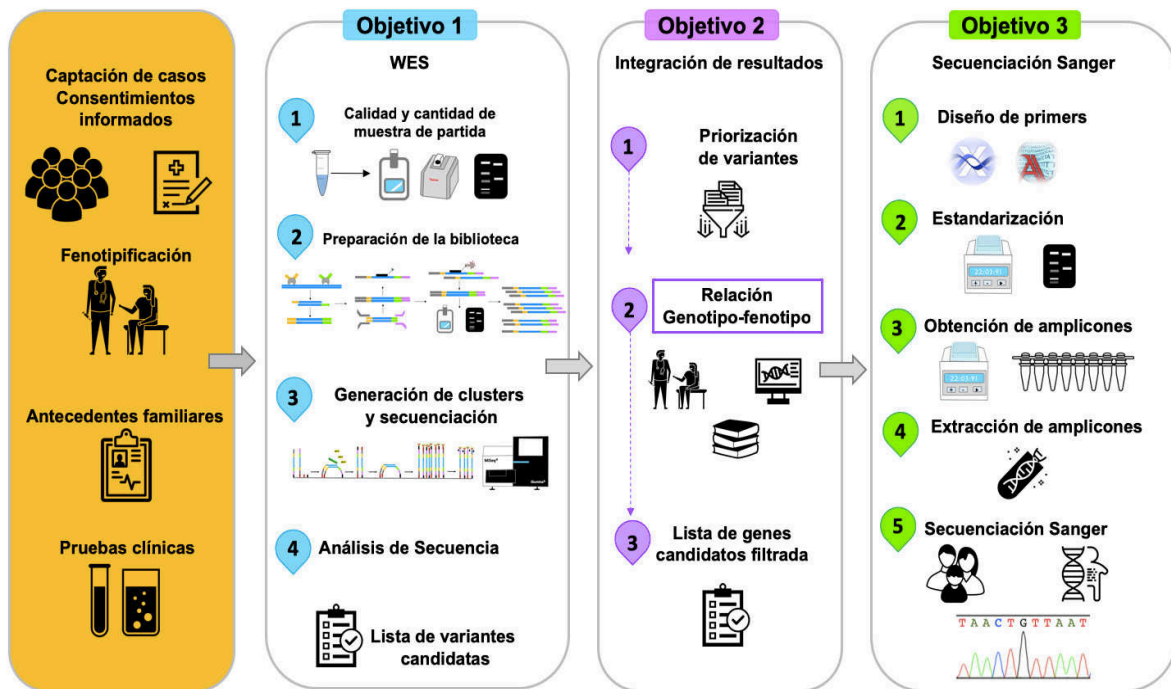
## Programas computacionales

Tabla 2. Programas computacionales y plataformas bioinformáticas.		
Nombre	Descripción	Liga
<b>OMIM</b>	Compendio de genes humanos y fenotipos genéticos que se actualiza constantemente. Contiene información sobre todos los trastornos mendelianos conocido	<a href="http://www.omim.org/">www.omim.org/</a>
<b>UCSC Genome Browser Home</b>	Navegador que ofrece acceso a datos de secuencia del genoma de una amplia variedad de especies, organismos modelo, integrado con una colección de anotaciones alineadas.	<a href="http://genome.ucsc.edu/">genome.ucsc.edu/</a>
<b>GeneCards</b>	Base de datos que proporciona información sobre todos los genes humanos anotados y predichos. Integra información genómica, transcriptómica, proteómica, genética, clínica y funcional.	<a href="http://www.genecards.org/">www.genecards.org/</a>
<b>ExAC Browser</b>	Exome Aggregation Consortium (ExAC) es una coalición de investigadores que buscan agregar y armonizar los datos de secuenciación del exoma de una amplia variedad de proyectos de secuenciación a gran escala.	<a href="http://exac.broadinstitute.org/">exac.broadinstitute.org/</a>
<b>1000 Genomes<sup>43</sup></b>	Base de datos que almacena el material genético de mil personas en todo el mundo, con la finalidad de obtener datos que permitan estudiar la variabilidad genética humana. <sup>43</sup>	<a href="http://www.internationalgenome.org/">www.internationalgenome.org/</a>
<b>PolyPhen-2</b>	Software que predice el posible impacto de las sustituciones de aminoácidos en la estabilidad y función de las proteínas humanas utilizando consideraciones evolutivas comparativas y estructurales. <sup>44</sup>	<a href="http://genetics.bwh.harvard.edu/pph2/">genetics.bwh.harvard.edu/pph2/</a>
<b>SIFT</b>	Predice si una sustitución de aminoácidos afecta la función de la proteína en función de la homología de secuencia y las propiedades físicas de los aminoácidos.	<a href="http://sift.bii.a-star.edu.sg/">sift.bii.a-star.edu.sg/</a>
<b>PANTHER</b>	Sistema que clasifica las proteínas y sus genes con el fin de facilitar el análisis de alto rendimiento. <sup>46</sup>	<a href="http://www.pantherdb.org/">www.pantherdb.org/</a>
<b>Uniprot</b>	Recurso integral para la secuencia de proteínas y datos de anotación. <sup>48</sup>	<a href="http://www.uniprot.org/">www.uniprot.org/</a>
<b>Reactome</b>	Base de datos de código abierto, curada manualmente y revisada por pares. Proporciona herramientas de bioinformática intuitivas para la visualización, interpretación y análisis del conocimiento de las vías para apoyar la investigación básica y clínica, el análisis del genoma y biología de sistemas.	<a href="https://reactome.org/">https://reactome.org/</a>

# CAPÍTULO VII

## MÉTODOS

### Esquema general



**Figura 1. Estrategia general de trabajo.** La estrategia general fue dividida en 4 bloques: captación de casos y obtención de consentimientos informados, secuenciación masiva de exoma, integración de resultados. y validación de mutaciones.

## **Naturaleza de las muestras**

Las muestras biológicas utilizadas en este estudio corresponden a muestras de DNAg de 5 tríos, conformados por el probando y sus padres biológicos. Los probandos contaban con un diagnóstico genético inconcluso por técnicas de rutina y moleculares. Las muestras fueron almacenadas a -20°C hasta su uso.

## **WES**

### **Comprobación de la calidad, cantidad e integridad de las muestras**

Para verificar la cantidad y pureza de los ácidos nucleicos presentes en las muestras de partida, se cuantificó cada una de las muestras utilizando el equipo NanoDrop 2000 (Thermo Scientific), este equipo se rige bajo el principio de espectrofotometría de luz UV en donde se sugiere que una molécula presente en una solución en un estado basal puede ser excitada por una fuente de radiación electromagnética (luz) y dichas moléculas tienen la capacidad de absorber radiación a determinada longitud de onda provocando transiciones electrónicas que pueden ser cuantificadas.

Para el ensayo fue necesario utilizar 1 µL de muestra. Al efectuar las lecturas, se utilizó la longitud de onda de 260 nm y como blanco se usó 1 µL agua grado biología molecular. En este ensayo se tomó en cuenta la relación 260/280 para analizar la pureza de las muestras, considerando una pureza óptima entre el rango de 1.8-2.0, ideal para llevar a secuenciación.

Así mismo, se realizó una electroforesis en gel de agarosa al 1% para comprobar la integridad de los DNAg. Los gels se mantuvieron durante 50 minutos a 90 volts en

la cámara de electroforesis. Posteriormente se visualizaron en el equipo ChemiDoc XRS+, Bio-Rad. El resultado óptimo fue visualizado en forma de una sola banda en la parte superior del gel sin presencia de barrido.

Una vez realizadas estas dos pruebas se procedió a cuantificar de nueva cuenta las muestras, pero esta vez utilizando el equipo Qubit (Invitrogene). El cual es un método más exacto para medir ácidos nucleicos de doble hebra. Este equipo se rige por el principio de fluorescencia. Para el ensayo se preparó una solución de trabajo mezclando 1  $\mu\text{L}$  de reactivo Qubit (Invitrogene) y 199  $\mu\text{L}$  del Buffer Qubit, (Invitrogene) por ensayo. Posteriormente se prepararon dos estándares para realizar la curva de calibración, añadiendo a dos tubos Qubit assay 190  $\mu\text{L}$  de solución de trabajo y 10  $\mu\text{L}$  de solución estándar número 1 (0 ng/  $\mu\text{L}$ ) y estándar número 2 (10 ng/  $\mu\text{L}$ ). El tubo que contenía la muestra fue preparado utilizando 198  $\mu\text{L}$  de la solución de trabajo y 2  $\mu\text{L}$  de muestra. Las soluciones se agitaron en el vórtex de dos a tres segundos e incubaron a temperatura ambiente por dos minutos para posteriormente ser leídas en el equipo Qubit.

### **Preparación de la biblioteca**

La preparación de la biblioteca de cada una de las muestras se realizó utilizando el kit TruSeq Rapid Exome Library Prep de la marca comercial Illumina.

A partir de las concentraciones obtenidas en el NanoDrop y Qubit de las muestras de DNAG, se realizó una primera dilución a una concentración de 50 ng en un volumen final de 20  $\mu\text{L}$  obteniendo una concentración final de 1000 ng/  $\mu\text{L}$  de DNAG total, se utilizó agua libre de nucleasas como diluyente, esta dilución se re cuantificó utilizando el método de Qubit para conocer exactamente la concentración final.

Después, partiendo de esta misma dilución, la muestra se normalizó a una concentración de 5 ng en un volumen final de 10  $\mu$ L obteniendo una concentración fina de 50 ng/  $\mu$ L totales de DNAg, misma que fue corroborada mediante el método de Qubit.

Una vez obtenida la concentración adecuada de DNAg (50 ng/  $\mu$ L), se procedió a realizar el proceso de tagmentación. Este Kit utiliza la tecnología de Nextera, el cual es un sistema de preparación de bibliotecas de DNA que corresponde a un proceso de fragmentación y etiquetado simultáneos por transposición in vitro. La reacción utilizada es una reacción modificada conocida como tagmentación en donde participa el complejo transposoma el cual está conformado por una transposasa y un transposón. Cuando este complejo se incuba con DNA bicatenario el objetivo es fragmentado de forma aleatoria y marcado simultáneamente en los extremos 5' del fragmento diana con secuencias adaptadoras específicas, después mediante una PCR de ciclos cortos las secuencias son complementadas en la porción 3' con las secuencias faltantes. Una vez finalizado este proceso la muestra se colocó en una placa cónica de 96 pozos para la purificación del DNA tagmentado. En este proceso se utilizaron perlas de purificación SPB (Sample Purification Beans, Illumina) necesarias para la generación de grupos y la secuenciación.

El siguiente paso consistió en dos ciclos de hibridación y captura de sondas biotiniladas. Se utilizó el reactivo CEX (Coding Exome Oligos, Illumina) el cual contiene sondas de captura exónicas que incluye >340,000 sondas de 95 bases construidas a partir del genoma de referencia humano hg19/NCBI37. El conjunto de sondas fue diseñado para enriquecer > 200,000 exones, abarcando 20,794 genes

de interés. Si bien la longitud total de estas sondas es de 32 Mb, el kit se enfoca en realidad en 62 Mb del genoma humano (117.5 Mb si se capturan las regiones de 150 bp corriente arriba y corriente abajo del objetivo). Además las secuencias enriquecidas también incluyen ~ 28 Mb de regiones no traducidas (UTR) <sup>51</sup>. Una vez hibridadas las sondas a sus regiones diana, la captura de las sondas se llevó a cabo con perlas magnéticas de estreptavidina (SMB, Illumina) las cuales se fueron extraídas de la solución utilizando una placa magnética. Posteriormente se realizó una amplificación final de la biblioteca utilizando el reactivo PPC (PCR Primer Cocktail) y EAM (Enrichment Amp Mix).

La biblioteca final fue visualizada mediante un gel de agarosa al 4% en donde se determinó el tamaño promedio de los fragmentos que conforman la biblioteca (300pb aproximadamente), además la biblioteca fue cuantificada utilizando el equipo Qubit. Estos dos resultados fueron utilizados para calcular la concentración de fragmentos de DNA en nM (Ecuación 1).

$$\frac{(\text{Concentración en ng/}\mu\text{l})}{(660 \text{ g/mol} \times \text{Tamaño promedio de la biblioteca})} \times 10^6 = \text{Concentración en nM}$$

\*660g/mol son equivalentes al PM de un par de bases.

**Ecuación 1.** Conversión de ng/  $\mu$ L a nM

Después se realizó una dilución de la biblioteca utilizando RSB como diluyente hasta llegar a una concentración de 4nM. Con la biblioteca a esta concentración se procedió a realizar el protocolo de desnaturalización y dilución de la biblioteca. Se

utilizó una solución fresca de NaOH 0.2 N y se realizaron diluciones de 4nM a 20pM y de 20pM a 12.5 pM utilizando como diluyente el reactivo TH1 (Hybridization Buffer). Este mismo proceso fue realizado con la biblioteca de referencia PhiX. Por último, ambas bibliotecas fueron mezcladas utilizando 594 µL de biblioteca de muestra y 6 µL biblioteca PhiX.

### **Generación de grupos y secuenciación**

Para llevar a cabo la generación de grupos y secuenciación, se utilizó un cartucho precargado de secuenciación incluido dentro del kit MiSeq Reagent v3.0 de la marca Illumina. Este cartucho contenía los reactivos necesarios para llevar a cabo la secuenciación masiva. Para su activación el cartucho fue descongelado utilizando un baño de agua desionizada a temperatura ambiente durante 60 minutos. Una vez que se descongelaron los reactivos contenidos en el cartucho se mezclaron varias veces con la finalidad de incorporar los reactivos contenidos. Una vez preparado el cartucho se cargó la biblioteca en el espacio de carga de muestra.

Para llevar a cabo la limpieza de la celda de flujo, la celda se extrajo de la solución de almacenamiento con la ayuda de unas pinzas de plástico. La celda se enjuagó ligeramente con agua destilada estéril y etanol al 80% hasta que se eliminó el exceso de sales tanto de los componentes plásticos como del cristal. En seguida se procedió a limpiar el exceso de alcohol que pudiera haber quedado en el cartucho utilizando un paño sin pelusa.

Por último, se cargó el cartucho, la celda de flujo y el buffer PR2 (incorporation Buffer) dentro del equipo MiSeq. En el software se introdujo un archivo de hoja de

muestras (simple sheet) el cual contuvo la información general de las muestras y el código de barras del cartucho cargado en el instrumento. Una vez introducidos todos los componentes se comenzó la secuenciación.

### **Análisis de secuencia**

La generación de grupos y secuenciación tuvo una duración promedio de 6 horas. Una vez secuenciada la biblioteca se produjo la llamada base, que es la inferencia de las bases a partir de las señales físicas obtenidas de la liberación de la fluorescencia de los ddNTP al ser incorporados a la cadena naciente. Esta llamada base fue analizada con el software MiSeq Control (MCS) v2.4., el cual es una interfaz de control que proporciona control para configurar el instrumento, monitorear ejecuciones y realizar procedimientos de mantenimiento. Las lecturas en bruto fueron analizadas utilizando el programa FastQ. El primer punto de control de calidad se realizó a nivel de pre-alineación obteniendo puntuaciones Q o puntuaciones de calidad, que incluirán datos relacionados con el % de error, puntuaciones de calidad por base y % de duplicación. En este punto solo aquellas secuencias que obtuvieron valores de  $Q \geq 30$  (probabilidad de error de 0.001), es decir, la probabilidad de encontrar un error cada 1000pb, pasaron el control de calidad.

Las secuencias fueron alineadas utilizando el genoma humano de referencia hg38 (GRCh38) para lo cual fue utilizado el algoritmo BWA Enrichment v.2.1.0.0.<sup>52</sup> El segundo punto de control de calidad se realizó a nivel de pos-alineación, se utilizó el programa FASTQC para calcular los valores de perfil de alineación. El programa proporcionó información referente a la proporción de lecturas alineadas con el



genoma de referencia, profundidad promedio, cobertura promedio, secuencias sobrerrepresentadas, tasa de error y tasa de duplicación. La cobertura óptima se estableció en  $\geq 95\%$  y la profundidad en  $\geq 20X$  para todos los análisis.

Una vez que se realizó el control de calidad apropiado y la alineación de las secuencias, se llevó a cabo el proceso de determinar con precisión las variaciones que existen entre el probando y el genoma de referencia. Este proceso se denomina asignación e identificación de variantes. Las variantes se presentaron en forma de variantes de nucleótido únicos (SNV), variantes de número de copias (CNV), inserciones o deleciones pequeñas, menos de 50 pb (INDELS) y variantes de significado incierto (VUS).

El siguiente paso realizado fue la asignación de información funcional a las variantes identificadas. Este proceso tenía como objetivo proporcionar un vínculo entre las variantes identificadas y el fenotipo del paciente. Como primer paso se utilizaron las anotaciones genéticas conocidas para clasificar las variantes de acuerdo con su categoría funcional (por ejemplo, sinónimo, cambio de sentido, sin sentido, splicing, etc.).

Después, se utilizaron reservorios de datos públicos como dbSNP, The 1000 Genomes Project Consortium (1000G) y ExAC, para analizar la frecuencia alélica del alelo menor. Además, las variantes fueron analizadas utilizando las herramientas bioinformáticas SIFT o Poly-Phen para predecir el efecto de la variante sobre la proteína.

## CAPÍTULO VIII

### RESULTADOS

#### Descripción de casos

##### Caso 1 - Probando LCM121

##### Diagnostico presuntivo: Nefropatía

Probando femenino de 11 años, originaria de Ixtlán de Juárez, Oaxaca. Proveniente de padres aparentemente sanos, niegan consanguinidad y toxicomanías, así como antecedentes de evento similar en ambas familias. Fue el producto de la primera gesta con atención prenatal adecuada.

Su desarrollo fue normal hasta el año 2014, después fue diagnosticada con urolitiasis (creatinina 1.8 mg/dL), un mes más tarde mediante tomografía axial computarizada (TAC) se detectó litos coraliformes confirmandos mediante uteroscopia izquierda a lo cual la paciente fue tratada con un procedimiento de fragmentación endoscópica por litotripsia con láser, con la consecuente colocación de un catéter uretral doble J. Al tercer mes se le realizó un recambio de catéter uretral y presento una urgencia dialítica (Urea 354 mg/dL) derivando en una peritonitis.

En 2015 la paciente reingresó y fue diagnosticada por segunda ocasión con urolitiasis recurrente mediante TAC, se realizó procedimiento de fragmentación endoscópica con la colocación de catéter uretral doble J. Fue evaluada por diferentes servicios, en donde oftalmología la describió como aparentemente sana, cardiología la diagnosticó con hipertrofia ventricular izquierda leve e insuficiencia

leve funcional de la válvula tricúspide. Tres meses después fue retirado el catéter uretral y se diagnosticó con reflujo vesicoureteral grado III del lado derecho.

En el 2016 la probando reingresa. Mediante una TAC de abdomen se detectaron riñones atróficos con densidades muy altas por calcificación parenquimatosa y litos calciales sin dilatación de sistemas colectores. Además, se reportaron porcentajes altos de oxalato de calcio (43%) y fosfatos (32%). Seis meses después se realizó una biopsia hepática por punción transcutánea en donde los hallazgos sugieren enfermedad hepática crónica con actividad moderada a severa con un Índice de METAVIR F3. Bajo los hallazgos patológicos de la biopsia, el abordaje que se utilizó fue en sospecha de algún problema autoinmune, para lo cual se amplió el abordaje. Se realizaron pruebas de: Anticuerpos de tipo 1 microsomales de hígado y riñón (Ac-LKM1), anticuerpos antinucleares, anticuerpos anti-mitocondriales y anticuerpos anti-Smith con resultando negativo, motivo por el cual fue descartada la etiología autoinmune.

Debido a la afección hepática el abordaje que se utilizó fue de etiología bioquímica, sin embargo, solo se reporta una elevación persistente, aunque discreta de TGO, que pudiera correlacionar con la afección en el riñón.

En el 2017 la probando reingreso y se detectó oxalatos en sangre (13.9 mM/L). Oftalmología detectó opacidades corticales sutiles que no afectan agudeza visual. La paciente presentó crisis convulsivas secundarias a hipertensión arterial (HTA). Cuatro meses después fue diagnosticada con polineuropatía involucrando trastorno

de la marcha y hemicuerpo derecho mixto con hipotonía e hiporreflexia, además de ratiquismo hiperfosfaturico: rodilla izquierda: Valgo izquierdo y rodilla derecha: Varo, así como rotoescoliosis curva. Meses más tarde la probando presentó una sobrecarga hídrica, derrame pericárdico y HTA no controlada.

## **Caso 2- Probando LCM124**

### **Diagnostico presuntivo: Síndrome FATCO**

Probando femenino de 1 año 4 meses de edad, originaria de Jamay, Jalisco, fue referida debido a alteraciones en miembros inferiores y superiores por sindáctila y oligodactilia en mano derecha y defecto terminal en piernas.

Proviene de padres aparentemente sanos no expuestos a agentes tóxicos; no se refirió ninguna consanguinidad y no hubo ningún otro miembro de la familia afectado con anomalías congénitas.

Fue valorada por traumatología y cirugía pediátrica en el hospital general de zona número 6 en Ocotlán, Jalisco. La exploración física (EF) revelo una talla 69 cm y peso 10.3 k. Cráneo con metópica prominente, fisuras palpebrales hacia arriba, puente nasal bajo, punta de la nariz bulbosa, narinas antevértidas, cavidad oral integra con paladar alto. Pabellones auriculares con lóbulo pequeño y adosado, cuello móvil, tórax normolíneo, área cardiaca rítmica sin agregados, abdomen sin visceromegalias, genitales de acuerdo con edad y sexo, extremidades mano izquierda 5 dedos y pliegues de flexión respetados, mano derecha con oligodactilia solo 3 dedos, cirugía para liberación de sindactilia entre dedos 2 y 3, extremidades

inferiores en región media de pierna aprecio curvatura anterior palpable, ambos pies con oligodactilia, solo 4 dedos, así mismo mostró agenesia bilateral de fíbula.

### **Caso 3 - Probando LCM129**

#### **Diagnostico presuntivo: Síndrome FATCO**

Paciente masculino de 9 años, originario de Guadalajara, Jalisco, acudió a una evaluación por abscesos en la región cervical, además de oligodactilia y agenesia de peronés.

La exploración física realizada por el servicio de genética revelo un perímetro cefálico (PC) 50 cm (Z -2.5), talla (T) 104 CM (Z -5.38), peso (P) 18.2 (Z -4.33), normocefalia, frente amplia, fisuras palpebrales inclinadas hacia arriba, puente nasal bajo, punta nasal bulbosa, cavidad oral integra con paladar hendido submucoso con úvula bífida, hiperplasia amigdalina, pabellones auriculares antevertidos, lóbulos hipoplásicos, tórax normolineo, área cardiaca rítmica sin agregados, abdomen sin visceromegalias, genitales de acuerdo a edad y sexo, extremidades superiores con mesomelia, mano derecha solo dos dedos con función de pinza, mano izquierda con oligo-sindactilia de dedos 2 y 3, limitación a la flexión de dedos de mano, en extremidades inferiores mesomelia con desviación lateral de ambos pies, oligodactilia de pie derecho (solo dos dedos) y en pie izquierdo oligosindactilia de dedos 2 y 3.

#### **Caso 4 - Probando LCM119**

##### **Diagnostico presuntivo:** Enfermedad mitocondrial

Paciente femenina de 25 años, originaria de Poncitlán, Jalisco. Proviene de padres aparentemente sanos que niegan consanguinidad. La paciente presentaba dismotilidad gastrointestinal, parestesias de miembros pélvicos inferiores, pseudo obstrucción intestinal, disfagia y gastroparesia.

#### **Caso 5 - Probando LCM13**

##### **Diagnostico presuntivo:** Trisomía 21 y Síndrome de Marfan

Paciente masculino de 20 años de edad diagnosticado con Síndrome de Down según su cariotipo mos47,XY,+21 (20)/ 48,XY,+21+mar(3). Presentaba talla baja, braquicefalia, ojos con epicanto bilateral, puente nasal y narinas simétricas, filtrum corto, comisuras labiales dirigidas hacia abajo, cavidad y mandíbula sin alteraciones, pabellones auriculares con implantación baja. Cuello tórax y abdomen simétrico, genitales masculinos con resección del testículo izquierdo, extremidades superiores con braquidactilia bilateral, extremidades inferiores con piernas incurvadas y presencia de pie plano bilateral.

Además de estas características el probando contaba con cataratas, luxación del cristalino y dilatación de la raíz aortica detectada por angiotomografía.

Características propias de síndrome de Marfan.

## **Caso 6 - Paciente LCM108**

### **Diagnostico presuntivo:** Trastorno del espectro autista (TEA)

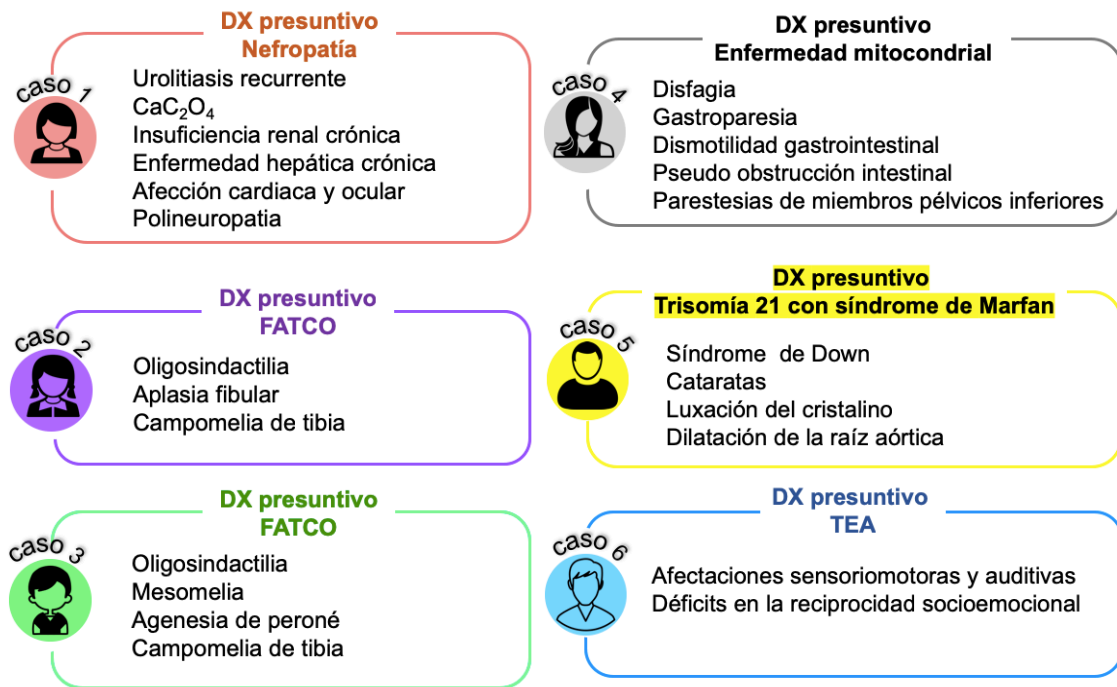
Probando masculino de 10 años, fue referido debido al diagnóstico de TAE. Proviene de padres aparentemente sanos, al momento de la concepción; no se refirió ninguna consanguinidad y no hubo ningún otro miembro de la familia afectado con anomalías neuropsicológicas.

Su desarrollo psicomotor fue normal hasta los 16 meses cuando exhibió episodios de convulsiones y regresión del lenguaje. Poco después, mostró comportamiento agresivo y retracción social, reacción exagerada al ruido, actividades estereotipadas y ritualizadas, a la edad de 3 años, sufrió una disfunción de la deglución y fue diagnosticado con TEA según los criterios del DSM-IV-TR y ADOS. Fue hospitalizado a los 6 años por pseudo-obstrucción intestinal. Posteriormente, fue evaluado por el Servicio de Neurología Pediátrica y fue tratado con risperidona, valproato de magnesio, levetiracetam y metilfenidato.

Se le realizó un electroencefalograma en donde se exhibió actividad subcortical irritativa con ondas bilaterales agudas lentas, potenciales auditivos evocados con hipoacusia izquierda para tonos agudos y conducción central derecha retrasada en la porción inferior del tronco cerebral. La detección metabólica y tiroidea del recién nacido fue normal. Se realizó un diagnóstico DSM-IV de TEA basado en los déficits en la reciprocidad socioemocional con una participación reducida en los intereses, las emociones o el afecto y la incapacidad de iniciar o responder a las interacciones sociales a los 12 meses, la falta de expresiones faciales y la comunicación no verbal a la edad 18 meses seguidos por dificultades para ajustar el comportamiento a

diferentes contextos sociales; dificultades para compartir juegos imaginativos o para hacer amigos; movimientos motores estereotipados y uso de objetos, angustia extrema por pequeños cambios, dificultad con las transiciones, patrones de pensamiento rígidos.

### Signos pivote exhibidos



**Figura 2. Signos pivote.** Signos pivote exhibidos por cada uno de los probandos que componen el estudio.



**Objetivo 1: Secuenciación de exoma de pacientes con diagnóstico de ER.**

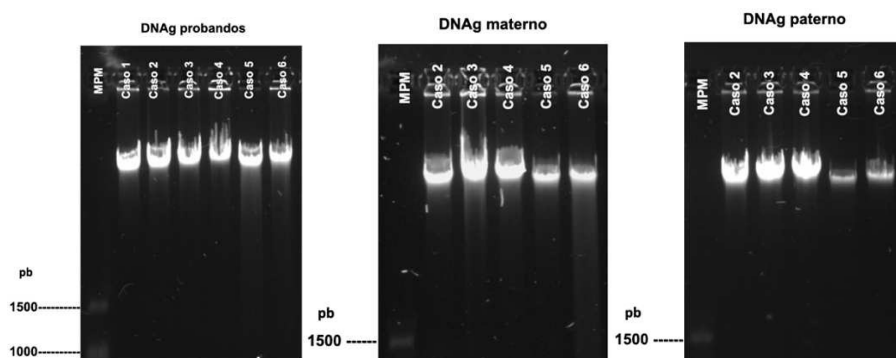
**Comprobación de cantidad, calidad e integridad de las muestras de partida.**

Para evaluar la cantidad y calidad del DNAg de las muestras se utilizó el equipo y protocolo de cuantificación de ácidos nucleicos de NanoDrop 2000. Con este, se verifico que las muestras presentaban una concentración de DNAg suficiente para llevar acabo la preparación de bibliotecas, además se verificó la calidad de la extracción por medio de la relación de absorbancia 260/280, en donde una relación de 1.8 a 2.0 es aceptada generalmente como una extracción con pureza óptima.

<b>Tabla 3. Cuantificación de muestras por NanoDrop.</b>			
<b>Caso</b>	<b>Muestras</b>	<b>Concentración ng/μl</b>	<b>260/280</b>
<b>1</b>	<b>LCM121*</b>	475.6	1.89
<b>2</b>	<b>LCM122</b>	440.7	1.89
	<b>LCM123</b>	394.4	1.89
	<b>LCM124*</b>	476.2	1.89
<b>3</b>	<b>LCM127</b>	860.5	1.86
	<b>LCM128</b>	540.9	1.9
	<b>LCM129*</b>	396.6	1.9
<b>4</b>	<b>LCM119*</b>	349.2	1.89
	<b>LCM125</b>	614.2	1.9
	<b>LCM 126</b>	551.2	1.9
<b>5</b>	<b>LCM13*</b>	434.8	1.9
	<b>LCM14</b>	106.2	1.87
	<b>LCM15</b>	144.2	1.89
<b>6</b>	<b>LCM107</b>	154.3	1.89
	<b>LCM108*</b>	99.1	1.9
	<b>LCM109</b>	137.7	1.87
<b>Requerido</b>		<b>Mínimo 50 ng</b>	<b>1.8-2.0</b>

Como resultado se obtuvieron concentraciones de DNAg que van de 99.1 a 860.5 ng/ $\mu$ l con valores 260/280 de 1.87 a 1.9. Los valores pueden ser observados en la tabla 3.

Para evaluar la calidad e integridad de las muestras se hizo uso de la técnica de electroforesis en gel de agarosa. Estas, fueron distribuidas, cargadas y corridas en tres geles de agarosa al 1% (imagen 1). En las imágenes se puede apreciar el patrón de migración electroforético de las muestras, en donde la mayor parte del DNAg se encuentra en la zona de alto peso molecular cercano al pozo de carga, además se observó que no había degradación de ácidos nucleicos que comprometa la muestra.



**Imagen 1. Calidad e integridad del DNAg.** Geles de agarosa al 1% que contiene las muestras de DNAg del caso 1 (probando LCM121) y de los tríos del caso 2 (probando LCM124, papá LCM123, mamá LCM LCM122) caso 3 (probando LCM129, papá LCM128 y mamá LCM127), cas 4 (probando LCM 119, papá LCM Y mamá LCM), caso 5 (probando LCM13, papá LCM114y mamá LCM115 y caso 6 (probando LCM108, papá LCM107 y mamá LCM109). MPM: marcador de peso molecular.

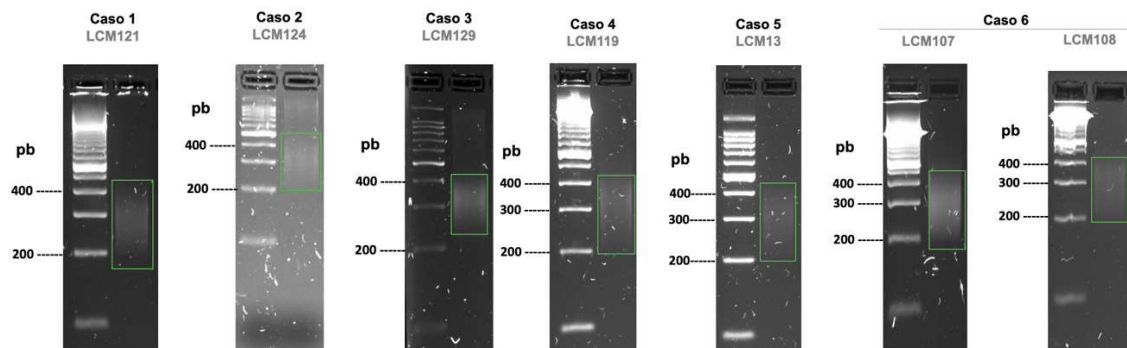
Con los resultados obtenidos en ambos ensayos se comprobó que todas las muestras tenían la cantidad, calidad e integridad suficiente para realizar los siguientes experimentos. Además, estos datos se utilizaron para normalizar el DNAg

de las muestras, utilizando como diluyente agua libre de nucleasas. Obteniendo un volumen final de 10µl a una concentración de 5 ng/µl (50 ng total).

### Generación de bibliotecas

Una vez confirmada la cantidad, calidad e integridad del DNAG de las muestras, se procedió a preparar las bibliotecas. El protocolo utilizado como guía para la generación de bibliotecas (TruSeq Rapid Exome Library Prep, Illumina) esta optimizado para obtener fragmentos de DNA hibrido con un tamaño de 200 a 500 pb, pudiéndose identificar en forma de un ligero barrido. Debido a esta cualidad se hizo uso de la técnica de electroforesis en gel de agarosa.

En el ensayo las bibliotecas fueron cargadas y corridas en geles de agarosa al 4% (imagen 2) además se utilizó como referencia un marcador de peso molecular (100 pb ladder, Promega). Como resultado, al comparar las bandas del marcador con la biblioteca se obtuvieron fragmentos de DNA hibrido de aproximadamente 300 a 350 pb, este resultado está acorde con el protocolo antes mencionado.



**Imagen 2. Bibliotecas.** Geles de agarosa al 4% que contiene las bibliotecas de los probandos así como la biblioteca del padre del caso 6. **MPM:** marcador de peso molecular (100pb Ladder, Promega). **PM:** marcador de peso molecular (100 bp DNA Ladder, invitrogen).

Después, se llevó a cabo la comprobación de las bibliotecas utilizando el ensayo Quant-i™ dsDNA HS en el fluorómetro Qubit®. Como resultado se obtuvieron valores en un rango de 14.0 a 25.1 ng/μl, mismos que están por encima del valor mínimo requerido para el ensayo (3.5 ng/μl) según el protocolo. Estos resultados junto con el tamaño de biblioteca obtenido mediante electroforesis en gel de agarosa sirvieron para calcular la concentración en nM de biblioteca generada. El cálculo se realizó utilizando la ecuación 1, los valores obtenidos se muestran en la tabla 4.

<b>Tabla 4. Comprobación de bibliotecas enriquecidas.</b>				
<b>CASO</b>	<b>ID</b>	<b>Concentración Qubit (ng/μl)</b>	<b>Tamaño de biblioteca (pb)</b>	<b>Concentración (nM)</b>
1	LCM121	24.02	300	121.31
2	LCM124	18.1	300	91.41
3	LCM129	14.0	350	60.60
4	LCM119	25.1	350	108.65
5	LCM13	14.7	300	74.24
6	LCM107	22.5	300	113.6
	LCM108	15.7	300	79.29
<b>Requerido</b>		<b>Mínimo 3.5 ng/μl</b>		

En seguida, las bibliotecas fueron normalizadas a una concentración de 4nM y se procedió a realizar el protocolo de desnaturalización y dilución de las librerías utilizando como control Phix. Al final, se obtuvieron 7 bibliotecas en un volumen de 600 μl a una concentración de 12.5 pM.

## Secuenciación y análisis de Secuencia

Para llevar a cabo la secuenciación, las 7 bibliotecas fueron cargadas de forma independiente en cartuchos de secuenciación v3 de 150 ciclos. La secuenciación se llevó a cabo en la plataforma MiSeq (Illumina). Las secuencias fueron alineadas con el programa de alineación y detección de variantes Enrichment 3.1.0., este asigna las lecturas utilizando el alineador iSAAC-03.16.02.20. Además, se realizaron las llamadas de variantes utilizando Starling Variant Caller 2.4.6 para SNV, Manta 0.28.0 para SV y Canvas 1.13.0 para CNV. Como motor de alineación se utilizó Illumina Annotation Engine 1.3.5.2 y para las métricas de alineación se utilizó el programa PUMA (Tabla 5).

Tabla 5. Versiones de software.	
Aplicación	Programas
Alineación	Enrichment 3.1.0
	iSAAC-03.16.02.20
Llama SNV	Starling Variant Caller 2.4.6
Llama SV	Manta 0.28.0
Llama CNV	Canvas 1.13.0
Motor de anotación	Illumina Annotation Engine 1.3.5.2
Cálculo de alineación	PUMA 00.15.11.02
Cálculo de métricas	PUMA Metrics 1.0.13.0

El reporte de enriquecimiento de secuencia reveló que el promedio del total de lecturas que pasaron el filtro de calidad fueron 46,644,476 millones de lecturas, en donde la biblioteca de LCM124 resulto ser el valor más bajo con un total de 24,206,648 millones de lecturas, en contraparte la biblioteca de LCM108 proporciono el mayor número de lecturas con 55,638,610 millones, así mismo, se obtuvo el promedio de lecturas únicas que pasaron el filtro, que fueron alineadas y

no se marcaron como duplicados, resultando en un promedio de 41,245,386 millones de lecturas, con valores de lecturas máximas de 50,493,624 y mínimos de 22,016,102 . En cuanto a la precisión de la llamada base, medida por el puntaje de calidad de Phred o puntuación Q, se encontró que, en todos los casos el porcentaje fue superior al 96%, es decir el 96% de bases secuenciadas se encontraban por encima del nivel Q30. Además, el promedio de la profundidad media de la cobertura objetivo resulto en 37.05X (Tabla 6).

Tabla 6. Información de muestra.					
Caso	Muestra	Total de lecturas que PF	Lecturas únicas que PF	Porcentaje Q30	Profundidad media
1	LCM121	53,945,724	47,935,396	96.10%	41.8
2	LCM124	24,206,648	22,016,102	98.69%	20.5
3	LCM129	52,443,984	45,397,702	96.05%	41.5
4	LCM119	49,752,554	42,770,002	96.06%	40.7
5	LCM13	43,879,336	38,859,492	97.07%	35
6	LCM108	55,638,610	50,493,624	96.30%	42.8
	<b>Promedio</b>	46,644,476	41,245,386	96.71%	37.05

**Total de lecturas que PF:** Número de lecturas que pasan el filtro para la muestra. **Lecturas únicas que PF:** Número de lecturas únicas que pasaron el filtro para la muestra y no se marcan como duplicados. **Porcentaje Q30:** Porcentaje de bases con una puntuación de calidad 30 o superior. **Profundidad media de la cobertura objetivo:** Número total de bases objetivo dividido por el tamaño de la región objetivo.

## Resumen de enriquecimiento

El análisis de enriquecimiento a nivel de lecturas reveló que en promedio el número total de lecturas alineadas (número total de lecturas que pasan el filtro presente en el conjunto de datos que se alinearon con el genoma de referencia) fue 40,116,953 lecturas con un porcentaje promedio de 94.33% lecturas. Además, se obtuvieron en promedio 26,690,143 lecturas alineadas con la secuencia blanco (número de

lecturas que se alinearon con el objetivo) con un porcentaje de enriquecimiento del 66.77% (tabla 7).

Tabla 7. Enriquecimiento a nivel de lecturas.					
Caso	Muestra	Total de lecturas alineadas	Porcentaje de lecturas alineadas	Lecturas blanco alineadas	Porcentaje de enriquecimiento
1	LCM121	44,835,802	93.53%	28,842,931	64.33%
2	LCM124	20,572,167	93.44%	14,120,779	68.64%
3	LCM129	52,346,114	99.80%	35,824,263	68.40%
4	LCM119	39,507,932	92.37%	27,720,858	70.17%
5	LCM13	36,463,105	93.83%	24,271,478	66.56%
6	LCM108	46,976,600	93.03%	29,360,551	62.50%
	<b>Promedio</b>	40,116,953	94.33%	26,690,143	66.77%

**Estadística a nivel de enriquecimiento de lecturas:** **Total de lecturas alineadas:** Número total de lecturas que pasaron el filtro y fueron alineadas con el genoma de referencia. **Porcentaje de lecturas alineadas:** Porcentaje de lecturas que pasan el filtro que se alinearon y no se marcaron como duplicados. **Lecturas blanco alineadas:** Número de lecturas que se alinearon con el objetivo y no se marcaron como duplicados. **Porcentaje de enriquecimiento:**  $100 * (\text{lecturas blanco alineadas} / \text{total de lecturas alineadas})$ .

El análisis de enriquecimiento a nivel de bases reveló un promedio de 2,345,742,672 bases alineadas con el genoma de referencia, con un porcentaje promedio de 94.23% de bases alineadas. Además, de este conjunto de datos en promedio se alinearon 1,370,539,754 bases con la región objetivo. Por último, se obtuvo el promedio del porcentaje de enriquecimiento, calculado a partir del total de bases alineadas en las regiones blanco / total de bases alineadas, resultando en 60.31% de enriquecimiento.

Para el cálculo de enriquecimiento a nivel de objetivo se consideró el número de regiones específicas proporcionadas por el formato TruSeq\_Exome\_TargetedRegions\_v1.2.txt. el cual toma en cuenta una longitud de

referencia de 214,126 pb. El promedio del número de regiones específicas cubiertas de extremo a extremo fue de 199,832 lo que representa en promedio el 93.33% de regiones específicas cubiertas de extremo a extremo.

Caso	Muestra	Total de bases alineadas	Porcentaje de bases alineadas	Bases blanco alineadas	Porcentaje de enriquecimiento
1	LCM121	3,303,276,246	93.35%	1,894,851,329	57.36%
2	LCM124	1,543,599,880	93.32%	927,957,500	60.12%
3	LCM129	52,346,114	99.80%	35,824,263	68.40%
4	LCM119	2,943,323,667	92.28%	1,842,569,084	62.60%
5	LCM13	2,728,537,638	93.68%	1,585,261,702	58.10%
6	LCM108	3,503,372,489	92.96%	1,936,774,643	55.28%
	<b>Promedio</b>	2,345,742,672	94.23%	1,370,539,754	60.31%

**Estadística a nivel de enriquecimiento de base.** *Total de bases alineadas: número total de bases presentes en el conjunto de datos que se alinearon con el genoma de referencia. **Porcentaje de bases alineadas:** Porcentaje de bases alineadas de lecturas no marcadas como duplicados. **Bases blanco alineadas:** Total de bases alineadas en la región objetivo no marcadas como duplicadas. **Porcentaje de enriquecimiento de bases:**  $100 * (\text{total de bases alineadas dirigidas} / \text{total de bases alineadas})$ .*

Caso	Muestra	Número de regiones específicas cubiertas de extremo a extremo	Porcentaje de regiones específicas cubiertas de extremo a extremo
1	LCM121	201,302	94.01%
2	LCM124	187,802	87.71%
3	LCM129	204,508	95.51%
4	LCM119	200,605	93.69%
5	LCM13	202,568	94.60%
6	LCM108	202,209	94.43%
	<b>Promedio</b>	199,832	93.33%

**Número de regiones específicas cubiertas de extremo a extremo:** Número de regiones objetivo que tienen al menos 1X de cobertura en todas las posiciones de referencia que no son N. **Porcentaje de regiones específicas cubiertas de extremo a extremo:** Porcentaje de regiones específicas que tienen al menos 1X de cobertura en todas las posiciones de referencia que no son N.



El análisis de variantes arrojó un promedio de 30,295 SNVs presentes en el conjunto de datos, de los cuales el 92.84% fueron identificados previamente en la base de datos de los 1000 genomas (dbSNP). Además, fueron clasificadas como inserciones o deleciones, obteniendo un promedio de 794 inserciones y 949 deleciones (tabla 10).

Tabla 10. Resumen de variantes.							
Caso	Muestra	SNVs	dbSNP	Inserciones	dbSNP inserciones	Deleciones	dbSNP deleciones
1	LCM121	29,915	92.16%	796	93.97%	898	89.42%
2	LCM124	28,171	90.69%	591	94.25%	698	85.10%
3	LCM129	31,236	93.67%	849	93.52%	954	90.57%
4	LCM119	30,770	92.81%	784	93.37%	923	91.55%
5	LCM13	31,192	92.41%	808	94.68%	976	89.75%
6	LCM108	30,295	92.84%	794	93.58%	949	92.20%
	<b>Promedio</b>	30,295	92.84%	794	93.58%	949	92.20%

**Estadística: SNVs:** Número total de SNVs presentes en el conjunto de datos que pasaron los filtros de calidad. **dbSNP SNVs:** Porcentaje de SNVs encontrado en la base de datos dbSNP calculado a partir de la fórmula  $100 \times (\text{número de variantes en dbSNP} / \text{Número de SNV})$ . **Inserciones:** Número total de inserciones presentes en el conjunto de datos que pasaron los filtros de calidad. **dbSNP inserciones:** Porcentaje de inserciones encontrado en dbSNP calculado a partir de la ecuación  $100 \times (\text{número de inserciones en dbSNP} / \text{Número de inserciones})$ . **Deleciones:** Número total de deleciones presentes en el conjunto de datos que pasan los filtros de calidad. **dbSNP deleciones:** Porcentaje de deleciones encontrado en dbSNP calculado a partir de la ecuación  $100 \times (\text{número de deleciones en dbSNP} / \text{Número de deleciones})$ .

### Variantes por contexto de secuencia

Además de identificar las variantes estas fueron ubicadas dentro de la secuencia genómica utilizando conceptos como gen, exón, región codificante, región UTR y sitio de empalme. Los resultados indicaron que la mayor cantidad de SNV estaban ubicadas dentro de genes (28,145), en cuanto a las regiones codificantes que son

de interés para esta tesis se encontró en promedio 19,684 SNVs, 195 inserciones y 232 deleciones (tabla 11).

### **Variantes por consecuencia**

De igual forma el conjunto de datos fueron analizados acorde a la consecuencia tomando en cuenta conceptos como: cambio del marco de lectura, variantes sinónimas o no sinónimas, así como ganancia o pérdida de codones de terminación. En cuanto al fenómeno de cambio del marco de lectura, se encontró que era provocado de igual forma por inserciones y deleciones. Para las variantes sinónimos en promedio se encontraron 10,293 variantes producidas por SNVs, para las no sinónimas, se encontró que las SNVs (9,298) generan más variantes sinónimas respecto a las inserciones (86.5) y las deleciones (122.3). En cuanto a la generación de codones de terminación en las 6 bibliotecas hubo mayor ganancia de estos provocados por SNVs (74.5) respecto a inserciones y deleciones (tabla 12)

Tabla 11. Variantes por contexto de secuencia.

Número de variantes ubicadas dentro de:																
Caso	Muestra	Gen			Exon			Región de codificación			UTR			Sitio de empalme		
		SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones
1	LCM121	27,819	725	820	20,585	258	295	19,427	204	238	1,158	54	57	1,166	53	70
2	LCM124	26,253	544	639	19,579	204	257	18,616	158	213	963	46	44	1,017	35	44
3	LCM129	28,967	774	869	21,367	289	295	20,198	222	229	1,169	67	66	1,205	58	72
4	LCM119	28,643	717	840	21,145	257	299	19,969	195	237	1,176	62	62	1,188	44	74
5	LCM13	29,011	742	889	21,398	254	316	20,234	190	242	1,164	64	74	1,225	50	69
6	LCM108	28,177	725	865	20,760	255	290	19,662	201	233	1,098	54	57	1,186	52	72
	Promedio	28,145	704.5	820.3	20,806	252.8	252.8	19,684	195	232	1,121	57.8	60	1,165	48.7	66.8

**Gen:** Número de variantes que caen en un gen (incluye exones, intrones y regiones UTR). **Exones:** Número de variantes que caen en un exón (exones incluyen regiones de codificación y UTR). **Región de codificación:** Número de variantes que caen en una región de codificación. **UTR:** Número de variantes que caen en una región no traducida (UTR, incluyen las regiones 5' y 3' UTR). **Sitio de empalme:** Número de variantes que caen en una región de sitio de empalme 8 incluyen regiones anotadas como aceptador de empalme, donante de empalme, sitio de empalme o región de empalme).

Tabla 12. Variantes por consecuencia.																
Caso	Muestra	Cambio del marco			Sinónimo			No sinónimo			Codones de terminación					
		SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones	Ganancia			Perdida		
											SNVs	Inserciones	Delecciones	SNVs	Inserciones	Delecciones
1	LCM121	-	115	103	10,215	-	-	9,124	89	131	70	3	0	18	0	1
2	LCM124	-	88	119	9,586	-	-	8,923	70	92	84	1	0	23	0	1
3	LCM129	-	119	103	10,588	-	-	9,530	103	124	56	3	0	24	0	1
4	LCM119	-	103	101	10,436	-	-	9,421	92	134	90	3	0	22	0	1
5	LCM13	-	109	108	10,642	-	-	9,491	81	132	79	4	0	23	0	2
6	LCM108	-	117	109	10,291	-	-	9,280	84	121	68	6	0	23	0	1
	Promedio	-	108.5	107.2	10,293	-	-	9,298	86.5	122.3	74.5	3.3	0	22.2	0	1.2

**Cambio del marco:** Número de variantes que causan un cambio de marco de lectura. **Sinónimo:** Número de variantes que están dentro de una región de codificación, pero que no causan un cambio de aminoácidos. **No sinónimo:** Número de variantes que causan un cambio de aminoácido en una región de codificación. **Ganancia de codones de terminación:** Número de variantes que causan un codón de parada adicional. **Perdida de codones de terminación:** Número de variantes que causan la pérdida de un codón de parada.

En cuanto a la cobertura, el promedio de la profundidad media con la región blanco fue de 37.05X con un promedio de uniformidad de cobertura de 87.49%.







Tabla 13. Resumen de cobertura			
Caso	Muestra	Profundidad media de la cobertura blanco	Uniformidad de cobertura (Pct> 0.2 * media)
1	LCM121	41.8 X	86.92%
2	LCM124	20.5 X	82.87%
3	LCM129	41.5 X	89.68%
4	LCM119	40.7 X	86.99%
5	LCM13	35 X	90.80%
6	LCM108	42.8 X	87.68%
	<b>Promedio</b>	37.05 X	87.49%

**Profundidad media de la cobertura del objetivo:** Número total de bases alineadas a la región objetivo dividido por el tamaño de la región objetivo. **Uniformidad de cobertura (Pct> 0.2 \* media).** Porcentaje de posiciones base orientadas en las que la profundidad de lecturas es superior a 0.2 veces la profundidad media de cobertura objetivo de la región.

**Objetivo 2: Análisis integral de los datos e identificación de variantes que puedan explicar el fenotipo clínico.**

Para la identificación de variantes se utilizó el software Variant Interpreter. En la plataforma se programó un filtro básico, que fue aplicado a todos los metadatos de los seis casos. El filtro consistió en seleccionar variantes que presentaron predicción patogénica, probablemente patogénica, con significancia clínica incierta, que afectara la codificación en sus diferentes formas, con una frecuencia poblacional <0.01, con preferencia por variantes homocigotas y/o heterocigotas y que además cumplieran con las métricas de control de calidad establecidas.

**Tabla 14. Aplicación de filtros.**

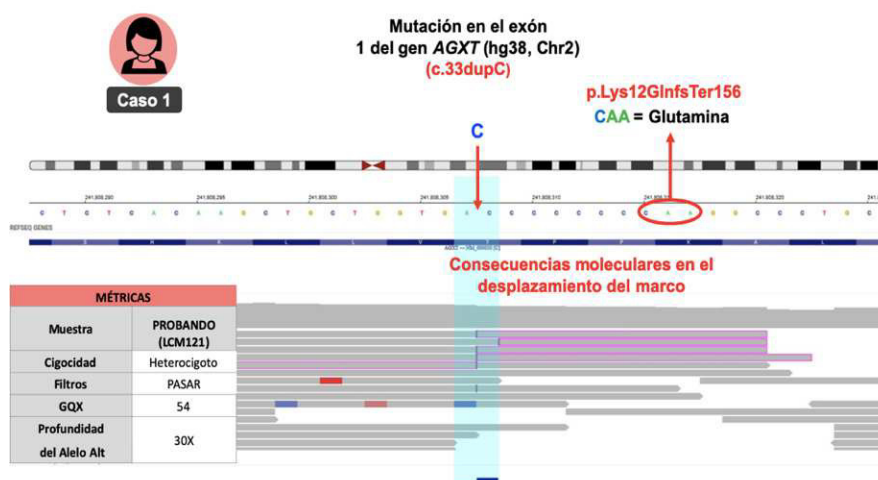
Casos	 Nefropatía	 FATCO	 FATCO	 Mitocondrial	 SD / S. de Marfan	 TEA
<b>Variantes filtradas</b>	567	833	553	549	595	430
<b>Genes afectados</b>	431	765	466	499	533	265
<b>Vías implicadas</b>	240	313	129	222	210	630
<b>Variantes filtradas</b>	15	53	26	12	6	17
<b>Genes candidatos</b>	1	5	4	3	1	10

**Caso 1**

Como resultado, para el conjunto de datos del caso 1 se obtuvieron un total de 567 variantes contenidas en 431 genes, implicados en 240 vías (tabla 14). Después de realizar la priorización de genes con base en el fenotipo clínico del probando, se identificó la variante heterocigota c.33dupC ubicada en el gen **AGXT**. Esta variante

ha sido previamente reportada en la base de datos de los 1000 genomas (dbSNP, rs180177201) como una mutación patogénica. La mutación, implica la duplicación de una citosina en el codón 33 a nivel de DNAg, generando consecuencias moleculares en el desplazamiento del marco de lectura. Este cambio implica a nivel de proteína el intercambio del aminoácido glutamina por lisina en la posición 12 y la generación de un codón de terminación en la posición 156. La variante tuvo una calidad de genotipo (GQX) = 54, profundidad de variante alternativa = 30X y una profundidad de lectura total = 94 puntuaciones (Fig. 3).

La mutación fue analizada en diferentes plataformas bioinformáticas, encontrando que esta posee una frecuencia alélica 0.0004656 (48/103082 alelos) en la base de datos ExAC y una frecuencia de 0.0001588 (43/270822) en gmoAD, sin reportes de homocigotos en ambas bases de datos.



**Figura 3. Alineamientos de lecturas gen AGXT.** Visualización de los alineamientos de las lecturas que se utilizaron para llamar la variante c.33dupC ubicada en el gen AGXT utilizando como referencia el genoma hg38. Métricas: Métricas de calidad obtenidas al llamar la variante.

### Caso 2 y 3

Respecto al caso 2 y 3, tras la aplicación de los filtros básicos se identificaron un total de 833 y 553 variantes, llegando a afectar 765 y 466 genes, implicados en 313 y 129 vías, respectivamente. Al realizar la priorización de genes con base en el fenotipo clínico de ambos probandos, se generó una lista de 53 y 26 genes candidatos, a su vez, estos fueron analizados detalladamente utilizando como referencia la bibliografía medica disponible, encontrando en cada caso un total de 5 genes potencialmente responsables de generar el fenotipo según la actividad y vías en las que intervienen. Para el caso 2, se destacan las variantes ubicadas en los genes **HOXD12**, **HOXC9**, **HOXD3**, **FGFR4** y **PLXNB1**. Para el caso 3, los genes **HOXC12**, **TBX6**, **BMP15**, **GDF5** Y **WNT3A** (tabla 15). Sin embargo, dentro de los metadatos tratados con el filtro básico no se encontraron variantes y/o genes compartidos entre los probandos que respondan por las manifestaciones clínicas exhibidas.

Las variantes fueron analizadas utilizando diferentes plataformas bioinformáticas, encontrando que solo 6 han sido reportadas en dbSNP y presentan frecuencia poblacional <0.01 según ExAC. En cuanto a la patogenicidad, Polyphen las clasifica como probablemente dañinas con excepción de la variante de **BMP15** (chrX:50658861), mientras que SIFT las predice como deletéreas con excepción de las variantes ubicada en el gen **GDF5** (chr20:34025116) y **HOXD12** (chr2:176964765).



Tabla 15. Identificación de genes y variantes, caso 2 y 3.

Caso	Gen	Nombre	Función	Variante	Vía
2	<b>FGFR4</b>	Receptor 4 del factor de crecimiento de fibroblastos	Receptor de superficie celular para factores de crecimiento de fibroblastos. Regula la proliferación, diferenciación y migración celular. <sup>53</sup>	Cambio de sentido c.1287C>A p.(Ser429Arg)	Vía de señalización de FGF
	<b>HOXC9</b>	Proteína homeobox Hox-C9	Factor de transcripción específico de secuencia que forma parte de un sistema que regula el desarrollo y proporciona a las células identidad posicional con especificidad por en el eje antero-posterior.	Cambio de sentido c.596A>G p.(Tyr199Cys)	Activación de genes HOX durante la diferenciación.
	<b>HOXD12</b>	Proteína homeobox Hox-D12	Regula la morfogénesis de las extremidades proporcionando a las células identidades posicionales específicas en el eje anterior-posterior.	Cambio de sentido c.236G>A p.(Gly79Asp)	Activación de genes HOX durante la diferenciación.
	<b>HOXD3</b>	Proteína homeobox Hox-D3	Regula la morfogénesis de las extremidades proporcionando a las células identidades posicionales específicas en el eje anterior-posterior.	Cambio de sentido c.257T>C p.(Leu86Pro)	Activación de genes HOX durante la diferenciación.
	<b>PLXNB1</b>	Plexin-B1	Receptor para SEMA4D. Juega un papel en la guía de axones, la organogénesis, la migración celular y la progresión tumoral. <sup>54</sup>	Cambio de sentido c.4880G>A p.(Arg1627His)	Inhibición de unión celular y migración mediada por Sema4D
3	<b>GDF5 (BMP14)</b>	Factor de diferenciación del crecimiento 5	Involucrado en la diferenciación del tejido condrogénico por dos vías: <b>Regulación positiva-</b> diferenciación del tejido condrogénico a través de la unión con BMPR1B y BMPR1A, lo que conduce a la inducción por fosforilación de SMAD1-SMAD5-SMAD8 y posterior transducción de señales. <sup>55</sup> <b>Regulación negativa-</b> diferenciación condrogénica a través de la interacción con NOG. <sup>56</sup>	Cambio de sentido c.593T>A p.(Leu198Gln)  Cambio de sentido c.536C>T p.(Thr179Met)	Vía de señalización BMP
	<b>BMP15</b>	Proteína morfogenética ósea 15	Involucrado en procesos de desarrollo celular, transducción de señales a través de la vía SMAD y regulación de la cascada MAPK.	Cambio de sentido c.433C>T p.(Arg145Cys)	Vía de señalización BMP
	<b>HOXC12</b>	Proteína homeobox Hox-C12	Factor de transcripción que proporciona a las células identidades posicionales específicas en el eje anterior-posterior.	Cambio de sentido c.139C>T p.(Arg47Cys)	Activación de genes HOX durante la diferenciación.
	<b>TBX6</b>	Factor de transcripción TBX6	Factor de transcripción que desempeña un papel esencial en la morfogénesis y desarrollo del mesodermo.	Cambio de sentido c.484G>A p.(Gly162Ser)	Factor de transcripción
	<b>WNT3A</b>	Proteína Wnt-3a	Desempeña funciones pleiotrópicas en la regulación de la diferenciación celular durante la osteogénesis en mamíferos. <sup>57</sup>	Cambio de sentido c.860C>G p.(Thr287Arg)	Vías de señalización Wnt

**Tabla 16. Frecuencia poblacional y predicción de la patogenicidad in silico, caso 2 y 3.**

Caso	Gen	Coordenada genómica	ExAC	dbSNP	Polyphen	SIFT
2	<i>HOXD12</i>	NM_021193.3 chr2:176964765	-	-	Probablemente dañino (0.915)	Tolerado (0.41)
	<i>HOXD3</i>	NM_006898.4 chr2:177034099	-	-	Probablemente dañino (0.98)	Deletéreo (0.02)
	<i>FGFR4</i>	NM_002011.3 chr5:176520442	0.0005126 (62/120950)	rs75935301	Posiblemente dañino (0.8)	Deletéreo (0.02)
	<i>HOXC9</i>	NM_006897.1 chr12:54396271	-	-	Probablemente dañino (1)	Deletéreo
	<i>PLXNB1</i>	NM_001130082.1. chr3:48454004	8.314e-06 (1/120272)	rs746314397	Probablemente dañino (1)	Deletéreo (0.02)
3	<i>HOXC12</i>	NM_173860.1 chr12:54348852	8.923e-06 (1/112068)	rs745654830	Probablemente dañino (0.985)	Deletéreo (0.01)
	<i>TBX6</i>	NM_004608.3. chr16:30100401	0.00455 (546/119996)	rs56098093	Probablemente dañino (1)	Deletéreo (0.02)
	<i>BMP15</i>	NM_005448.2 chrX:50658861	3.441e-05 (3/87173)	rs782369768	Benigno (0.395)	Deletéreo (0.01)
	<i>GDF5 (BMP14)</i>	NM_000557.2. chr20:34025116.	-	-	Probablemente dañino (0.995)	Tolerado (0.28)
	<i>GDF5 (BMP14)</i>	NM_000557.2 chr20:34025173	8.274e-06 (1/120858)	rs371529960	Probablemente dañino (0.995)	Deletéreo
	<i>WNT3A</i>	NM_033131.3. chr1:228246967	-	-	Probablemente dañino (0.99)	Deletéreo (0.02)

#### Caso 4

Para el caso 4, se obtuvo un total de 549 variantes que afectan a 499 genes implicados en 222 vías (tabla 14), tras realizar la priorización de genes con base en el fenotipo de enfermedad mitocondrial exhibido por el probando, se produjo una lista de 12 genes candidatos, a su vez estos fueron analizados utilizando como apoyo la bibliografía medica disponible encontrando un total de 2 genes potencialmente relacionados con el fenotipo clínico (tabla 17). El gen **NDUFS7** que participa en el ciclo del ácido cítrico y cadena de transporte de electrones y el gen **PRKAG2** que participa en la activación de AMPK.

Tabla 17. Identificación de genes y variantes, caso 4.			
Gen / Nombre	Función	Variante	Vía
<p><b>NDUFS7 / PSST</b></p> <p><i>NADH - deshidrogenasa ubiquinona proteína Fe-S 7</i></p>	<p>Codifica para una proteína que es parte del brazo periférico del complejo I (NADH: ubiquinona oxidoreductasa) de la cadena de transporte de electrones. Este complejo participa en la transferencia de electrones de NADH a la cadena respiratoria, y se cree que la ubiquinona es el receptor de electrones inmediato para la enzima.</p>	<p>Cambio de sentido c.434G&gt;A p.(Arg145His)</p>	<p>Biogénesis del complejo I de la cadena de transporte de electrones</p>
<p><b>PRKAG2</b></p> <p>Subunidad gamma-2 de la proteína quinasa activada por 5'-AMP</p>	<p>Codifica para la subunidad gamma reguladora no catalítica de la proteína quinasa activada por AMP (AMPK). Está, es una importante enzima censora de energía que controla el estado y las funciones de la energía celular al inactivar las enzimas clave involucradas en la regulación de la biosíntesis de novo de ácidos grasos y colesterol.<sup>58</sup></p>	<p>Cambio de sentido c.1040A&gt;G p.(Glu347Gly)</p>	<p>Activación de PPARGC1A (PGC-1alpha) por fosforilación</p>

De las dos variantes seleccionadas solo **PRKAG2** (chr16:23546355) presenta frecuencia poblacional reportada en ExAC <0.01. Las dos variantes no poseen número de identificación en dbSNP. Polyphen reporta a **NDUFS7** como probablemente dañina, mientras que SIFT las detecta como deletéreas.

Tabla 18. Frecuencia poblacional y predicción de la patogenicidad in silico, caso 4.					
Gen	Coordenada genómica	ExAC	dbSNP	Polyphen	SIFT
<b>PRKAG2</b>	NM_016203.3 chr7:151269761	8.26e-06 (1 / 121064)	-	Benigno (0.07)	Deletéreo (0.02)
<b>NDUFS7</b>	NM_001542.3 chr1:117146504	-	-	Probablemente dañino (0.99)	Deletéreo (0)

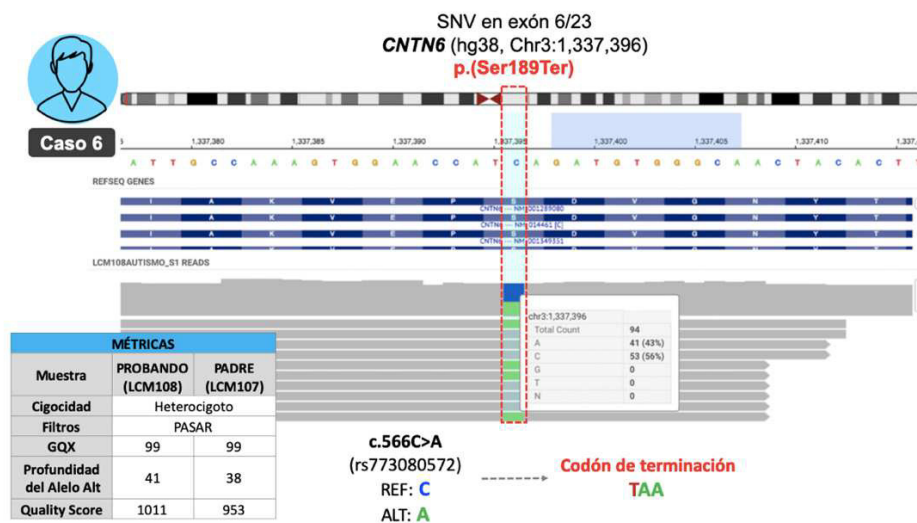
## Caso 5

Después de aplicar el filtro a los metadatos del caso 5 se obtuvieron un total de 595 variantes afectando a 533 genes implicados en 240 vías. Tras realizar la priorización de genes con base en el diagnóstico presuntivo de síndrome de Marfan exhibido por el probando, se identificó un grupo de genes involucrados en la biosíntesis, organización y degradación del tejido conectivo que podrían responder al fenotipo (tabla 19), sin embargo, el gen **FBN3** resulta un candidato interesante al proporcionar instrucciones para codificar a una proteína que pertenece a la familia de fibrilinas.

Tabla 19. Genes involucrados en la biosíntesis organización y degradación de tejido conectivo.	
Gen	Nombre de la vía (Reactome)
<b>COL18A1;COL11A1</b>	Biosíntesis de colágeno y enzimas modificadoras. Trimerización de la cadena de colágeno. Degradación del colágeno. Montaje de fibrillas de colágeno Organización de la matriz extracelular. Degradación de la matriz extracelular.
<b>GUSB</b>	Mucopolisacaridosis, captación y degradación del hialuronano.
<b>SCUBE3;CAPN2;NID1</b>	Degradación de la matriz extracelular.
<b>FLNA</b>	Interacciones de la matriz extracelular de la célula.
<b>SCUBE3;CAPN2;NID1</b>	Degradación de la matriz extracelular.
<b>SCUBE3;VWF;CAPN2;NID1;DSPP</b>	Organización de la matriz extracelular.

## Caso 6

De igual forma, a los metadatos del caso 6 se le aplicaron los filtros básicos. El filtro arrojó 430 variantes contenidas en 265 genes, implicados en 630 vías. Tras realizar la priorización de genes con base en el fenotipo clínico de TEA en el probando, se identificó una mutación heterocigota en el exón 6 del gen **CNTN6** cuya consecuencia implica la ganancia de un codón de terminación (c.566C> A, p.Ser189Ter; calidad de genotipo (GQX) = 99, profundidad de variante alternativa = 41 y profundidad de lectura total = 94 puntuaciones) (Fig. 4).



**Figura 4. Visualización de alineamientos gen CNTN6.** Visualización de los alineamientos de las lecturas que se utilizaron para llamar la variante c.566C> A ubicada en el gen CNTN6 utilizando como referencia el genoma hg38. **Métricas:** Métricas de calidad obtenidas al llamar la variante en el probando y el padre.

Además, se predijo que esta variante interrumpe la transcripción del gen **CNTN6** en el segundo dominio de tipo Ig produciendo una proteína corta con una longitud <200 aminoácidos (189/1028). La SNV ha sido reportada previamente en la base de datos de los 1000 genomas (dbSNP, rs773080572), sin embargo, su frecuencia y

validación aún no se había registrado en otras bases de datos, como ExAC, gnomAD, ClinVar y Exome Variant Server, debido a esto, la SNV fue registrada en la base de datos ClinVar con el número de entrada SCV000786639.1.

**Objetivo 3. Validación de variantes encontradas potencialmente dañinas mediante secuenciación Sanger.**

**Diseño y estandarización de cebadores.**

Una vez seleccionados los genes candidatos se procedió a validar las variantes por secuenciación Sanger. Los ensayos de secuenciación de Sanger se diseñaron para cubrir las regiones genómicas interrogadas. Para llevar a cabo este objetivo, se utilizaron las coordenadas genómicas de las 15 variantes candidatas.

Tabla 20. Diseño de cebadores.					
Caso	Gen	Ensamblaje Coordenada	Cebador 5' - 3'	Cebador 3' - 5'	Amplicón (pb)
2	<i>HOXD12</i>	NM_021193.3 chr2:176964765	tatgtgggctcgcttctgaatctg	gtagtcatactggccgctttgag	387
	<i>HOXD3</i>	NM_006898.4 chr2:177034099	ggaggctatggctacagcaaaact	tgacctccaagggttcattccag	565
	<i>FGFR4</i>	NM_002011.3 chr5:176520442	aggccaggtatacggacatca	gaactcccatagtggtcgaga	370
	<i>HOXC9</i>	NM_006897.1 chr12:54396271	actctgaagctgaggaccagact	ccgacggctccctggttaatacat	392
	<i>PLXNB1</i>	NM_001130082.1 chr3:48454004	gacacctctactcccaacttc	gccgcgtagctgtactgaa	444
3	<i>HOXC12</i>	NM_173860.1 chr12:54348852	tgttatgtcccgttactggtgctg	ttggtgagcttgacacgatgga	365
	<i>TBX6</i>	NM_004608.3. chr16:30100401	gctaaacctctgctctgttcct	agctgagataccatctgcttgcg	342
	<i>BMP15</i>	NM_005448.2 chrX:50658861	gctaaacctctgctctgttcct	agctgagataccatctgcttgcg	534
	<i>GDF5</i>	NM_000557.2 chr20:34025116 chr20:34025173	gatctgtcccagctcctt	ttgaaagcccctccattc	289
	<i>WNT3A</i>	NM_033131.3. chr1:228246967	catcgggtgactcctcaaggacaa	aggagtactgccccgttaggt	481
4	<i>PRKAG2</i>	NM_016203.3 7:151269761	ccgcctcattttctaattgc	catagctttcccgatactaaaaga	217
	<i>NDUFS7</i>	NM_001542.3 1:117146504	ggcacactaccaacaagat	gcactcacgctacgtctgtc	201
5	<i>FBN3</i>	NM_032447.3 19:8175991	gacagggatgaatgtgccgagaa	tcacagctcccaaagcacaga	296
6	<i>CNTN6</i>	NM_014461.2 chr3:1337396	tacgtccaagaggacaataggc	ccagaaaccagatggcttgga	387

Como resultado, se diseñaron 14 juegos de cebadores (tabla 20). El diseño de estos fue elaborado siguiendo las pautas del programa primer-BLAST del Centro Nacional de Información Biotecnológica de la Biblioteca Nacional de Medicina de EE.UU.<sup>59</sup>

La estandarización de los cebadores se llevó a cabo en reacciones de 25  $\mu$ l (Tabla 21) y consistió en la utilización de un programa de PCR cualitativa convencional el cual fue adaptado para cada cebador en el proceso de hibridación, esto según la  $t_m$  proporcionada por el proveedor. Para establecer la temperatura óptima se realizó un gradiente de temperatura tomando como base la  $t_m \pm 1.5^\circ\text{C}$ .

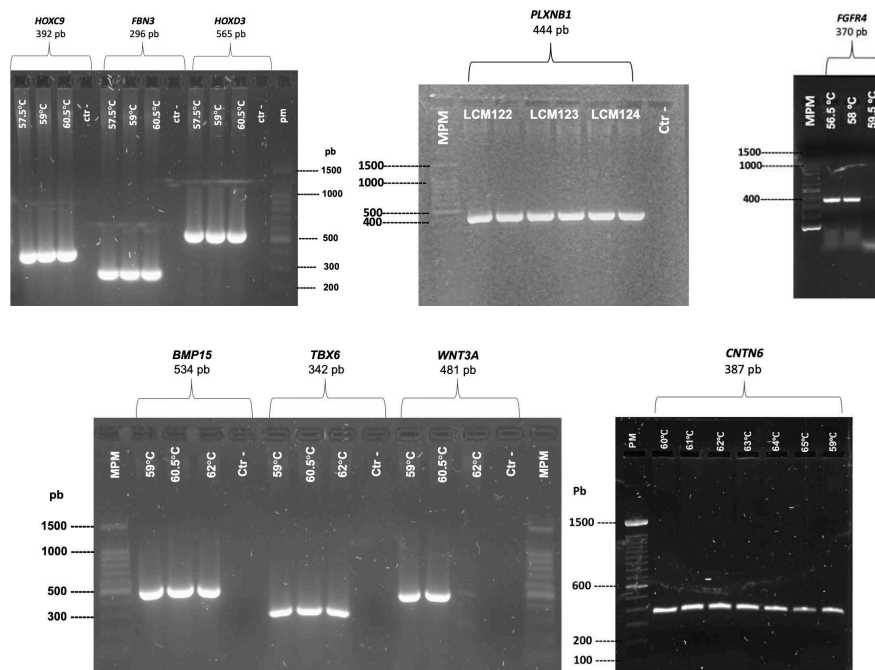
**Tabla 21. Mezcla de reacción para el ensayo de amplificación.**

<b>Reactivo</b>	<b>Ci</b>	<b>Cf</b>	<b>Vol. para una reacción de 25 <math>\mu</math>l</b>
Agua grado biología molecular	-	-	8.5 $\mu$ l
GoTaq® Green Master Mix	2 X	1 X	12.5 $\mu$ l
Iniciador sentido	10 $\mu$ M	150 nM	1 $\mu$ l
Iniciador antisentido	10 $\mu$ M	150 nM	1 $\mu$ l
DNA molde	50 ng/ $\mu$ L	100 ng/ $\mu$ L	2 $\mu$ l
Volumen total	-	-	25 $\mu$ l

Ci = Concentración inicial; Cf = concentración final.



Los productos de amplificación fueron visualizados por medio de electroforesis en gel de agarosa al 1.5% (Fig. 3). Como guía se utilizó el patrón de electroforesis e intensidad de la banda para determinar la temperatura óptima en la cual la reacción es más efectiva. Así mismo, se comprobó que todos los cebadores generaron amplicones acorde al diseño previo. Desafortunadamente, la estandarización de los genes *HOXD12* y *HOXC12* no mostro resultados positivos.



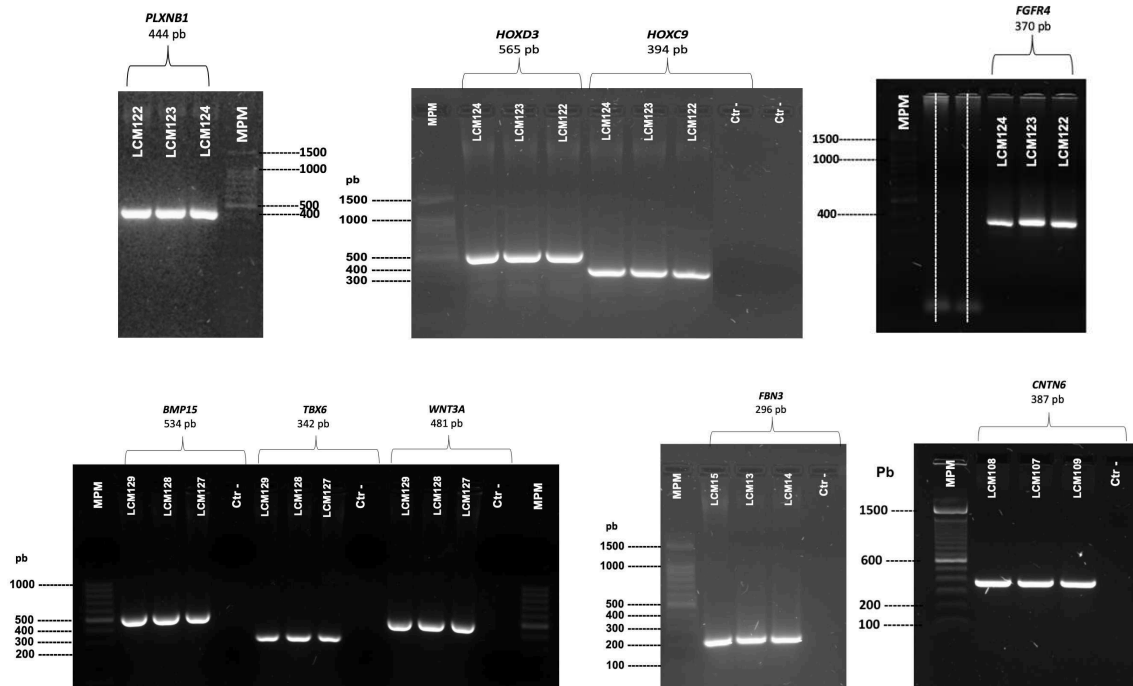
**Imagen 3. Estandarización de cebadores.** Selección de temperatura de hibridación. **MPM:** marcador de peso molecular (100pb Ladder, Promega). **PM:** marcador de peso molecular (100 bp DNA Ladder, invitrogen).

Tabla 22. Condiciones experimentales de PCR cualitativa.										
Caso	Gen	Desnaturalización	Hibridación	Extensión; número de ciclos						
2	<i>HOXD12</i>	94°C, 30 seg	61°C, 45 seg*	72°C, 1 min; 35 ciclos						
	<i>HOXD3</i>		59°C, 45 seg							
	<i>FGFR4</i>		61°C, 45 seg							
	<i>HOXC9</i>		59°C, 45 seg							
	<i>PLXNB1</i>		60°C, 45 seg							
3	<i>HOXC12</i>		94°C, 30 seg		60°C, 45 seg*	72°C, 1 min; 35 ciclos				
	<i>TBX6</i>				60.5°C, 45 seg					
	<i>BMP15</i>				60.5°C, 45 seg					
	<i>GDF5</i>				60°C, 45 seg					
	<i>WNT3A</i>				60.5°C, 45 seg					
4	<i>PRKAG2</i>				94°C, 30 seg		60°C, 45 seg	72°C, 1 min; 35 ciclos		
	<i>NDUFS7</i>						60°C, 44 seg			
5	<i>FBN3</i>						94°C, 30 seg		59°C, 45 seg	72°C, 1 min; 35 ciclos
6	<i>CNTN6</i>								61°C, 45 seg	

\*Temperaturas propuestas por el proveedor.

### Obtención y extracción de amplicones

Una vez establecidas las condiciones experimentales de la PCR cualitativa se llevó a cabo la obtención de los amplicones en un volumen de reacción de 25  $\mu$ L en gel de agarosa al 1.5%. La distribución del gel consistió en cargar un carril con el marcador de peso molecular y tres carriles ocupados por los amplicones de cada una de las muestras pertenecientes al trio (Fig. 4).



**Imagen 4. Obtención de amplicones.** Obtención de amplicones para los genes *HOXC9*, *HOXD3*, *PLXNB1*, *FGFR4* (caso 2), *BMP15*, *TBX6* Y *WNT3A* (caso 3), *FBN3* (caso 5) y *CNTN6* (caso 6).

## Secuenciación Sanger

### Caso 2

Para el caso 2 solo las variantes ubicadas en los genes *FGFR4* (c.1287C>A) y *PLXNB1* (c.4880G>A) fueron validadas mediante secuenciación Sanger. En ambos casos se encontraron padres homocigotos para el alelo silvestre y madres heterocigotas, es decir la mutación fue heredada maternalmente (Fig. 5).

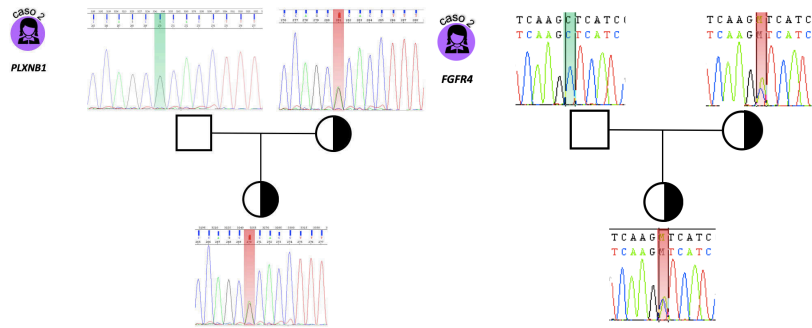


Figura 5. Validación de variantes gen PLXNB1 y FGFR4.

### Caso 3

Para el caso 3, de las 6 candidatas solo se fueron validadas las variantes ubicadas en el gen **TBX6** (c.484G>A) encontrando que la madre es homocigota para el alelo silvestre, mientras que el padre es heterocigoto, al igual que el probando. Para la variante c.433C>T ubicada en el gen **BMP15** la validación confirmo la cigosidad hemicigotica en el probando, detectándolo como hemicigoto para el alelo alternativo, mientras que el padre resulto ser homocigoto para el alelo silvestre y la madre heterocigota. Respecto a la variante c.860C>G ubicada en el gen **WNT3A** sorpresivamente esta no se encontró presente en ninguno de los casos (Fig.6).

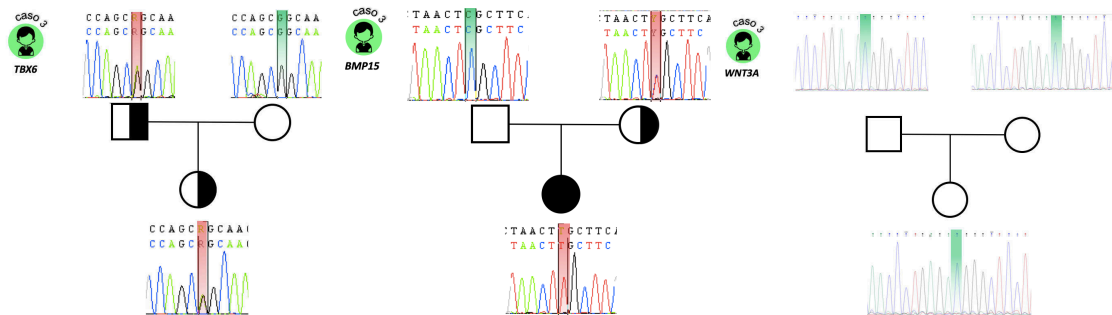


Figura 6. Validación de variantes encontradas en gen TBX6, BMP15 y WNT3A.

#### Caso 4

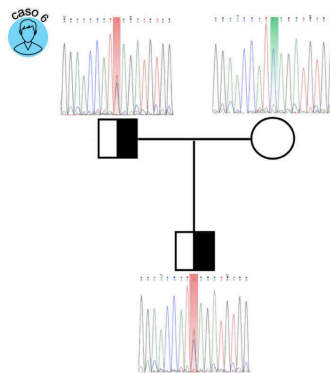
Para el caso número 4, fueron validadas las variantes encontradas en los genes **PRKAG2** (c.1040A>G) y **NDUFS7** (c.434G>A).

#### Caso 5

Para este caso, la secuenciación de Sanger descartó la presencia de la mutación ubicada en el gen **FBN3** tanto en el probando como en sus padres.

#### Caso 6

Para el caso 6, la secuenciación de Sanger confirmó la variante en el gen **CNTN6** en el probando, así como en su padre (Fig. 7). Debido a este resultado decidimos analizar el DNAg del padre mediante WES (Imagen 2). Una vez secuenciado el genoma del padre fueron comparadas las variantes de riesgo de TEA encontradas en el paciente, con los metadatos del padre. Dentro los datos descubrimos, que además de la variante **CNTN6**, se encontró una segunda variante compartida **COL11A1** (Tabla 1, nota al pie de página). Además de estos no se observó otro SNV relacionado con TEA truncante en el exoma del paciente. Adicional a esto, el análisis en la aplicación de Enrichment reveló una inversión de 824 pb que involucraba parcialmente al exón 4 del gen **CNTN5**.



**Figura 7. Validación de variante encontrada en el gen **CNTN6**.**

**Tabla 23. Identificación de genes caso 6.**

Gen / Nombre	Función molecular	Variante	Vía
<b>CNTN6</b> Contactina 6	El producto de este gene es una proteína de membrana neuronal anclada en glicosilfosfatidilinositol (GPI) que funciona como una molécula de adhesión celular. Jugar un papel en la formación de conexiones de axonales en el sistema nervioso en desarrollo. <sup>60</sup>	Codón de paro c.566C>A p.(Ser189Ter) Exón: 6/23	Interacciones CHL1. Guía de axón. Interacción L1CAM.
<b>DNAH10</b> Cadena pesada de la dineína 10 axonemal	Proteína motoras asociadas a microtúbulos, dependiente de ATP, dirigida hacia el extremo negativo.	Cambio de sentido c.6117G>C p.(Met2039Ile) Exón: 36/78	Proteína que media la motilidad hacia el extremo negativo
<b>DNAH7</b> Cadena pesada de la dineína 7, axonemal	Proteína motoras asociadas a microtúbulos, dependiente de ATP, dirigida hacia el extremo negativo.	Cambio de sentido c.11720A>G p.(Tyr3907Cys) Exón: 63/65	Proteína que media la motilidad hacia el extremo negativo
<b>DNAJC5</b> Miembro 5 de la subfamilia C homóloga DnaJ	Actúa como una chaperona general en la exocitosis regulada. Tiene un papel importante en la función presináptica y estar involucrado en la liberación de neurotransmisores dependientes de calcio en las terminaciones nerviosas. <sup>61</sup>	Cambio de sentido c.350C>T p.(Thr117Met) Exón: 4/5	Transmisión a través de sinapsis químicas
<b>KMT2C</b> Histona-lisina N-metiltransferasa 2C	Histona que metila 'Lys-4' de la histona H3; La metilación H3 'Lys-4' representa una etiqueta específica para la activación transcripcional epigenética. Componente central del complejo MLL2 / 3, complejo coactivador de receptores nucleares, implicado en la coactivación transcripcional. KMT2C / MLL3 puede ser una subunidad catalítica de este complejo.	Cambio de sentido c.2578C>T p.(Pro860Ser) Exón: 15/59. c.2645T>C p.(Ile882Thr) Exón: 15/59	Activación de genes hox en el desarrollo del cerebro posterior durante la embriogénesis.
<b>MTNR1A</b> Receptor de melatonina tipo 1A	Receptor de alta afinidad por la melatonina. Media acciones reproductivas y circadianas de la melatonina.	Cambio de sentido c.287G>A p.(Gly96Asp) Exón: 2/2	Receptor acoplado a proteína G, inhibe la adenilato-ciclasa
<b>NEO1</b> Neogenina 1	Receptor multifuncional de la superficie celular para los miembros de las familias de BMP, netrina y molécula de guía repulsiva (RGM) que regula la adhesión celular en diversos procesos de desarrollo, donde se incluyen la formación del tubo neural. Las interacciones Netrin-Neogenin dan como resultado una respuesta de guía de axones quimioattractiva y una adhesión célula-célula.	Cambio de sentido c.1112C>T p.(Thr371Ile) Exón: 6/29	Regulación de forma positiva la vía de señalización BMP
<b>NRXN2</b> Neurexin 2	El productos de este gene funcionan como molécula de adhesión celular y receptor en el sistema nervioso de vertebrados. <sup>62</sup>	Cambio de sentido c.1975C>T p.(Arg659Trp) Exón: 10/23	Interacciones proteína-proteína en las sinapsis. Neurexinas y neuroliginas
<b>SLC45A1</b> Transportador de azúcar asociado a protones A	Codifica un transportador neuronal de glucosa que se expresa predominantemente en el cerebro en desarrollo y en el adulto. <sup>63</sup>	Cambio de sentido c.2080G>C p.(Val694Leu) Exón: 8/8	Transportador neuronal de glucosa

## CAPÍTULO IX

### DISCUSIÓN

#### CASO 1

En el caso 1, se presenta una paciente con signos y síntomas consistentes con insuficiencia renal progresiva, urolitiasis recurrente, acumulación de oxalato, nefrolitiasis, nefrocalcinosis, polineuropatía, afección cardíaca y ocular. Mediante WES se identificó una mutación puntual (c.33dupC) en el gen **AGXT** presente de manera heterocigota. La duplicación trae consigo consecuencias moleculares en el desplazamiento del marco de lectura, ocasionando el cambio del aminoácido lisina por glicina en la posición 12, generando a su vez un codón de terminación en la posición 156 (p.Lys12GlnfsTer156). Según la plataforma ClinVar la duplicación es clasificada con significancia clínica patogénica asociándola directamente a la enfermedad de hiperoxaluria primaria tipo 1 (PH tipo1, MIM 259900).

Al identificar esta mutación mediante pruebas moleculares se confirmó uno de los posibles diagnósticos reservados para esta paciente, comprobando que el fenotipo es compatible con la enfermedad de PH 1.

PH 1 es una enfermedad rara con una prevalencia que varía de 1 / 333,000 a 1 / 1,000,000 en donde los valores más altos se reportan en poblaciones específicas con una alta tasa de consanguinidad.<sup>64</sup> Posee una importante heterogeneidad en la presentación clínica y el inicio de los síntomas. Además, es una enfermedad autosómica recesiva causada por la deficiencia de la enzima peroxisomal hepática

alanina: glioxilato-aminotransferasa (AGT, EC 2.6.1.44), que cataliza la conversión de glioxilato en glicina. Cuando la actividad de AGT está ausente, el glioxilato se convierte en oxalato, que posteriormente forma cristales de oxalato de calcio insolubles que se acumulan en el riñón y otros órganos. Si ésta no es diagnosticada adecuadamente puede producirse una disminución progresiva de la función renal, resultado en depósitos de oxalato de calcio en el tejido renal y a su vez la nefrolitiasis puede progresar a la deposición tisular de oxalato de calcio conocida como oxalosis sistémica y muerte posterior.<sup>65</sup>

En nuestro estudio, no fue posible validar la mutación en el trio por secuenciación Sanger debido a la imposibilidad de obtener las muestras de DNAg de ambos padres. Sin embargo, algo que nos llama la atención en este caso es la detección en heterocigosis de la mutación c.33dupC en un probando con signos y síntomas consistentes con PH 1, enfermedad a la cual se le ha atribuido un modelo de herencia autosómico recesivo. Este hecho nos llevó a cuestionarnos si es posible encontrar un heterocigoto con PH 1.

Para entender el suceso es necesario tomar en cuenta algunos antecedentes referentes a la enzima AGT. Uno de estos y tal vez el más relevante, es el trabajo realizado Zhang X. y cols., en el 2003 al dilucidar la estructura cristalina de la AGT en complejo con el inhibidor del ácido aminooxiacético.<sup>66</sup> La estructura reveló que la AGT es una enzima homodimérica, en donde cada subunidad contiene 392 residuos de aminoácidos, que cataliza la transaminación (desintoxicación) del metabolito glioxilato al metabolito de dos carbonos a glicina. Es dependiente de su



cofactor piridoxal 5'-fosfato (PLP). Además, cada subunidad está compuesta por una extensión N-terminal (residuos 1–21) que envuelve la superficie de la otra subunidad, un gran dominio N-terminal (residuos 22–282) que contiene la mayor parte del sitio activo y la interface de dimerización, y un dominio-C terminal (residuos 283–392) que contiene una secuencia de direccionamiento peroxisomal PTS1.

En nuestro caso, la mutación identificada (c.33dupC) al desplazar el marco de lectura da lugar a la reorganización de los codones, culminando con la generación de un codón de paro en el residuo 156, no es difícil imaginar que este evento afecte por completo el orden de los residuos alterando la conformación del monómero. Sumado a esto hay que recordar que para que la enzima lleve a cabo su función es necesaria la interacción y conformación de homodímero, proceso en el cual interactúan residuos altamente conservados.<sup>67</sup>

Un antecedente importante a la causa es la identificación de pacientes heterocigotos compuestos. Se estima, que la prevalencia de estos representa 27% de los pacientes, según lo informado por un estudio basado en una gran cohorte europea.<sup>68</sup> Este fenómeno representa una fuente general de variación fenotípica en muchas enfermedades genéticas que involucran proteínas multiméricas, como fenilcetonuria<sup>69</sup> deficiencia de argininosuccinato liasa.<sup>70</sup> En estos casos, lo que podría ocurrir es una complementación interalelo (CI) que determina las características metabólicas de los pacientes heterocigotos compuestos.

Sin embargo, en nuestro caso solo se identificó la mutación c.33dupC en heterocigosis. En el 2014 Montioli R. y cols., en un intento por explicar el fenómeno de CI en PH 1 dieron a conocer un posible mecanismo que explica esta interacción, dicho mecanismo se retoma aquí para proponer un posible evento que dé explicación a nuestros resultados. Este modelo involucra la co-expresión de las dos variantes alélicas, es decir la variante mutada (p.Lys12GlnfsTer156) y la variante silvestre en condiciones homocigotas y heterocigotas. Al interaccionar ambas variantes, se esperaría obtener tres especies de fenotipos enzimáticos y clínicos: el homodímero para la variante silvestre, homodímero para la variante p.Lys12GlnfsTer156 y el heterodímero p.Lys12GlnfsTer156/ variante silvestre. Dado que no existe información sobre la cantidad relativa de estas especies, el supuesto invocado en estas situaciones es una asociación de subunidades aleatorias que conduciría a generar una distribución del 25% de homodímeros para la variante silvestre, 25% de homodímeros para la variante p.Lys12GlnfsTer156 y 50% para el heterodímero p.Lys12GlnfsTer156/ variante silvestre. Estas combinaciones de subunidades de monómeros como resultado producirán especies con posibles efectos positivos, negativos o nulos, siendo estos dos últimos expresados en un mayor porcentaje. Como fenómeno responsable a las variantes polimórficas podrían atribuírseles a un efecto nulo por deterioros en la estabilidad y/o tasa de dimerización de los monómeros.<sup>71</sup>

Una de las problemáticas que rodean el diagnóstico definitivo de PH 1 es que, para establecerse es necesario realizar la estimación de la actividad de la AGT en una muestra de biopsia de hígado. Esta técnica resulta incómoda y sumamente invasiva,

además, la muestra debe enviarse congelada a algún laboratorio certificado para su análisis<sup>72</sup>. Al aplicar un enfoque genético molecular como lo es WES se pretende superar estas limitaciones generando un diagnóstico mínimamente invasivo cuando exista sospecha, signos clínicos y se hayan excluido causas secundarias.

## **CASO 2 Y 3**

Respecto a los casos 2 y 3 ambos probandos fueron diagnosticados presuntivamente con síndrome de FATCO (OMIM # 246570). El diagnóstico se basó en la presencia de oligosindactilia, aplasia fibular y campomelia de tibia para el caso 2 y oligosindactilia, mesomelia, agenesia de peroné y campomelia de tibia para el caso 3.

FATCO es una afección sindrómica extremadamente rara con una prevalencia < 1/1 000 000.<sup>73</sup> El síndrome involucra malformaciones óseas tales como: aplasia de peroné, campomelia de tibia y oligosindactilia.<sup>74</sup> Fue descrito por primera vez en 1981 por Hecht y Scott en el caso de dos medias hermanas con malformaciones osteomusculares<sup>75</sup>. Más adelante, en 1994, Capece y cols., reportaron 1 caso más<sup>76</sup> y en 2003, Huber y cols., reportaron 2 casos nuevos.<sup>77</sup> Sin embargo, fue hasta el año 2005 en donde Courtens y cols., reportaron un caso nuevo, y tras realizar una revisión de reportes de casos, lograron identificar los 5 reportes anteriores que incluían los mismos hallazgos patológicos debido a esto, propusieron el acrónimo FATCO para este síndrome<sup>78</sup>. Desde ese momento, se han descrito más de 12 de casos con características similares<sup>79,80,89,90,81-88</sup>, sin embargo, a la fecha el síndrome presenta una herencia y base genética desconocidas debido a la falta de

información sobre la transmisión familiar, además, aún no ha sido posible asociar ningún gen causante de las alteraciones. Por otro lado, no hay asociaciones consistentes con las exposiciones ambientales.

Dentro de nuestros resultados, describimos un conjunto de variantes ubicadas en genes pertenecientes a las familias FGF (**FGFR4**, c.1287C>A), genes Hox (**HOXC9**, c.596A>G) ( **HOXD12**, c.236G>A), ( **HOXD3**, c.257T>C) para el caso 2 y para el caso 3, miembros de la superfamilia del factor de crecimiento transformante beta (TGF- $\beta$ ), como los factores de crecimiento y diferenciación (GDF) (**GDF5**, c.536C>T) y BMP (**BMP15**, c.433C>T), genes Hox (**HOXC12**, c.139C>T), factores de transcripción de la familia T-box (**TBX6**, c.484G>A) y morfógenos pertenecientes a la familia Wnt (**WNT3A**, c.860C>G). Estas variantes fueron predichas *in silico* por al menos un programa como probablemente dañinas o deletéreas.

Uno de los signos más característicos de este síndrome es la aplasia o ausencia de peroné, este es el hueso largo ausente congénitamente más común. La aplasia puede ocurrir de manera unilateral o bilateral, esporádica o familiar, aislada o sindrómica.<sup>91</sup> Sin embargo, se desconoce la etiología exacta de esta deficiencia, considerándose un evento esporádico o aislado. Esta puede ser parte de algún síndrome de malformación.<sup>92</sup>

Existen algunos otros trastornos, que dentro de sus manifestaciones involucran agenesia de peroné tales como: síndrome fémur-peroné-cúbito (OMIM 228200),<sup>93</sup> síndrome de Fuhrmann (OMIM 228200),<sup>94</sup> displasia campomélica (OMIM

114290),<sup>95</sup> ausencia de cúbito y peroné con deficiencia severa de extremidades (OMIM 276820),<sup>69</sup> braquidactilia-ectrodactilia con hipoplasia o aplasia de peroné (OMIM 113310),<sup>96</sup> síndrome oral-facial-digital con aplasia del peroné (OMIM 165590)<sup>97</sup> y síndrome de Weismann Netter (OMIM 112350),<sup>98</sup> sin embargo, estos fueron excluidos del diagnóstico debido a que difieren de la condición presentada por los probandos. Principalmente por la presencia de alteraciones en pelvis, alteraciones en cabeza, cuello, alteraciones neurológicas y endocrinas entre otras, ninguna presente en los probandos.

Otra característica relevante en este síndrome es la edad de inicio la cual ha sido identificada como prenatal o neonatal. Considerando estas características se ha propuesto la teoría de interrupción del campo de desarrollo de las extremidades durante la embriogénesis.<sup>99</sup> Esta teoría se sustenta en varios estudios embriológicos y genéticos en modelos de vertebrados, en donde fue posible establecer modelos que definen las interacciones celulares y moleculares que regulan el patrón y el desarrollo esquelético, además, de establecer genes clave que controlan la formación de las extremidades.<sup>100</sup> Estos modelos describen que, a final de la sexta semana de gestación, comienzan a desarrollarse los miembros inferiores. Estos surgen como pequeños brotes que sobresalen del cuerpo y contienen células mesodérmicas (células mesenquimales morfológicamente homogéneas) que están revestidas de ectodermo a las que se denominan esbozos de miembros inferiores.<sup>101</sup>

Además, se han establecido los principales componentes moleculares que coordinan el crecimiento de la extremidad en sus tres dimensiones el cual está

determinado por genes y factores consecuentes. Los ejes por considerar son: el eje próximo-distal que está bajo el control de los factores de crecimiento de fibroblastos (FGF) y la cresta ectodérmica apical (AER, por sus siglas en inglés: apical ectodermal ridge). El eje anteroposterior que está bajo el control de Sonic hedgehog (SHH) del mesénquima posterior y el eje dorsoventral controlado por proteínas morfogénicas óseas (BMP), el factor de transcripción Engrailed1 (En1) del ectodermo ventral y por el gen **WNT7A** del ectodermo dorsal.<sup>101</sup>

Un cuarto grupo de genes que ha sido asociado a este fenómeno son los genes Hox, esta es una amplia familia de genes conocidos como reguladores maestros del desarrollo embrionario, su función es proporcionar información posicional a las células no comprometidas para asegurar que se diferencien en el tejido apropiado para su correcta ubicación dentro del embrión en desarrollo. De este modo, los grupos de células, conocidos como dominios funcionales, se comprometen a formar estructuras corporales como extremidades y órganos.<sup>102</sup>

De este conjunto de genes el único asociado a trastornos relacionados con malformaciones óseas en miembros inferiores y reportado en la base de datos OMIM es el gen **GDF5**, este ha sido asociados a hipoplasia de peroné en el síndrome Du Pan (OMIM 228900), además en éste síndrome los afectados presentaron otras alteraciones como rótula desplazada, braquidactilia, equinovalgo y uñas hipoplásicas. Dichas alteraciones no son compatibles con el fenotipo reportado para el caso 3.

#### **CASO 4**

En el caso 4, la probando fue diagnosticada presuntivamente con enfermedad mitocondrial, esto según las características clínicas exhibidas (dismotilidad gastrointestinal, parestesias de miembros pélvicos inferiores, pseudo obstrucción intestinal, disfagia y gastroparesia).

En el análisis, se identificaron dos variantes de cambio de sentido ubicadas en dos genes de DNA nuclear (DNAn). **NDUFS7** (c.434G>A) y **PRKAG2** (c.1040A>G) que podrían responder al diagnóstico de enfermedad mitocondrial. Las variantes fueron predichas como probablemente dañinas y deletéreas por las plataformas SIFT y Polyphen-2. Estas no poseen un registro de identificación en la base de datos de los 1000 genomas y solo la variante ubicada en **PRKAG2** pose frecuencia alélica reportada según ExAC (8.26e-06 1/121064)

Los dos genes han sido asociados a enfermedades mitocondriales, sin embargo, ninguno de los fenotipos reportados se asemeja al fenotipo exhibido por la paciente.

Tras realizar el diagnóstico se identificó que las características atípicas presentes en la paciente se asemejan al síndrome de encefalopatía neurogastrointestinal mitocondrial (MNGIE, por sus siglas en inglés; OMIM 603041). Este se caracteriza por graves trastornos gastrointestinales (cólicos, vómitos, diarrea, pseudo obstrucción intestinal, disfagia y gastroparesia) producidos por una motilidad intestinal anormal. Los trastornos gastrointestinales pueden progresar gradualmente hacia una pseudo-obstrucción crónica que conduce a caquexia.

Además de estas afecciones también se reportan alteraciones neurológicas incluyendo neuropatía periférica sensitivo-motora. Sin embargo, hay algunos signos que no fueron reportados en la paciente tales como: oftalmoplejía crónica progresiva con o sin ptosis, leucodistrofia, sordera, retinopatía pigmentaria y daños a nivel de cerebelo.

En cuanto a la relación fenotipo-genotipo se ha señalado como responsable al gen ***TYMP*** en conjunto con un modelo de herencia autosómico recesivo. Sin embargo, también se han reportado formas menos comunes de MNGIE causada por mutaciones recesivas en el gen ***RRM2*** (OMIM 612075) y ***POLG*** (OMIM 613662).

El gen ***NDUFS7*** o ***PSST*** (homologo bovino) codifica para la proteína NADH: deshidrogenasa ubiquinona proteína Fe-S 7 la cual es parte del brazo periférico del complejo I (NADH ubiquinona oxidorreductasa, EC 1.6.5.3) de la cadena de transporte de electrones. Este complejo es crucial para la respiración aeróbica. En las mitocondrias, se encarga de oxidar el NADH del ciclo del ácido tricarboxílico y la  $\beta$ -oxidación, también participa reduciendo la ubiquinona y transporta protones a través de la membrana interna, lo que contribuye a la fuerza motriz protónica.<sup>103</sup>

Mediante la estructura cristalina se ha determinado la topología de la subunidad encontrando que el complejo en mamíferos es una molécula en forma de L con un brazo incrustado en la membrana interna mitocondrial y un brazo que sobresale de la matriz denominado brazo periférico.<sup>104</sup> Se determinó también que está conformada a partir de 46 subunidades polipeptídicas, siete codificadas por el DNA



mitocondrial y el resto por el genoma nuclear. Además, la enzima bovina se fraccionó en tres partes funcionales: la parte que contiene flavin mononucleótido (FMN) como grupo prostético, la parte que contiene hierro-azufre y el dominio incrustado en la membrana.<sup>105</sup>

Además, mediante varios modelos se ha intentado determinar la vía de ensamblaje de estas subunidades encontrando que la subunidad PSST se localiza asociada al grupo N2 del hierro-azufre<sup>105</sup> y es necesaria para ensamblar el brazo periférico<sup>106</sup> estos hechos sugieren que PSST se localiza entre el brazo periférico y de membrana siendo requerida para la estabilización del brazo periférico.<sup>107</sup>

## **CASO 5**

Para el caso 5, el paciente fue diagnosticado con síndrome de Down según su cariotipo mos47,XY,+21 (20)/ 48,XY,+21+mar(3), con presencia de alteraciones asociadas al síndrome de Marfan (dilatación de la raíz aortica, cataratas y luxación del cristalino). El síndrome de Marfan (SMF) es una condición autosómica dominante, causada en la mayoría de los casos por mutaciones en el gen **FBN1** ubicado en el cromosoma 15 que codifica la proteína fibrilina. Este defecto resulta en un conjunto en diversas manifestaciones que involucran a varios órganos y sistemas, siendo las alteraciones musculoesqueléticas, cardiovasculares y oftálmicas las más notorias<sup>108</sup>. El síndrome tiene una incidencia estimada de 2-3 afectados por cada 10.000 habitantes.<sup>109</sup>

En el análisis realizado no se identificaron mutaciones en el gen ***FBN1***. Debido a esto, el análisis se centró en la búsqueda de nuevos genes que pudieran estar relacionados con el padecimiento. En los metadatos se identificó una mutación en el gen ***FBN3*** (c.4161T>A) que involucra un codón de paro en la posición 1387 (p.Cys1387Ter) a nivel de aminoácidos. Sin embargo, la mutación no pudo ser validada mediante secuenciación Sanger. Creemos que el hecho obedece a la baja frecuencia de lecturas reportadas tanto para el alelo alternativo (2) como para el alelo silvestre (9). Este resultado nos guió a la búsqueda de un segundo grupo de genes que pudieran estar generando el fenotipo.

Además del gen ***FBN3***, varios estudios han proporcionado evidencia convincente de que las mutaciones en la fibrilina están asociadas con una deficiente regulación de la señalización de TGF- $\beta$  en la matriz, lo que conduce a niveles excesivos de TGF- $\beta$  bioactivo en el microentorno tisular.<sup>110</sup> Debido a este evidente aumento del ligando de TGF- $\beta$ , se han reportado mutaciones en otros componentes de la vía (receptores, Smads, Schnurri) de TGF- $\beta$  que pueden dar lugar a trastornos relacionados con fenotipos atenuados en relación con el síndrome de Marfan, denominados colectivamente síndromes similares a SMF.<sup>111,112</sup>

En el análisis identificamos a los genes ***MAML3*** y ***NOTCH4*** como posibles candidatos. Estos dos genes interactúan para formar el complejo co-activador NOCH4, con el objetivo de incitar la transcripción de objetivos NOTCH bien establecidos de manera dependiente del contexto celular.<sup>113</sup> Además, ***NOCH4*** inhibe la señalización canónica de TGF- $\beta$  a través de la degradación dependiente

de la proteína SMAD3 fosforilada.<sup>114</sup> Sin embargo, ninguno de estos genes ha sido asociado directamente a las patologías presentadas por el probando. Por otro lado, la proteína SMAD3 ha sido asociada a aneurisma aórtico (dilatación aórtica),<sup>115</sup> en el síndrome Loeys-Dietz (# 609192), síndrome que muestra una importante superposición de características clínica con el SMF<sup>116</sup>. Con estos hechos resulta tentador especular sobre la participación del gen **NOTCH4** en el desarrollo de una de las afecciones presentes en el paciente; sin embargo, este gen no puede por sí solo explicar el fenotipo completo del paciente o por lo menos no la ectopia lentis.

Debido a esto, la búsqueda se centró en otros genes que afecten el tejido conectivo centrándose en algunos componentes de la matriz extracelular. Encontrando un grupo de 15 genes asociados a diferentes procesos relacionados con la biosíntesis, organización y degradación de esta (tabla 19).

## **CASO 6**

Respecto al caso 6, el paciente presenta un diagnóstico de TEA con base en alteraciones en la deficiencia de comunicación social recíproca y en patrones de conducta restringidos y repetitivos, además de estos fenotipos de comportamiento, el paciente presenta alteraciones sensoriales y motoras incluyendo hiperacusia, indiferencia ante el dolor y fascinación por las luces y el movimiento.

Dentro de los metadatos obtenidos por WES en la muestra del probando identificamos un conjunto de genes que han sido asociados a riesgo de desarrollo de TEA (tabla 23). Dentro de estos encontramos al gen **CNTN6** que codifica la molécula de adhesión de células neuronales contactina 6, la cual es miembro de una subfamilia de proteínas que contienen dominios de inmunoglobulina neural (Ig). Esta es una proteína altamente expresada en diferentes tejidos cerebrales, como corteza, hipocampo y cerebelo, además se sabe que participa promoviendo el crecimiento de las neuritas en las vías neuronales sensoriales-motoras.<sup>117</sup> Esta es regulada por la proteína T-Brain-1,<sup>118</sup> e interactúa con la molécula de adhesión celular L1-like,<sup>119</sup> ambas proteínas han sido asociadas a riesgo de TEA y discapacidad intelectual respectivamente.

Se ha reportado que SNV de pérdida de función heredados o *de novo* en este gen parecen tener un papel patógeno central en un reducido grupo de pacientes con TEA con alteraciones sensoriomotora y auditivas.<sup>60</sup> Sin embargo, como ocurre con varios otros SNV relacionadas con el TEA, su impacto clínico real es a menudo inconsistente.

A pesar de las conductas sensoriales y motoras y los problemas de audición en el presente paciente con diagnóstico de TEA son comparables con tales hallazgos, nos llama la atención la presencia de la misma variante trunca (**CNTN6**, c.566C>A) en el padre, el cual carece de cualquier síntoma. Este resultado es inesperado e incluso contradice el impacto sugerido de tales mutaciones en relación con el TEA.<sup>120,121</sup>

Se ha sugerido que varias mutaciones en **CNTN6** no son completamente penetrantes (incluso algunas relacionadas con deleciones), y requieren la presencia de variantes concomitantes, que en algunos casos pueden estar en otros genes relacionados con contactinas, para causar TEA.<sup>60,120,122</sup>

En este sentido, nuestra búsqueda de variantes en otros genes de riesgo TEA ya informados demostró, otras posibles variantes patógenas en los genes **CNTN5**, **DNAH10**, **NEO1** y **NRXN2** presentes en el paciente, y ausentes en su padre, así como en otros datos de secuenciación masiva de otros sujetos sanos y sin TEA que poseemos en el laboratorio de citogenómica y microarreglos (n = 15, datos no mostrados). Nuestros resultados concuerdan con algunos estudios genéticos respecto a la heredabilidad poligénica del TEA<sup>123,124</sup>

Al igual que **CNTN6**, **CNTN5** tiene funciones en el desarrollo de vías neuronales sensoriales-motoras; por lo tanto, las alteraciones de este gen también pueden llevar a un fenotipo de TEA que involucre afectaciones sensori-motoras y auditivas.

<sup>60</sup> La variante encontrada en este gen se encontró en forma de SV sin embargo la patogenicidad de esta no está clara. Esta fue descrita previamente por DGV (dgv705n106 / nsv1078300) en al menos dos pacientes sin datos neurológicos disponibles, el daño previsto en el exón 4 / intrón 4 generando una proteína aberrante y sumándose al efecto en el gen **CNTN6**.

Por otro lado, **DNAH10** codifica una dineína expresada en cerebro (UniProtKB, ID Q8IVF4). Las dineínas participan en el establecimiento y mantenimiento de la

compleja morfología de los axones y las dendritas.<sup>125</sup> En 2017, Mercati y cols., Describieron un hecho similar el presentado aquí en el sentido que describieron a dos hermanos con TEA que heredaron de su madre afectada una SNV trunca en el gen **CNTN6** y un VUS en **DNAH10**.

Por su parte **NEO1** y **NRXN2** codifican proteínas de superficie celular neuronal; **NEO1** ha participa en la adhesión celular, la migración neural y la guía de los axones, mientras que **NRXN2** participa en el reconocimiento y la adhesión celular (UniProtKB, ID Q9P2S2)<sup>126,127</sup> Otro gen que alberga una posible variante patógena, pero que aún no se ha informado como factor de riesgo a TEA, es el gen **DNAH7**.

Si bien la mayoría de los genes relacionados con TEA mencionados en este informe parecen tener una penetración incompleta o una asociación menos definida,<sup>60,128</sup> es tentador especular sobre la participación de más de uno de estos y quizás otros genes (probablemente aún no relacionados con el desarrollo de TEA) en la patogenia de este paciente, principalmente a través de sutiles efectos aditivos en el funcionamiento de axones y / o dendritas. Esto es consistente con la superposición funcional de los genes de riesgo de autismo por ejemplo, los que codifican contactinas y neurexinas implicados en la regulación de la estabilidad estructural de las neuronas.<sup>129</sup> Los datos y resultados obtenidos en este caso fueron publicados en el Journal of Autism and Developmental Disorders.<sup>130</sup>

## CAPÍTULO X

### CONCLUSIONES

#### Caso 1

Logramos identificar una variante de pérdida de función en el gen **AGXT** asociada previamente al desarrollo de PH 1 en una paciente con signos y síntomas consistentes con el desarrollo de la enfermedad.

#### Casos 2 y 3

Identificamos un grupo de mutaciones patógenas en genes involucrados en el desarrollo de miembros óseos que actúan durante el proceso de embriogénesis. Sin embargo, tales alteraciones no son concluyentes en el fenotipo exhibido por ambos pacientes.

#### Caso 4

Identificamos dos mutaciones en genes de DNA nuclear que podrían explicar el fenotipo de la probando sin, embargo la mutación en el gen **NDUFS7** parece jugar un papel protagónico en el desarrollo de las manifestaciones clínicas, por tal razón proponemos a la variante como causante del fenotipo clínico.

#### Caso 5

Identificamos un grupo de mutaciones patogénicas en genes relacionados con la biosíntesis, organización y degradación de matriz extracelular que podrían explicar parcialmente el fenotipo del probando.

## Caso 6

Logramos identificar una variante de pérdida de función en el gen **CNTN6** relacionada con un fenotipo de TEA. Creemos que este hallazgo refuerza el vínculo entre las mutaciones del gen **CNTN6** y el TEA con alteraciones de la percepción sensorial, además, creemos que para que se manifieste este fenotipo es necesaria la presencia de posibles variantes patógenas concomitantes, incluida una SV que afecta a otra contactina que están estrechamente relacionada. Por lo tanto, nuestros hallazgos respaldan aún más la afirmación de que el TEA relacionado con el gen **CNTN6** requiere un fondo genético específico (y / o incluso epigenético) para manifestarse.

En todos los casos se identificaron mutaciones relacionadas con el fenotipo de los probandos, estableciendo una relación genotipo-fenotipo, sin embargo, dado el diseño del proyecto solo podemos proporcionar respuestas a 4 de los 6 casos, consideramos oportuna la utilización de otras técnicas para dar solidez a los resultados en estos dos casos restantes.

La hipótesis se cumplió en 4 de los 6 casos.



## CAPÍTULO XI

### PERSPECTIVAS

- Validar las variantes candidatas de los casos no resueltos.
- Realizar ensayos *in silico* e *in vivo* para determinar el impacto de las mutaciones encontradas en la estructura de las proteínas.

## CAPÍTULO XII

### BIBLIOGRAFÍA

1. Herder M, Krahn TM. Some numbers behind Canada's decision to adopt an orphan drug policy: US orphan drug approvals in Canada, 1997-2012. *Healthc Policy*. 2016. doi:10.12927/hcpol.2016.24626
2. RARE DISEASES ACT OF 2002.; 2000. <https://history.nih.gov/research/downloads/PL107-280.pdf>. Accessed September 9, 2018.
3. Enfermedades raras - Comisión Europea. [https://ec.europa.eu/health/non\\_communicable\\_diseases/rare\\_diseases\\_en](https://ec.europa.eu/health/non_communicable_diseases/rare_diseases_en). Accessed September 9, 2018.
4. Senado de la República. <http://www.senado.gob.mx/index.php?watch=11&sm=3&id=40533>. Accessed August 16, 2018.
5. INSERM. Prevalence and incidence of rare diseases: Bibliographic data. *Orphanet Rep Ser Rare Dis Collect*. 2015.
6. OMIM Gene Map Statistics. <http://omim.org/statistics/geneMap>. Accessed August 14, 2018.
7. Field MJ (Marilyn J, Boat TF, Institute of Medicine (U.S.). Committee on Accelerating Rare Diseases Research and Orphan Product Development. *Rare Diseases and Orphan Products: Accelerating Research and Development*. National Academies Press; 2010.
8. Dudding-Byth T. A powerful team: The family physician advocating for patients with a rare disease. *Aust Fam Physician*. 2015.
9. Dharssi S, Wong-Rieger D, Harold M, Terry S. Review of 11 national policies for rare diseases in the context of key patient needs. *Orphanet J Rare Dis*. 2017. doi:10.1186/s13023-017-0618-0
10. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an Online catalog of human genes and genetic disorders. *Nucleic Acids Res*. 2015. doi:10.1093/nar/gku1205
11. ProductDevelopment I of M (US) C on ARDR and O, Field MJ, Boat TF. Profile of Rare Diseases. 2010. <https://www.ncbi.nlm.nih.gov/books/NBK56184/#ch2.s4>. Accessed July 31, 2019.
12. Nambot S, Thevenon J, Kuentz P, et al. Clinical whole-exome sequencing for the diagnosis of rare disorders with congenital anomalies and/or intellectual disability: substantial interest of prospective annual reanalysis. *Genet Med*. 2018;20(6):645-654. doi:10.1038/gim.2017.162
13. *Rare Diseases Europe WHAT IS A RARE DISEASE? Eurordis-Rare Diseases Europe Www.Eurordis.Org-Eurordis@eurordis.Org SURVEY OF THE DELAY IN DIAGNOSIS FOR 8 RARE DISEASES IN EUROPE ('EURORDISCARE 2')*

- Diseases and Countries Number of Respondents by Country.* www.eurordis.orgwww.eurordis.org. Accessed September 25, 2018.
14. Clark MM, Stark Z, Farnaes L, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *npj Genomic Med.* 2018. doi:10.1038/s41525-018-0053-8
  15. Inoue S, Mangat C, Rafe'e Y, Sharman M. Forme Fruste of HLH (haemophagocytic lymphohistiocytosis): diagnostic and therapeutic challenges. *BMJ Case Rep.* 2015;2015. doi:10.1136/bcr-2014-206190
  16. Xu J, Murphy SL, Kochanek KD, Arias E. Mortality in the United States, 2015. *NCHS Data Brief.* 2016;(267):1-8. <http://www.ncbi.nlm.nih.gov/pubmed/27930283>. Accessed November 17, 2018.
  17. Miller NA, Farrow EG, Gibson M, et al. A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. *Genome Med.* 2015;7:100. doi:10.1186/s13073-015-0221-8
  18. Saunders CJ, Miller NA, Soden SE, et al. Rapid Whole-Genome Sequencing for Genetic Disease Diagnosis in Neonatal Intensive Care Units. *Sci Transl Med.* 2012;4(154):154ra135-154ra135. doi:10.1126/scitranslmed.3004041
  19. Sobreira NL, Valle D. Lessons learned from the search for genes responsible for rare Mendelian disorders. *Mol Genet Genomic Med.* 2016. doi:10.1002/mgg3.233
  20. Lawce HJ, Brown MG. Cytogenetics: an overview. In: *The AGT Cytogenetics Laboratory Manual.* ; 2017. doi:10.1111/j.1744-7429.1998.tb00373.x
  21. Gersen SL, Keagle MB. *The Principles of Clinical Cytogenetics, Third Edition.*; 2013. doi:10.1007/978-1-4419-1688-4
  22. Slater HR, Bailey DK, Ren H, et al. High-resolution identification of chromosomal abnormalities using oligonucleotide arrays containing 116,204 SNPs. *Am J Hum Genet.* 2005;77(5):709-726. doi:10.1086/497343
  23. Choi HMT, Beck VA, Pierce NA. Next-generation in situ hybridization chain reaction: Higher gain, lower cost, greater durability. *ACS Nano.* 2014. doi:10.1021/nn405717p
  24. Ratan ZA, Zaman S Bin, Mehta V, Haidere MF, Runa NJ, Akter N. Application of Fluorescence In Situ Hybridization (FISH) Technique for the Detection of Genetic Aberration in Medical Science. *Cureus.* 2017;9(6):e1325. doi:10.7759/cureus.1325
  25. Kallioniemi A, Kallioniemi O, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science (80-).* 1992;258(5083):818-821. doi:10.1126/science.1359641
  26. Kallioniemi A. CGH microarrays and cancer. *Curr Opin Biotechnol.* 2008. doi:10.1016/j.copbio.2007.11.004
  27. Weiss MM, Hermsen MAJA, Meijer GA, et al. Demystified . . . Comparative genomic hybridisation. *J Clin Pathol - Mol Pathol.* 1999. doi:10.1136/mp.52.5.243
  28. Inazawa J, Inoue J, Imoto I. Comparative genomic hybridization (CGH)-arrays pave the way for identification of novel cancer-related genes. *Cancer Sci.* 2004;95(7):559-563. doi:10.1111/j.1349-7006.2004.tb02486.x

29. de Leeuw N, Dijkhuizen T, Hehir-Kwa JY, et al. Diagnostic interpretation of array data using public databases and internet sources. *Hum Mutat.* 2012. doi:10.1002/humu.22049
30. Bertier G, Héту M, Joly Y. Unsolved challenges of clinical whole-exome sequencing: a systematic literature review of end-users' views. *BMC Med Genomics.* 2016. doi:10.1186/s12920-016-0213-6
31. Chilamakuri CSR, Lorenz S, Madoui M-A, et al. Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics.* 2014;15(1):449. doi:10.1186/1471-2164-15-449
32. Dillioтт AA, Farhan SMK, Ghani M, et al. Targeted Next-generation Sequencing and Bioinformatics Pipeline to Evaluate Genetic Determinants of Constitutional Disease. *J Vis Exp.* 2018;(134). doi:10.3791/57266
33. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci U S A.* 2009;106(45):19096-19101. doi:10.1073/pnas.0910672106
34. McInerney-Leo AM, Marshall MS, Gardiner B, et al. Whole exome sequencing is an efficient, sensitive and specific method of mutation detection in osteogenesis imperfecta and Marfan syndrome. *Bonekey Rep.* 2013;2:456. doi:10.1038/bonekey.2013.190
35. Wang Z, Liu X, Yang B-Z, Gelernter J. The role and challenges of exome sequencing in studies of human diseases. *Front Genet.* 2013;4:160. doi:10.3389/fgene.2013.00160
36. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med.* 2013;369(16):1502-1511. doi:10.1056/NEJMoa1306555
37. Kaname T, Yanagi K, Naritomi K. A commentary on the promise of whole-exome sequencing in medical genetics. *J Hum Genet.* 2014;59(3):117-118. doi:10.1038/jhg.2014.7
38. Goh G, Choi M. Application of Whole Exome Sequencing to Identify Disease-Causing Variants in Inherited Human Diseases. *Genomics Inform.* 2012. doi:10.5808/GI.2012.10.4.214
39. Dharssi S, Wong-Rieger D, Harold M, Terry S. Review of 11 national policies for rare diseases in the context of key patient needs. *Orphanet J Rare Dis.* 2017;12(1):63. doi:10.1186/s13023-017-0618-0
40. Chong JX, Buckingham KJ, Jhangiani SN, et al. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am J Hum Genet.* 2015. doi:10.1016/j.ajhg.2015.06.009
41. Collins FS. Positional cloning moves from perditional to traditional. *Nat Genet.* 1995. doi:10.1038/ng0495-347
42. Korbel JO, Urban AE, Affourtit JP, et al. Paired-end mapping reveals extensive structural variation in the human genome. *Science (80- ).* 2007. doi:10.1126/science.1149504
43. Boycott KM, Rath A, Chong JX, et al. International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *Am J Hum Genet.* 2017. doi:10.1016/j.ajhg.2017.04.003
44. Hamilton A, Tétreault M, Dymont DA, et al. Concordance between whole-exome sequencing and clinical Sanger sequencing: implications for patient

- care. *Mol Genet genomic Med*. 2016;4(5):504-512. doi:10.1002/mgg3.223
45. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA - J Am Med Assoc*. 2014. doi:10.1001/jama.2014.14601
  46. Cooper DN, Chen J-M, Ball E V, et al. Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum Mutat*. 2010. doi:10.1002/humu.21260
  47. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci*. 2009. doi:10.1073/pnas.0910672106
  48. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: Successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med*. 2011. doi:10.1097/GIM.0b013e3182088158
  49. Rangel-Sosa MM, Figuera-Villanueva LE, González-Ramos IA, et al. Exome sequencing reveals three homozygous missense variants in *SNRPA* in two sisters with syndromic intellectual disability. *Clin Genet*. 2018;93(6):1229-1233. doi:10.1111/cge.13235
  50. Córdova-Fletes C, Becerra-Solano LE, Rangel-Sosa MM, et al. Uncommon runs of homozygosity disclose homozygous missense mutations in two ciliopathy-related genes ( *SPAG17* and *WDR35* ) in a patient with multiple brain and skeletal anomalies. *Eur J Med Genet*. 2018;61(3):161-167. doi:10.1016/j.ejmg.2017.11.011
  51. Chen R, Im H, Snyder M. Whole-exome enrichment with the illumina truseq exome enrichment platform. *Cold Spring Harb Protoc*. 2015. doi:10.1101/pdb.prot084863
  52. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010. doi:10.1093/bioinformatics/btp698
  53. Ornitz DM, Itoh N. The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip Rev Dev Biol*. 2015;4(3):215-266. doi:10.1002/wdev.176
  54. Soong J, Chen Y, Shustef EM, Scott GA. Sema4D, the ligand for Plexin B1, suppresses c-Met activation and migration and promotes melanocyte survival and growth. *J Invest Dermatol*. 2012;132(4):1230-1238. doi:10.1038/jid.2011.414
  55. Bai X, Xiao Z, Pan Y, et al. Cartilage-derived morphogenetic protein-1 promotes the differentiation of mesenchymal stem cells into chondrocytes. *Biochem Biophys Res Commun*. 2004;325(2):453-460. doi:10.1016/J.BBRC.2004.10.055
  56. Schwaerzer GK, Hiepen C, Schrewe H, et al. New insights into the molecular mechanism of multiple synostoses syndrome (SYNS): Mutation within the GDF5 knuckle epitope causes noggin-resistance. *J Bone Miner Res*. 2012;27(2):429-442. doi:10.1002/jbmr.532
  57. Fushimi S, Nohno T, Nagatsuka H, Katsuyama H. Involvement of miR-140-3p in Wnt3a and TGF $\beta$ 3 signaling pathways during osteoblast differentiation in MC3T3-E1 cells. *Genes to Cells*. 2018;23(7):517-527. doi:10.1111/gtc.12591
  58. Xu Y, Gray A, Hardie DG, et al. A novel, de novo mutation in the *PRKAG2* gene: infantile-onset phenotype and the signaling pathway involved. *Am J*

- Physiol Heart Circ Physiol.* 2017;313(2):H283-H292. doi:10.1152/ajpheart.00813.2016
59. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* 2012;13:134. doi:10.1186/1471-2105-13-134
  60. Mercati O, Huguet G, Danckaert A, et al. CNTN6 mutations are risk factors for abnormal auditory sensory perception in autism spectrum disorders. *Mol Psychiatry.* 2017;22(4):625-633. doi:10.1038/mp.2016.61
  61. Fernández-Chacón R, Südhof TC. GENETICS OF SYNAPTIC VESICLE FUNCTION: Toward the Complete Functional Anatomy of an Organelle. *Annu Rev Physiol.* 1999;61(1):753-776. doi:10.1146/annurev.physiol.61.1.753
  62. Born G, Grayton HM, Langhorst H, et al. Genetic targeting of NRXN2 in mice unveils role in excitatory cortical synapse function and social behaviors. *Front Synaptic Neurosci.* 2015;7:3. doi:10.3389/fnsyn.2015.00003
  63. Srouf M, Shimokawa N, Hamdan FF, et al. Dysfunction of the Cerebral Glucose Transporter SLC45A1 in Individuals with Intellectual Disability and Epilepsy. *Am J Hum Genet.* 2017. doi:10.1016/j.ajhg.2017.03.009
  64. Cochat P, Rumsby G. Primary Hyperoxaluria. Ingelfinger JR, ed. *N Engl J Med.* 2013;369(7):649-658. doi:10.1056/NEJMra1301564
  65. Milliner DS, Harris PC, Cogal AG, Lieske JC. *Primary Hyperoxaluria Type 1*; 1993. <http://www.ncbi.nlm.nih.gov/pubmed/20301460>. Accessed May 25, 2019.
  66. Zhang X, Roe SM, Hou Y, et al. Crystal structure of alanine:glyoxylate aminotransferase and the relationship between genotype and enzymatic phenotype in primary hyperoxaluria type 1. *J Mol Biol.* 2003;331(3):643-652. <http://www.ncbi.nlm.nih.gov/pubmed/12899834>. Accessed May 25, 2019.
  67. Zhang X, Roe SM, Hou Y, et al. Crystal structure of alanine:glyoxylate aminotransferase and the relationship between genotype and enzymatic phenotype in primary hyperoxaluria type 1. *J Mol Biol.* 2003. doi:10.1016/S0022-2836(03)00791-5
  68. Mandrile G, van Woerden CS, Berchiolla P, et al. Data from a large European study indicate that the outcome of primary hyperoxaluria type 1 correlates with the AGXT mutation type. *Kidney Int.* 2014;86(6):1197-1204. doi:10.1038/KI.2014.222
  69. Subhani M, Akangire G, Kulkarni A, Wilson GN. Al-Awadi/Raas-Rothschild/Schinzel (AARRS) phocomelia syndrome: Case report and developmental field analysis. *Am J Med Genet Part A.* 2009;149A(7):1494-1498. doi:10.1002/ajmg.a.32890
  70. Howell PL, Turner MA, Christodoulou J, et al. Intragenic complementation at the argininosuccinate lyase locus: Reconstruction of the active site. *J Inherit Metab Dis.* 1998;21(1+):72-85. doi:10.1023/A:1005361724967
  71. Oppici E, Montioli R, Cellini B. Liver peroxisomal alanine:glyoxylate aminotransferase and the effects of mutations associated with Primary Hyperoxaluria Type I: An overview. *Biochim Biophys Acta - Proteins Proteomics.* 2015. doi:10.1016/j.bbapap.2014.12.029
  72. Rumsby G, Williams E, Coulter-Mackie M. Evaluation of mutation screening as a first line test for the diagnosis of the primary hyperoxalurias11The

- nomenclature used in this paper is based on that recommended by Antonarakis SE, and the Nomenclature Working Group (1998): Recommendations for a nomenclature system for human gene mutations. *Hum Mutat* 11:1–3, 1998, where “c” denotes cDNA sequence and nucleotide numbering uses the “A” of the ATG translation initiation start site as +1. *Kidney Int.* 2004;66(3):959-963. doi:10.1111/j.1523-1755.2004.00842.x
73. Orphanet: síndrome de FATCO. [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=EN&Expert=2492](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=EN&Expert=2492). Accessed May 31, 2019.
  74. Courtens W, Jespers A, Harrewijn I, Puylaert D, Vanhoenacker F. Fibular aplasia, tibial campomelia, and oligosyndactyly in a male newborn infant: A case report and review of the literature. *Am J Med Genet Part A.* 2005;134A(3):321-325. doi:10.1002/ajmg.a.30441
  75. Hecht JT, Scott CI. Limb deficiency syndrome in half-sibs. *Clin Genet.* 2008;20(6):432-437. doi:10.1111/j.1399-0004.1981.tb01054.x
  76. Capece G, Fasolino A, Monica M Della, Lonardo F, Scarano G, Neri G. Prenatal diagnosis of femur-fibula-ulna complex by ultrasonography in a male fetus at 24 weeks of gestation. *Prenat Diagn.* 1994;14(6):502-505. doi:10.1002/pd.1970140616
  77. Huber J, Volpon JB, Ramos ES. Fuhrmann syndrome: two Brazilian cases. *Clin Dysmorphol.* 2003;12(2):85-88. doi:10.1097/01.mcd.0000059772.40218.96
  78. Courtens W, Jespers A, Harrewijn I, Puylaert D, Vanhoenacker F. Fibular aplasia, tibial campomelia, and oligosyndactyly in a male newborn infant: A case report and review of the literature. *Am J Med Genet Part A.* 2005;134A(3):321-325. doi:10.1002/ajmg.a.30441
  79. Kitaoka T, Namba N, Kim JY, et al. A Japanese Male Patient with ‘Fibular Aplasia, Tibial Campomelia and Oligodactyly’: An Additional Case Report. *Clin Pediatr Endocrinol.* 2009;18(3):81-86. doi:10.1297/cpe.18.81
  80. Isik E, Atik T, Ozkinay F. The first report of fibular agenesis, tibial campomelia, and oligosyndactyly syndrome with hydrocephaly. *Clin Dysmorphol.* 2019;28(1):38-40. doi:10.1097/MCD.0000000000000247
  81. Abdalla E, El-Beheiry A. Overlap between Fibular Aplasia, Tibial Campomelia, and Oligosyndactyly and Fuhrmann’s Syndromes in an Egyptian Female Infant. *J Pediatr Genet.* 2017;06(02):118-121. doi:10.1055/s-0036-1597931
  82. Ahmad K, Ahmad Malla H, Dawood S. FATCO Syndrome (Fibular Aplasia, Tibial Campomelia, Oligosyndactyly with Talar Aplasia). A Case Study. *Ortop Traumatol Rehabil.* 2017;19(1):75-78. doi:10.5604/15093492.1235280
  83. D’Amato Gutiérrez M, Palacio Díaz FA. Reporte de caso de síndrome FATCO: aplasia bular, campomelia de tibia y oligosindactilia. *Arch Argent Pediatr.* 2016;114(3):e167-70. doi:10.5546/aap.2016.e167
  84. Smets G, Vankan Y, Demeyere A. A Female Newborn Infant with FATCO Syndrome Variant (Fibular Hypoplasia, Tibial Campomelia, Oligosyndactyly) – A Case Report. *J Belgian Soc Radiol.* 2016;100(1):41. doi:10.5334/jbr-btr.929
  85. Goyal N, Kaur R, Gupta M, Bhatti S, Paul R. FATCO Syndrome Variant - Fibular Hypoplasia, Tibial Campomelia and Oligosyndactyly — A Case Report. *J Clin DIAGNOSTIC Res.* 2014;8(9):LD01-2. doi:10.7860/JCDR/2014/9275.4787

86. Sezer O, Gebesoglu I, Yuan B, Karaca E, Gokce E, Gunes S. Fibular aplasia, tibial campomelia, and oligosyndactyly. *Clin Dysmorphol.* 2014;23(4):121-126. doi:10.1097/MCD.0000000000000051
87. Bieganski T, Jamsheer A, Sowinska A, et al. Three new patients with FATCO: Fibular agenesis with ectrodactyly. *Am J Med Genet Part A.* 2012;158A(7):1542-1550. doi:10.1002/ajmg.a.35369
88. Ekbote A V., Danda S. A Case Report of Fibular Aplasia, Tibial Campomelia, and Oligosyndactyly (FATCO) Syndrome Associated With Klinefelter Syndrome and Review of the Literature. *Foot Ankle Spec.* 2012;5(1):37-40. doi:10.1177/1938640011422594
89. Vyskocil V, Dortova E, Dort J, Chudacek Z. FATCO syndrome - fibular aplasia, tibial campomelia and oligosyndactyly. *Jt Bone Spine.* 2011;78(2):217-218. doi:10.1016/j.jbspin.2010.08.013
90. Karaman A, Kahveci H. A male newborn infant with fatco syndrome (fibular aplasia, tibial campomelia and oligodactyly): a case report. *Genet Couns.* 2010;21(3):285-288. <http://www.ncbi.nlm.nih.gov/pubmed/20964118>. Accessed May 24, 2019.
91. Courtens W, Jespers A, Harrewijn I, Puylaert D, Vanhoenacker F. Fibular aplasia, tibial campomelia, and oligosyndactyly in a male newborn infant: A case report and review of the literature. *Am J Med Genet Part A.* 2005;134A(3):321-325. doi:10.1002/ajmg.a.30441
92. Choi IH, Kumar SJ, Bowen JR. Amputation or limb-lengthening for partial or total absence of the fibula. *J Bone Joint Surg Am.* 1990;72(9):1391-1399. <http://www.ncbi.nlm.nih.gov/pubmed/2229119>. Accessed May 26, 2019.
93. Lenz W, Zygulska M, Horst J. FFU complex: an analysis of 491 cases. *Hum Genet.* 1993;91(4):347-356. doi:10.1007/bf00217355
94. OMIM Entry - 228200 - FEMUR-FIBULA-ULNA SYNDROME. <https://www.omim.org/entry/228200?search=OMIM228200&highlight=228200> omim. Accessed August 1, 2019.
95. Matsushita M, Kitoh H, Kaneko H, et al. A novel SOX9 H169Q mutation in a family with overlapping phenotype of mild campomelic dysplasia and small patella syndrome. *Am J Med Genet Part A.* 2013;161A(10):n/a-n/a. doi:10.1002/ajmg.a.36134
96. Evans JA, Reed MH, Greenberg CR. Fibular aplasia with ectrodactyly. *Am J Med Genet.* 2002;113(1):52-58. doi:10.1002/ajmg.10754
97. OMIM Entry - 165590 - OROFACIODIGITAL SYNDROME X; OFD10. <https://www.omim.org/entry/165590?search=OMIM165590&highlight=165590> omim. Accessed August 1, 2019.
98. Peippo M, Valanne L, Perhomaa M, Toivanen L, Ignatius J. Weismann-Netter syndrome and mental retardation: A new patient and review of the literature. *Am J Med Genet Part A.* 2009;149A(11):2593-2601. doi:10.1002/ajmg.a.33019
99. Lewin SO, Opitz JM. Fibular a/hypoplasia: review and documentation of the fibular developmental field. *Am J Med Genet Suppl.* 1986;2:215-238. <http://www.ncbi.nlm.nih.gov/pubmed/3146293>. Accessed May 24, 2019.
100. Niswander L. Pattern formation: old models out on a limb. *Nat Rev Genet.* 2003;4(2):133-143. doi:10.1038/nrg1001



101. Niswander L. Pattern formation: Old models out on a limb. *Nat Rev Genet.* 2003. doi:10.1038/nrg1001
102. Lappin TRJ, Grier DG, Thompson A, Halliday HL. HOX genes: seductive science, mysterious mechanisms. *Ulster Med J.* 2006;75(1):23-31. <http://www.ncbi.nlm.nih.gov/pubmed/16457401>. Accessed May 24, 2019.
103. Hirst J. Mitochondrial Complex I. *Annu Rev Biochem.* 2013;82(1):551-575. doi:10.1146/annurev-biochem-070511-103700
104. Grigorieff N. Structure of the respiratory NADH:ubiquinone oxidoreductase (complex I). *Curr Opin Struct Biol.* 1999;9(4):476-484. doi:10.1016/S0959-440X(99)80067-0
105. Nehls U, Friedrich T, Schmiede A, Ohnishi T, Weiss H. Characterization of assembly intermediates of NADH: Ubiquinone oxidoreductase (complex I) accumulated in *Neurospora* mitochondria by gene disruption. *J Mol Biol.* 1992;227(4):1032-1042. doi:10.1016/0022-2836(92)90519-P
106. Duarte M, Pópulo H, Videira A, Friedrich T, Schulte U. Disruption of iron-sulphur cluster N2 from NADH: ubiquinone oxidoreductase by site-directed mutagenesis. *Biochem J.* 2002;364(Pt 3):833-839. doi:10.1042/BJ20011750
107. Ugalde C, Vogel R, Huijbens R, van den Heuvel B, Smeitink J, Nijtmans L. Human mitochondrial complex I assembles through the combination of evolutionary conserved modules: a framework to interpret complex I deficiencies. *Hum Mol Genet.* 2004;13(20):2461-2472. doi:10.1093/hmg/ddh262
108. Araújo MR, Marques C, Freitas S, Santa-Bárbara R, Alves J, Xavier C. Marfan Syndrome: new diagnostic criteria, same anesthesia care? Case report and review. *Brazilian J Anesthesiol (English Ed.)* 2016;66(4):408-413. doi:10.1016/J.BJANE.2014.09.004
109. Ammash NM, Sundt TM, Connolly HM. Marfan Syndrome—Diagnosis and Management. *Curr Probl Cardiol.* 2008;33(1):7-39. doi:10.1016/J.CPCARDIOL.2007.10.001
110. Dietz HC. Potential Phenotype–Genotype Correlation in Marfan Syndrome. *Circ Cardiovasc Genet.* 2015;8(2):256-260. doi:10.1161/CIRCGENETICS.115.001040
111. Mátyás G, Arnold E, Carrel T, et al. Identification and in silico analyses of novel TGFBR1 and TGFBR2 mutations in Marfan syndrome-related disorders. *Hum Mutat.* 2006;27(8):760-769. doi:10.1002/humu.20353
112. Horbelt D, Guo G, Robinson PN, Knaus P. Quantitative analysis of TGFBR2 mutations in Marfan-syndrome-related disorders suggests a correlation between phenotypic severity and Smad signaling activity. *J Cell Sci.* 2010;123(24):4340-4350. doi:10.1242/jcs.074773
113. Lin S-E, Oyama T, Nagase T, Harigaya K, Kitagawa M. Identification of New Human Mastermind Proteins Defines a Family That Consists of Positive Regulators for Notch Signaling. *J Biol Chem.* 2002;277(52):50612-50620. doi:10.1074/jbc.M209529200
114. Grabias BM, Konstantopoulos K. Notch4-dependent antagonism of canonical TGF- $\beta$ 1 signaling defines unique temporal fluctuations of SMAD3 activity in sheared proximal tubular epithelial cells. *Am J Physiol Physiol.* 2013;305(1):F123-F133. doi:10.1152/ajprenal.00594.2012

115. MacFarlane EG, Parker SJ, Shin JY, et al. Lineage-specific events underlie aortic root aneurysm pathogenesis in Loeys-Dietz syndrome. *J Clin Invest*. 2019;129(2):659-675. doi:10.1172/JCI123547
116. Schepers D, Tortora G, Morisaki H, et al. A mutation update on the LDS-associated genes TGFB2/3 and SMAD2/3. *Hum Mutat*. 2018;39(5):621-634. doi:10.1002/humu.23407
117. Hu J, Liao J, Sathanoori M, et al. CNTN6 copy number variations in 14 patients: a possible candidate gene for neurodevelopmental and neuropsychiatric disorders. *J Neurodev Disord*. 2015;7(1):26. doi:10.1186/s11689-015-9122-9
118. Chuang H-C, Huang T-N, Hsueh Y-P. T-Brain-1 - A Potential Master Regulator in Autism Spectrum Disorders. *Autism Res*. 2015;8(4):412-426. doi:10.1002/aur.1456
119. Petrone A, Schachner M, Wang C, et al. Receptor protein tyrosine phosphatase is essential for hippocampal neuronal migration and long-term potentiation. *EMBO J*. 2003;22(16):4121-4131. doi:10.1093/emboj/cdg399
120. van Daalen E, Kemner C, Verbeek NE, et al. Social responsiveness scale-aided analysis of the clinical impact of copy number variations in autism. *Neurogenetics*. 2011;12(4):315-323. doi:10.1007/s10048-011-0297-2
121. Cheng H, Dharmadhikari A V, Varland S, et al. Truncating Variants in NAA15 Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies. *Am J Hum Genet*. 2018;102(5):985-994. doi:10.1016/j.ajhg.2018.03.004
122. Oguro-Ando A, Zuko A, Kleijer KTE, Burbach JPH. A current view on contactin-4, -5, and -6: Implications in neurodevelopmental disorders. *Mol Cell Neurosci*. 2017;81:72-83. doi:10.1016/J.MCN.2016.12.004
123. Gaugler T, Klei L, Sanders SJ, et al. Most genetic risk for autism resides with common variation. *Nat Genet*. 2014;46(8):881-885. doi:10.1038/ng.3039
124. Huguet G, Ey E, Bourgeron T. The Genetic Landscapes of Autism Spectrum Disorders. *Annu Rev Genomics Hum Genet*. 2013;14(1):191-213. doi:10.1146/annurev-genom-091212-153431
125. Kapitein LC, Schlager MA, Kuijpers M, et al. Mixed microtubules steer dynein-driven cargo transport into dendrites. *Curr Biol*. 2010;20(4):290-299. doi:10.1016/j.cub.2009.12.052
126. Stessman HAF, Xiong B, Coe BP, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat Genet*. 2017;49(4):515-526. doi:10.1038/ng.3792
127. Siu W-K, Lam C-W, Gao W-W, Vincent Tang H-M, Jin D-Y, Mak CM. Unmasking a novel disease gene NEO1 associated with autism spectrum disorders by a hemizygous deletion on chromosome 15 and a functional polymorphism. *Behav Brain Res*. 2016;300:135-142. doi:10.1016/J.BBR.2015.10.041
128. Yuen RKC, Thiruvahindrapuram B, Merico D, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med*. 2015;21(2):185-191. doi:10.1038/nm.3792
129. Lin Y-C, Frei JA, Kilander MBC, Shen W, Blatt GJ. A Subset of Autism-Associated Genes Regulate the Structural Stability of Neurons. *Front Cell*

*Neurosci.* 2016;10:263. doi:10.3389/fncel.2016.00263

130. García-Ortiz JE, Zarazúa-Niño AI, Hernández-Orozco AA, et al. Case Report: Whole Exome Sequencing Unveils an Inherited Truncating Variant in CNTN6 (p.Ser189Ter) in a Mexican Child with Autism Spectrum Disorder. *J Autism Dev Disord.* March 2019. doi:10.1007/s10803-019-03951-z

## CAPÍTULO XIII

### ANEXOS

#### PROTOCOLO PREPARACIÓN DE LA BIBLIOTECA

##### 1. TAGMENTACIÓN DEL DNA GENÓMICO.

En este proceso se marca el DNAg mediante el transposoma de Nextera. El transposoma de Nextera fragmenta de manera simultánea y aleatoria el DNAg y añade secuencias de adaptadores a los extremos, lo que permite la amplificación mediante PCR en los procesos posteriores.

##### Consumibles.

- **SPB**- Sample Purification Beads (Perlas de purificación de muestra).
- **RSB**- Resuspension Buffer (Buffer de resuspensión).
- **ST**- Stop Tagment Buffer (Buffer de detención de Tagmentación).
- **TD** -Tagment DNA Buffer (Buffer de Tagmentación de DNA).
- **TDE2** -Tagment DNA Enzyme 2 (Enzima de Tagmentación de DNA 2).
- **DNAg** - (50ng por muestra).
- Tubo de PCR de pared delgada de 0,2 ml.

##### Acerca de Reactivos.

- Para SPB utilizar vórtex antes de cada uso.
- Para SPB utilizar vórtex con frecuencia para asegurarse de que las cuentas se distribuyan de manera uniforme.
- Aspirar y dosificar SPB lentamente debido a la viscosidad de la solución.

##### Preparación.

Preparar los siguientes consumibles.

Reactivo	Almacenamiento	Instrucciones
<b>DNAg</b>	-25°C to -15°C	Descongelar en hielo. Invertir suavemente los tubos descongelados 3-5 veces, y luego centrifugar brevemente.
<b>RSB</b>	-25°C to -15°C	Descongelar a temperatura ambiente. Almacenar a 2 °C a 8 °C después de la descongelación inicial para su uso posterior en el protocolo.
<b>TD</b>	-25°C to -15°C	Descongelar en hielo. Invertir suavemente los tubos descongelados 3-5 veces, y luego centrifugar brevemente.
<b>TDE2</b>	-25°C to -15°C	Descongelar en hielo. Invertir suavemente los tubos descongelados 3-5 veces, y luego

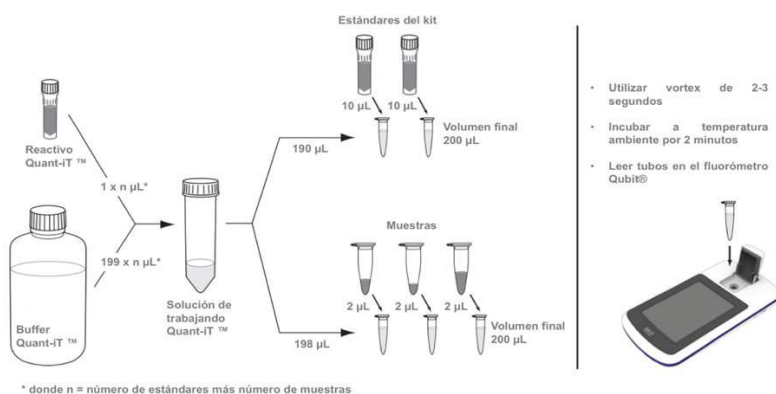
		centrifugar brevemente. Ponga a un lado en el hielo.
<b>SPB</b>	2°C to 8°C	Dejar reposar durante 30 minutos para llevar a temperatura ambiente. Mantener a temperatura ambiente para su uso posterior en el protocolo.
<b>ST</b>	15°C to 30°C	Comprobar si hay precipitados. Si está presente, agite en el vórtex hasta que todas las partículas estén resuspendidas.

2. Guarde el siguiente programa **TAG58** en el termociclador:

- Elija la opción de tapa de precalentamiento y ajústela a 100 °C.
- 58 ° C durante 10 minutos.
- Mantener a 10 °C.
- Cada tubo contiene 50µL.
- Guarde el siguiente programa **TAG60** en el termociclador:
- Elija la opción de tapa de precalentamiento y ajústela a 100 °C.
- 60 °C durante 5 minutos.
- Mantener a 10 °C.
- Cada tubo contiene 65µL.

### Procedimiento.

- Cuantificación y normalización de DNAg.
- Cuantificación del DNA utilizando un método espectroscópico NanoDrop.
- Realizar una primera dilución del DNAg a un volumen final de 20µL en 50ng/µL (1000ng total).
- Cuantificar el DNAg utilizando un método fluorométrico Qubit.



### Visión general para usar el ensayo Quant-iT™ dsDNA HS en Qubit

- Normalizar el DNAg a 10ng/µl utilizando agua libre de nucleasas.
- Recuantificar el DNAg normalizado utilizando el equipo Qubit.

- Diluir el DNAg normalizado en agua libre de nucleasas a un volumen final de **10µL en 5ng/µL** (50ng total) y cuantificar nuevamente (contemplar 2µl para la cuantificación).

### **Tagmentación del DNA.**

- Añadir los siguientes elementos en el orden indicado a un tubo de PCR nuevo de 0.2 ml.
- TD (25µL).
- DNAg normalizado (10µL).
- TDE2 (15µL).
- Mezclar bien por pipeteo.
- Centrifugar brevemente.
- Colocar en el termociclador preprogramado y ejecutar el programa **TAG58**.

**NOTA:** Después de terminar el programa TAG58 inmediatamente agregar el reactivo ST para inactivar el trasposoma.

- Agregue 15µl de ST al tubo.
- Mezclar bien por pipeteo.
- Centrifugar brevemente.
- Colocar en el termociclador preprogramado y ejecutar el programa **TAG60**.

## **2. LIMPIEZA DEL DNA TAGMENTADO**

En este proceso se purifica el DNA marcado por el transposoma de Nextera utilizando perlas de purificación de muestra (SPB). Este paso es crítico porque el transposoma de Nextera se puede unir fuertemente a los extremos del DNA e interferirá en los procesos descendentes si no se elimina.

### **Consumibles.**

- **RSB-** Resuspensión Buffer (Tampón de resuspensión).
- **SPB-** Sample Purification Beads (Perlas de purificación de muestra).
- **Etanol al 80%** recién preparado (EtOH).
- Placa midi de 96 pocillos.
- Sellos adhesivos Microseal 'B'.

### **Acerca de los reactivos.**

- Para SPB utilizar vórtex antes de cada uso.
- Para SPB utilizar vórtex con frecuencia para asegurarse de que las cuentas estén distribuidas uniformemente.
- Aspire y dispense SPB lentamente debido a la viscosidad de la solución.
- Reduzca al mínimo las salpicaduras y la formación de espuma durante la mezcla para evitar la transferencia del tampón ST a la amplificación por PCR. El arrastre puede reducir el rendimiento final antes del enriquecimiento.

### Preparación.

Prepare los siguientes consumibles:

Reactivo	Almacenamiento	Instrucciones
RSP	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
SPB	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.

Prepare EtOH fresco al 80%.

### Procedimiento.

- Transfiera todo el sobrenadante al pozo correspondiente de una nueva placa midi.
- Agregue 52µl de SPB a cada pocillo o tubo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 5 minutos.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Transfiera 98µl de sobrenadante a un nuevo pozo de una placa midi.
- Agregue 137µl de SPB al pocillo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 5 minutos.
- Coloque sobre un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Retire y deseche todo el sobrenadante del pocillo.
- Lave 2 veces de la siguiente manera.
- Agregue 200µl de EtOH al 80% recién preparado al pocillo.
- Incube en el soporte magnético durante 30 segundos.
- Retire y deseche todo el sobrenadante de cada pocillo.
- Con una pipeta de 20µl, elimine el EtOH residual del pocillo.
- Seque al aire sobre el soporte magnético durante 5 minutos.
- Retire del soporte magnético la placa midi.
- Agregue 22.5µl de RSB al pocillo, y luego mezcle bien por pipeteo hasta que el pellet se resuspenda.
- Incubar a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Transfiera 20µl del sobrenadante a un nuevo tubo de PCR de 0.2 ml.

### 3. AMPLIFICACIÓN DE DNA TAGMENTADO.

Este proceso amplifica el DNA marcado purificado mediante un programa de PCR de 10 ciclos. También añade las secuencias del índice 1 (i7) y el índice 2 (i5) necesarias para la secuenciación, así como los adaptadores comunes (P5 y P7) necesarios para la generación de grupos y la secuenciación.

### Consumibles.

- Adaptadores de índice 1 (i7) y tapas de tubo naranja.
- Adaptadores de índice 2 (i5) y tapas de tubo blanco.
- LAM (Library Amplification Mix).
- Tubo de PCR de 0,2 ml.

### Preparación.

Prepare los siguientes consumibles.

Reactivo	Almacenamiento	Instrucciones
Adaptadores de índice (i5 y i7)	-25 ° C a -15 ° C	Descongelar a temperatura ambiente durante 20 minutos. Agite en el Vórtex cada tubo para mezclar.
LAM	-25 ° C a -15 ° C	Descongelar en hielo.

- Guarde el siguiente programa **LAM AMP** en el termociclador:
- Elija la opción de tapa de precalentamiento y ajústela a 100 °C.
- 72 °C durante 3 minutos.
- 98 °C durante 30 segundos.
- **10 ciclos de:**
- 98 °C por 10 segundos.
- 60 °C por 30 segundos.
- 72 °C durante 30 segundos.
- 72 °C durante 5 minutos.
- Mantener a 10 °C.

### Procedimiento.

- Agregue 5µl de adaptador Index 1 (i7) y coloque una nueva tapa naranja en cada tubo adaptador i7.
- Agregue 5µl de cada adaptador Index 2 (i5) y coloque una nueva tapa blanca en cada tubo adaptador i5.
- Agregue 20µl de LAM al tubo, y luego mezcle bien por pipeteo.
- Centrifugar brevemente.
- Coloque en el termociclador preprogramado y ejecute el programa **LAM AMP**.

**PUNTO DE PAROSEGURO-** Si desea detenerse, selle el tubo y almacene de 2 °C a 8 °C hasta por 2 días.

## 4. PRIMERA LIMPIEZA DEL DNA AMPLIFICADO

En este proceso se emplean perlas de purificación de muestras (SPB) para purificar el DNA de la biblioteca y eliminar los productos no deseados.

### Consumibles.

- **RSB** (tampón de resuspensión).
- **SPB** (Perlas de purificación de muestra).



- **Etanol** al 80% recién preparado (EtOH).
- Placa midi de 96 pocillos.
- Sellos adhesivos Microseal 'B'.

#### **Acerca de reactivos**

- Agitar en vórtex las perlas de purificación antes de cada uso para asegurarse de que las cuentas estén distribuidas uniformemente.
- Aspirar y dispensar las SPB lentamente debido a la viscosidad de la solución.

#### **Preparación.**

Prepare los siguientes consumibles.

<b>Reactivo</b>	<b>Almacenamiento</b>	<b>Instrucciones</b>
<b>RSB 2</b>	2 °C a 8 °C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
<b>SPB 2</b>	2 °C a 8 °C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.

Prepare EtOH fresco al 80%.

#### **Procedimiento.**

- Centrifugar brevemente el tubo.
- Transfiera 50µl de volumen total a un nuevo pocillo de una placa midi.
- Agregue 90µl de SPB a cada pocillo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 5 minutos.
- Centrifugar brevemente.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descartar el sobrenadante de cada pocillo.
- Lave 2 veces de la siguiente manera.
- Agregue 200µl de EtOH al 80% recién preparado al pocillo.
- Incube en el soporte magnético durante 30 segundos.
- Descarte el sobrenadante del pocillo.
- Con una pipeta de 20µl, elimine el EtOH residual del pocillo.
- Seque al aire sobre el soporte magnético durante 5 minutos.
- Retire del soporte magnético.
- Agregue 17µl de RSB al pocillo, y luego mezcle bien por pipeteo.
- Incubar a temperatura ambiente durante 2 minutos.
- Centrifugar a 280x g por 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Transfiera 15µl del sobrenadante a un nuevo tubo de PCR de 0,2 ml o a un tubo de microcentrifuga de 0.5 ml si se desea almacenar.
- Cuantifique la biblioteca utilizando un método fluorométrico (Qubit).

**PUNTO DE PARADA SEGURO-** Si desea detenerse, selle el tubo y almacene de -25 °C a -15 °C por hasta 14 días.

## 5. PRIMERA HIBRIDACIÓN DE SONDAS

Este proceso mezcla la biblioteca de DNA con sondas de captura en regiones objetivo de interés. El tiempo de hibridación recomendado garantiza que las regiones objetivo se unan con las sondas de captura completamente.

### Consumibles

- **BLR** (Bloqueador).
- **CEX** (Codificación Exoma Oligos).
- **EHB1** (Tampón de enriquecimiento de hibridación 1).
- **EHB2** (Tampón de enriquecimiento de hibridación 2).
- **RSB** (Tampón de resuspensión).
- **SPB** (Perlas de purificación de muestra).
- Etanol al 80% recién preparado (EtOH).
- Placa midi de 96 pocillos.
- Sello adhesivo Microseal 'B'.

### Acerca de reactivos.

- Antes de usar BLR, agite en vórtex para resuspender la solución. Asegúrate de que no haya estructuras de cristal presentes. Si se observan cristales y turbidez, agite en vórtex hasta que la solución esté clara.
- Antes de utilizar SPB utilice vórtex.
- Antes de utilizar SPB asegurarse de que las cuentas estén distribuidas uniformemente.
- Aspire y dispense SPB lentamente debido a la viscosidad de la solución.

### Preparación.

Prepare los siguientes consumibles.

Reactivo	Almacenamiento	Instrucciones
<b>BLR</b>	-25°C a -15°C	Descongele a temperatura ambiente.
<b>CEX</b>	-25°C a -15°C	Descongele a temperatura ambiente.
<b>EHB1</b>	-25°C a -15°C	Descongele a temperatura ambiente.
<b>RSB</b>	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
<b>SPB</b>	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
<b>EHB2</b>	5°C a 30°C	Eliminar del almacenamiento.

- Guarde el programa TRE HYB en el termociclador:
- Elija la opción de tapa de precalentamiento y ajústela a 100 °C.
- 95 °C durante 10 minutos.

- 58 °C durante 30 minutos.
- Cada pozo o tubo contiene 10µl.

**Procedimiento:** Asegurarse de tener una librería a una **concentración de 500ng en 30µl** acorde al resultado obtenido en el ensayo de Qubit. Utilizar RSB para realizar la dilución.

**Utilizar la siguiente formula para realizar la dilución:**

Supongamos que al cuantificar la biblioteca utilizando el Qubit obtenemos una concentración de 33.2 ng.

Entonces:

$$1\mu l > 33.2ng$$

$$X < 500ng$$

$$X = 15.06\mu l \text{ de biblioteca de DNA}$$

Diluir en 14.94µl de RSB

para obtener una librería a una concentración

de 500ng en un volumen de 30µl.

- Una vez obtenida la librería a una concentración de 500ng en 30µl, agregue los siguientes reactivos en el orden indicado a un pocillo de una placa midi.
- Biblioteca de DNA (30µl).
- BLR (10µl).
- CEX (10µl).
- Mezcle bien de por pipeteo.
- Centrifugar a 280 x g por 1 minuto.
- Agregue 125µl de SPB al pocillo, y luego mezcle bien por pipeteo.
- Incubar a temperatura ambiente durante 10 minutos.
- Centrifugar a 280 x g por 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descarte todo el sobrenadante del pocillo.
- Lave 2 veces de la siguiente manera.
- Agregue 200µl de 80% de EtOH recién preparado al pocillo.

- Incube en el soporte magnético durante 30 segundos.
- Descarte el sobrenadante del pocillo.
- Con una pipeta de 20µl, elimine los residuos de EtOH del pocillo.
- Seque al aire sobre el soporte magnético durante 10 minutos.
- Retire del soporte magnético.
- Agregue 7.7µl de EHB1 al pocillo, y luego mezcle por pipeteo.
- Incube a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (~ 2 minutos).
- Transfiera 7.5µl del sobrenadante a un nuevo tubo de PCR.
- Agregue 2,5µl de EHB2 a un tubo nuevo de PCR, y luego mezcle bien por pipeteo.
- Centrifugar brevemente.
- Coloque en el termociclador preprogramado y ejecute el programa **TRE HYB**.

## 6. PRIMERA CAPTURA DE SONDAS HÍBRIDAS.

Este proceso emplea perlas magnéticas de estreptavidina (SMB) para capturar sondas hibridadas en regiones objetivo de interés. Se realizan dos procedimientos de lavado caliente para eliminar la unión no específica de las perlas. La biblioteca enriquecida se eluye de las perlas y se prepara para una segunda ronda de hibridación. Es necesario el uso del equipo y las temperaturas adecuadas para garantizar la extracción de DNA no específico, así como la retención de las regiones objetivo.

### Consumibles.

- **EE1** - Enrichment Elution Buffer (Tampón de elución de enriquecimiento 1).
- **ET2** - Elute Target Buffer 2.
- **EEW** - Enhanced enrichment wash solution (solución mejorada de lavado enriquecido).
- **HP3** - (NaOH 2 N).
- **RSB** - Resuspensión buffer (Tampón de resuspensión).
- **SMB** - Streptavidin Magnetic Beads (Perlas magnéticas de estreptavidina).
- Placas midi de 96 pocillos.
- Tubos de microcentrifuga de 1.5 ml.
- Sellos adhesivos Microseal 'B'.

### Acerca de reactivos.

- EEW debe estar a temperatura ambiente para su uso.
- Proteger EEW de la luz.
- Asegúrese de utilizar SMB (tubo de 2 ml) y no SPB (tubo de 15 ml) para este procedimiento.
- Invierte SMB para mezclar antes de usar.
- Deseche la premezcla de elución después del uso.

### Preparación.

Prepare los siguientes consumibles.

Reactivo	Almacenamiento	Instrucciones
EE1	-25°C a -15°C	Descongele a temperatura ambiente. Almacene después de su uso.
EEW	-25°C a -15°C	Descongele a temperatura ambiente. Almacene después de su uso.
HP3	-25°C a -15°C	Descongele a temperatura ambiente. Almacene después de su uso.
ET2	2°C a 8°C	Deje reposar a temperatura ambiente. Regrese al almacenamiento después del uso.
SMB	2°C a 8°C	Deje reposar por 30 minutos para llevar a temperatura ambiente. Regrese al almacenamiento después del uso.
RSB	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.

- Precalentar el sistema de microcalentamiento con inserto para placa midi a 50 °C.
- **Procedimiento.**
- Primer enlace.
- Centrifugar brevemente.
- Transferir todo el contenido (~ 10 µl) a un pocillo nuevo de una placa midi.
- Agregar 250µl de SMB a cada pocillo, y luego mezcle bien utilizando la pipeta.
- Incubar a temperatura ambiente durante 25 minutos.
- Centrifugar a 280 x g por 1 minuto.
- Colocar en un soporte magnético y esperar hasta que el líquido esté claro (2-5 minutos).
- Descartar todo el sobrenadante del pocillo.
- Retirar del soporte magnético.

### Primer lavado.

- Agregar 200µl de EEW al pocillo, y luego resuspender el pellet de perlas utilizando la micropipeta.

**NOTA** - Se requiere una resuspensión adecuada para garantizar la eliminación eficiente de DNA inespecífico de la reacción, que de lo contrario resulta en pobres estadísticas de enriquecimiento.

- Sellar la placa con los adhesivos Microseal 'B' e incubar a 50 °C la placa en el sistema de microcalentamiento Hybex durante 30 minutos.
- Colocar inmediatamente sobre un soporte magnético y esperar hasta que el líquido esté claro (~ 2 minutos).
- Descartar todo el sobrenadante del pocillo.

- Retirar del soporte magnético.
- Repetir los pasos 1-5 para un total de 2 lavados.

**Primera elución.**

- Crear la premezcla de elución en un tubo de microcentrífuga de 1,5 ml y luego mezclar brevemente en vórtex.
- EE1 (28.5µl).
- HP3 (1.5µl).
- Agregar 23µl de premezcla de elución al pocillo, y luego mezclar bien por pipeteo.
- Incubar a temperatura ambiente durante 2 minutos.
- Centrifugar a 280x g por 1 minuto.
- Colocar en un soporte magnético y espere hasta que el líquido esté claro (~ 2 minutos).
- Transferir 21µl de sobrenadante a un nuevo pozo de una placa midi o tubo de microcentrífuga de 0.5 ml si se desea hacer uso del punto de paro.
- Agregar 4µl de ET2 al pocillo o tubo, y luego mezclar por pipeteo.
- Agregar 5µl de RSB a cada pocillo o tubo, y luego mezcle bien por pipeteo.
- Centrifugar brevemente.

**PUNTO DE PARO SEGURO** - Si se detiene, tape el tubo y almacene de -25 °C a -15 °C hasta por 7 días.

**7. SEGUNDA HIBRIDACIÓN.**

Este proceso combina la biblioteca de DNA eluida de la primera ronda de enriquecimiento con sondas de captura adicionales en regiones objetivo de interés. Esta segunda hibridación es necesaria para garantizar la alta especificidad de las regiones capturadas.

**Consumibles.**

- Antes de usar BLR, agite en vórtex para resuspender la solución. Asegúrate de que no haya estructuras de cristal presentes. Si se observan cristales y turbidez, agite en vórtex hasta que la solución esté clara.
- Utilice vórtex SPB antes de cada uso.
- Utilice vórtex SPB con frecuencia para asegurarse de que las cuentas estén distribuidas uniformemente.
- Aspirar y dispensar SPB lentamente debido a la viscosidad de la solución.

Reactivo	Almacenamiento	Instrucciones
<b>BLR</b> (Bloqueador)	-25°C a -15°C	Descongele a temperatura ambiente.
<b>CEX</b> (Codificación Exoma Oligos)	-25°C a -15°C	Descongele a temperatura ambiente.

<b>EHB1</b> (Tampón de hibridación de enriquecimiento 1)	-25°C a -15°C	Descongele a temperatura ambiente.
<b>RSB</b>	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
<b>SPB</b> (Perlas de purificación de muestra)	2°C a 8°C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
<b>EHB2</b> (Tampón de hibridación de enriquecimiento 2)	2°C a 8°C	Descongele a temperatura ambiente.
<b>Etanol al 80%</b>	-	-

### Consumibles.

- Placa midi de 96 pocillos.
- Sellos adhesivos Microseal 'B'.
- Tubo de PCR de 0.2ml.

### Procedimiento.

- Agregar los siguientes reactivos en el orden indicado a un nuevo pocillo de una placa midi.
- Biblioteca de DNA (30µl).
- BLR (10µl).
- CEX (10µl).
- Mezcle bien por pipeteo.
- Centrifugar brevemente.
- Agregue 125µl de SPB al pozo y mezclar por pipeteo.
- Incubar a temperatura ambiente durante 10 minutos.
- Centrifugar brevemente (280 × g durante 1 minuto).
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descartar todo el sobrenadante.
- Lavar 2 veces de la siguiente manera.
  - Agregue 200µl de 80% de EtOH.
  - Incube en el soporte magnético durante 30 segundos.
  - Retire y deseche todo el sobrenadante de cada pocillo o del tubo.
- Con una pipeta de 20µl, elimine los residuos de EtOH 80% del pocillo.
- Seque al aire sobre el soporte magnético durante 10 minutos.
- Agregue 7.7µl de EHB1 al pocillo, y luego mezcle con la pipeta hasta que el pellet se resuspenda.

- Retire del soporte magnético.
- Incube a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (~ 2 minutos).
- Transfiera 7.5µl de sobrenadante a un nuevo tubo de PCR de pared delgada de 0.2ml.
- Agregue 2,5µl de EHB2 al tubo y mezcle por pipeteo.
- Centrifugar brevemente.
- Coloque en el termociclador preprogramado y ejecute el programa **TRE HYB**.  
95°C 10 min  
58°C 30 min
- **Volumen final de la reacción: 10µl**

## 8. SEGUNDA CAPTURA DE SONDAS HÍBRIDAS

Este paso usa Perlas magnéticas de streptavidina (SMB) para capturar las sondas hibridadas en las regiones objetivo de interés. Son necesarios dos lavados calientes para eliminar la unión no específica de las perlas. La biblioteca enriquecida se eluye de las cuentas y se prepara una segunda ronda de hibridación.

### Consumibles y equipos

- **EE1** (Tampón de elución de enriquecimiento 1).
- **ET2** (Elute Target Buffer 2).
- **EEW** (Solución de lavado de enriquecimiento mejorado).
- **HP3** (NaOH 2 N).
- **RSB** (tampón de resuspensión).
- **SMB** (Streptavidin Magnetic Beads).
- Placa midi de 96 pocillos.
- Tubos de microcentrifuga de 1.5 ml.
- Sellos adhesivos Microseal 'B'.
- Sistema de microcalentamiento a 50 °C (Hybex a 50°C).

### Acerca de reactivos.

- **EEW** debe estar a temperatura ambiente para su uso.
- Proteger EEW de la luz.
- Deseche la premezcla de elución después del uso.

### Preparación.

Prepare los siguientes consumibles.

Reactivo	Almacenamiento	Instrucciones
<b>EE1</b>	-25°C to -15°C	Descongele a temperatura ambiente. Almacene después de su uso.
<b>EEW</b>	-25°C to -15°C	Descongele a temperatura ambiente. Almacene después de su uso.



<b>HP3</b>	-25°C to -15°C	Descongele a temperatura ambiente. Almacene después de su uso.
<b>ET2</b>	2°C to 8°C	Deje reposar a temperatura ambiente. Regrese al almacenamiento después del uso.
<b>SMB</b>	2°C to 8°C	Deje reposar por 30 minutos para llevar a temperatura ambiente. Regrese al almacenamiento después del uso.

- Precalentar el sistema Hybex a 50 °C.

#### **Procedimiento.**

##### **Segunda unión.**

- Dar un spin.
- Transferir el contenido del tubo a un pocillo de una placa mid (10µl).
- Agregue 250µl de SMB y mezcle bien por pipeteo.
- Incubar a temperatura ambiente durante 25 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descartar todo el sobrenadante del pocillo.
- Retire del soporte magnético.

##### **Segundo lavado.**

- Agregue 200µl de EEW al pocillo, y luego mezcle bien por pipeteo hasta deshacer el pellet de perlas.

**NOTA** - Se requiere una resuspensión adecuada para garantizar la eliminación eficiente de DNA inespecífico de la reacción, de lo contrario resultaría en pobres estadísticas de enriquecimiento.

- Tapar con sellos adhesivos la placa. Colocar en el sistema Hybex a 50 °C durante 30 minutos.
- Colocar inmediatamente sobre un soporte magnético y espere hasta que el líquido esté claro (~ 2 minutos).
- Descartar el sobrenadante del pocillo.
- Retire del soporte magnético.
- Repita los pasos del 1 al 5 para un total de 2 lavados.

##### **Segunda elución**

- Cree la premezcla de elución en un tubo de microcentrífuga de 1,5 ml y mezclar en vórtex.
- EE1 (28.5µl).
- HP3 (1.5µl).
- Agregue 23µl de premezcla de elución al pocillo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.

- Coloque en un soporte magnético y espere hasta que el líquido esté claro (~ 2 minutos).
- Transfiera 21µl de sobrenadante a un nuevo pocillo de una placa midi.
- Agregue 4µl de ET2 a cada pocillo y luego mezcle bien por pipeteo.
- Centrifugar a 280 × g durante 1 minuto.

## 9. LIMPIEZA DE LA BIBLIOTECA CAPTURADA

Este paso usa perlas de purificación de muestra (SPB) para purificar la biblioteca capturada antes de la amplificación en PCR.

### Consumibles.

- RSB (Tampón de resuspensión).
- SPB (Perlas de purificación de muestra).
- Etanol al 80% recién preparado (EtOH).
- Elija de los siguientes contenedores:
- [Placa] 96-well Hard-Shell 0.3 ml placa de PCR.
- [Placa] Sellos adhesivos Microseal 'B'.

### Acerca de reactivos.

- Utilice vórtex antes de usar SPB para asegurarse de que las perlas estén distribuidas uniformemente.
- Aspire y dispense SPB lentamente debido a la viscosidad de la solución.

### Preparación.

Prepare los siguientes consumibles.

Reactivos	Almacenamiento	Instrucciones
RSB 2	2 a 8 °C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
SPB	2 a 8 °C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.

Prepare EtOH al 80% fresco.

### Procedimiento.

- Utilizar vórtex para distribuir el reactivo SPB hasta que esté bien disperso.
- Agregue 45µl de SPB a cada pocillo o tubo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 5 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descartar todo el sobrenadante de cada pocillo.
- Lavar 2 veces de la siguiente manera.
- Agregue 200µl de EtOH al 80% al pocillo.
- Incubar en el soporte magnético durante 30 segundos.
- Retire y deseche todo el sobrenadante de cada pocillo.
- Use una pipeta de 20µl para eliminar el EtOH residual del pocillo.

- Secar al aire sobre el soporte magnético (~ 5 minutos).
- Remover de la placa magnética.
- Agregue 27.5µl de RSB al pocillo, y luego mezcle bien por pipeteo.
- Incubar a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque sobre un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Transferir 25µl de sobrenadante a un tubo de PCR de 0.2 ml.

**PUNTO DE PARO SEGURO** - Si desea detenerse, selle la placa o tape el tubo y almacene a -25 °C a -15 °C hasta 7 días.

### 10. AMPLIFICACIÓN DE LA BIBLIOTECA ENRIQUECIDA.

En este paso se usa un programa de PCR de 10 ciclos para amplificar la biblioteca enriquecida.

#### Consumibles.

- **EAM** (Mezcla de amplificación de enriquecimiento).
- **PPC** (PCR Primer Cocktail).

#### Preparación

Prepare los siguientes consumibles:

Reactivo	Almacenamiento	Instrucciones
<b>EAM</b>	-25 ° C a -15 ° C	Descongelar en hielo.
<b>PPC</b>	-25 ° C a -15 ° C	Descongelar en hielo.

- Guarde el siguiente programa **AMP10** en el termociclador:
- Elija la opción de tapa de precalentamiento y ajústela a 100 ° C
- 98 ° C durante 30 segundos
- 10 ciclos de:
- 98 ° C por 10 segundos
- 60 ° C por 30 segundos
- 72 ° C durante 30 segundos
- 72 ° C durante 5 minutos
- Mantener a 10 °C
- **Volumen final: 50µl**

#### Procedimiento.

- Agregue 5µl de PPC al tubo.
- Agregue 20µl de EAM al tubo, y luego mezcle bien.
- Centrifugar brevemente.
- Coloque en el termociclador preprogramado y ejecute el programa **AMP10**.

**PUNTO DE PARO SEGURO** - Si deseas detenerte, almacene el tubo de 2 a 8 °C durante un máximo de 2 días.

## 11. LIMPIEZA DE LA BIBLIOTECA AMPLIFICADA Y ENRIQUECIDA.

Este paso utiliza perlas de purificación de muestra (SPB) para purificar la biblioteca enriquecida y eliminar productos no deseados.

### Consumibles:

- **RSB** (Tampón de resuspensión).
- **SPB** (Perlas de purificación de muestra).
- **Etanol al 80%** recién preparado (EtOH).  
Elija de los siguientes contenedores:
- [Placa] placa midi de 96 pocillos y placa de PCR de 96 pocillos con carcasa dura de 0.3 ml.
- [Tubo] Tubos de microcentrifuga de 1.5 ml o tiras de 8 tubos.
- [Placa] Sellos adhesivos Microseal 'B'.

### Acerca de reactivos.

- Utilice vórtex SPB antes de cada uso.
- Utilice vórtex SPB con frecuencia para asegurarse de que las cuentas estén distribuidas uniformemente.
- aspire y dispense SPB lentamente debido a la viscosidad de la solución.

### Preparación.

Prepare los siguientes consumibles:

Reactivos	Almacenamiento	Instrucciones
RSB	2 ° C a 8 ° C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.
SPB	2 ° C a 8 ° C	Deje reposar durante 30 minutos para llevar a temperatura ambiente.

Prepare EtOH fresco al 80%.

### Procedimiento.

- Centrifugar brevemente el tubo de PCR que contiene el producto de amplificación de la biblioteca enriquecida.
- Utilizar vórtex hasta que las SPB esté bien dispersas.
- Transferir 50µl del resultado de la PCR a un nuevo pocillo de una placa midi.
- Agregue 50µl de SPB a un pocillo, y luego mezclar bien por pipeteo.
- Incube a temperatura ambiente durante 5 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Descarte todo el sobrenadante del pocillo.
- Lave 2 veces de la siguiente manera.
- Agregue 200µl de EtOH al 80% al pocillo.
- Incube en el soporte magnético durante 30 segundos.
- Descartar todo el sobrenadante del pocillo.

- Use una pipeta de 20µl para eliminar el EtOH residual de cada pocillo.
- Seque al aire sobre el soporte magnético (~ 5 minutos).
- Retire del soporte magnético.
- Agregue 20µl de RSB a cada pocillo o tubo, y luego mezcle bien por pipeteo.
- Incube a temperatura ambiente durante 2 minutos.
- Centrifugar a 280 × g durante 1 minuto.
- Coloque en un soporte magnético y espere hasta que el líquido esté claro (2-5 minutos).
- Transfiera 20µl de sobrenadante a un nuevo tubo de microcentrífuga de 0.5 ml. Este tubo contendrá la biblioteca enriquecida.

**PUNTO DE PARO SEGURO** - Si desea detenerse, almacene de -25 a -15 °C hasta por 7 días.

## 12. COMPROBACIÓN DE LAS BIBLIOTECAS ENRIQUECIDAS.

Realice los siguientes procedimientos para verificar la calidad de la biblioteca enriquecida.

### Cuantificación de las bibliotecas.

- Cuantifique con precisión las bibliotecas de DNA para garantizar las densidades de clúster óptimas en la celda de flujo.
- Cuantifique la biblioteca post enriquecida usando el kit de ensayo Qubit dsDNA BR, en donde se obtendrá una concentración en ng/µl.
- Correr un gel de agarosa al 4% para determinar el tamaño promedio de la biblioteca, utilizar 6µl de muestra.
- Use la siguiente fórmula para convertir de ng/µl a nM.

$$\frac{\text{(Concentración en ng/}\mu\text{l)}}{\text{(660 g/mol} \times \text{Tamaño promedio de la biblioteca)}} \times 10^6 = \text{Concentración en nM}$$

**\*660g/mol son equivalentes al PM de un par de bases.**

- Utilizar la formula  $C1.V1=C2.V2$  para realizar una dilución de la biblioteca utilizando RSB como diluyente hasta llegar a una concentración de 4nM, contemplando tomar como mínimo un volumen de 2µl de la biblioteca.

$$\frac{\text{Concentración en nM} \times \text{2}\mu\text{l de la biblioteca}}{\text{4nM}} = \text{Volumen final de la dilución 4nM}$$

Componentes de la dilución 4nM

Volumen de muestra (biblioteca) = 2µl

Volumen de RSB = Volumen final de dilución 4nM - Volumen de muestra (2µl)

- Al finalizar esta dilución, se tendrá la biblioteca a una concentración 4nM, y estará lista para comenzar con el protocolo del manual PhiX.

**NOTA: MINIMO 33ng/µl**

## PROTOCOLO DE DESNATURALIZACIÓN Y DILUCIÓN DE LA BIBLIOTECA

### Protocolo A: Método de normalización estándar.

Química: MiSeq Reagent Kit v3 desnaturalización compatible a 4nM con resultados de 6-20pM.

#### Preparación de Reactivos:

- Preparación de una dilución fresca de NaOH.
- Combinar los siguientes volúmenes en un tubo de microcentrifuga:
- Agua grado biología molecular (800µl).
- Solución stock de NaOH 1.0 N (200µl).
- Mezclar por vórtex.
- Como resultado se obtendrá 1 ml de NaOH 0,2 N.

**NOTA:** Usar la dilución fresca dentro de un lapso no mayor a 12 horas.

#### Desnaturalización y dilución de la biblioteca.

- Desnaturalización y dilución de la biblioteca de 4nM a 20pM.
- Combine los siguientes volúmenes en un tubo de microcentrifuga de 1.5 ml.
- 5µl de la biblioteca a 4nM.
- 5µl de NaOH 0.2 N recién diluido.
- Utilizar vórtex para mezclar.
- Centrifugar.
- Incubar por 5 minutos a temperatura ambiente.
- Agregar 990µl de TH1.
- Utilizar vórtex para mezclar (30 segundos).
- Centrifugar.
- Incubar en hielo.
- El resultado es una biblioteca desnaturalizada a una concentración de 20pM en 1ml.

#### Dilución de la biblioteca de 20pM desnaturalizada a 12.5pM.

- Combine los siguientes volúmenes en un tubo de microcentrifuga de 1.5 ml.
- 375µl de la biblioteca a 20pM desnaturalizada.
- 225µl de HT1.
- Utilizar vórtex para mezcla.
- Centrifugar.
- Incubar en hielo.
- El resultado es una dilución de 12.5pM en un volumen de 600µl.

## DESNATURALIZACIÓN Y DILUCIÓN DE PHIX.

### Dilución de PhiX a 4nM.

- Combinar los siguientes volúmenes en un tubo de microcentrífuga de 1.5 ml.
- 2µl de biblioteca PhiX a 10nM.
- 3µl de RSB.

#### **Desnaturalización de PhiX.**

- Combinar los siguientes volúmenes en un tubo de microcentrífuga de 1.5 ml.
- 5µl de PhiX 4nM.
- 5µl de NaOH 0.2 N recién diluido.
- Utilizar vórtex para mezclar.
- Centrifugar.
- Incubar por 5 minutos a temperatura ambiente.

#### **Dilución de PhiX desnaturalizado a 20pM.**

- Combinar los siguientes volúmenes en un tubo de microcentrífuga de 1.5 ml.
- 10µl de PhiX desnaturalizado.
- 990µl de HT1.
- Utilizar vórtex para mezclar.
- Centrifugar.
- Incubar en hielo.
- El resultado es una dilución de PhiX 20pM desnaturalizado en 1ml.

#### **Dilución de PhiX desnaturalizado 20pM a 12.5pM.**

- Combine los siguientes volúmenes en un tubo de microcentrífuga de 1.5 ml.
- 375 µl de PhiX desnaturalizado 20pM.
- 225µl de HT1.
- Utilizar vórtex para mezclar.
- Centrifugar.
- Incubar en hielo.
- El resultado es una dilución de 12.5pM en un volumen de 600µl.

#### **Combinar la biblioteca de la muestra con el PhiX.**

- Retirar 6µl de la biblioteca de muestra 12.5pM.
- Añadir a la biblioteca de la muestra 12.5pM, 6µl de PhiX 12.5pM para tener un volumen final de 600µl. Mantener en hielo.

## **PREPARACIÓN DE GEL DE AGAROSA**

### **Preparación de gel de agarosa al 1%:**

1. Pesar 0.35 g de agarosa ultrapura (invitrogene) y diluirla en 35 ml de buffer TBE 1X.
2. Calentar por 1 minuto o hasta que la agarosa este completamente disuelta.
3. Agregar 3.5µl de SYBER Safe (invitrogene) y agitar brevemente hasta que se incorpore.
4. Vaciar la agarosa en la cámara de electroforesis.

Para la preparación de gel de agarosa al 4% seguir el procedimiento antes descrito modificando la cantidad de agarosa (2g) y diluirlos en 50 ml de TBE 1X, usar 6µl de SYBER Safe.

### **BUFFER TBE 10X**

1. Pesar 108 g de Tris base.
  2. Pesar 55 g de ácido bórico.
  3. Medir 40 ml de EDTA 0.5 M pH 8.0.
- Al final se obtendrán 40 ml de TBE 10X, para preparar buffer TBE 1X es necesario realizar una dilución 1:10 del buffer TBE 10X recién preparado con agua miliQ.