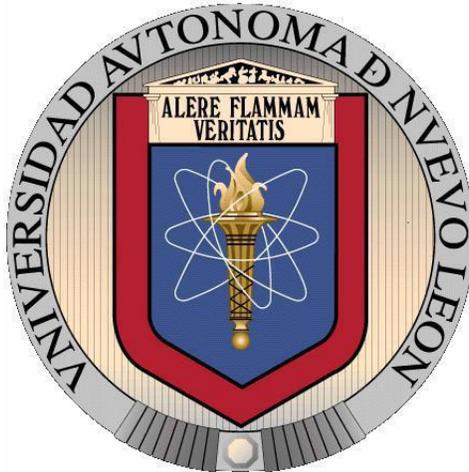


**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE MEDICINA**



**ANÁLISIS DEL EXOMA DE PACIENTES CON ENFERMEDADES
CONGÉNITAS IDIOPÁTICAS MEDIANTE SECUENCIACIÓN DE NUEVA
GENERACIÓN**

POR

LBT. MARTHA MONTSERRAT RANGEL SOSA

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

AGOSTO, 2017

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE MEDICINA
DEPARTAMENTO DE BIOQUÍMICA Y MEDICINA MOLECULAR**



**ANÁLISIS DEL EXOMA DE PACIENTES CON ENFERMEDADES
CONGÉNITAS IDIOPÁTICAS MEDIANTE SECUENCIACIÓN DE NUEVA
GENERACIÓN**

POR

LBT. MARTHA MONTSERRAT RANGEL SOSA

**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, 2017

El presente 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, bajo la dirección del Dr. C. Carlos Córdova Fletes y la co-dirección de la Dra. C. Viviana Chantal Zomosa Signoret y la Dra. C. Ana María Guadalupe Rivas Estilla.

Vo. Bo.



Dr. C. Carlos Córdova Fletes
Director de Tesis

Análisis del exoma de pacientes con enfermedades congénitas idiopáticas mediante secuenciación de nueva generación.

Aprobación de la tesis:



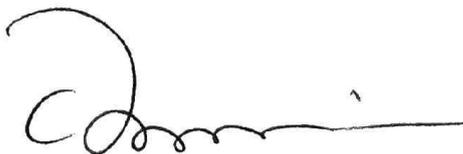
Dr. C. Carlos Córdova Fletes
Director de Tesis



Dra. C. Viviana Chantal Zomosa Signoret
Co-Directora de Tesis



Dra. C. Ana María Guadalupe Rivas Estilla
Miembro de la Comisión de Tesis



Dr. med. Felipe Arturo Morales Martínez
Subdirector de Estudios de Posgrado

COLABORACIONES

Dr. C. José Elías García Ortiz

División de Genética
Centro de Investigación Biomédica de Occidente
Instituto Mexicano del Seguro Social

Dr. C. Luis Eduardo Figuera Villanueva

División de Genética
Centro de Investigación Biomédica de Occidente
Instituto Mexicano del Seguro Social

FINANCIAMIENTO

Convocatoria Apoyo Infraestructura Científica 2016, proyecto No. 268123

Responsable técnico: Dr. Carlos Córdova Fletes

APROBACIÓN DEL COMITÉ DE ÉTICA

Este trabajo fue aprobado por el Comité de Ética de la Facultad de Medicina de la Universidad Autónoma de Nuevo León.

No. BI15-003

Agradecimientos

A mi familia, a mis papás porque aunque sé que están felices y orgullosos por estos dos años fuera de casa, fue difícil no tenerme allá en todo momento, y fue difícil no tenerlos aquí. Gracias por dejarme y motivarme a crecer, creer y confiar en mí. A mi hermana que es la más maravillosa del mundo, por hacer que me sintiera muy muy querida cada vez que nos veíamos, por extrañarme y por apoyarme siempre. A mis hermanos, que aunque siempre que iba a Aguascalientes nunca estaban en la casa, sé que me extrañaban. Y obviamente, a Bonis.

A Stephen, por compartir momentos buenos y malos, momentos de estrés y felicidad, por estar conmigo en todo este camino y mostrarme que hay un mundo lleno de posibilidades y oportunidades, por siempre motivarme a salir adelante y esforzarme más, por soportar todos mis momentos pre y de malas, por todas las comidas y los viajes maravillosos. Has marcado mi formación y la persona que he llegado a ser, gracias por todo estudiante destacado de Inmunología.

A, literalmente mi familia en Mty, Stephen, Aidé, Frexia, Juanito, Florencia, Alejandro, Iris, Enrique y bebecito. A todos ustedes niños, gracias por convertirse en mi familia, porque había momentos en los que me daba cuenta de que era muy feliz, y todo era gracias a ustedes, sus locuras y su compañía. Gracias por aguantar también mis momentos pre, con cara de pedo, berrinches, necedades, estrés, etc. y aun así seguirme amando. Gracias por que vivimos de todo juntos: viajes, comidas diferentes, bebidas, lugares nuevos, pláticas de temas sensibles e importantes y luego otras de lo más vulgares, embarazos, casi morir en matacanes, despilfarro de dinero y luego pobreza, inundaciones, dolores, muelas, migrañas, nostalgia, pérdidas. Hace poco leí: “Cuando no tienes que tener cuidado con lo que dices, sabes que estas con las personas correctas”, ustedes son mis personas correctas. Los voy a extrañar mucho.

Al Dr. Carlos, por darme un lugar en su laboratorio, por las clases y el conocimiento que nos impartió, permitirme manipular y aprender técnicas muy delicadas que no hubiera aprendido en otra parte, por tener paciencia cuando las cosas salían mal, celebrar y reconocer los logros, los días de pizza y las buenas referencias.

A la Dra. Viviana y a la Dra. Ana María por estar al pendiente de mi trabajo y revisarme todos los escritos y presentaciones siempre.

A Ana, por adoptarnos en su familia, ser mi hermana de laboratorio, los regalos más bonitos, por pasar y solucionar todo juntas, quejarnos, comer y apoyarnos en muchísimas ocasiones. A Piqué por siempre estar molestando y hacer los días de lo más divertidos (ahora valoro todas las peleas y lanzadas de alcohol), por ser un buen guía turístico y por tratarme como señora del seven el último mes. A Papá Alex, por ser mi maestro y enseñarme muchas de las cosas que aprendí en el laboratorio, porque era la persona con la que podía discutir mis ideas y mis dudas, por animarme, escucharme y convertirse en un amigo con el que podía hablar de absolutamente todo. Son el mejor grupo de trabajo del mundo, hicieron que me gustara mucho ir al laboratorio, que me divirtiera, y que el trabajo fuera mil veces más liviano. Remember: Friendships can end, girlfriends/boyfriends can end, only lab partner has no end jajaja ya los estoy extrañando.

A todos los compañeros de laboratorio con los que me tocó convivir, Priscila, Tony, Elda, Paúl, Carlos y Yanelli.

A Cecy por estar conmigo y ayudarme siempre desde aquel verano de investigación, por más que una compañera de laboratorio, convertirse en una gran amiga y confidente. A Alex Torres, porque aún a distancia me animó y escuchó muchas quejas, por compartir tonterías cuando las cosas salían mal y por coincidir en mi seminario de terminación.

Al Dr. Daniel Cervantes, por ser un excelente maestro y tutor de tesis de licenciatura, por inculcarme el gusto por la biología molecular, la ingeniería genética y la investigación, por ser el primero en enseñarme a razonar y planear los experimentos. Al Dr. Augusto Rojas, por aceptarme en aquel verano de la ciencia que hizo reafirmara mis ganas de venir a estudiar aquí, gracias por que en el tiempo que estuve en su laboratorio aprendí muchísimas cosas, nuevas técnicas, razonamientos y diseños de experimentos, por hacer que creciera mi gusto por la terapia génica e inculcarme el interés por la inmunoterapia, gracias por seguir al pendiente y en contacto cuando ya no nos veíamos con mucha frecuencia.

A todo el personal de Bioquímica, en especial a Juany y Adriana por siempre estar al pendiente de nosotros y ayudarnos tanto.

Contenido

Índice de tablas	X
Índice de figuras	XI
Abreviaturas	XII
Resumen	XIII
Capítulo I: Introducción	1
1.1 Anomalías congénitas.....	1
1.1.1 Periodos de susceptibilidad en el desarrollo anormal	1
1.1.2 Mortalidad	2
1.1.3 Prevalencia	3
1.1.4 Clasificación	4
1.1.5 Malformaciones múltiples	6
1.1.6 Etiología de las anomalías congénitas	6
1.2 Discapacidad intelectual	7
1.2.1 Etiología de la discapacidad intelectual	8
1.3 Métodos de diagnóstico convencionales.....	9
1.3.1 Secuenciación de exoma	11
Capítulo II: Antecedentes	13
Capítulo III: Justificación	17
Capítulo IV: Objetivos	18
4.1 Objetivo general.....	18
4.2 Objetivos específicos	18
Capítulo V: Diseño experimental	19
5.1 Diseño del experimento	19
5.2 Universo de estudio	19
Capítulo VI: Metodología	21
6.1 Estrategia experimental.	21
6.2 Naturaleza de la muestra.....	21
6.3 Evaluación de las pérdidas de heterocigocidad de los microarreglos de los pacientes.	21
6.4 Secuenciación de exoma.....	22
6.4.1 Comprobación de la calidad, cantidad e integridad de las muestras.....	22
6.4.2 Preparación de la biblioteca.	23

6.4.3 Secuenciación	24
6.5 Análisis integral de los datos.....	24
Capítulo VII: Materiales y Equipos	26
7.1 Materiales y reactivos	26
7.1.1 Material biológico	26
7.1.2 Material de consumo	26
7.1.3 Materiales y reactivos para secuenciación masiva	27
7.2 Equipos.....	27
7.2.1 Equipo de uso general.....	27
7.2.2 Equipos para secuenciación masiva.....	28
7.3 Programas computacionales y plataformas bioinformáticas.....	28
Capítulo VIII: Resultados	30
8.1 Descripción de casos	30
8.1.1 Pacientes 1 y 2.....	30
8.1.3 Paciente 3.	31
8.1.4 Paciente 4.	32
8.2 Naturaleza de la muestra.....	32
8.3 Evaluación de las pérdidas de heterocigocidad de los microarreglos de los pacientes.	33
8.3.1 Pacientes 1 y 2.....	33
8.4 Secuenciación de exoma.....	35
8.4.1 Comprobación de la calidad, cantidad e integridad de las muestras.....	35
8.4.2 Generación de bibliotecas.	36
8.5 Análisis integral de los datos.....	39
8.5.1 Pacientes 1 y 2.....	39
8.5.2 Paciente 3.	42
8.5.2 Paciente 4.	47
Capítulo IX: Discusión	51
Capítulo X: Conclusiones	60
Capítulo XI: Perspectivas	61
Capítulo XII: Bibliografía.....	62
Capítulo XIII: Anexos	70

Índice de tablas

Tabla 1. Clasificación de la discapacidad intelectual de acuerdo al IQ, según la OMS.	8
Tabla 2. Genes dentro de la región de LOH de las pacientes 1 y 2.....	33
Tabla 3. Genes relacionados a desarrollo neuronal u otros procesos celulares.	34
Tabla 4. Genes dentro de la región de LOH del paciente 3.....	34
Tabla 5. Cuantificación de las muestras por Qubit y NanoDrop.	35
Tabla 6. Concentraciones de las bibliotecas medidas por el equipo Qubit.	36
Tabla 7. Número de clústers obtenidos por muestra.	37
Tabla 8. Enriquecimiento a nivel lectura, valor Q30 y profundidad promedio.	38
Tabla 9. Enriquecimiento a nivel de pares de bases.	38
Tabla 10. Resumen de variantes según su localización.....	38
Tabla 11. Resumen de variantes según su consecuencia.....	39
Tabla 12. Genes con mutaciones recesivas (Frec. Pob.<0.01) en pacientes 1 y 2.....	40
Tabla 13. Análisis de las variantes p.Ile33Ala y p.Phe34Ile de las pacientes 1 y 2.	42
Tabla 14. Genes con mutaciones recesivas (Frec. Pob.<0.01) del paciente 3.	43
Tabla 15. Cuadro clínico de la displasia cráneo-ectodérmica.....	45
Tabla 16. Análisis de las variantes <i>SPAG17</i> p.Asp357His y <i>WDR35</i> p.Arg472Gln.	47
Tabla 17. Diagnóstico diferencial de afecciones similares a Marfan.....	49
Tabla 18. Genes con mutaciones recesivas (Frec. Pob.<0.01) del paciente 4.	50

Índice de figuras

Figura 1. Periodos y grados de susceptibilidad de los órganos de un embrión.....	2
Figura 2. Mortalidad mundial por anomalías congénitas.....	3
Figura 3. Prevalencia nacional de las anomalías congénitas.....	4
Figura 4. LOH de copia neutral del cromosoma 19 de las pacientes 1 y 2.....	30
Figura 5. LOH de copia neutral de 11 cromosomas afectados en el paciente 3.....	31
Figura 6. Microarreglo aCGH del paciente 4.....	32
Figura 7. Gel agarosa al 1% del DNA de las pacientes 1 y 2 y sus padres.....	36
Figura 8. Gel de agarosa al 1% del DNA de los pacientes 3 y 4.....	36
Figura 9. Gel de agarosa al 4% de las bibliotecas de los pacientes.....	37
Figura 10. Análisis de las mutaciones del gen <i>SNRPA</i>	41
Figura 11. Alineamiento de las mutaciones <i>SPAG17</i> c.1069G>C y <i>WDR35</i> c.1415G>A..	46
Figura 12. Predicción de la estructura secundaria de la proteínas WT y mutadas de los genes <i>SPAG17</i> y <i>WDR35</i>	46

Abreviaturas

AC	Anomalías congénitas
ACM	Anomalías congénitas múltiples
aCGH	Microarreglo de hibridación genómica comparativa
CED	Displasia cráneo-ectodérmica
CN-LOH	Pérdida de heterocigocidad de copia neutral
CNV	Variación en el número de copias
DI	Discapacidad intelectual
DNA	Ácido desoxirribonucleico
EVC	Síndrome de Ellis–Van Creveld
FISH	Hibridación fluorescente <i>in situ</i>
Gbp	Gigapares de base
GWAS	Estudio de asociación de genoma completo
IGV	Integrative Genome Viewer
IQ	Coeficiente de inteligencia
Kb	Kilobases
LOH	Pérdida de heterocigocidad
mar	Marcador cromosómico
Mb	Megabases
μL	Microlitro
mL	Mililitro
MLPA	Amplificación de sondas dependiente de ligandos múltiples
MM	Malformaciones múltiples
mos	Mosaico
PCR	Reacción en cadena de la polimerasa
RD	Retraso en el desarrollo
SNP	Polimorfismos de nucleótido único
SNV	Variantes de nucleótido único
SRP	Síndrome de Polidactilia-Costillas Cortas
WES	Secuenciación de exoma completo
WGS	Secuenciación de genoma completo

Resumen

Análisis del exoma de pacientes con enfermedades congénitas idiopáticas mediante secuenciación de nueva generación.

Nombre: LBT. Martha Montserrat Rangel Sosa.

Institución: Universidad Autónoma de Nuevo León, Facultad de Medicina.

Localización: Monterrey, NL. México.

Fecha de titulación: Agosto 2017.

Número de páginas: 80.

Introducción: Las anomalías congénitas son defectos estructurales, funcionales o bioquímicos que ocurren de manera intrauterina, en 2015 fallecieron 303,000 niños en periodo neonatal por anomalías congénitas. Debido a las limitaciones de los métodos convencionales, alrededor del 40-60% de los casos son idiopáticos. Frente a esto, la secuenciación de exoma completo representa un costo-beneficio importante en el diagnóstico ya que contiene el 85% de las mutaciones causantes de enfermedades. **Objetivo:** Identificar y caracterizar las alteraciones genómicas en pacientes con AC idiopáticas mediante secuenciación de exoma. **Material y Métodos:** Se prepararon las bibliotecas del exoma de 4 pacientes con anomalías congénitas y los padres de dos de estos pacientes (hermanas) utilizando el kit TruSeq Rapid Exome. Se realizó secuenciación paired-end de 76 ciclos en el equipo MiSeq (Illumina). Para el análisis de los datos se utilizó el genoma de referencia hg19 y las plataformas BaseSpace, BWA Enrichment, GATK, VariantStudio y VariantInterpreter (beta), entre otros. **Resultados:** Se lograron identificar 3 mutaciones puntuales homocigotas en el gen *SNRPA* en las pacientes 1 y 2 que ocasionan un cambio en dos aminoácidos (p.Ile33Ala y p.Phe34Ile), ambos cambios fueron predichos como patogénicos por distintas plataformas bioinformáticas, debido a que ocasionan la ganancia de una región α -hélice; ambos padres fueron heterocigotos para estas 3 mutaciones. *SNRPA* no ha sido asociado a enfermedades, sin embargo otros genes del spliceosoma se han relacionado con desórdenes craneofaciales y esqueléticos. Se encontraron 2 mutaciones probablemente responsables del fenotipo del paciente 3, la primera, c.1069G>C (p.Asp357His) en el gen *SPAG17* y la segunda, c.1415G>A (p.Arg472Gln) en el gen *WDR35*. Ambas modifican la estructura secundaria de la proteína (perdiendo y ganando una porción α -hélice, respectivamente) y se predice que son patogénicas. Mutaciones en *WDR35* se han asociado a displasia craneoectodérmica, mientras que *SPAG17* no se ha asociado a enfermedades; ambos genes están involucrados en las estructuras ciliares. Se descartó la presencia del Síndrome de Marfan en el paciente 4, ya que no presentó mutaciones patogénicas en *FBN1*, sin embargo presenta una delección del marco de lectura (p.Gln262del) en el gen *MED15*, el cual regula la vía de TGF- β , asociada a daños en tejido conectivo y ojos; también se encontró una mutación en un sitio aceptor de splicing del gen *ADAMTS4* (c.1736-2A>G), la cual probablemente está afectando el splicing del pre-RNA y por lo tanto la proteína; mutaciones en la familia ADAMTS y en ADAMTSL4 han sido asociadas a ectopia lentis, la cual está presente en el paciente. **Conclusión:** Se realizó la secuenciación de exoma de 4 pacientes con anomalías congénitas, pudiendo determinar la causa del fenotipo de tres de ellos.

Firma del asesor:



Dr. C. Carlos Córdova Fletes
Director de Tesis

Capítulo I: Introducción

1.1 Anomalías congénitas.

Las anomalías congénitas (AC), también conocidas como defectos de nacimiento, desordenes congénitos o malformaciones congénitas (1, 2), son defectos estructurales, funcionales o bioquímico moleculares que ocurren de manera intrauterina (2, 3). Estos pueden ser identificados de manera prenatal, al momento del nacimiento o en etapas posteriores (2). Las anomalías congénitas pueden tener un impacto en la salud, el desarrollo o la supervivencia del afectado (4).

Las AC varían sustancialmente en gravedad, algunas de estas se asocian con abortos espontáneos y muerte al nacimiento o en periodos pre- y post-natales. Estas afecciones abarcan una amplia gama de patologías, las más frecuentes son anomalías congénitas del corazón (60/10,000), defectos del tubo neural (12/10,000), síndrome de Down (23/10,000) y re-acomodos cromosómicos desequilibrados (4, 5).

1.1.1 Periodos de susceptibilidad en el desarrollo anormal.

Existen ciertos periodos críticos durante el embarazo en donde los embriones son más susceptibles a que agentes o factores causen anomalías congénitas. Es poco probable que daños al embrión durante las primeras 3 semanas de embriogénesis (periodo temprano antes de que comience la organogénesis) causen un desarrollo anormal ya que, o el daño es tan fuerte que el embrión no logra sobrevivir, o es compensado por las potentes propiedades reguladoras del embrión en esas etapas. El periodo de máxima susceptibilidad para que se produzca un desarrollo anormal persistente, ocurre entre la tercer y la octava semana, ya que en este periodo comienza la organogénesis y se establecen la mayoría de las regiones del cuerpo. Es poco probable que después de la octava semana de embarazo se desarrollen anomalías estructurales mayores, debido a que para este punto, la mayoría de los órganos ya han sido establecidos; las anomalías derivadas del tercer al noveno mes tienden a ser funcionales, por ejemplo, discapacidad intelectual (DI), o implican alteraciones en el crecimiento de las partes ya formadas (6).

Los órganos tienen diferentes periodos de susceptibilidad durante el embarazo. Aquellos que se desarrollan de manera temprana, como el corazón, tienden a ser más sensibles en

las primeras semanas; órganos complejos como el cerebro u órganos sensoriales presentan periodos prolongados de susceptibilidad. El desarrollo de un órgano presenta una curva de susceptibilidad a las influencias teratógenas. Durante los primeros días del periodo crítico, la susceptibilidad (medida como la incidencia o severidad de la malformación) incrementa bruscamente y luego declina durante un periodo prolongado donde presenta una menor influencia en el desarrollo (6).

Algunos daños como la exposición a toxinas o enfermedades intrauterinas pueden resultar en la destrucción de partes o todas las estructuras que ya habían sido formadas (6).

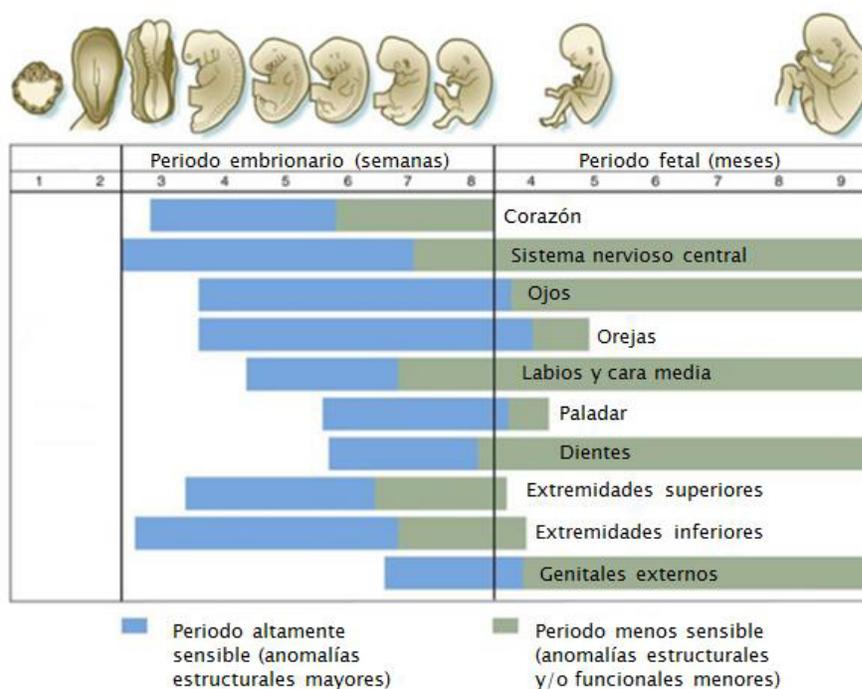


Figura 1. Periodos y grados de susceptibilidad de los órganos de un embrión (Modificado de Carlson, 2008) (6).

1.1.2 Mortalidad.

Las AC son una importante causa de enfermedades crónicas, admisión a hospitales, discapacidad y muerte en la infancia (7, 8). En el año 2015 murieron 2.68 millones de niños en el periodo neonatal, 303,000 de estas muertes tuvieron como causa algún tipo de anomalía congénita, representando el 11% del total (2). La causa más importante de

muerte neonatal en 2013 fue encefalopatía, seguida de sepsis, anomalías congénitas e infecciones respiratorias de las vías aéreas inferiores (5).

La tasa de mortalidad a nivel mundial varía fuertemente de país a país, esto puede ser explicado por factores sociales, raciales, ecológicos y económicos (7). El país con más alta prevalencia es Angola con 19.62 muertes por cada 100,000 habitantes, mientras que el que presenta una menor mortalidad es Luxemburgo con 0.75 por cada 100,000 (2).

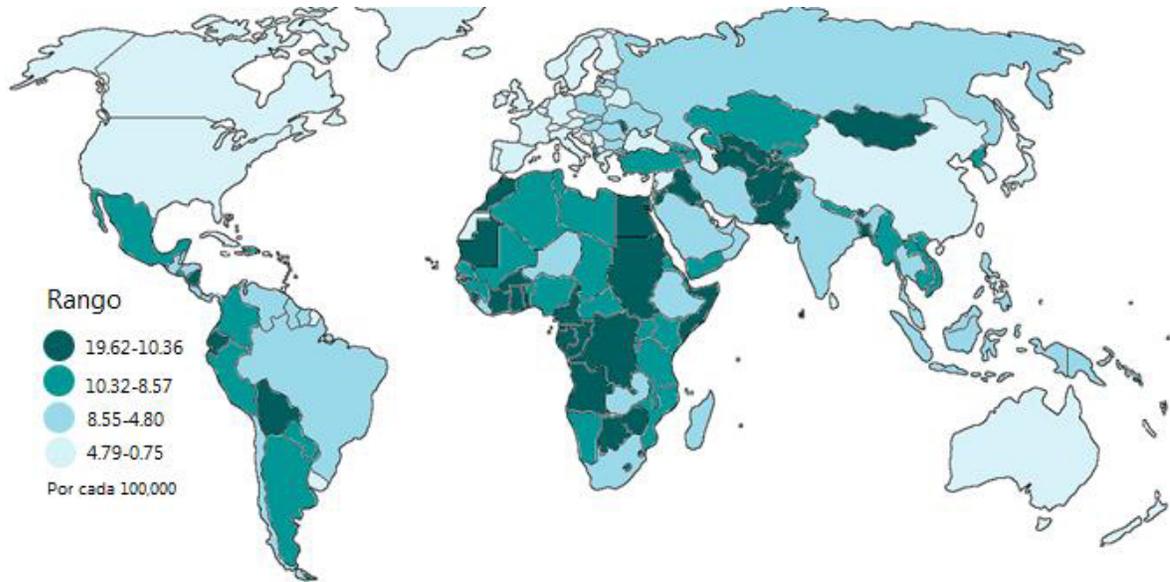


Figura 2. Mortalidad mundial por anomalías congénitas (WHO, 2014).

1.1.3 Prevalencia.

Las AC afectan a una proporción remarcable de recién nacidos (8). Cada año un estimado de 7.9 millones de infantes (6% de los nacimientos a nivel mundial) nacen con alguna malformación congénita. Aunque algunos de estos padecimientos pueden ser controlados y tratados, se estima que 3.2 millones de estos infantes presentarán alguna discapacidad de por vida (9).

En México, se han realizado varios estudios para determinar la frecuencia de las malformaciones congénitas: en Guadalajara, Jalisco, la incidencia en una población de 7,791 recién nacidos vivos consecutivos fue de 12.8/1,000; en el Distrito Federal (ahora Ciudad de México) se encontró una incidencia de 1.2% de malformaciones congénitas mayores y 2.1% de malformaciones congénitas menores en 3,283 nacidos vivos. También en el Distrito Federal se detectó una incidencia de 2.6% en 12,659 nacidos vivos, y de

8.7% en 208 muertes fetales; en Monterrey, Nuevo León, la incidencia fue de 2.31% en 9,675 nacidos vivos (10).

Las tasas de mayor prevalencia se concentraron principalmente en los estados del centro y dos del sur de la república: Aguascalientes, Colima, Querétaro, Distrito Federal, Tlaxcala, Yucatán, Hidalgo, Puebla, Guanajuato, Michoacán y Tabasco. La tasa de prevalencia observada por Navarrete-Hernández y colaboradores a nivel nacional fue de 73.9/10,000 nacimientos (11).



Figura 3. Prevalencia nacional de las anomalías congénitas (Modificado de Navarrete-Hernández y cols. 2013) (11).

1.1.4 Clasificación.

Las AC pueden ser estructurales o funcionales. Las AC estructurales son aquellas que involucran alteraciones morfológicas. Es decir, que afectan algún tejido, órgano o conjunto de órganos del cuerpo. Algunos ejemplos son hidrocefalia, espina bífida, fisura de labio y/o paladar y cardiopatías congénitas. A su vez, entre las AC estructurales pueden encontrarse anomalías mayores y menores (12).

Son clasificadas como mayores si tienen un impacto significativo en la calidad de vida del paciente o requiere tratamiento médico/cirugía (13). Las anomalías mayores se consideran letales si causan el nacimiento de un feto muerto, la muerte de un infante, o el

embarazo es terminado después del diagnóstico prenatal en más del 50% de los casos, ejemplo de estas son la anencefalia o el síndrome de corazón izquierdo hipoplásico. Las anomalías mayores se consideran severas cuando, si no se cuenta con intervención médica, pueden causar discapacidades o la muerte, ejemplo de estas son el labio y paladar hendido o la estenosis pilórica congénita (3).

Las anomalías menores son aquellas que no requieren cuidado médico o éste es menor, no afectan alguna función importante del cuerpo y tienen una buena expectativa de vida, aunque pueden tener alguna consecuencia cosmética. Las AC menores generalmente no son reportadas a menos que co-ocurran con una anomalía mayor. Ejemplos de estas son la hernia umbilical, la sindactilia parcial entre dos dedos de los pies o manos, el pliegue palmar único, pliegues epicánticos, entre otros (3).

Las anomalías mayores y menores no están universalmente definidas, la distinción generalmente depende de la opinión subjetiva del médico tratante (13). La prevalencia de las anomalías mayores es del 3% de los nacidos vivos, más un 3% adicional que se diagnostica en edades posteriores. Las anomalías menores se presentan en el 10% de los nacidos vivos (14).

La mayoría de la población tiene una o dos anomalías menores heredadas de los padres. La presencia de tres o más, las cuales no están presentes en otros miembros de la familia, puede indicar que ese individuo tiene un síndrome y/o una malformación mayor. Los genetistas usan las anomalías menores como pistas para establecer un diagnóstico en los pacientes (15).

Las AC funcionales son aquellas que interrumpen procesos biológicos sin implicar un cambio macroscópico de forma; involucran alteraciones metabólicas, hematológicas, del sistema inmune, entre otras. Algunos ejemplos son: hipotiroidismo congénito, discapacidad intelectual, tono muscular disminuido, ceguera, sordera y convulsiones de inicio neonatal, entre otras (12).

Cuando varios órganos están involucrados, el conjunto de anomalías puede ser declarado como un síndrome, indicando que existe una reconocida combinación de defectos, los cuales se considera que están relacionados. Si se sospecha de la relación, pero no ha sido confirmada, la combinación de defectos puede llamarse asociación. Con el

incremento del conocimiento en la genética, existen más combinaciones de anomalías que han sido reconocidas y promovidas a síndromes (13).

1.1.5 Malformaciones múltiples.

Las AC se pueden clasificar, según el número de ellas presentes en el recién nacido, como: 1) defectos únicos (defectos aislados) y 2) malformaciones múltiples (MM). En este último grupo se pueden encontrar: a) aquellos cuyos defectos correspondan con alguno de los síndromes ya descritos en la literatura, los síndromes generalmente suelen tener la misma causa y el mismo riesgo familiar; b) pacientes con defectos congénitos que afectan a diferentes estructuras corporales aparentemente no relacionadas, en quienes no se ha podido diagnosticar ningún síndrome conocido y no se ha identificado un agente causal; c) pacientes con un único defecto del desarrollo, pero en quienes, ese defecto, ha dado lugar secundaria y secuencialmente, a alteraciones de otras estructuras cuyo desarrollo morfológico era normal. Este tipo de eventos se denominan “secuencias” y estas pueden ser producidas por muchos agentes causales, pueden manifestarse en pacientes polimalformados sin etiología establecida, como síndromes o como una secuencia aislada. Las MM, también son conocidas como anomalías congénitas múltiples (ACM). Para que las malformaciones congénitas sean consideradas como MM, el paciente debe presentar dos o más malformaciones mayores (ej. defecto del tubo neural, defectos cardíacos) y/o tres o más malformaciones menores (ej. sindactilia, pabellones auriculares de forma anormal, etc.) (16).

Los pacientes con malformaciones múltiples generalmente están relacionados a un reacomodo cromosómico ya que este puede alterar varios genes contiguos, causando una perturbación en más de un órgano (17, 18).

1.1.6 Etiología de las anomalías congénitas.

Las principales causas de las AC pueden ser clasificadas en tres grandes grupos:

- a) Genéticas: Son responsables del 15-25% del total de los casos. Estas causas incluyen aberraciones cromosómicas, (ej. Síndrome de Down) o defectos en un único gen (también conocidas como enfermedades mendelianas, ej. Acondroplasia o el Síndrome Holt-Oram). Existen dos condiciones que pueden ser responsables de la alta prevalencia de anomalías congénitas con origen genético: Mujeres

embarazadas después de los 35 años de edad y la alta tasa de matrimonios consanguíneos.

- b) Ambientales: Son responsables del 8-12% de los casos. Estas incluyen enfermedades infecciosas (ej. Rubeola), enfermedades maternas (ej. Diabetes mellitus), exposición a algún teratógeno como drogas, alcohol, humo de cigarro, químicos o algunos contaminantes.
- c) Complejas (multifactoriales): Son responsables del 20-25% de los casos. Son causadas por una interacción gen-medio ambiente, el paciente presenta predisposición a una enfermedad que se desencadena por un factor de riesgo ambiental. La mayoría de las anomalías congénitas comunes, como los defectos del tubo neural, las malformaciones cardiovasculares, la dislocación congénita de la cadera, etc., pertenecen a este grupo etiológico.
- d) Desconocidas: Aproximadamente en el 60% de los casos, las causas se mantienen desconocidas (3, 7).

Es importante mencionar que las anomalías genéticas son la causa más frecuente de discapacidad intelectual (7).

1.2 Discapacidad intelectual.

La discapacidad intelectual (DI) se caracteriza por limitaciones significativas en las habilidades mentales y funciones adaptativas que surgen durante el desarrollo de un niño (ej. Habilidades cognitivas, motoras, sociales o de lenguaje). Las limitaciones deben ser evidentes cuando se compara contra personas de la misma edad, género y entorno socio cultural. Cuando la persona afectada es menor a 5 años, se utiliza el término retraso en el desarrollo (RD), ya que en estos pacientes no se pueden realizar todavía el tipo de pruebas que se utiliza para determinar la DI. Los individuos que presentan DI muestran problemas para percibir y procesar la información, aprender de manera rápida y eficiente, aplicar su conocimiento y habilidades para resolver problemas, pensar creativa y flexiblemente, y responder de una manera rápida y precisa. En niños mayores a 5 años, el funcionamiento intelectual es medido usando una prueba de inteligencia estandarizada y administrada de manera individual. El puntaje del coeficiente de inteligencia (IQ) normalmente tiene una media de 100, y una desviación estándar de 15. Un $IQ < 70$ (dos desviaciones estándar por debajo de la media) pueden indicar una deficiencia significativa en la función intelectual. Aproximadamente el 2.5-3% de la poblaciones presenta valores

de IQ menores a 70. Las personas con DI muestran dificultades en el funcionamiento adaptativo, este se refiere a que tan efectivos son los individuos para enfrentarse con los problemas comunes y que tan bien cumplen con los estándares de independencia para una persona de su misma edad, grupo social y comunidad (19).

Tabla 1. Clasificación de la discapacidad intelectual de acuerdo al IQ, según la OMS.

1. Leve	IQ entre 50-69
2. Moderada	IQ entre 35-49
3. Grave	IQ entre 20-34
4. Profunda	IQ inferior a 20

La DI puede ocurrir como un fenómeno aislado, o acompañada de malformaciones, signos neurológicos, deterioro de los sentidos, convulsiones y trastornos del comportamiento. El determinar si la DI está asociada a malformaciones o AC puede ser útil ya que en algunos casos sugiere un diagnóstico clínico sindrómico que permite al médico seleccionar un tratamiento para el paciente. Se estima que del 2 al 3% de la población presenta DI (17). Dentro de estos, la DI leve, moderada, severa y profunda afecta a cerca del 85%, 10%, 4% y 2% de las personas que la padecen, respectivamente (20). Existen pocos estudios de incidencia, sin embargo la prevalencia de la DI calculada a partir de un meta-análisis de 52 estudios fue de 10.37/1000 habitantes (20).

1.2.1 Etiología de la discapacidad intelectual.

La DI puede ser causada por factores no genéticos (14% de los casos), como infecciones o asfixia perinatal, aunque se cree que las formas más graves son debidas a factores genéticos (25% de los casos). Las causas genéticas de la DI incluyen desbalances cromosómicos visibles al microscopio, variación en el número de copias (ganancias o pérdidas de material genético) o mutaciones puntuales (Indels o sustituciones). Aunque hay un número de síndromes genéticos bien definidos, que surgen por una delección o duplicación, la mayoría de los casos de DI no tienen una causa genética definida. Por lo que en el 60% de los casos, las causas de este padecimiento se mantienen inconclusas (21).

Estudios recientes indican que existe una mayor cantidad de mutaciones *de novo* en personas con DI que en controles sanos, remarcando la importancia clínica de estas

mutaciones. El que la DI sea esporádica en la mayoría de los casos (sin factores ambientales o familiares obvios que pudieran causarla) hace más fuerte la hipótesis de que una gran proporción de casos son causados por mutaciones *de novo*. Se ha estimado que mutaciones en más de 1000 genes diferentes pueden causar DI. Las alta probabilidad de mutaciones raras *de novo* en esta gran cantidad de genes pueden ser compensadas por la baja tasa de reproducción de pacientes con DI, manteniendo la incidencia de este padecimiento estable en la población general (21).

El entendimiento de las causas genéticas de la DI puede beneficiar a los pacientes y sus familias, ya que el diagnóstico puede proveer información acerca del pronóstico, impide nuevas pruebas innecesarias invasivas y puede conducir a una terapia más adecuada. Además, un diagnóstico a menudo facilita el acceso al cuidado médico adecuado. Los familiares pueden beneficiarse ya que conocen el riesgo de recurrencia del padecimiento, pueden solicitar consejo genético reproductivo y tener un posible diagnóstico prenatal, si fuera el caso (21). Sin embargo, el diagnóstico de las AC o la DI es difícil, debido a que muchos de los reacomodos cromosómicos responsables de la condición del paciente no pueden ser detectados debido a las limitaciones de los métodos que se utilizan en la clínica (18).

1.3 Métodos de diagnóstico convencionales.

El cariotipo es la técnica de rutina para el diagnóstico, sin embargo, tiene una resolución muy gruesa, pudiendo detectar únicamente reacomodos largos (>5 Mb) visibles al microscopio. Es probable que una elevada proporción de niños con AC y DI/RD tengan algún tipo de reacomodo cromosómico submicroscópico que no sea evidente por cariotipo. Debido a esto se diseñaron otras técnicas complementarias más eficientes, como la hibridación fluorescente *in situ* (FISH), la amplificación de sondas dependiente de ligandos múltiples (MLPA), la hibridación genómica comparativa (CGH) en metafases y basada en arreglos (aCGH), siendo la más nueva y específica la secuenciación de nueva generación (NGS) (18). Se estima que usando todos los análisis, la tasa global de detección de las anomalías genómicas en infantes y niños con AC y/o DI/RD es de 12-18%, 3-5% detectadas por bandeo en cariotipo, 5-6% detectadas por FISH subtelomérico pero no por cariotipo y 4-7% detectadas por aCGH (17). La NGS ofrece una tasa de diagnóstico mucho mayor, del 27% si se analiza el exoma y de 42% si se analiza el genoma completo (22).

La FISH es una técnica de citogenética molecular que permite la detección de anomalías cromosómicas que no pueden ser apreciadas por cariotipo (ej. Síndromes de microdeleciones). En esta técnica, los cromosomas en metafase o núcleos en interfase son desnaturalizados en un portaobjetos, donde se añade una sonda fluorescente marcada, esta sonda hibrida en su región blanco y se visualiza con un microscopio de fluorescencia (23). De esta manera, es posible localizar secuencias de DNA en regiones del genoma, pudiendo detectar microdeleciones, microduplicaciones, repeticiones, fusiones de genes, translocaciones, entre otras (23, 24). Una desventaja del método, es que el éxito depende de la estabilidad de la doble hélice del DNA, aparte de que es necesario que el análisis lo realice una persona experta (17, 25).

La técnica MLPA es una variación de la PCR multiplex que permite la amplificación de varios blancos empleando sondas que hibridan en regiones de interés y un solo par de primers. Una sonda contiene la secuencia reconocida por el primer forward y la otra contiene la secuencia reconocida por el reverse, solo cuando ambas sondas hibridan en sus respectivos blancos pueden ser ligadas, convirtiéndolas en una sola, la cual es amplificada. La ventaja de añadir las sondas divididas es que solo se amplifican las que fueron ligadas, indicando proximidad en cierta región genómica (21).

La CGH ha sido usada para la detección de cambios en el número de copias en tumores sólidos usando cromosomas en metafase, a pesar de que la CGH es una técnica robusta para la identificación de desbalances cromosómicos de gran escala, no es eficiente para identificar cambios genómicos de menos de 5-10 Mb. Esto condujo al desarrollo de los aCGH, la cual permite una resolución de hasta 4 Kb. Esta técnica consiste en marcar un genoma problema con un fluorocromo y un genoma de referencia con otro, ambos se combinan y compiten para hibridarse con sondas específicas presentes en una laminilla. Esta técnica también es conocida como “cariotipo molecular” (26).

La mayor ventaja de los aCGH es la habilidad de realizar el equivalente a cientos o incluso miles de FISH y MLPA en una sola prueba, además de que no se requieren células en división, se puede comenzar a partir de muy poca cantidad de DNA, es posible la detección de mosaicismo y consume un quinto del tiempo requerido para el cariotipo/FISH. Sin embargo, pese a estas ventajas, presenta algunas desventajas como que no puede detectar desequilibrios de pocas bases ni reacomodos equilibrados (ej. Inversiones) (17). Los estudios de microarreglos, y más recientemente, la secuenciación

de exoma han demostrado la importancia de las variaciones en el número de copias (CNVs) *de novo* y las variaciones de nucleótido único (SNVs) en DI/RD y anomalías congénitas, así como en otras enfermedades complejas como el cáncer.

1.3.1 Secuenciación de exoma.

Probablemente, el método más utilizado de secuenciación dirigida es la secuenciación de exoma (regiones codificantes de proteínas).

La secuenciación de exoma completo (WES, por sus siglas en inglés *Whole-Exome Sequencing*) consiste en la captura, secuenciación y análisis de todos los exones de los genes que codifican para proteínas en el genoma humano. En lugar de analizar el genoma completo, compuesto de 3 mil millones de pb, la WES se enfoca en aproximadamente 30 millones de pb (27) (menos del 2% del genoma (28, 29)) que se traducen a proteínas funcionales; debido a esto, si ocurren mutaciones en esta región es más probable que desencadenen una consecuencia fenotípica grave (27). Se estima que el 85% de las mutaciones que causan enfermedades se localizan en el exoma, por esta razón la WES tiene el potencial de descubrir las causas de un gran número de enfermedades genéticas raras, principalmente monogénicas (29, 30).

Esta tecnología ha emergido como una herramienta de diagnóstico exitosa en el estudio de enfermedades genéticas y ha sido particularmente efectiva en la identificación de nuevos genes asociados a enfermedades. Además ha mostrado una importante utilidad diagnóstica en la clínica. Estudios recientes reportan una tasa de diagnóstico del 27% en una población heterogénea de 250 pacientes, aunque otros estudios reportan tasas más altas (31).

La WES representa una alternativa rentable a la secuenciación del genoma completo (28), en comparación con otros métodos como los estudios de asociación del genoma completo (GWAS, por sus siglas en inglés *Genome-Wide Association Studies*) o la secuenciación del genoma completo (WGS, por sus siglas en inglés *Whole-Genome Sequencing*), se considera un método mucho menos costoso y más eficiente para identificar todas las posibles mutaciones en los genes. La integración de WES en el cuidado médico ya está en marcha, contribuyendo al desarrollo de la medicina “traslacional” personalizada. Analizar el exoma del paciente a través de una única prueba, es ahora menos costoso

que realizar pruebas para un número específico de genes, especialmente cuando se tiene poco conocimiento acerca de los antecedentes genéticos de la enfermedad (27).

En comparación, el éxito diagnóstico de la WES es mayor que el de aCGH. Gillissen y cols. reportaron que de 1,489 pacientes con DI a los cuales se les realizó aCGH, solo el 12% (173 casos) pudo obtener un diagnóstico concluyente; a 100 de los pacientes con diagnóstico inconcluso por aCGH se les realizó secuenciación de exoma completo, gracias a esta técnica se pudo establecer el diagnóstico del 27% de estos pacientes (27 casos) (22). La ventaja de usar WES es que se combina el análisis de mutaciones puntuales (SNPs), Indels, aneuploidías y CNVs para incrementar la resolución y la tasa de detección con una sola prueba. Además esta técnica permite el análisis de SNPs en alelos cuyo par presenta alguna microdelección o microduplicación, por lo que proporciona información sobre la región genómica y los efectos clínicos (32).

Capítulo II: Antecedentes

La tecnología de NGS ha sido exitosa en encontrar las causas de enfermedades mendelianas y raras. Para el estudio de estas enfermedades es deseable (pero no determinante) contar al menos con dos individuos afectados con la enfermedad que no estén relacionados o evidencia de heredabilidad en al menos una familia (33). Hay un incremento en la suposición de que las mutaciones *de novo* pueden ser las causas subyacentes de las condiciones genéticas no diagnosticadas en la clínica. Estos pacientes, típicamente presentan un amplio rango de características clínicas y se mantienen sin un diagnóstico por medio de técnicas que se utilizan en la valoración de enfermedades mendelianas (33). Así, la secuenciación de exoma de un pequeño número de individuos afectados no relacionados es una poderosa y eficiente estrategia para la identificación de genes causantes de enfermedades mendelianas raras que probablemente transformará el análisis genético de los desórdenes monogénicos (34).

En el 2010, Bamshad y cols. demostraron la primera aplicación exitosa de la secuenciación de exoma con el descubrimiento de un gen asociado a una rara enfermedad mendeliana sin causa conocida, el síndrome Miller. A cuatro individuos afectados en tres diferentes familias se les realizó WES, cubriendo el 97% de estas secuencias con una profundidad media de 40X. Filtraron las mutaciones de los pacientes comparándolas con bases de datos de SNPs y contra 8 exomas del HapMap, encontrando dos variantes previamente no reportadas en cada uno de los 4 individuos en un solo gen candidato, *DGODH*, el cual codifica para una enzima en el vía de la síntesis *de novo* de pirimidinas. La secuenciación Sanger confirmó la presencia de estas mutaciones en tres pacientes adicionales con el síndrome Miller (34).

Bale y cols. en el 2016 realizaron un estudio en seis individuos no relacionados con un fenotipo común, el cual consistía en desórdenes en el neurodesarrollo, retraso severo en el lenguaje, hipotonía y dimorfismos faciales. A estos pacientes y a sus padres no afectados se les realizó WES y se identificaron variantes heterocigotas *de novo* posiblemente patogénicas en el gen *KAT6A* no reportadas previamente. Este gen codifica para una histona acetiltransferasa ya conocida por su participación en la leucemia mieloide aguda, pero nunca antes asociada con trastornos congénitos. El ortólogo de este

gen en modelos animales está relacionado con la regulación transcripcional durante el desarrollo (35).

Rump y cols, en el 2016 investigaron la utilidad de la WES para establecer un diagnóstico molecular de un grupo altamente heterogéneo de pacientes con DI variada y microcefalia. Se realizó la secuenciación a 38 pacientes, incluyendo a tres parejas de hermanos. En 10 de las 35 familias estudiadas (29%) se pudo establecer un diagnóstico genético, ya que se encontraron mutaciones en genes ya conocidos por causar una enfermedad en la cual la microcefalia es la característica principal. De estos, dos pacientes no relacionados tuvieron una mutación en el gen *ASPM*; en otros 7 pacientes se encontraron mutaciones en los genes *RAB3GAP1*, *RNASEH2B*, *KIF11*, *ERCC8*, *CASK*, *DYRK1A* y *BRCA2*; y en uno de los pares de hermanos se encontraron mutaciones en el gen *RTTN*. Siete de estos pacientes mostraron una herencia recesiva (36).

La WES también se puede utilizar para identificar el gen causante de una enfermedad en una familia con varios miembros afectados. Semsarian y cols. en el 2014 describieron una familia con una enfermedad cardíaca en la cual existe una marcada heterogeneidad clínica entre los individuos que la padecen. Los miembros afectados mostraban diferentes fenotipos cardíacos, mientras algunos individuos eran asintomáticos, otros presentaban ausencia de compactación de la ventrícula izquierda, paro cardíaco por fibrilación ventricular idiopática, cardiomiopatía dilatada e incluso una muerte repentina inexplicable. En ausencia de un diagnóstico clínico, se utilizó la WES para identificar la mutación causal, encontrando la mutación Ala119Thr en el gen de la alfa actinina 2 (*ACTN2*) que segregaba con la enfermedad. Análisis de los haplotipos de una segunda familia (previamente descrita) con enfermedades cardíacas clínicamente diversas mostraron la presencia de la misma mutación. Los autores proponen que esta mutación fue heredada de un antepasado en común entre ambas familias (37).

La mayoría de los reportes que existen acerca de secuenciación de nueva generación analizan pacientes que comparten un mismo fenotipo clínico o que es muy parecido. Sin embargo el grupo de trabajo de Golstein y cols. decidió realizar un estudio para predecir la tasa de éxito en el diagnóstico en un entorno clínico que incluyera pacientes con una amplia gama de presentaciones fenotípicas, los autores no buscaron pacientes con fenotipos similares, sino que reclutaron cualquier paciente no diagnosticado donde la causa de su enfermedad parecía ser genética, lo cual sería más realista y asemejaría a lo

que se enfrentaría una típica clínica genética. Los autores presentaron un programa piloto donde realizaron WES a 12 pacientes con diferentes condiciones que se sospechaba tenían origen genético y a sus padres no afectados. Gracias a este estudio se diagnosticaron 6 de los 12 pacientes, en 4 de estos se encontraron mutaciones en genes ya asociados con enfermedades mendelianas (*TDF4*, *EFTUD2*, *SCN2A* y *SMAD4*) y en uno de los pacientes, se encontró una mutación en el gen *NGLY1* que interacciona con genes ya asociados a enfermedades. Es importante mencionar que al momento del estudio el gen *EFTUD2* aún no se establecía como causante de una enfermedad mendeliana, pero fue nominado como tal ya que se encontraron mutaciones *de novo* en este gen en otros dos sujetos afectados no relacionados (38).

Posteriormente en el 2015, este mismo grupo de investigación realizó la secuenciación de 119 tríos (nuevamente con una variedad de enfermedades heterogéneas), en los cuales los pacientes no contaban con un diagnóstico genético aún después de haber pasado por varias técnicas de rutina en una clínica genética (31).

Es importante mencionar que en muchas ocasiones no se establece un diagnóstico genético porque, aunque se encuentren mutaciones en un gen que pueda explicar el fenotipo del paciente, si este no ha sido previamente reportado y asociado con una enfermedad, muchas veces no es tomado en cuenta para el diagnóstico final (31). El grupo de Golstein y cols., quiso confrontar este paradigma, implementando nuevas estrategias bioinformáticas que pueden ser de utilidad en la identificación de nuevas asociaciones gen-enfermedad. En el estudio con los 119 tríos, se diagnosticaron 29 pacientes (24%) por WES. En otros 21 pacientes se identificó un gen que fue un fuerte candidato para ser el responsable del fenotipo del paciente y aunque hubo algunos tríos donde no se determinó un diagnóstico genético, el estudio si logró identificar genes candidatos para los pacientes. Algunos tríos tenían una mutación en un gen que ya había sido asociado a una enfermedad en la base de datos OMIM (Online Mendelian Inheritance in Man), sin embargo, debido a que esta asociación seguía en estudio (ya que no se tenían completamente descritas las características que desencadenaban estas mutaciones), en estos tríos no se dictaminó un diagnóstico genético. Interesantemente, 4 tríos tuvieron mutaciones en genes de interés (*ASCL2*, *SLC9A1*, *IL1RAPL2* y *SON*) que no habían sido reportados previamente en la base de datos OMIM. Este grupo de investigación reportó por primera vez la asociación de los genes *SLC9A1* y *SON* con una enfermedad mendeliana (31).

La identificación de las variables responsables del fenotipo de los pacientes en este estudio, se realizó con base en la comparación de que tan patogénicas se predecían que eran las mutaciones *de novo* encontradas en los pacientes, en comparación con las encontradas en un grupo de sujetos control (31).

Así, la secuenciación de exoma es una herramienta muy completa en el análisis de diversas enfermedades, permitiendo encontrar mutaciones responsables del fenotipo de los pacientes que previamente no habían sido detectadas por otras técnicas.

Capítulo III: Justificación

Los estudios en AC y DI están limitados a técnicas de baja y mediana resolución, siendo ineficientes para identificar alteraciones submicroscópicas o crípticas. La aplicación de tecnologías de nueva generación para la caracterización molecular de este tipo de mutaciones permite establecer una relación genotipo-fenotipo más precisa, identificando genes asociados a la etiología de estas enfermedades.

El entendimiento de las causas genéticas de las AC y DI puede beneficiar a los pacientes y sus familias, brindando información diagnóstica más precisa y el asesoramiento respecto a cuestiones genéticas y clínicas, incrementando la posibilidad de elegir una terapia específica y eficiente para los pacientes.

En conjunto, los resultados de este estudio pretenden acortar la brecha entre la investigación básica y la clínica, y establecer los estudios genómicos como abordaje especializado en nuestro sistema de salud.

Hipótesis

El exoma de estos pacientes con diagnóstico molecular genético parcial, que presentan anomalías congénitas y discapacidad intelectual, contiene mutaciones en regiones codificantes que explican su fenotipo.

Capítulo IV: Objetivos

4.1 Objetivo general.

Identificar y caracterizar las alteraciones genómicas en los pacientes con enfermedades congénitas idiopáticas mediante la secuenciación de su exoma.

4.2 Objetivos específicos.

1. Re-evaluar sistemáticamente los resultados de pérdidas de heterocigocidad de los microarreglos de los pacientes para obtener regiones o genes candidatos para la secuenciación de nueva generación.
2. Realizar la secuenciación del exoma completo de los pacientes.
3. Realizar un análisis integral de los datos para identificar las mutaciones que puedan explicar el fenotipo clínico.

Capítulo V: Diseño experimental

5.1 Diseño del experimento.

La presente investigación es un estudio descriptivo, transversal, observacional y retrospectivo con seguimiento. El protocolo fue sometido y aprobado por el Comité de Ética de la Subdirección de Investigación de la Facultad de Medicina, con el número de registro BI15-003.

5.2 Universo de estudio.

El universo de estudio estuvo constituido por individuos con AC y RD/DI de etiología desconocida. Esta muestra se obtuvo de pacientes que acudieron a la consulta de genética del Centro de Investigación Biomédica de Occidente del IMSS en Guadalajara, Jalisco.

5.3 Muestra de estudio.

Se analizaron muestras de DNA de pacientes con AC (así como de sus padres biológicos en caso de ser posible), que previamente habían sido referidos con un médico genetista y continuaban con un diagnóstico genético inconcluso después haberseles realizado técnicas como cariotipo y aCGH.

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

Criterios de inclusión.

- ✓ Pacientes con AC y/o RD/DI referidos por las instituciones participantes.
- ✓ Cariotipo normal o cariotipo alterado sin delineación de la alteración.
- ✓ Que ya se les haya realizado aCGH y el diagnóstico molecular siga siendo inconcluso.

- ✓ Que cuenten con evaluación clínica integral (diferentes especialistas y/o abordajes):
- ✓ Que se hayan descartado factores extrínsecos, como agentes teratógenos, infecciones y exposición a tóxicos como causantes del fenotipo.
- ✓ Padres firmen consentimiento informado.

Criterios de no inclusión.

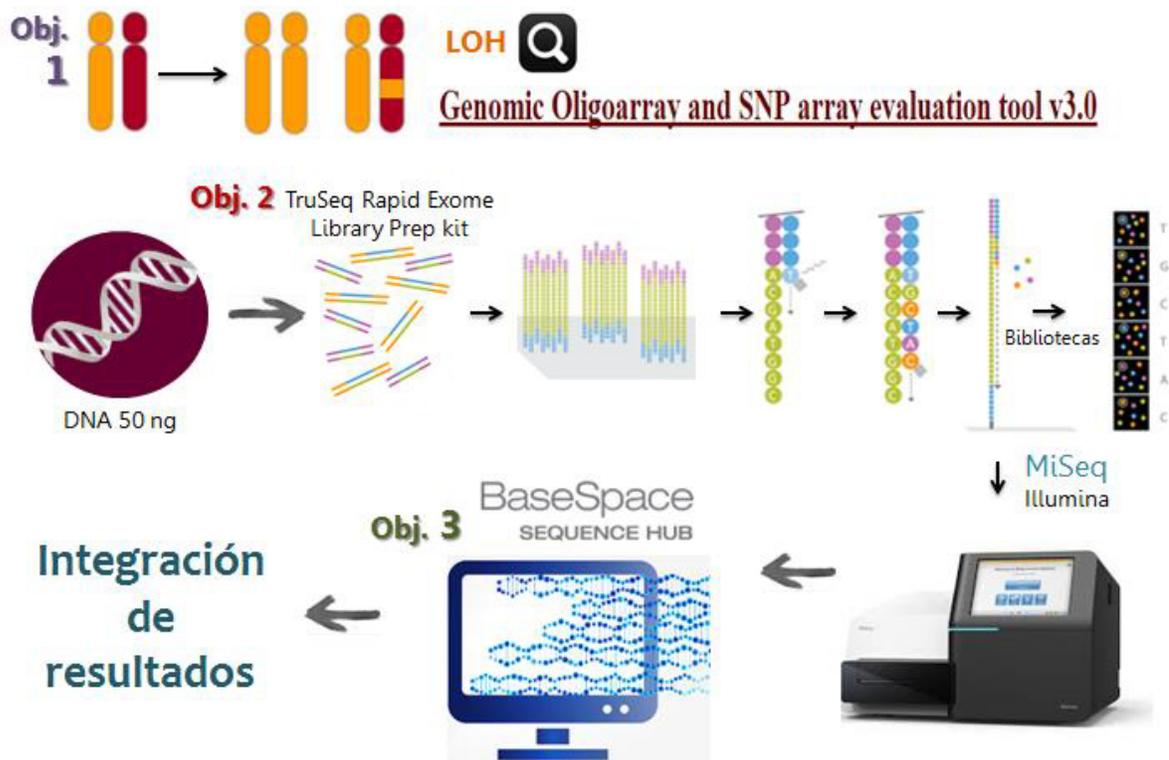
- ✗ Pacientes con MM y RD/DI asociado a un síndrome y alteración genética conocidos.

Criterios de eliminación.

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

Capítulo VI: Metodología

6.1 Estrategia experimental.



6.2 Naturaleza de la muestra.

Las muestras utilizadas en este estudio fueron DNA genómico de 4 pacientes con AC idiopáticas que contaban con diagnóstico genético parcial. Las pacientes 1 y 2 eran hermanas, y también se pudo contar con la muestra de sus ambos padres, por lo que se secuenció la familia completa. Los pacientes 3 y 4 fueron independientes, con diferentes afecciones, y no se secuenció ningún otro miembro de su familia. Las muestras fueron almacenadas a -20°C hasta su uso.

6.3 Evaluación de las pérdidas de heterocigocidad de los microarreglos de los pacientes.

A partir de los reportes de los microarreglos de los pacientes, se identificaron las regiones donde estos presentaron pérdida de heterocigocidad (LOH). Se accedió a la plataforma

Genomic Oligoarray and SNP Array evaluation tool v3.0 (39) y se introdujeron las regiones cromosómicas de interés, se eligió el genoma de referencia GRCh37 y Pérdida de heterocigocidad como el tipo de consulta. Dependiendo del caso, se seleccionó en qué tipo de genes se estaba interesado: todos, solo los registrados en OMIM, registrados en OMIM asociados a enfermedades, o registrados en OMIM con herencia dominante o recesiva. De esta plataforma se obtuvo una lista de genes con su número de registro y el fenotipo al que están asociados. Se compararon estas características con las reportadas en la historia clínica del paciente, y en caso de coincidir, ese gen era considerado como candidato para ser el responsable (o uno de los responsables) del fenotipo del paciente.

6.4 Secuenciación de exoma.

6.4.1 Comprobación de la calidad, cantidad e integridad de las muestras.

Para comprobar la calidad de las muestras, se cuantificó 1 μ l en el equipo NanoDrop 2000 (Thermo Scientific), era necesario que las muestras tuvieran una relación 260/280 de 1.8-2.0 (o muy cercana) para poder ser secuenciadas. Se realizó un gel de agarosa al 1% para comprobar la integridad del DNA genómico, este se corrió por 50 minutos a 90 volts y posteriormente fue visualizado con el equipo ChemiDoc XRS+(Bio-Rad), esperando observar una sola banda en la parte superior del gel sin la presencia de barrido.

Se cuantificaron las muestras con el fluorómetro Qubit (Invitrogen), para esto, se preparó una solución de trabajo, colocando 1 μ l del reactivo Qubit Reagent (Invitrogen) y 199 μ l del reactivo Qubit Buffer (Invitrogen). Posteriormente se prepararon 2 estándares para realizar la curva de calibración, añadiendo a dos tubos 190 μ L de solución de trabajo y 10 μ l de la solución estándar #1 o #2 (0 ng/ μ l y 10 ng/ μ l, respectivamente). En tubos Qubit assay, se realizaron las diluciones de las muestras añadiendo 198 μ l de la solución de trabajado y 2 μ l de muestra a cada tubo, es necesario cuidar que la cantidad de DNA que se añade al tubo Qubit assay sea menor a 100 ng, ya que este es el límite de detección del equipo. Se midieron los estándares con el fluorómetro Qubit y posteriormente las muestras.

6.4.2 Preparación de la biblioteca.

Se realizó la preparación de la biblioteca de cada una de las muestras con el kit TruSeq Rapid Exome Library Preparation (Illumina), el cual incluye >340,000 95mer sondas para la captura de 214,405 exones (Anexo 1). Primeramente a partir de una muestra de gDNA a 50 ng/μL se hizo una dilución a 5 ng/μl en un volumen final de 12 μL con agua libre de nucleasas, esta dilución fue cuantificada con el equipo Qubit. Posteriormente se realizó la tagmentación del DNA, este proceso consiste en la fragmentación del DNA en fragmentos de 300 pb y su etiquetado con secuencias adaptadoras en un solo paso, realizado por medio de un complejo enzimático llamado transposoma (transposasa acoplada a un transposón), una vez finalizado este proceso se colocó la muestra en una placa de 96 pocillos para la purificación del DNA tagmentado utilizando perlas de purificación SPB (del inglés Sample Purification Beads, Illumina) las cuales se unen al DNA por diferencia de cargas, el DNA fue eluído en 22.5 μL de buffer RSB (Resuspension Buffer, Illumina). Se realizó una PCR para la amplificación del DNA tagmentado y para la adición de secuencias índices (i5 y i7), secuencias adaptadoras (P5 y P7) y secuencias complementarias a los primers de secuenciación.

El paso siguiente consistió en dos hibridaciones de las muestras con sondas biotiniladas de captura CEX (Coding Exome Oligos, Illumina) dirigidas contra los exones. Para la captura de las regiones de interés hibridadas a las sondas biotiniladas se utilizaron perlas SMB con estreptavidina (Streptavidin Magnetic Beads, Illumina), las cuales fueron separadas de la solución utilizando una placa magnética. Posteriormente se realizó una amplificación final de la biblioteca utilizando PPC (PCR Primer Cocktail) y EAM (Enrichment Amp Mix).

La biblioteca final se visualizó en un gel de agarosa al 4% para determinar el tamaño promedio de los fragmentos y se cuantificó con el equipo Qubit. Se realizó el cálculo para convertir la concentración de ng/μl a nM utilizando el tamaño promedio de la biblioteca y el peso molecular de un par de bases (Ecuación 1). La biblioteca se diluyó a una concentración de 4 nM, utilizando buffer RSB.

$$\frac{\text{Concentracion en ng}/\mu\text{l}}{\left(660 \frac{\text{g}}{\text{mol}} * \text{tamaño promedio de libreria}\right)} \times 10^6 = \text{concentracion en nM}$$

*Los 660 g/mol = PM de 1 par de bases.

Ecuación 1. Conversión de ng/μl a nM.

6.4.3 Secuenciación.

Se descongeló el cartucho de secuenciación que contenía los reactivos (Reagent Cartridge) usando un baño de agua desionizada a temperatura ambiente, se sumergió la base del cartucho por 60-90 minutos. A la biblioteca a 4 nM se le realizaron dos diluciones seriadas a 20 pM y 12.7 pM usando buffer HT1 (Hybridization Buffer). A la par, se realizaron diluciones de la secuencia control PhiX, a 4 nM, 20 pM y 12.7 pM. Finalmente se preparó la muestra a secuenciar, mezclando 594 µl de biblioteca de la muestra 12.7 pM con 6 µl del control PhiX 12.7 pM.

Se cargó esta última muestra en el cartucho. Se realizó el lavado de la celda de flujo con agua destilada estéril y etanol al 80%. Se colocó la celda de flujo, el cartucho y el buffer PR2 (Incorporation Buffer) en el equipo MiSeq para comenzar la secuenciación pair-end de 75 pb.

6.5 Análisis integral de los datos.

Las secuencias fueron alineadas contra el genoma de referencia GRCh37/Hg19 con el algoritmo BWA Enrichment v.2.1.0.0, y la llamada de variantes se realizó con el software GATK 1.6 (40). Los datos fueron importados a las plataformas Variant Studio v1.0.0 y Variant Interpreter (beta) v.1.0.15015 (Illumina) para su análisis. En el software Variant Interpreter, se aplicaron diversos filtros para que la aplicación solo mostrara las variantes que se predecía eran patogénicas, probablemente patogénicas o con significancia clínica incierta (VUS), homocigotas recesivas con una frecuencia poblacional <0.01 y que pasaran los controles de calidad de la secuenciación ($>Q30$, profundidad mayor a 10). Las variantes que pasaron estos filtros, fueron analizadas con las plataformas Genome Browser (41), OMIM (42), GeneCards y MalaCards para estudiar los genes involucrados, y en que procesos o vías estaban relacionados; así como con las plataformas 1000 Genomes (43), ExAC (44) y NHLBI (45) para determinar su frecuencia poblacional. Después de este segundo filtro, obtuvimos las mutaciones candidatas que tenían alta probabilidad de estar relacionadas con el fenotipo del paciente y una baja frecuencia poblacional.

Estas mutaciones fueron analizadas con las plataformas SIFT (46), PANTHER (47), Mutation Taster 2 (48), PhD-SNP (49), PROVEAN (50) y PolyPhen-2 (51) las cuales

predecían si las variantes eran toleradas (neutras) o patogénicas. Se analizó la conservación de los nucleótidos afectados con las plataformas PhyloP (52), PhastCons (53) y GERP. Finalmente, se analizaron los cambios en la estructura secundaria de las proteínas debidas a estas mutaciones con la plataforma CFSSP de ExPASy (54) y se utilizó la plataforma MuPro (55) para determinar si la estabilidad de la proteína se encontraba afectada.

En caso de que la mutación estuviera localizada en un sitio de splicing, se utilizaron las plataformas ESEfinder (56) y Human Splicing Finder (57) para su análisis.

Capítulo VII: Materiales y Equipos

7.1 Materiales y reactivos.

7.1.1 Material biológico.

Se utilizaron 6 muestras de DNA genómico extraído a partir de sangre periférica.

7.1.2 Material de consumo.

- Micropipeta de 100 a 1000 μL , marca Eppendorf Research Plus.
- Micropipeta de 20 a 200 μL , marca Eppendorf Research Plus.
- Micropipeta de 10 a 100 μL , marca Eppendorf Research Plus.
- Micropipeta de 2 a 20 μL , marca Eppendorf Research Plus.
- Micropipeta de 0.5 a 10 μL , marca Eppendorf Research Plus.
- Micropipeta de 0.1 a 2.5 μL , marca Eppendorf Research Plus.
- Tubos de 15 mL, marca Corning.
- Tubos de 50 mL, marca Corning.
- Toallas desechables de papel, marca Kimberly Clark.
- Puntillas desechables para micropipeta de 1000 μL marca Axygen.
- Puntillas desechables para micropipeta de 200 μL marca Axygen.
- Puntillas desechables para micropipeta de 10 μL marca Axygen.
- Puntillas con filtro desechables para micropipeta de 1000 μL marca Axygen.
- Puntillas con filtro desechables para micropipeta de 200 μL marca Axygen.
- Puntillas con filtro desechables para micropipeta de 10 μL marca Axygen.
- Microtubos de 2 mL, marca Axygen.
- Microtubos de 1.5 mL, marca Axygen.
- Microtubos de 0.5 mL, marca Axygen.
- Microtubos de 0.2 mL, marca Axygen.
- Guantes de nitrilo, marca microflex.
- Tubos Qubit assay 0.6 ml, marca life technologies.
- Marcador de peso molecular 100 pb, marca Invitrogen.
- Agarosa Ultra pura, marca Invitrogen.

- Ácido bórico, marca Amresco.
- EDTA, marca J. T. Baker.
- Tris-Base, marca Sigma-Aldrich.

7.1.3 Materiales y reactivos para secuenciación masiva.

- Kit Qubit dsDNA HS Assay, marca Invitrogen.
- Kit TruSeq Rapid Exome Library Prep, marca Illumina.
- Kit MiSeq Reagent v3.0, marca Illumina.
- Soporte magnético, marca Invitrogen.
- Cubiertas autoadheribles para placas de 96 pocillos, marca Bio-Rad.
- Placas cónicas de 96 pocillos de 0.8 ml, marca Thermo Scientific.
- Tapas de repuestos para los índices, marca Illumina.
- Tris-HCl, marca Sigma-Aldrich.
- Etanol 80% grado biología molecular, marca Sigma-Aldrich.
- PhiX, marcar Illumina.

7.2 Equipos.

7.2.1 Equipo de uso general.

- Minicentrífuga *centrifuge* 5415 C, marca Eppendorf.
- Minicentrífuga, marca Labnet.
- Plancha de calentamiento, marca Thermolyne.
- Vortex, marca Thermolyne.
- Balanza analítica BM-500, marca FraLib.
- Refrigerador 4°/-20°C, marca Thermo Scientific.
- Ultracongelador -80°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.

7.2.2 Equipos para secuenciación masiva.

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

7.3 Programas computacionales y plataformas bioinformáticas.

- Genomic Oligo array and SNP array Evaluation tool (firefly.ccs.miami.edu/cgi-bin/ROH/ROH_analysis_tool.cgi) (39)
- OMIM (omim.org/) (42)
- BaseSpace Sequence Hub (Basespace.illumina.com)
 - BWA Enrichment 2.1.0.0 (basespace.illumina.com/apps/2906905/BWA-Enrichment)
 - Isis 2.5.41.27
 - SAMtools 0.1.19-isis-1.0.3
 - BWA 0.7.7-isis-1.0.0
 - Picard 1.79(1282)
 - GATK (Variant Caller) v1.6-23-gf0210b3 (40)
 - IAS v3
- UCSC Genome Browser (genome.ucsc.edu) (41)
- IGV (basespace.illumina.com/apps/1886885/Integrative-Genomics-Viewer)
- VariantStudio (basespace.illumina.com/apps/639639/VariantStudio)
- VariantInterpreter (variantinterpreter.informatics.illumina.com/registry/cases)
- GeneCards (genecards.org)
- MalaCards (malacards.org)
- ExAC Browser (Beta) Exome Aggregation Consortium (exac.broadinstitute.org) (44)
- 1000 Genomes (internationalgenome.org/1000-genomes-browsers) (43)
- NHLBI Exome Variant Server (evs.gs.washington.edu/EVS) (45)
- PolyPhen-2 (genetics.bwh.harvard.edu/pph2) (51)
- SIFT (sift.bii.a-star.edu.sg) (46)

- PROVEAN (-13 a 4, patogénico<0) (provean.jcvi.org) (50)
- PANTHER (pantherdb.org) (47)
- Mutation Taster 2 (mutationtaster.org) (48)
- PhD-SNP(snps.biofold.org/phd-snp/phd-snp.html) (49)
- PhyloP (puntajes de -14 a 3, conservado>0) (52)
- GERP (mendel.stanford.edu/SidowLab/downloads/gerp) (puntajes de -12.3 a 6.17, conservado>0)
- PhastCons ((puntajes de 0 a 1, conservado=cercano a 1) (53)
- CFSSP de ExPASy (biogem.org/tool/chou-fasman) (54)
- MuPro (mupro.proteomics.ics.uci.edu) (55)
- ESE Finder (56)
- Human Splicing Finder (57)
- Uniprot (58)

Capítulo VIII: Resultados

8.1 Descripción de casos.

8.1.1 Pacientes 1 y 2.

Pacientes femeninas de 21 y 18 años de edad provenientes de padres que niegan consanguinidad y aparentemente sanos. Ambas pacientes presentan el mismo cuadro clínico que consta de retraso psicomotor, DI grave, problemas de lenguaje, talla baja (<3ra percentil) y dismorfias; sin embargo, el fenotipo es más agresivo en la paciente 1. Esta presentó retraso psicomotor desde la lactancia, un problema cardíaco que remitió espontáneamente, bipedestación a los 3 años, anguiloglosia, edad ósea retrasada, prognatismo, hipertriosis, retraso en el desarrollo sexual secundario y problemas para socializar. El fenotipo es menos agresivo en la paciente 2.

A ambas pacientes se les realizó previamente cariotipo y tamizaje metabólico y los resultados fueron completamente normales. Debido a esto se les realizó un microarreglo de CGH, encontrándose que ambas pacientes presentaban en el cromosoma 19 una pérdida de heterocigocidad de copia neutral (CN-LOH) prácticamente en la misma región (Paciente 1: chr19=36,713,905-46,545,925; Paciente 2: chr19=36,715,020-46,952,402) (Figura 4). Pese a este hallazgo, las pacientes continúan con un diagnóstico genético parcial, por lo que en este estudio se realizó la secuenciación de su exoma.

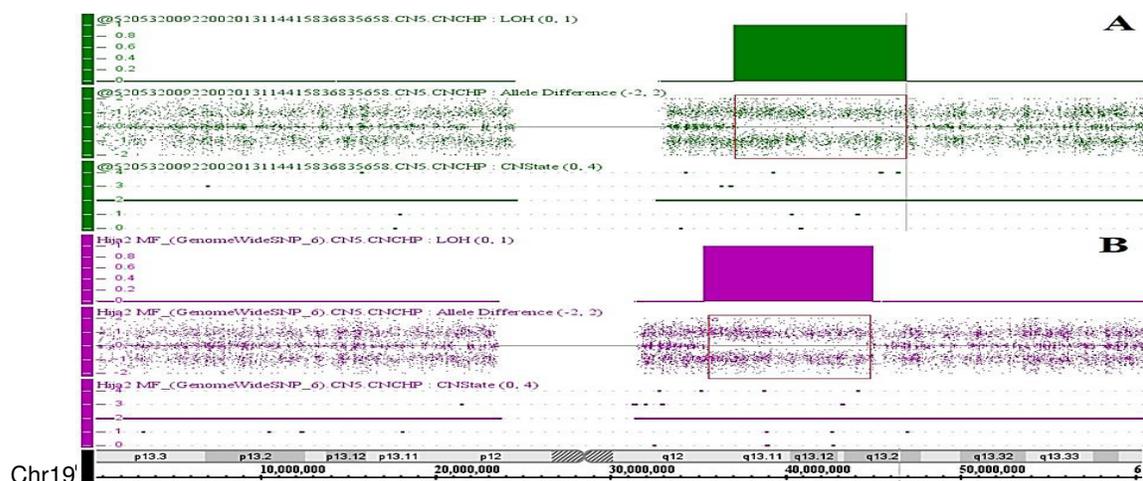


Figura 4. LOH de copia neutral del cromosoma 19 de las pacientes 1 (A) y 2 (B). Plataforma Affymetrix. En este gráfico se observan las regiones de LOH identificadas en las pacientes 1 y 2. Los bloques en verde y en fucsia indican la LOH. Las barras horizontales en verde y fucsia indican el número de copias (2) para el cromosoma 19; no se observan deleciones equivalentes a las regiones de LOH.

8.1.3 Paciente 3.

Paciente masculino de 8 años, estudiado en el primer año de vida por presentar RD grave y AC múltiples; debido a que es adoptado, no se conocen los antecedentes familiares. El paciente fue prematuro, presenta DI, microcefalia, peso, talla y perímetro cefálico en percentiles menores a 3; facies progeroide, escaso cabello, manos con braquidactilia y pliegues redundantes; estenosis valvular pulmonar, cifoescoliosis, retraso en el cierre de la fontanela anterior, edad ósea retrasada, microdoncia, oligodoncia, maloclusión, cuello y manos cortas.

El paciente presentó un cariotipo normal. Se encontró por medio de un aCGH que el paciente presenta largos bloques de CN-LOH con sondas ininterrumpidas, en algunos casos, estos bloques abarcaban prácticamente todo el cromosoma (Figura 5), lo cual es sumamente inusual. Y, aunque se determinó esta característica en el paciente, este continúa con un diagnóstico genético inconcluso, por lo que en este estudio se realizó la secuenciación de su exoma.

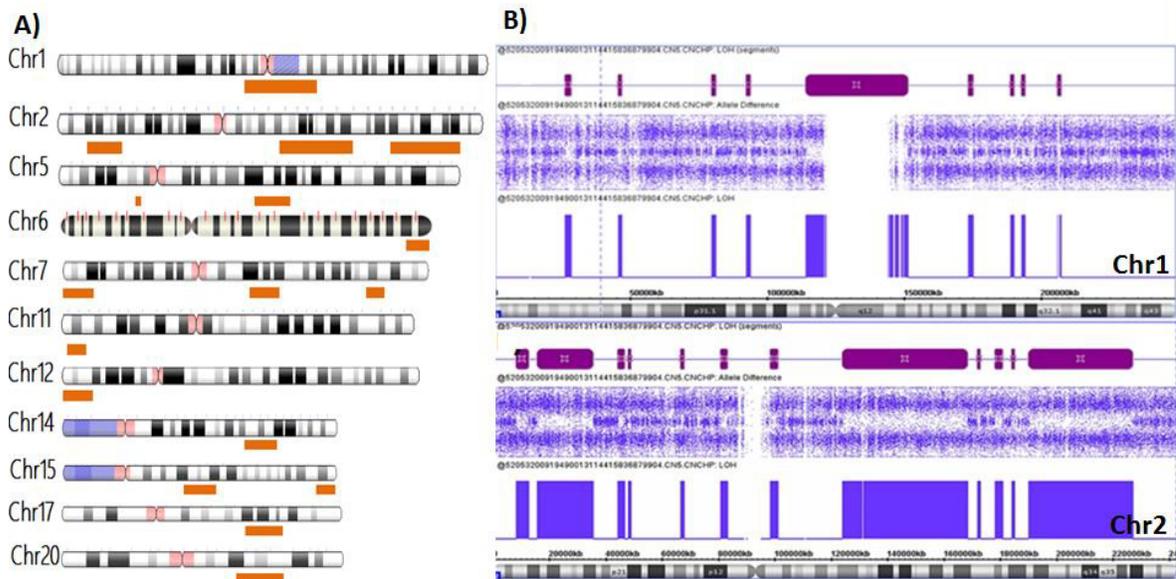


Figura 5. A) LOH de copia neutral de 11 cromosomas afectados en el paciente 3 (bloques naranjas). B). Plataforma Affymetrix. En bloques morados se indican las CN-LOH de los cromosomas 1 y 2 del paciente, observando que se encuentra afectado la mayoría del cromosoma 2. No se observan deleciones equivalentes a las regiones de LOH.

8.3 Evaluación de las pérdidas de heterocigocidad de los microarreglos de los pacientes.

8.3.1 Pacientes 1 y 2.

La región de CN-LOH que compartían las pacientes 1 y 2 era de 10.23 Mb. Esta se analizó con la plataforma Genomic Oligoarray and SNP array Evaluation tool (41), encontrándose 348 genes involucrados, de estos, 41 ya estaban asociados a enfermedades y 27 presentaban un tipo de herencia recesiva. Se seleccionaron 8 por concordar parcialmente con el fenotipo de ambas pacientes: *PEPD*, *GPI*, *MAG*, *SDHAF1*, *WDR62*, *SARS2*, *PLEKHG2* y *MEGF8* (Tabla 2). Sin embargo ninguno de los genes candidatos explicaba por completo todo el cuadro clínico, aparte de estar asociados a otras características clínicas distintivas que no estaban presentes en las pacientes. Por esto, se analizaron los otros 348 genes con el objetivo de identificar genes relacionados a desarrollo neuronal o a procesos celulares de importancia. Se encontraron 31 genes que cumplían con estas características, los cuales no habían sido asociados a enfermedades (Tabla 3).

Tabla 2. Genes dentro de la región de LOH de las pacientes 1 y 2 relacionados a enfermedades recesivas, las cuales comparten algunas características con el cuadro clínico de las pacientes (palomeadas).

Característica	PEPD	GPI	MAG	SDHAF1	WDR62	SARS2	PLEKHG2	MEGF8
Retraso psicomotor severo	✓	✓	✓	✓	✓	✓	✓	
Problemas de lenguaje								
Talla baja								
Dismorfias	✓							
Problema cardiaco al nacimiento que remitió espontáneamente				✓				✓
Anquiloglosia								
Edad ósea retrasada								
Cabello grueso y abundante								
Prognatismo								
Hipertriosis								
OTRAS	Úlceras	Anemia	Paraplejía espástica	Fallo en hígado y riñones	Microcefalia	Fallo en el riñón	Microcefalia	Polisindactilia

Tabla 3. Genes relacionados a desarrollo neuronal o a procesos celulares de importancia, seleccionados como posibles candidatos responsables del fenotipo de las pacientes.

Ciclo celular	Transportadores de membrana/ canales iónicos	Modificaciones postraduccionales	Factores de transcripción/ transcripción	Crecimiento/ motilidad/ desarrollo	Remodelación de la cromatina
<i>PDCD5</i>	<i>SLC7A10</i>	<i>UBA2</i>	<i>USF2</i>	<i>UPK1A</i>	<i>KMT2B</i>
<i>HAUS5</i>		<i>FBXO17</i>	<i>ZNF382</i>	<i>CAPN12</i>	<i>SIRT2</i>
		<i>FBXO27</i>	<i>HKR1</i>	<i>NFKBIB</i>	<i>PAF1</i>
		<i>NCCRP1</i>	<i>HNRNPL</i>	<i>PAK4</i>	
			<i>MED29</i>	<i>GMFG</i>	
			<i>SERTAD3</i>	<i>RAB4B</i>	
Traducción	Metabolismo	Mitocondria/ cadena respiratoria	Placenta	Señalización	
<i>U2AF1L4</i>	<i>PRODH2</i>	<i>TIMM50</i>	<i>LGALS14</i>	<i>ARHGEF1</i>	
<i>eIF3K</i>				<i>GSK3A</i>	
<i>MRPS12</i>					
<i>EXOSC5</i>					

8.3.2 Paciente 3.

Las regiones de CN-LOH del paciente 3 abarcaron más de 11 cromosomas, afectando 247 Mb. Estuvieron involucrados 3246 genes, 333 asociados a enfermedades y de estos, 192 con herencia recesiva. Se seleccionaron 4, *WDR35*, *IRC4*, *C12orf57* y *TTC21B*, por concordar en varias características con el fenotipo del paciente (Tabla 4), contemplándolos como buenos candidatos para ser los responsables.

Tabla 4. Genes dentro de la región de LOH del paciente 3 relacionados a enfermedades recesivas, las cuales comparten algunas características con el cuadro clínico del paciente.

Característica	WDR35	ORC4	C12ORF57	TTC21B
Estenosis de la válvula pulmonar	✓			✓
Convulsiones			✓	
Discapacidad intelectual			✓	✓
Enfermedad de reflujo esofágico		✓		
Constipación crónica				✓
Gastritis				
Retraso en el cierre de la fontanela anterior (2.3 años, normal: 12-18 meses)	✓			
Déficit de atención				
Hipertonia				
Dolicocefalia				
Cifoesciosis				
Edad ósea retrasada				✓
Cuello y manos cortas	✓			✓
Microcefalia	✓	✓	✓	
Talla baja		✓		

Por otro lado, dentro de los 247 Mb de LOH se encontraron 11 genes relacionados a DI, 33 a microcefalia, 21 a estenosis valvular pulmonar, 57 a retraso en el cierre de la fontanela anterior, 5 a dolicocefalia y 3 a cifoescoliosis, por el amplio espectro que presenta el paciente es probable que más de un gen sea el responsable de su fenotipo. La plataforma utilizada permite calcular el coeficiente de endogamia cuando la CN-LOH es mayor a 200 Mb (39), ya que nuestro paciente contaba con 247 Mb afectadas, se obtuvo un valor de 7.81% lo cual nos indica que muy probablemente los padres de este paciente hayan sido consanguíneos, pudiendo explicar los grandes bloques de LOH. Se esperaría que para una persona común el coeficiente de endogamia fuera cercano a cero.

El paciente 4 no presentó CN-LOH.

8.4 Secuenciación de exoma.

8.4.1 Comprobación de la calidad, cantidad e integridad de las muestras.

Se obtuvo la calidad de las muestras y se cuantificaron en el equipo NanoDrop 2000, donde se esperaba una relación 260/280 óptima de 1.8-2.0, podemos observar en la tabla 5 que todas las muestras cumplen con este requisito o se encuentran muy cercanas a este rango, por lo que podemos decir que presentan una buena calidad. Se cuantificaron también por medio del equipo Qubit, donde se determinó la concentración del dsDNA de las muestras (Tabla 5), para comenzar la secuenciación se requerían 50 ng totales, y podemos apreciar como todas las muestras cumplen este requisito. Se corrió un gel de agarosa al 1% para comprobar la integridad de las muestras, como podemos observar en las figuras 7 y 8, todas las muestras de los pacientes, y los padres de las pacientes 1 y 2 cuentan con buena integridad, ya que se observa una banda fuerte en la parte superior del gel, sin ningún barrido por debajo de estas bandas.

Tabla 5. Cuantificación de las muestras por Qubit y NanoDrop.

Muestra	Concentración Qubit ng/μl	Relación 260/280 NanoDrop
Paciente 1	29.6	2.05
Paciente 2	100	2.08
Mamá	85	1.96
Papá	80	1.98
Paciente 3	89.2	1.83
Paciente 4	59.0	1.89
Requerido	Mínimo 50 ng totales	1.8-2.0

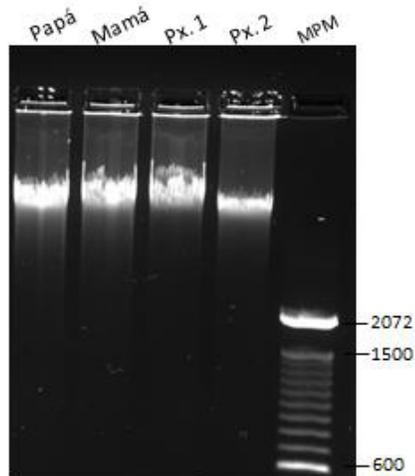


Figura 7. Gel agarosa al 1% del DNA de las pacientes 1 y 2 y sus padres. MPM: marcador de peso molecular.

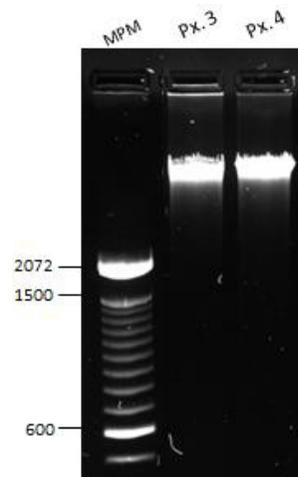


Figura 8. Gel de agarosa al 1% del DNA de los pacientes 3 y 4. MPM: marcador de peso molecular.

8.4.2 Generación de bibliotecas.

Una vez preparada la biblioteca de los pacientes, esta se cuantificó en el equipo Qubit (Thermo Fisher). Se esperaban concentraciones mayores a 3.5 ng/μl y como podemos observar en la tabla 6, todas nuestras muestras cumplieron con este requisito. Se realizó un gel de agarosa al 4% para observar el tamaño de las bibliotecas, esperando un bandeo de las 200 a las 500 pb; podemos observar en la figura 9 que todas las bibliotecas cumplieron con este tamaño de fragmento, se seleccionó un tamaño promedio de biblioteca de 300 pb. Se utilizó la ecuación 1 para el cálculo de la concentración de las bibliotecas en nanomoles (Tabla 6).

Tabla 6. Concentraciones de las bibliotecas medidas por el equipo Qubit.

Muestra	Concentración Qubit ng/μl	Concentración nM
Paciente 1	20.26	102.32
Paciente 2	25.5	128.20
Mamá	11.0	55.55
Papá	27.7	139.89
Paciente 3	12.3	62.12
Paciente 4	14.7	74.24
Requerido	Mínimo 3.5 ng/μl	

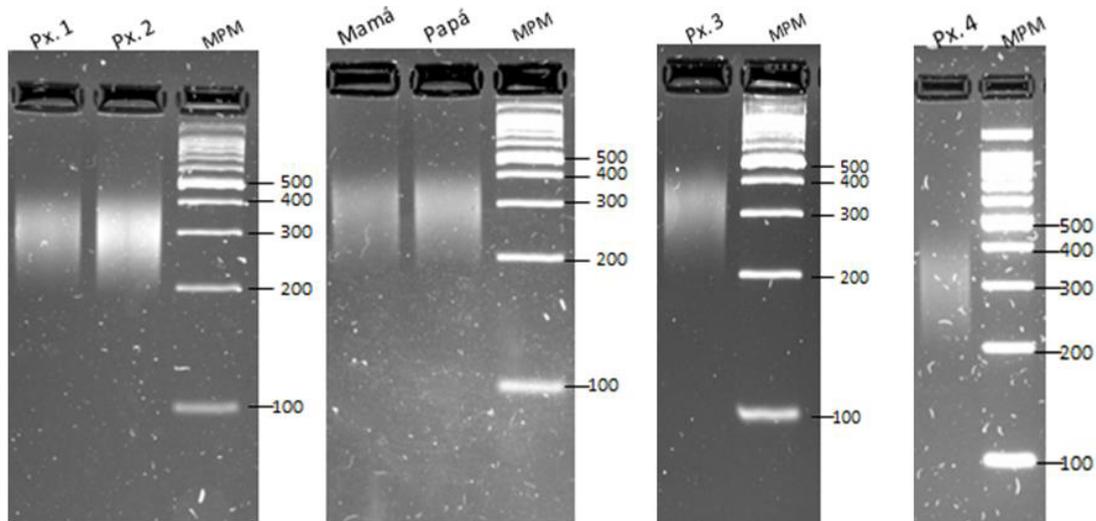


Figura 9. Gel de agarosa al 4% de las bibliotecas de los pacientes. A) Pacientes 1 y 2. B) Padres de los pacientes 1 y 2. C) Paciente 3. D) Paciente 4. MPM: marcador de peso molecular.

8.4.2 Secuenciación.

Se cargó en un cartucho de secuenciación cada una de las bibliotecas a una concentración 12.7 pM, estandarizada como óptima para obtener el número de clústers deseados (entre 1000-1200). Se esperaba que la mayoría de los clústers pasaran el filtro de lectura, esto quiere decir que tuvieran una adecuada dispersión y pudieran emitir una señal fuerte y reconocible, y hubiera un rendimiento cercano a los 4 Gbp (Tabla 7). Se esperaba que la mayoría de las lecturas pasara el valor Q30, esto es que de cada 1000 bases secuenciadas el equipo comete un error en una de ellas, podemos observar en la tabla 8 que más del 96% de las lecturas cumplieron con este valor, e incluso hubo algunas que estuvieron cerca del Q40 (datos no mostrados). Se obtuvo una profundidad promedio para las secuencias de 46X.

- Métricas

Tabla 7. Número de clústers obtenidos por muestra.

Muestra	Densidad Clusters (K/mm ²)	Clusters que pasaron el filtro	Rendimiento
Paciente 1 (Hija1)	1104	93.7%	4.22 Gbp
Paciente 2 (Hija2)	1230	94.6%	4.13 Gbp
Mamá	1302	92.4%	4.30 Gbp
Papá	1132	93.6%	4.23 Gbp
Paciente 3 (Ab)	1024	94.5%	3.95 Gbp
Paciente 4 (T21)	987	94.77%	3.77 Gbp
Óptimo	1000-1200		

Tabla 8. Enriquecimiento a nivel lectura, valor Q30 y profundidad promedio.

Muestra	>Q30	Total de lecturas alineadas	Porcentaje de lecturas alineadas	Lecturas blanco alineadas	Porcentaje de enriquecimiento	Profundidad promedio
Paciente 1	96.2%	52,005,125	99.9%	32,665,139	62.8%	47.6X
Paciente 2	96.4%	54,771,414	99.9%	35,435,646	64.7%	51.6X
Mamá	96.1%	45,082,332	99.9%	30,037,030	66.6%	43.6X
Papá	96.7%	49,210,831	99.9%	31,313,314	63.6%	45.7X
Paciente 3	97.2%	45,262,758	99.9%	31,408,487	69.4%	45.3X
Paciente 4	97.1%	43,831,663	99.9%	29,496,671	67.3%	42.6X

Tabla 9. Enriquecimiento a nivel de pares de bases.

Muestra	Total de bases alineadas	Porcentaje de bases alineadas	Bases blanco alineadas	Porcentaje de enriquecimiento
Paciente 1	3,884,822,837	99.5%	2,157,233,518	55.5%
Paciente 2	4,099,468,680	99.6%	2,340,289,880	57.1%
Mamá	3,349,893,658	99.6%	1,976,001,968	59.0%
Papá	3,677,947,796	99.5%	2,070,726,818	56.3%
Paciente 3	3,396,388,366	99.6%	2,051,705,274	60.4%
Paciente 4	3,275,354,720	99.6%	1,932,539,070	59.0%

En la tabla 10 se muestran los SNVs encontrados por muestra, obteniendo en promedio 27,766, de estos cerca del 98% ya habían sido reportados en dbSNP. En la misma tabla se muestra el número de SNV e INDELS encontrados en genes, regiones codificantes, UTRs y sitios de empalme exónico (splicing). Las variantes fueron divididas según su consecuencia en SNVs sinónimos o no sinónimos, ganancia o pérdida de codones STOP, e inserciones o deleciones causantes de un cambio en el marco de lectura (Tabla 11).

Tabla 10. Resumen de variantes según su localización.

Muestra	SNVs	Porcentaje encontrado en dbSNP	En Genes		Regiones codificantes		UTRs		Sitios de splicing	
			SNV	INDELS	SNV	INDELS	SNV	INDELS	SNV	INDELS
Px 1	28,550	98.2%	27,814	1,506	18,870	298	1,091	113	1,055	131
Px 2	28,923	98.2%	28,328	1,596	19,080	291	1,111	110	1,069	139
Mamá	26,690	98.2%	26,158	1,598	17,530	252	973	116	986	156
Papá	27,959	98.2%	27,474	1,492	18,430	292	1,057	100	1,041	112
Px 3	26,231	97.9%	25,723	1,505	17,355	289	1016	110	983	129
Px 4	28,248	98.1%	27,699	1,444	18,795	297	1,076	134	1,057	113

Tabla 11. Resumen de variantes según su consecuencia.

Muestra	SNVs		Cambio de lectura		Codones STOP	
	Sinónimos	No sinónimos	Inserciones	Deleciones	Ganados	Perdidos
Px 1	9,626	9,141	75	67	83	20
Px 2	9,788	9,185	76	65	85	22
Mamá	9,039	8,397	61	61	79	15
Papá	9,425	8,910	84	65	73	22
Px 3	8,819	8,436	77	68	79	20
Px 4	9,006	9,693	83	70	75	21

8.5 Análisis integral de los datos.

8.5.1 Pacientes 1 y 2.

Se utilizó el software Variant Interpreter para la identificación de las variantes. Después de aplicar los filtros establecidos (mutaciones patogénicas, probablemente patogénicas, con significancia clínica incierta, homocigotas recesivas, frecuencia poblacional <0.01 y que pasaran los filtros de la secuenciación) se obtuvieron 22 variantes para cada una de las hermanas. De estas, las de interés para este estudio eran las que compartían ambas, ya que las dos pacientes presentan el mismo fenotipo. Se encontró que las pacientes compartían 14 variantes involucrando 9 genes (Tabla 12). Se descartaron las mutaciones en *ANKRD36B*, *ZNF717*, *MAML3*, *HLA-DRB5*, *MAP3K4*, *MEOX2*, *ASPN*, *GOLGA6L2* debido a que estaban presentes de manera homocigota también en uno o ambos padres. Concluido esto, resaltaron 3 mutaciones de cambio de sentido casi contiguas en el gen *SNRPA* como posibles candidatas, estas estuvieron presentes en ambas hermanas de manera homocigota, mientras que los padres fueron heterocigotos, interesantemente, estas 3 mutaciones se localizan en la región de pérdida de heterocigocidad de ambas hermanas. Ninguna de estas mutaciones ha sido previamente reportada en las plataformas ExAC, dbSNP, NHLBI.

Tabla 12. Genes con mutaciones recesivas (Frec. Pob.<0.01) presentes en ambas pacientes.

Gen	Nombre	Mutación	Asociado a enfermedades	Herencia
ANKRD36B S P	Ankyrin Repeat Domain 36B	Cambio de sentido c.3127C>G p.Arg1043Gly	No	
ANKRD36B S P	Ankyrin Repeat Domain 36B	Cambio de sentido c.3125A>G p.Tyr1042Cys	No	
ZNF717 S P	Zinc Finger Protein 717	Cambio de sentido c.764T>G p.Val255Gly	No	
ZNF717 S P	Zinc Finger Protein 717	Cambio de sentido c.191A>G p.Tyr64Cys	No	
MAML3	Mastermind Like Transcriptional Coactivator 3	Del. Marco lectura p.Gln507_ Gln510del	No	
HLA-DRB5 S P	Major Histocompatibility Complex, Class II, DR Beta 5	Cambio de sentido c.118G>C p.Asp40His	Fiebre hemorrágica con síndrome renal y Pitiriasis rosada	
HLA-DRB5 S P	Major Histocompatibility Complex, Class II, DR Beta 5	Cambio de sentido c.112C>G p.Gln38Glu	Fiebre hemorrágica con síndrome renal y Pitiriasis rosada	
MAP3K4	Mitogen-Activated Protein Kinase Kinase Kinase 4	Del. Marco lectura c.3596_3598delCTG p.Ala1199del	Cáncer de endometrio	
MEOX2	Mesenchyme Homeobox 2	Del. Marco de lectura c.228_230delCCA p.His80del	Incontinencia de estrés femenina	
ASPN	Abnormal Spindle Microtubule Assembly	Del. Marco de lectura c.153_155delTGA p.Asp51del	Microcefalia, degeneración de disco lumbar, susceptibilidad a osteoartritis.	
GOLGA6L2	Golgin A6 Family-Like 2	No overlapping canonical transcript	-	
SNRPA S P	Small Nuclear Ribonucleoprotein Polypeptide A (U1 SnRNP A)	Cambio de sentido c.97A>G p.Ile33Val	No	
SNRPA S P	Small Nuclear Ribonucleoprotein Polypeptide A (U1 SnRNP A)	Cambio de sentido c.98T>C p.Ile33Thr	No	
SNRPA S P	Small Nuclear Ribonucleoprotein Polypeptide A (U1 SnRNP A)	Cambio de sentido c.100T>A p.Phe34Ile	No	

Las mutaciones c.A97G, c.T98C y c. T100A del gen *SNRPA* de ambas hermanas se alinearon contra el genoma de referencia Hg19 utilizando el software IGV (Figura 10. A). Las 3 mutaciones ocasionaban el cambio de dos aminoácidos, p.Ile33Ala y p.Phe34Ile (Figura 10. B). Con la plataforma CSSFP de ExPASy (54) se logró determinar la estructura secundaria de la proteína WT y de la mutada, encontrando que en esta última había una ganancia de una región rica en hélices- α , la cual no está presente en la proteína WT.

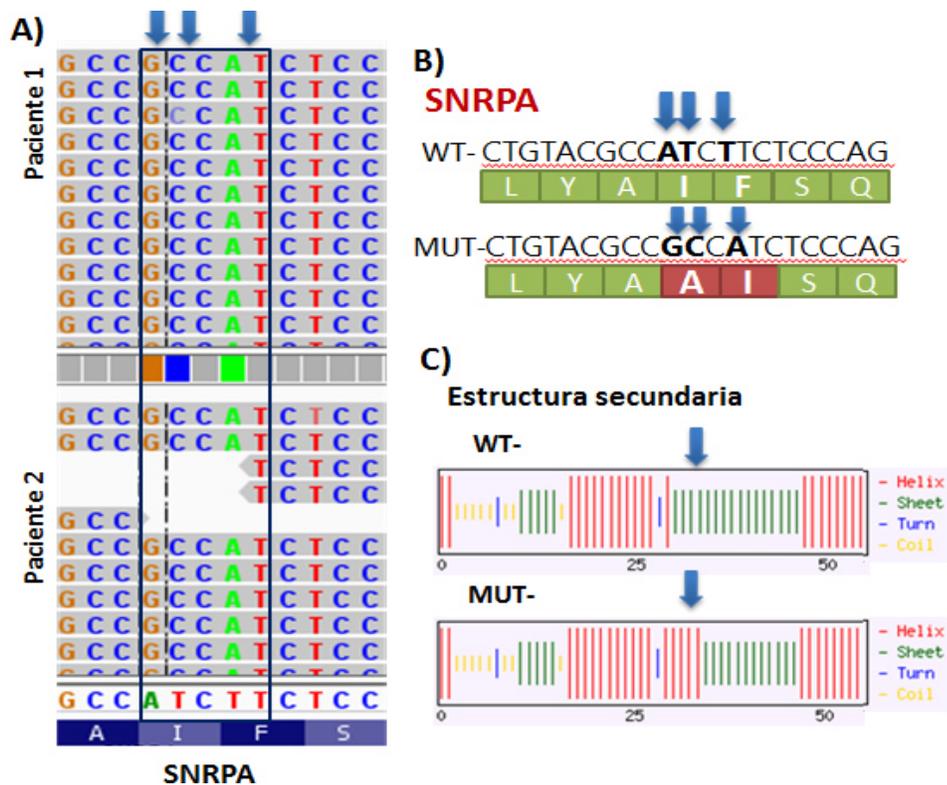


Figura 10. Análisis de las mutaciones Las mutaciones c.A97G, c.T98C y c. T100A del gen *SNRPA*. A) Alineamiento de las lecturas contra el genoma de referencia Hg19. B) Las 3 mutaciones puntuales ocasionan el cambio de dos aminoácidos contiguos. C) Las mutaciones ocasionan la ganancia de una región rica en α -hélices que no está presente en la proteína WT.

Se analizaron ambas mutaciones por separado utilizando 7 diferentes plataformas bioinformáticas, las cuales predicen si las mutaciones son toleradas o patogénicas, encontrando que el cambio en los aminoácidos p.Ile33Ala y p.Phe34Ile es patogénico. Gracias a la plataforma PROVEAN (54) se pudo evaluar la sustitución múltiple, obteniendo un puntaje más patogénico (-8.53) que cuando se evaluaron las dos mutaciones por separado (-3.46 y -5.06). En cuanto a la conservación de los nucleótidos

donde ocurrieron las mutaciones, se encontró que estos eran altamente conservados, y se obtuvieron puntajes de 7.59 en PhyloP, 5.8 en GERP y 1 en PhansCons (Tabla 13).

Tabla 13. Análisis de las variantes p.Ile33Ala y p.Phe34Ile presentes en las pacientes 1 y 2.

Plataforma	Predicción/puntaje	
	Missense C.97AT>GC p.Ile33Ala	Missense C.100T>A p.Phe34Ile
Frecuencia poblacional ExAC	No esta	No esta
Predicción		
PROVEAN	Deletérea (-3.5)	Deletérea (-5.0)
SIFT	Patogénica	Deletérea
PolyPhen-2	Posiblemente patogénica	Deletérea
Mutation taster 2	Patogénica	Patogénica
PhD-SNP	Relacionada a enfermedades	Relacionada a enfermedades
PANTHER	Probablemente patogénica	Probablemente patogénica
Conservación		
PhyloP	7.59	7.59
GERP	5.8	5.8
PhastCons	1	1
Estructura secundaria		
MuPro	Decrece estabilidad	Incrementa estabilidad
CFSSP ExPASy	Cambio de una región con B-plegada a a-hélice	

Se propone que el fenotipo de las pacientes 1 y 2 es causado por estas 3 mutaciones puntuales, ocasionando el cambio de dos aminoácidos, las cuales se encuentran presentes de manera homocigota en ambas hermanas, y de manera heterocigota en los padres no afectados.

8.5.2 Paciente 3.

Se utilizó el software Variant Interpreter para la identificación de las variantes. Después de aplicar los filtros establecidos (mutaciones patogénicas, probablemente patogénicas, con significancia clínica incierta, homocigotas recesivas, frecuencia poblacional <0.01 y que pasaran los filtros de la secuenciación) se obtuvieron 30 variantes (Tabla 14). Se analizaron los genes afectados con el fin de identificar si estaban relacionados a enfermedades que condujeran a un cuadro clínico similar al de nuestros pacientes.

Tabla 14. Genes con mutaciones recesivas (Frec. Pob.<0.01) presentes en el paciente 3.

Gen	nombre	Mutación	Asociado a enfermedades
TAS2R7	Taste receptor, type 2, member 7	Dup. marco de lectura c.703_709dupAAAGCTG p.Val237GlufsTer37	No
SPAG17 S P	Sperm-associated antigen 17	Cambio de sentido c.1069G>C p.Asp357His	No
WDR35 (IFT121) S P	WD repeat domain 35	Missense c.1415G>A p.Arg472Gln	Displasia Craneoectodérmica 2 (#613610 OMIM), Displasia torácica de tronco corto con o sin polidactilia (#614091 OMIM).
PLB1 S P	Phospholipase b1	Cambio de sentido c.3967C>G p.Pro1323Ala	Artritis reumatoide, Esclerosis lateral amiotrófica 3.
CLIP4 S P	Cbl-interacting protein 4	Cambio de sentido c.80C>G p.Ser27Cys	No
ARHGAP15 S P	Rho gtpase-activating protein 15	Cambio de sentido c.185A>G p.Asn62Ser	No
XIRP2 P	Xin repeat protein 2, cardiomyopathy-associated protein 3	Cambio de sentido c.310C>T p.Arg104Trp	Miopatías miofibrilares, osteocondrosis.
CPS1	Carbamoyl phosphate synthetase	Región de splicing Intron c.1183-5T>C	Deficiencia de carbamoil fosfato sintetasa (#237300 OMIM), susceptibilidad a hipertensión pulmonar neonatal (#615371).
FAM157A	Family with sequence similarity 157 member A	Del. En marco de lectura c.246_251delGCAGCA p.Gln83_Gln84del	No
MAML3	Mastermind-like 3	Del. En marco de lectura c.1515_1526delGCAGCAGCAGCA p.Gln507_Gln510del	No
SRA1 S P	Serine- and arginine-rich pre-mrna splicing factor.	Cambio de sentido c.328G>C p.Val110Leu	No
HLA-DRB5 S P	Major histocompatibility complex, class ii, dr beta-5	Cambio de sentido c.112C>G p.Gln38Glu	No
SERAC1	Serine active site-containing protein 1	Region de splicing Intron c.1684+7C>T	Aciduria 3-metilglutaconica con sordera, encefalopatía y síndrome tipo Leigh
TBP	Tata box-binding protein	Dup. Marco de lectura c.213_215dupGCA p.Gln95dup	Ataxia espinocerebral 17 (#607136 OMIM), Susceptibilidad a enfermedad de Parkinson (#168600 OMIM)
THSD7A S P	Thrombospondin type-1 domain-containing	Cambio de sentido c.2399A>G	No

	protein 7a	p.His800Arg	
VWDE S P	Von willebrand factor d and egf domains	Cambio de sentido c.2957A>G p.Asn986Ser	No
MEOX2	Mesenchyme homeobox 2	Del. En marco de lectura c.228_230delCCA p.His80del	Vejiga neurógena, incontinencia femenina por estrés.
ZFHX4	Zinc finger homeobox 4 (*606940 OMIM)	Del. En marco de lectura c.6162_6164delTCC p.Pro2061del	Ptosis congénita (#178300 OMIM).
ANKRD20A4 S P	Ankyrin repeat domain 20 family member a4	Cambio de sentido c.2088A>T p.Gln696His	No
NOP2 S P	Nop2 nucleolar protein	Cambio de sentido c.1279G>A p.Ala427Thr	Asociado a síndrome de Kabuki según GeneCards,
NOP9	Nop9 nucleolar protein	Dup. marco de lectura c.507_509dupGGA p.Glu169dup	No
GOLGA6L2	Golgina 6 family-like 2	No overlapping canonical transcript	No
FAM98B	Family with sequence similarity 98, member b	Del. En marco de lectura c.1269_1277delTGGTGGTGG p.Gly427_Gly429del	No
MEF2A	Mads box transcription enhancer factor 2, polypeptide a	Del. En marco de lectura c.1262_1267delAGCAGC p.Gln421_Gln422del	Enfermedad de arteria coronaria (#608320 OMIM)
RALY	Heterogeneous nuclear ribonucleoprotein associated with lethal yellow, mouse, homolog of, p542	Del. En marco de lectura c.690_691insAGC p.Ala230_Gly231insSer	No
COL18A1	Collagen, type xviii, alpha-1	Del. En marco de lectura c.3378_3386delCCCAGGCC p.Pro1127_Pro1129del	Síndrome Knobloch tipo 1 (#267750 OMIM)
TEX11 S P	Testis-expressed gene 11	Región de splicing c.292C>A p.His98Asn	Falla espermatogénica (#309120 OMIM)
CITED1 S P	Cbp/p300-interacting transactivator with glu/asp-rich c-terminal domain, 1	Cambio de sentido c.215A>G p.His72Arg	Melanoma maligno.
CXorf57 S P	Chromosome x open reading frame 57	Cambio de sentido c.1354G>C p.Val452Leu	No
MECP2 S	Methyl-cpg-binding protein 2	Cambio de sentido c.1315G>A p.Asp439Asn	Síndrome Rett (#312750 OMIM), Encefalopatía neonatal severa (#300673 OMIM), Discapacidad intelectual sindrómica ligada a X 13

Posterior al exhaustivo análisis de las mutaciones encontradas, se seleccionó la mutación en el gen *WDR35* (p.Arg472Gln) como candidata para ser la responsable del fenotipo del paciente, ya que este presentaba varias características encontradas en enfermedades asociadas a este gen (Tabla 15). Mutaciones en *WDR35* se han asociado con ciliopatías, en especial del Síndrome de Polidactilia-Costillas Cortas (SRP), síndrome de Ellis–Van Creveld (EVC) y la Displasia Cráneo-Ectodérmica (CED), distinguiéndose esta última por la presencia de dolicocefalia, la cual está es una de las características del sujeto de estudio. Debido a esto se propone que el paciente presenta CED debida a una mutación en el gen *WDR35*. Una vez que se concluyó que el paciente presentaba una ciliopatía, también se consideró la mutación en el gen *SPAG17* (p.Asp357His) como un posible *loci* modificador, contribuyendo al fenotipo del paciente, ya que también está relacionado a las estructuras ciliares.

Tabla 15. Cuadro clínico de la displasia cráneo-ectodérmica, se resaltan con verde las características que se comparten con el paciente.

WDR35		Cambio de sentido p.Arg472Gln		
Frecuentes >75%	Características faciales (Frente ancha)	✓	Sindactilia	✓
	Braquidactilia	✓	Uñas anormales	
	Tórax estrecho	✓	Retraso en el desarrollo	✓
	Dolicocefalia	✓	Daño a corazón	✓
	Extremidades cortas	✓	Elasticidad en piel	
Comunes 50-75%	Anomalías dentales	✓	Polidactilia	
	Estatura corta	✓	Hernias	
	Poco y/o delgado cabello en frente	✓	Distrofia retinal	
	Nefronoptisis		Higroma quístico	
	Otras características esqueléticas	✓	Displasia de cadera	
	Articulaciones flexibles		Craneosinostosis	
	Daño a hígado			

Menos comunes 25-50%
Poco frecuente <25%

Ambas mutaciones caen dentro de las regiones de CN-LOH del paciente (Figura 5.B). Se muestra a continuación el alineamiento de ambas mutaciones contra el genoma de referencia Hg19 (Figura 11). Por medio de la plataforma MuPro se predijo que ambas mutaciones decrementaban la estabilidad de la proteína por la adición de una porción rica en α -hélices (*WDR35*) o por la ausencia de una de estas regiones (*SPAG17*).

Se analizaron ambas mutaciones utilizando 7 diferentes plataformas bioinformáticas, encontrando que para *SPAG17* todas predecían que esta mutación era patogénica, mientras que para *WDR35*, esto sucedía con 4 plataformas. En cuanto a la conservación de los nucleótidos donde ocurrieron las mutaciones, se encontró que ambos eran altamente conservados. La frecuencia poblacional de la mutación de *SPAG17* es de 0.00019, habiendo solo 22 individuos heterocigotos reportados con esta mutación, para *WDR35*, solo existen 5 individuos heterocigotos para esta mutación (Tabla 16). No se ha reportado ningún individuo homocigoto para estas mutaciones, nuestro paciente sería el primero.

Tabla 16. Análisis de las variantes *SPAG17* p.Asp357His y *WDR35* p.Arg472Gln presentes en el paciente 3.

Plataforma	Predicción/puntaje	
	WDR35 p.Arg472Gln	SPAG17 p.Asp357His
Frecuencia poblacional ExAC	5/121400= 0.0000419	22/114272=0.000192
Predicción		
PROVEAN	Neutral	Deletérea
SIFT proteína	No tolerada	N/A
SIFT dbSNP	Tolerada	Deletérea
PolyPhen-2	Posiblemente patogénica	Deletérea
Mutation taster 2	Patogénica	Patogénica
PhD-SNP	Neutral	Relacionada a enfermedades
PANTHER	Probablemente patogénica	Probablemente patogénica
Conservación		
PhyloP	4.78	6.38
GERP	5.65	4.73
PhastCons	1	1
Estructura secundaria		
MuPro	Decrece estabilidad	Decrece estabilidad
CFSSP ExpASy	+ α -hélice	-- α -hélice

8.5.2 Paciente 4.

Debido a que el paciente presentaba afecciones propias del síndrome de Marfan, se buscaron mutaciones en el gen *FBN1*, principal responsable de este síndrome. Se encontraron 15 mutaciones en este gen, sin embargo, 11 de estas se localizaron en intrones y 3 fueron sinónimas, por lo que se descartaron; se encontró una mutación de cambio de sentido (c.1415G>A, p.Cys472Tyr), sin embargo esta también se descartó por tener una frecuencia poblacional de 1.

El síndrome de Marfan también se asocia, aunque en menor proporción, a la señalización por parte de receptor de TGF- β (MalaCards). Se buscaron mutaciones en el gen *TGFBR2*, encontrándose solo una (c.530-4T>A), sin embargo se descartó por tener una frecuencia poblacional de 0.45. Se analizaron mutaciones en los genes *FBN2* y *FBN3* debido a su relación con *FBN1*, siendo solo significativa una mutación heterocigota que introducía un codón STOP (posición 2803 de 2809) en el gen *FBN3*, sin embargo esta mutación presentó muy pocas lecturas, únicamente 2 para el alelo normal y 1 para el mutado, por lo que no se consideró como posible candidata.

Se realizó un diagnóstico diferencial y se analizaron mutaciones en síndromes parecidos a Marfan (Tabla 17). Se encontró una mutación de cambio de sentido (Frec. Pob.<0.01) en el gen *COL11A2*, sin embargo este gen se asocia con el Síndrome de Stickler tipo 4 el cual no es propio de anomalías en el corazón ni en los ojos, por lo que esta mutación se descartó.

Se buscaron mutaciones en todos los genes de las familias *FBN*, *TGF- β* , *SMAD*, *LBTP*, *SKI*, *MYH11*, *ACTA2*, *MYLK*, *ATT*, *PLOD*, *CBS*, *ADAMTS* y *MMPs*, así como en los genes que codifican para colágeno. Se encontró una mutación (c.1736-2A>G) en el sitio aceptor de splicing dentro del intrón 6 del gen *ADAMTS4*. Esta mutación se analizó con la plataforma ESEfinder (56) y Human Splicing Finder (57), las cuales predijeron que la mutación era probablemente patogénica y que es altamente probable que el splicing de este gen este afectado. No se han relacionado enfermedades al gen *ADAMTS4*, sin embargo, mutaciones en la familia *ADAMTS* y mutaciones en el gen *ADAMTSL4* se han asociado con ectopia lentis, síntoma presente en el paciente. Con base en esto se propone la mutación en el gen *ADAMTS4* (c.1736-2A>G) como posible candidato para explicar el fenotipo del paciente, sin embargo es importante realizar más estudios para comprobar la patogenicidad de esta mutación.

Tabla 17. Diagnóstico diferencial de afecciones similares a Marfan.

Fenotipo	Genes relacionados/ Herencia	Característica distintiva	Comparación con el Paciente 4
Fenotipo MASS	FBN1 / AD	Dilatación de raíz aórtica (Z score <2), NO ectopia lentis	Si presenta ectopia lentis
Síndrome de prolapso de la válvula mitral	-	Prolapso de la válvula mitral, dilatación de raíz aórtica (Z score <2), NO ectopia lentis	No prolapso de válvula mitral. Si presenta ectopia lentis
Síndrome ectopia lentis	FBN1 / AD	Dilatación de raíz aórtica (Z score <2), no mutaciones patogénicas en FBN1 asociadas a otro síndrome.	
Síndrome Shprintzen-Goldberg (SGS)	SKI / AD (regulador negativo de TGF-B)	Mutaciones en SKI, el cual codifica para un regulador negativo de la vía TGF-B	Se descarta, presenta una mutación de cambio de sentido tolerada.
Síndrome Loey-Dietz (LDS)	TGFBR1, TGFBR2, SMAD3	Mutaciones en genes relacionados a la vía TGF-B	Mutaciones intrónicas en TGFBR1 y TGFBR2. Mutación benigna en UTR 5' de SMAD3
Aracnodactilia contractural congénita (CCA)	FBN2 / AD	Dedos más largos y delgados	No presenta mutaciones en FBN2
Diseción aórtica y aneurismo aórtico torácico familiar (TAAD)	TGFBR2, TGFBR1, MYH11, ACTA2, MYLK, SMAD3, AAT1, AAT2 / AD	Ninguna otra manifestación clínica más que dilatación aórtica	El paciente presenta Ectopia lentis, no mutaciones importantes en genes relacionados
Síndrome Ehlers-Danlos (EDS)	COL5A1, COL5A2 / AD	Hipermovilidad en articulaciones. Hiperelasticidad en la piel.	No mutaciones en esos genes
Homocistinuria	CBS / AR	Deficiencia en cistationina-β-sintetasa	No mutaciones en esos genes
Síndrome Stickler	COL2A1, COL11A1, COL11A2/AD COL9A1, COL9A2/AR	Anomalías en ojos, cataratas, labio y paladar hendido	Mutación de cambio de sentido p.Pro779Leu. Se descarta debido a que mutaciones en COL11A2 se asocian al Síndrome Stickler tipo 4, él se caracteriza por la ausencia de anomalías en ojos o en corazón
Síndrome X frágil	FMR1/ >número de CGG	Discapacidad intelectual, hipermovilidad en las articulaciones, características faciales	No mutaciones en FBN1.

AD: Autosómica dominante (puede ser causada por mutaciones heterocigotas); AR: Autosómica recesiva.

Por otro lado se analizaron las mutaciones que fueran patogénicas, probablemente patogénicas o con significancia clínica incierta, con herencia recesiva y frecuencia poblacional <0.01, encontrándose 21 variantes (Tabla 18). Se analizaron los genes involucrados en estas mutaciones con las plataformas OMIM (42), GeneCards y MalaCards para encontrar genes candidatos relacionados al fenotipo del paciente. Se destaca la mutación p.Gln261del del gen *MED15*, ya que este gen está relacionado a la vía de TGF- β , involucrada en el síndrome de Marfan. Aunque se predice que esta mutación es tolerada según la plataforma PROVEAN (50), es considerada como un gen candidato para explicar el fenotipo del paciente por su relación con la vía TGF- β .

Tabla 18. Genes con mutaciones recesivas (Frec. Pob.<0.01) presentes en el paciente 4.

Gen	Consecuencia	Gen	Consecuencia
CDCP2	Cambio marco lectura c.1224dupC p.Met409HisfsTer45	RBM23	Dup. Marco lectura p.Ala393dup
MIR205HG	Del. Marco lectura p.Ala94_Ala97del	CHRNA3	Del. Marco lectura p.Leu23del
OR2T35	Cambio marco lectura p.Ala317Thr	MPRIIP	Del. Marco lectura p.Ser189del
ANKRD36B	Cambio marco lectura p.Arg1043Gly	CD3EAP	Del. Marco lectura p.Lys428del
ANKRD36B	Cambio marco lectura p.Tyr1042Cys	COL18A1	Del. Marco lectura p.Pro1127_Pro1129del
ZNF717	Cambio marco lectura p.Val255Gly	XKR3	Cambio marco lectura p.Pro232Leu
ZNF717	Cambio marco lectura p.Tyr64Cys	MED15	Del. Marco lectura p.Gln262del
MAML3	Del. Marco lectura p.Gln507_Gln510del	STS	Cambio marco lectura p.Arg583His
MAP3K4	Del. Marco lectura p.Ala1199del	ARHGAP6	Cambio marco lectura p.Ala846Pro
ANKRD20A4	Cambio marco lectura p.Gln696His	BEX1	Cambio marco lectura p.His68Arg
		FLNA	Cambio marco lectura p.Arg563Cys

Capítulo IX: Discusión

Las AC son comunes, costosas en muchos aspectos y tienen consecuencias críticas, ocurren en uno de cada 33 nacimientos, lo que se estima en un aproximado de 7.9 millones de nacidos anualmente a nivel mundial. En Estados Unidos cada año se destinan cerca de 2.6 billones de dólares para su tratamiento, pero dicho monto no considera los costos indirectos para las familias, además de las afectaciones sociales de los pacientes (59).

El cariotipo es la técnica de rutina para el diagnóstico genético de pacientes con AC y DI/RD; sin embargo, esta técnica cuenta con una resolución de aproximadamente 5 Mb, lo que se refleja en una tasa de diagnóstico exitosa de solamente el 3% (60). Esta resolución se ha mejorado a lo largo de los años por medio de la implementación de técnicas que permiten el estudio del genoma o exoma completo, como aCGH y la NGS, pudiendo detectar variaciones de hasta un único nucleótido; ambas han permitido incrementar la tasa de diagnósticos exitosos considerablemente (61).

Según la Organización Mundial de la Salud (OMS), actualmente se desconocen las causas del 50% de los casos de AC. Esto probablemente se debe a que las tecnologías de alta resolución para el diagnóstico que detectan mutaciones submicroscópicas o crípticas, como aCGH o secuenciación, no se aplican en todos los pacientes; por lo tanto, si se implementara este tipo de diagnóstico en la clínica, el porcentaje de AC idiopáticas disminuiría. Un claro ejemplo de esto son los pacientes 1, 2 y 3 de este estudio, los cuales fueron sometidos a cariotipo y aCGH previamente, sin poder determinar un diagnóstico genético definitivo por estas técnicas; sin embargo, al estudiarse por medio de WES, fue posible la identificación de mutaciones clave, probablemente responsables de su fenotipo, demostrándose las claras ventajas y la alta resolución y tasa de diagnóstico de esta técnica.

En el caso de las pacientes 1 y 2, se lograron identificar 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 con significancia clínica incierta por la plataforma Variant Interpreter, probablemente debido a que no han sido reportadas previamente, siendo este el primer estudio que las reporta. Los análisis de predicción de cada mutación por separado en todas las plataformas bioinformáticas sugieren que las dos mutaciones son patogénicas; sin embargo, la única herramienta que permitió realizar el análisis de ambas sustituciones al mismo tiempo fue PROVEAN (50), obteniendo un puntaje más patogénico al que se obtuvo cuando se analizaron por separado. Se observa la ganancia de una región rica en α -hélices en la predicción de la estructura secundaria de la proteína mutada, y aunque los aminoácidos afectados (Ile y Phe) y mutados (Ala e Ile) son no polares, esto puede ser debido a la diferencia en la estructura de sus cadenas laterales, generando interacciones distintas con los aminoácidos cercanos.

El producto del gen *SNRPA* forma parte de la ribonucleoproteína U1, la cual se une al sitio de splicing 5' de los mRNAs precursores durante el proceso de corte y empalme. Los componentes de la maquinaria de splicing se unen a secuencias conservadas: el sitio donador (extremo 5' del intrón), el sitio de ramificación (cerca del extremo 3' del intrón) y el sitio aceptor (extremo 3' del intrón) para ensamblar el spliciosoma, un complejo multicomponente encargado del splicing. El spliciosoma incluye 5 ribonucleoproteínas nucleares pequeñas, las cuales se denominan snRNP U1, U2, U4, U5 y U6 (62, 63).

Las otras dos proteínas con las que SNRPA forma el complejo U1 son SNRP70 y SNRPC (42, 64). Según la plataforma Uniprot (58), 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.

Actualmente se tiene poco conocimiento de enfermedades congénitas que están relacionadas específicamente con *SNRPA*. Se han descrito apenas 3 pacientes con grandes desequilibrios cromosómicos involucrando este gen (ClinVar, 2017), los cuales presentan retraso psicomotor y/o del desarrollo, y un caso con una mutación de sitio de empalme que aparentemente ocasiona alteraciones musculares y nerviosas (DECIPHER).

Existen estudios de NGS donde mutaciones en genes de splicing se han asociado a diversas enfermedades, dentro de las cuales destacan desórdenes esqueléticos y craneofaciales. La disostosis mandibulofacial del tipo Guion-Almeida incluye malformaciones en oídos, microcefalia y DI; es causada por haploinsuficiencia, mutaciones heterocigotas de pérdida de función o deleciones en el gen *EFTUD2*, un componente de la snRNP U5. También se han reportado mutaciones en *SF3B4*, componente de U2, en pacientes con Síndrome Nager, el cual incluye defectos en las extremidades superiores y fenotipo craneofacial; este gen también ha sido asociado a disostosis acrofacial del tipo Rodríguez, que incluye fenotipos craneofaciales similares a los observados en el síndrome de Nager y fallo respiratorio que puede conducir a la muerte. Mutaciones heterocigotas en *SNRPB*, que forma parte de las proteínas Sm, causan Síndrome Cerebro-Costo-Mandibular; los individuos afectados presentan la secuencia Robin y defectos en las costillas. Asimismo, mutaciones en *SNRPE*, otra proteína Sm, causan un tipo dominante de hipotricosis aislada. La evidencia anterior lleva a pensar que, aunque tanto SNRPB como SNRPE forman parte de las proteínas Sm, mutaciones en cada una de las 7 diferentes proteínas que conforman el complejo Sm pueden conllevar a una amplia variedad de fenotipos diferentes (65). Se ha reportado también que una deleción de 6.4 Kb que abarca el exón 1 del gen *SNRPN* causa síntomas característicos del Síndrome de Prader-Willi, los cuales incluyen hipotonía, características dismórficas, DI y obesidad (66).

Además de su participación en el proceso de splicing, cuando la proteína SNRPA no se encuentra formando complejos, es capaz de fungir como un regulador de la poliadenilación de algunos RNAs mensajeros (67). Hasta ahora se ha encontrado que puede autoregular su propio pre-RNA_m (68), así como los pre-RNA_m de *COX-2* (69), *IgM* (70) y *SMN1* (67); dicha función probablemente también esté afectada de algún modo por las mutaciones presentes en las pacientes 1 y 2.

Es importante mencionar que SNRPA regula el pre-RNA_m de *SMN1*, un gen que está asociado a la atrofia muscular espinal (SMA). Al unirse SNRPA a este pre-RNA_m se bloquea el sitio de corte de la proteína CPSF73, ocasionando que no sea reclutada la polimerasa poli(A) e impidiendo por lo tanto la poliadenilación de este pre-RNA_m (67). La SMA causa degeneración y pérdida de neuronas motoras en el asta anterior de la médula espinal, lo que provoca debilidad muscular progresiva y, en casos graves, insuficiencia respiratoria y muerte (71). Las pacientes 1 y 2 comparten algunas características

presentes en SMA, como retraso psicomotor severo, problemas de lenguaje, talla baja y problemas cardiacos. Es posible que estas las mutaciones generen una proteína SNRPA con función aberrante, que pueda estar interactuando con el pre-RNA_m de *SMN1*, acumulando o disminuyendo los niveles de esta proteína, y con esto, ocasionando un fenotipo parecido al de SMA. Sin embargo, se desconoce si las mutaciones en *SNRPA* sean realmente las causantes del fenotipo o solo una de las varias causas del mismo.

La presencia de estas mutaciones en ambas pacientes es un fenómeno raro, y aunque la presencia de CN-LOH en el genoma puede ser común, estas últimas son extremadamente raras en el cromosoma 19 (72). Esto lleva a pensar que la CN-LOH de ambas hermanas es resultado de un fenómeno conocido como autociguidad (73, 74), esto significa que es probable que sus padres tengan un ancestro común no muy lejano y por esto, no ha sido posible romper mediante recombinación los bloques de homociguidad de su cromosoma 19 a lo largo del tiempo, teniendo entonces ambos padres los mismos alelos mutantes pero heterocigotos.

Se descartó el posible solapamiento del cuadro clínico de las pacientes con el fenotipo asociado al síndrome de microdelección 19q13, debido a que las pacientes no presentaron pérdida de material genético. También se descartó una CN-LOH por disomía uniparental, ya que al momento de analizar las sondas de SNPs de los microarreglos de las pacientes y los padres, se observó la presencia de algunos alelos con herencia paterna y otros con herencia materna. Serán necesarios estudios posteriores para esclarecer cómo afectan las mutaciones p.Ile33Ala y p.Phe34Ile a la proteína SNRPA, y comprobar que efectivamente afectan su función de reconocimiento y por lo tanto su capacidad de unión a diferentes RNAs mensajeros; dichos análisis permitirán concluir con mayor certeza si las mutaciones estudiadas son las responsables de desencadenar el desarrollo del fenotipo de las pacientes. Interesantemente, además de las consecuencias esperadas de dificultar el empalme exónico de sus genes blanco por no ser capaz de unirse a ellos, la potencial ausencia de su autoregulación podría sugerir un incremento en la producción de proteínas mutantes.

En el caso del paciente 3, se lograron identificar mutaciones en dos genes, *SPAG17* (p.Asp357His) y *WDR35* (p.Arg472Gln), como posibles responsables del fenotipo del paciente. Estas mutaciones se predijeron como patogénicas por diversas plataformas bioinformáticas, ya que afectan nucleótidos altamente conservados. La mutación

p.Asp357His en *SPAG17* solo ha sido reportada en 22 individuos de manera heterocigota, mientras que p.Arg472Gln en *WDR35*, se ha reportado únicamente en 5 personas heterocigotas (44). Nuestro paciente presenta ambas mutaciones de manera homocigota debido a la CN-LOH que presenta en gran parte de su genoma.

No se conocen los antecedentes familiares del paciente 3 debido a que es adoptado, sin embargo, el coeficiente de consanguinidad de 7.81% obtenido indica que es altamente probable que provenga de un matrimonio consanguíneo. El coeficiente de consanguinidad para un hijo de dos primos hermanos es de 3.12% (75), de modo que es posible inferir que los padres del paciente 3 tienen una consanguinidad mucho más cercana a este tipo de parentesco.

En el caso de *SPAG17* ocurre el cambio de un aminoácido ácido (Asp) el cual es hidrofílico, tiene carga negativa y normalmente se localiza en la superficie de las proteínas, por un aminoácido básico (His), el cual también es hidrofílico pero se encuentra cargado positivamente; este aminoácido desempeña un papel activo en reacciones enzimáticas que implican el intercambio de iones hidrógeno. Por otro lado, en *WDR35*, ocurre un cambio de un aminoácido básico (Arg), hidrofílico, cargado positivamente, por un aminoácido no polar (Gln), el cual no interactúa con el agua; este cambio puede repercutir en el funcionamiento de la proteína.

Mutaciones en *WDR35* se han asociado con las ciliopatías SRP, EVC y CED (76). Los cilios son proyecciones de la membrana celular de aproximadamente 200 nm de diámetro (77). Estas complejas y dinámicas estructuras se dividen en cilios móviles y no móviles (primarios); ambos comparten un citoesqueleto de 25 μm de diámetro, denominado axonema, compuesto de cientos de proteínas. El axonema contiene nueve dobletes de microtúbulos externos formados por un túbulo A completo, que contiene 13 protofilamentos y un túbulo B incompleto que contiene 10 u 11 protofilamentos (78). Los cilios primarios consisten únicamente en los dobletes de microtúbulos periféricos, teniendo una conformación "9+0", mientras que los cilios móviles contienen un par central de microtúbulos, los cuales están conectados a los microtúbulos periféricos mediante complejos de proteínas llamados espinas radiales, se dice que estos tienen una conformación "9+2" (79). Existen múltiples cilios móviles en algunas células especializadas, entre las que se encuentran las células epiteliales de las vías respiratorias, trompas de Falopio, y tráquea, entre otros; casi todas las células humanas

poseen un único cilio primario, involucrado en procesos de señalización y sensoriales, principalmente (78). Se han identificado más de 1000 diferentes polipéptidos en el proteoma de las estructuras ciliares (79).

Las ciliopatías son enfermedades asociadas con la disfunción ciliar (79), debido a que los cilios primarios se localizan en la mayoría de las células del cuerpo, las mutaciones en estos genes afectan a una gran variedad de órganos y tejidos (80). Se han relacionado más de 80 genes con al menos algún tipo de ciliopatía (81). Existen dos complejos de proteínas conocidos como transportadores intraflagelares (IFT); uno de los complejos está encargado del transporte de proteínas al ápice del cilio (transporte anterógrado, complejo IFT-B) para la síntesis de este, mientras que la función del otro es el transporte de proteínas hacia la base del cilio (transporte retrógrado, complejo IFT-A) para que posteriormente sean expulsadas hacia el citoplasma (81).

WDR35 también es conocido como *IFT121* y forma parte, junto con *IFT43*, del complejo IFT-A. Se han reportado 22 casos de ciliopatías relacionadas a mutaciones en este gen (76), siendo el paciente 3 el caso reportado número 23; no obstante, este es el único paciente que presenta una mutación en el exón 13. La displasia craneoéctodérmica, causada por mutaciones en el gen *WDR35*, se distingue de otras ciliopatías por la presencia de dolicocefalia, la cual está también en el paciente. Con base en las mutaciones encontradas y el cuadro clínico es posible sugerir que el paciente presenta CED. La proteína *WDR35* también está involucrada en el reconocimiento y regulación de los cargos para su transporte, la entrada y salida de proteínas del cilio y la organización del centriolo (82). Se ha observado que pacientes con mutaciones en *WDR35* tienen cilios más cortos comparados contra personas control, presentando complicaciones en el aclaramiento de las vías aéreas (83). Modelos *Knock down* con siRNA para *WDR35* muestran un déficit en la ciliogénesis, generándose cilios más cortos y menos móviles, lo que demuestra lo esencial de este gen en la función motora y en la estructura de los cilios; por otro lado, ratones *Knock out* para *WDR35* mueren a los 12 días, resaltando la importancia de este en la viabilidad de los organismos. No se han encontrado pacientes vivos que presenten mutaciones sin sentido o deleciones en este gen, probablemente porque no son compatibles con la vida, lo que habla de la alta penetrancia y la importancia del buen funcionamiento este (84). Como se comentó anteriormente, las diferentes mutaciones en *WDR35* presentan una variabilidad clínica considerable, esto debido al tipo de mutación ocurrida y/o a la presencia de *loci* modificadores, los cuales

son mutaciones en otros genes que pueden estar incrementando la severidad de la enfermedad. En este trabajo se ha considerado la mutación en *SPAG17* como un *loci* modificador o viceversa (que *SPAG17* sea el responsable del fenotipo y *WDR35* esté actuando como *loci* modificador).

SPAG17 codifica para una proteína presente en el par central de los axonemas, la cual es importante para la estructura e integridad de los mismos. Se ha observado que mutaciones presentes en *SPAG17* ocasionan un decremento en el movimiento de los cilios traqueales o nasales, acumulación de moco, muerte en 12 horas debida a fallo respiratorio y defectos en el par central, generándose en ocasiones solo un microtúbulo en vez de dos. Ratones *Knock out* para este gen presentan cilios pequeños e inmóviles y el 23.7% de los cilios sin par central (85).

SPAG17 se había asociado previamente sólo con los cilios móviles por su papel en el par central; sin embargo, Teves y cols. en el 2015, encontraron que ratones *Knock out* presentaban malformaciones esqueléticas (desarrollo alterado del esternón) y anomalías en hueso (fémur y tibia más cortos), como consecuencia de una disfunción en los condrocitos y osteoblastos ocasionadas por el deficiente y alterado desarrollo del cilio primario, el cual se observaba mucho más corto que en los controles. Debido a esto se sugirió que *SPAG17* también tiene un papel crucial en el mantenimiento de la estructura y en las funciones de señalización del cilio primario, por lo cual es esencial para su buen funcionamiento y formación (86, 87). Estos hallazgos complementan y/o confirman la asociación de SNVs de *SPAG17* con alteraciones en la altura de los individuos y lo sitúan como un gen importante en el desarrollo esquelético.

Es posible decir que debido al parecido del fenotipo del paciente con las ciliopatías y el análisis de los datos, es altamente probable que las mutaciones en *SPAG17* y *WDR35* sean las responsables del fenotipo del paciente.

El paciente 4, además del cuadro clínico del Síndrome de Down, presenta afecciones características del síndrome de Marfan, como dilatación de la raíz aórtica, cataratas y luxación del cristalino (ectopia lentis). El síndrome de Marfan es un desorden autosómico dominante que afecta al tejido conectivo, este es fundamental en el soporte de piel, huesos, vasos sanguíneos y otros órganos; las personas que presentan este síndrome tienen problemas en el corazón, huesos, ojos, piel, vasos sanguíneos, sistema nervioso y

pulmones (42). En la mayoría de los casos, el Síndrome de Marfan es causado por mutaciones en el gen *FBN1*, el cual codifica para la proteína fibrilina 1, principal constituyente de las microfibrillas de la matriz extracelular (88), sin embargo este síndrome también se asocia con alteraciones en la vía TGF- β .

TGF- β es producido en forma de dímero en las células y está unido a la proteína asociada a latencia (LAP) para formar el complejo latente pequeño (SLC). El SLC secretado se une extracelularmente a la proteína de unión a TGF- β latente (LTBP) para formar el complejo latente largo (LLC). El LLC se une posteriormente a las microfibrillas de fibrilina 1 mediante LTBP. La palabra latencia en el nombre de estas proteínas hace referencia a su función de mantener de forma inactiva a TGF- β . Cuando ocurren mutaciones en el gen de la fibrilina 1, la LLC se vuelve incapaz de unirse a las microfibrillas, por lo que la forma latente no es generada, resultando en un nivel elevado de TGF- β , el cual se une a su receptor y activa vías de señalización canónicas (Smad2/3, Smad4) y no canónicas (ERK1, JNK1, MAPK, MEK1, p38, Ras-GTP, ShcA, TAK1, TRAF6) (89). La activación de estas vías conlleva al incremento de varios genes blanco, como *CTGF*, *IL6*, *MCP1*, *MMPs*, *PAI-1* y TGF- β mismo (90). Mutaciones en *TGFBR2* también se han asociado con Síndrome de Marfan.

Al no encontrarse variantes en *FBN1* ni en *TGFBR2*, se analizaron las mutaciones en los genes relacionados a síndromes parecidos a Marfan, se encontraron mutaciones en los genes *COL11A2*, *COL18A1*, *ADAMTSL5* y *ADAMTS4*. De estas se descartaron las mutaciones en *COL11A2* (c.2336C>T p.Pro779Leu), *COL18A1* (c.3562G>A p.Val1188Ile) y *ADAMTSL5* (c.659G>T p.Arg220Leu) por no estar relacionadas a anomalías de corazón y ojos, por ser de herencia recesiva cuando el paciente es heterocigoto y por presentar una alta frecuencia poblacional respectivamente.

La mutación en *ADAMTS4* (c.1736-2A>G) se considera un candidato interesante ya que cae en un nucleótido altamente conservado, localizado en el sitio aceptor de splicing, el cual consta de tres pares de bases, [C/T]AG. El paciente 4 presenta una mutación que afecta este sitio, [C/T]GG, lo cual probablemente pueda estar causando un splicing aberrante de esta proteína. Ya se ha reportado que mutaciones en el sitio aceptor de splicing resultan en defectos en el procesamiento del RNAm afectado, generando deleciones de algunos exones (91). Mutaciones en genes de la familia ADAMTS y mutaciones en *ADAMTSL4* han sido asociadas a ectopia lentis, la cual está presente en el

paciente, apoyando la propuesta de que este gen es un candidato a ser el responsable de su fenotipo. Sin embargo es necesaria la generación de un modelo que permita evaluar y confirmar la patogenicidad de la mutación, para poder validar los resultados propuestos.

Analizando las mutaciones recesivas, destacó una delección en el gen *MED15* (c.784_786delCAG p.Gln262del); el producto de este gen es parte un complejo mediador que parece ser requerido para la regulación de la expresión de los transcritos de la RNA polimerasa II. Este hallazgo fue de interés por que se ha reportado que *MED15* juega un papel esencial en la vía de TGF- β /Smad2/3 (92). La mutación detectada está presente en una región rica en glutaminas, y aunque la plataforma PROVEAN (50) la marca como benigna, es notorio su parecido con la etiología de la enfermedad de Huntington, donde está involucrada la longitud de una región poliQ, por lo que tal vez delecciones o inserciones de glutaminas en esta región del gen *MED15* puedan dar lugar a una proteína no funcional; sin embargo también para esta variante quedan muchos estudios por hacer con el fin de esclarecer si la mutación es patogénica o no.

Aunque se encontraron estas dos mutaciones candidatas en el paciente 4, es necesario realizar más estudios para comprobar que puedan ser las causantes del fenotipo observado. Es necesario también analizar las regiones UTR y los microRNAs del paciente en busca de alguna anomalía, ya que son regiones genómicas reguladoras de la transcripción. Es importante también determinar el origen del marcador cromosómico que se encontró en el paciente por cariotipo; para poder realizar dicho análisis por secuenciación, es necesario contar con una mayor profundidad de muestra (aproximadamente 500X), ya que el marcador se presenta en mosaico, no estando en todas las células del cuerpo, lo que dificulta su detección por NGS.

Es importante destacar el papel de la WES en el diagnóstico de AC idiopáticas, ya que por medio de esta técnica fue posible brindar un diagnóstico genético a 3 de los 4 pacientes estudiados, lo cual no había sido posible previamente utilizando cariotipo y microarreglos de aCGH. Finalmente, vale la pena resaltar que las mutaciones encontradas en estos pacientes, están en proceso de ser agregadas a una base de datos internacional del NCBI (ClinVar), enfocada en almacenar mutaciones con interés clínico, como requisito para la publicación de futuros artículos en correcciones.

Capítulo X: Conclusiones

- Se propone que las mutaciones en el gen *SNRPA* son las causantes del fenotipo de las pacientes 1 y 2, ya que este gen regula el pre-RNA de *SMN1*, gen asociado a atrofia muscular espinal.
- El fenotipo del paciente 3 está relacionado a mutaciones homocigotas en los genes *WDR35* y *SPAG17*, las cuales están dentro de las regiones con LOH de copia neutral del paciente. Según las predicciones *in silico*, estas mutaciones patogénicas ocasionan cambios en la estructura y función de las proteínas.
- Se determinaron mutaciones en 2 genes candidatos que podrían explicar parcialmente el fenotipo del paciente 4.
- Se logró encontrar mutaciones relacionadas con el fenotipo de los pacientes, estableciendo una relación genotipo-fenotipo, en el 75% de los casos.
- La hipótesis se cumplió en el 75% de los casos.

Capítulo XI: Perspectivas

- Dilucidar como afectan las mutaciones de *SNRPA* en *SMN1*. Hacer ensayos inhibiéndolo y sobreexpresándolo, así como investigar qué otros RNAm regula.
- Realizar ensayos para esclarecer cómo afectan las mutaciones que nosotros encontramos en *WDR35* y *SPAG17* en las estructuras ciliares, y sus posibles consecuencias.
- Analizar secuencias reguladoras como UTRs y microRNAs y/o estudiar el patrón de metilación del paciente 4. Determinar el origen del marcador cromosómico.

Capítulo XII: Bibliografía

1. Hannon R PC, Porth C. **Concepts of Altered Health States.** *Lippincott Williams & Wilkins.* 2009. p. 1-5.
2. WHO. **Congenital Anomalies.** World Health Organization (WHO). 2016.
3. Czeizel AE. **Birth defects are preventable.** *Int J Med Sci.* 2005;2(3):91-2.
4. DeSilva M, Munoz FM, Mcmillan M, Kawai AT, Marshall H, Macartney KK, et al. **Congenital anomalies: Case definition and guidelines for data collection, analysis, and presentation of immunization safety data.** *Vaccine.* 2016;34(49):6015-26.
5. Abubakar I, Tillmann T, Banerjee A. **Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.** *Lancet.* 2015;385(9963):117-71.
6. Carlson BM. **Human Embryology and Developmental Biology.** Online Access: *Elsevier Health Sciences.* 2012.
7. Francine R, Pascale S, Aline1a H. **Congenital anomalies: prevalence and risk factors.** *Universal Journal of Public Health.* 2014;2(2):58-63.
8. Hossein MA, Kargar Maher MH, Afsharnia F, Dastgiri S. **Prevalence of congenital anomalies: a community-based study in the Northwest of Iran.** *ISRN pediatrics.* 2014.
9. Lobo I, Zhaurova K. **Birth defects: causes and statistics.** *Nature Education.* 2008;1(1):18.
10. Valdés-Hernández J, Canún-Serrano S, Reyes-Pablo AE, Navarrete-Hernández E. **Mortalidad por defectos al nacimiento en menores de 5 años de edad en México de 1998 a 2006.** *Salud pública de México.* 2009;51(5):381-9.
11. Navarrete Hernández E, Canún Serrano S, Reyes Pablo AE, Romero S, del Carmen M, Valdés Hernández J. **Prevalencia de malformaciones congénitas registradas en el certificado de nacimiento y de muerte fetal: México, 2009-2010.** *Boletín médico del Hospital Infantil de México.* 2013;70(6):499-505.
12. Argentina. MdSd. **Enfermedades poco frecuentes y anomalías congénitas.** Argentina [http://www.msal.gov.ar/congenitas/introduccion-a-la-genetica].

13. McIntyre S, Blair E, Goldsmith S, Badawi N, Gibson C, Scott H, et al. **Congenital anomalies in cerebral palsy: where to from here?** *Developmental Medicine & Child Neurology*. 2016.
14. RG. **Congenital malformations, Stillbirths and Perinatal Deaths.** *Johns Hopkins Bloomberg School of Public Health*. 2007 [<http://www.msal.gov.ar/congenitas/introduccion-a-la-genetica>].
15. Médicos DdS. **Anomalías menores TX, EE.UU.** *Texas Tech University*. 2016.
16. Barboza-Arguello MdIP, Umaña-Solís LM. **Análisis de diez años de registro de malformaciones congénitas en Costa Rica.** *Acta méd costarric*. 2008;50(4):221-9.
17. Stankiewicz P, Beaudet AL. **Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation.** *Current opinion in genetics & development*. 2007;17(3):182-92.
18. Shaikh TH. **Oligonucleotide arrays for high-resolution analysis of copy number alteration in mental retardation/multiple congenital anomalies.** *Genetics in Medicine*. 2007;9(9):617-25.
19. Weis R. **Introduction to abnormal child and adolescent psychology.** *Sage Publications*. 2017.
20. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. **Prevalence of intellectual disability: a meta-analysis of population-based studies.** *Research in developmental disabilities*. 2011;32(2):419-36.
21. De Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, et al. **Diagnostic exome sequencing in persons with severe intellectual disability.** *New England Journal of Medicine*. 2012;367(20):1921-9.
22. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, et al. **Genome sequencing identifies major causes of severe intellectual disability.** *Nature*. 2014.
23. Pathology Do. **FISH.** *Stanford Medicine*. 2013 [<http://cytogenetics.stanford.edu/fish.html>].
24. Chial H. **Cytogenetic methods and disease: Flow cytometry, CGH, and FISH.** *Nature Education*. 2008;1(1):76.
25. O'connor C. **Fluorescence in situ hybridization (FISH).** *Nature Education*. 2008;1(1):171.

26. de Vries BB, Pfundt R, Leisink M, Koolen DA, Vissers LE, Janssen IM, et al. **Diagnostic genome profiling in mental retardation.** *The American Journal of Human Genetics.* 2005;77(4):606-16.
27. Bertier G, Héту M, Joly Y. **Unsolved challenges of clinical whole-exome sequencing: a systematic literature review of end-users' views.** *BMC Medical Genomics.* 2016;9(1):52.
28. Illumina. **Introduction to Exome Sequencing.** 2016 [http://www.illumina.com/techniques/sequencing/dna-sequencing].
29. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, et al. **Genetic diagnosis by whole exome capture and massively parallel DNA sequencing.** *Proceedings of the National Academy of Sciences.* 2009;106(45):19096-101.
30. Rabbani B, Tekin M, Mahdieh N. **The promise of whole-exome sequencing in medical genetics.** *Journal of human genetics.* 2014;59(1):5-15.
31. Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu Y-F, McSweeney KM, et al. **Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios.** *Genetics in Medicine.* 2015;17(10):774-81.
32. Russo CD, Di Giacomo G, Cignini P, Padula F, Mangiafico L, Mesoraca A, et al. **Comparative study of aCGH and Next Generation Sequencing (NGS) for chromosomal microdeletion and microduplication screening.** *Journal of prenatal medicine.* 2014;8(3-4):57.
33. Illumina. **Genetic disease Review.** 2012.
34. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. **Exome sequencing identifies the cause of a mendelian disorder.** *Nature genetics.* 2010;42(1):30-5.
35. Millan F, Cho MT, Retterer K, Monaghan KG, Bai R, Vitazka P, et al. **Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder.** *American Journal of Medical Genetics Part A.* 2016;170(7):1791-8.
36. Rump P, Jazayeri O, van Dijk-Bos KK, Johansson LF, van Essen AJ, Verheij JB, et al. **Whole-exome sequencing is a powerful approach for establishing the etiological diagnosis in patients with intellectual disability and microcephaly.** *BMC medical genomics.* 2016;9(1):7.
37. Bagnall RD, Molloy LK, Kalman JM, Semsarian C. **Exome sequencing identifies a mutation in the ACTN2 gene in a family with idiopathic ventricular**

- fibrillation, left ventricular noncompaction, and sudden death.** *BMC medical genetics*. 2014;15(1):99.
38. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, McDonald MT, et al. **Clinical application of exome sequencing in undiagnosed genetic conditions.** *Journal of medical genetics*. 2012;jmedgenet-2012-100819.
39. Wierenga KJ, Jiang Z, Yang AC, Mulvihill JJ, Tsinoremas NF. **A clinical evaluation tool for SNP arrays, especially for autosomal recessive conditions in offspring of consanguineous parents.** *Genetics in Medicine*. 2012;15(5):354-60.
40. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. **The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.** *Genome research*. 2010;20(9):1297-303.
41. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. **The human genome browser at UCSC.** *Genome research*. 2002;12(6):996-1006.
42. **Online Mendelian Inheritance in Man, OMIM®.** Baltimore, MD: McKusick-Nathans Institute of Genetic Medicine, *Johns Hopkins University*; [<https://omim.org/>].
43. Consortium GP. **A global reference for human genetic variation.** *Nature*. 2015;526(7571):68-74.
44. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. **Analysis of protein-coding genetic variation in 60,706 humans.** *Nature*. 2016;536(7616):285-91.
45. **Exome Variant Server**, NHLBI Exome Sequencing Project (ESP) Seattle, WA2011.
46. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. **SIFT missense predictions for genomes.** *Nature protocols*. 2016;11(1):1.
47. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. **PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements.** *Nucleic Acids Research*. 2016:gkw1138.
48. Schwarz JM, Cooper DN, Schuelke M, Seelow D. **MutationTaster2: mutation prediction for the deep-sequencing age.** *Nature methods*. 2014;11(4):361-2.
49. Capriotti E, Calabrese R, Casadio R. **Predicting the insurgence of human genetic diseases associated to single point protein mutations with support**

- vector machines and evolutionary information.** *Bioinformatics*. 2006;22(22):2729-34.
50. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. **Predicting the functional effect of amino acid substitutions and indels.** *PloS one*. 2012;7(10):e46688.
51. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. **A method and server for predicting damaging missense mutations.** *Nature methods*. 2010;7(4):248-9.
52. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. **Detection of nonneutral substitution rates on mammalian phylogenies.** *Genome research*. 2010;20(1):110-21.
53. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. **Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes.** *Genome research*. 2005;15(8):1034-50.
54. Chou PY, Fasman GD. **Prediction of protein conformation.** *Biochemistry*. 1974;13(2):222-45.
55. Cheng J, Randall A, Baldi P. **Prediction of protein stability changes for single-site mutations using support vector machines.** *Proteins: Structure, Function, and Bioinformatics*. 2006;62(4):1125-32.
56. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. **ESEfinder: a web resource to identify exonic splicing enhancers.** *Nucleic acids research*. 2003;31(13):3568-71.
57. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. **Human Splicing Finder: an online bioinformatics tool to predict splicing signals.** *Nucleic acids research*. 2009;37(9):e67-e.
58. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. **UniProt: the universal protein knowledgebase.** *Nucleic acids research*. 2004;32(suppl 1):D115-D9.
59. Feldkamp ML, Carey JC, Byrne JL, Krikov S, Botto LD. **Etiology and clinical presentation of birth defects: population based study.** *BMJ*. 2017;357:j2249.
60. Riegel M. **Human molecular cytogenetics: from cells to nucleotides.** *Genetics and molecular biology*. 2014;37(1):194-209.
61. Gullotta F, Biancolella M, Costa E, Colapietro I, Nardone AM, Molinaro P, et al. **Prenatal diagnosis of genomic disorders and chromosome abnormalities**

- using array-based comparative genomic hybridization. *Journal of prenatal medicine*. 2007;1(1):16.
62. Chen M, Manley JL. **Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches.** *Nature reviews Molecular cell biology*. 2009;10(11):741-54.
63. Will CL, Lührmann R. **Spliceosome structure and function.** *Cold Spring Harbor perspectives in biology*. 2011;3(7):a003707.
64. Turunen JJ, Niemelä EH, Verma B, Frilander MJ. **The significant other: splicing by the minor spliceosome.** *Wiley Interdisciplinary Reviews: RNA*. 2013;4(1):61-76.
65. Lehalle D, Wieczorek D, Zechi-Ceide R, Passos-Bueno M, Lyonnet S, Amiel J, et al. **A review of craniofacial disorders caused by spliceosomal defects.** *Clinical genetics*. 2015;88(5):405-15.
66. Cao Y, AlHumaidi SS, Faqeih EA, Pitel BA, Lundquist P, Aypar U. **A novel deletion of SNURF/SNRPN exon 1 in a patient with Prader-Willi-like phenotype.** *European Journal of Medical Genetics*. 2017.
67. Workman E, Veith A, Battle DJ. **U1A regulates 3' processing of the survival motor neuron mRNA.** *Journal of Biological Chemistry*. 2014;289(6):3703-12.
68. Boelens WC, Jansen EJ, van Venrooij WJ, Stripecke R, Mattaj IW, Gunderson SI. **The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA.** *Cell*. 1993;72(6):881-92.
69. Hall-Pogar T, Liang S, Hague LK, Lutz CS. **Specific trans-acting proteins interact with auxiliary RNA polyadenylation elements in the COX-2 3'-UTR.** *RNA*. 2007;13(7):1103-15.
70. Phillips C, Jung S, Gunderson SI. **Regulation of nuclear poly (A) addition controls the expression of immunoglobulin M secretory mRNA.** *The EMBO Journal*. 2001;20(22):6443-52.
71. Simone C, Ramirez A, Bucchia M, Rinchetti P, Rideout H, Papadimitriou D, et al. **Is spinal muscular atrophy a disease of the motor neurons only: pathogenesis and therapeutic implications?.** *Cellular and molecular life sciences*. 2016;73(5):1003-20.
72. Ling-Hui L, Sheng-Feng H, Chen C-H, Chun-Yu W, Wan-Ching W, Li-Ying L, et al. **Long contiguous stretches of homozygosity in the human genome.** *Human mutation*. 2006;27(11):1115.

73. Monies D, Maddirevula S, Kurdi W, Alanazy MH, Alkhalidi H, Al-Owain M, et al. **Autozygosity reveals recessive mutations and novel mechanisms in dominant genes: implications in variant interpretation.** *Genetics in Medicine*. 2017.
74. Carr IM, Bhaskar S, O'Sullivan J, Aldahmesh MA, Shamseldin HE, Markham AF, et al. **Autozygosity mapping with exome sequence data.** *Human mutation*. 2013;34(1):50-6.
75. Rédei GP. **Encyclopedia of genetics, genomics, proteomics, and informatics.** Springer Science & Business Media. 2008.
76. Smith C, Lamont RE, Wade A, Bernier FP, Parboosingh JS, Innes AM. **A relatively mild skeletal ciliopathy phenotype consistent with cranioectodermal dysplasia is associated with a homozygous nonsynonymous mutation in *WDR35*.** *American Journal of Medical Genetics Part A*. 2016;170(3):760-5.
77. Mukhopadhyay S, Badgandi HB, Hwang S-h, Somatilaka B, Shimada IS, Pal K. **Trafficking to the primary cilium membrane.** *Molecular Biology of the Cell*. 2017;28(2):233-9.
78. Mitchison HM, Valente EM. **Motile and non-motile cilia in human pathology: from function to phenotypes.** *The Journal of pathology*. 2017;241(2):294-309.
79. Waters AM, Beales PL. **Ciliopathies: an expanding disease spectrum.** *Pediatric Nephrology*. 2011;26(7):1039-56.
80. Hildebrandt F, Benzing T, Katsanis N. **Ciliopathies.** *New England Journal of Medicine*. 2011;364(16):1533-43.
81. Braun DA, Hildebrandt F. **Ciliopathies.** *Cold Spring Harbor perspectives in biology*. 2017;9(3):a028191.
82. Bacino CA, Dhar SU, Brunetti-Pierri N, Lee B, Bonnen PE. ***WDR35* mutation in siblings with Sensenbrenner syndrome: a ciliopathy with variable phenotype.** *American journal of medical genetics Part A*. 2012;158(11):2917-24.
83. Li Y, Garrod AS, Madan-Khetarpal S, Sreedher G, McGuire M, Yagi H, et al. **Respiratory motile cilia dysfunction in a patient with cranioectodermal dysplasia.** *American journal of medical genetics Part A*. 2015;167(9):2188-96.
84. Arts H, Knoers N. **Cranioectodermal dysplasia.** *GeneReviews*. 2013.
85. Teves ME, Zhang Z, Costanzo RM, Henderson SC, Corwin FD, Zweit J, et al. **Sperm-Associated Antigen–17 Gene Is Essential for Motile Cilia Function and**

- Neonatal Survival.** *American journal of respiratory cell and molecular biology.* 2013;48(6):765-72.
86. Teves ME, Sundaresan G, Cohen DJ, Hyzy SL, Kajan I, Maczis M, et al. **Spag17 Deficiency results in skeletal malformations and bone abnormalities.** *PloS one.* 2015;10(5):e0125936.
87. Teves ME, Nagarkatti-Gude DR, Zhang Z, Strauss JF. **Mammalian axoneme central pair complex proteins: Broader roles revealed by gene knockout phenotypes.** *Cytoskeleton.* 2016;73(1):3-22.
88. Kaartinen V, Warburton D. **Fibrillin controls TGF- β activation.** *Nature genetics.* 2003;33(3):331-2.
89. Benke K, Ágg B, Szilveszter B, Tarr F, Nagy ZB, Pólos M, et al. **The role of transforming growth factor-beta in Marfan syndrome.** *Cardiol J.* 2013;20(3):227-34.
90. Levine RA, Hagège AA, Judge DP, Padala M, Dal-Bianco JP, Aikawa E, et al. **Mitral valve disease morphology and mechanisms.** *Nature Reviews Cardiology.* 2015.
91. Thongnoppakhun A, Rungroj N, Wilairat P, Vareesangthip K, Sirinavin C, Yenchitsomanus P-t. **A novel splice-acceptor site mutation (IVS13-2A> T) of polycystic kidney disease 1 (PKD1) gene resulting in an RNA processing defect with a 74-nucleotide deletion in exon 14 of the mRNA transcript.** *Human mutation.* 2000;15(1):115.
92. Kato Y, Habas R, Katsuyama Y, Näär AM, He X. **A component of the ARC/Mediator complex required for TGF β /Nodal signalling.** *Nature.* 2002;418(6898):641-6.

Capítulo XIII: Anexos

Anexo 1: Preparación de Bibliotecas (Kit TruSeq Rapid Exome Library Prep).

Tagmentación de DNA

Cuantificación y normalización de gDNA

1. Cuantificar el gDNA usando el equipo Qubit
2. Normalizar el gDNA a 10 ng/μl con tris-HCl 10 mM pH8.5 o agua libre de nucleasas.
3. Cuantificar el gDNA normalizado con el equipo Qubit.
4. Diluir el gDNA normalizado con Tris-HCl 10 mM pH 8.5 o agua libre de nucleasas a un volumen final de 10 μl a 5ng/μl (50 ng en total) y cuantificar nuevamente (contemplar 2 μl para la cuantificación).

Tagmentación de DNA

1. Añadir los siguientes reactivos en orden a un tubo de PCR.
 - a. TD 25 μl
 - b. gDNA normalizado 10 μl
 - c. TDE2 15 μl
2. Mezclar muy bien por pipeteo, dar un spin. Poner en el termociclador el programa TAG58:
 - a. 58°C 10 min
 - b. Sostener en 10°C
 - c. Volúmen final de la reacción: 50 μl
3. Añadir 15 μl de ST a cada tubo. Poner en el termociclador el programa TAG60:
 - a. 65°C 5 min
 - b. Sostener en 10°C
 - c. Volúmen final de la reacción: 65 μl

Lavado del DNA Tagmentado

Nota: Dar vortex a las perlas hasta que estén homogéneas. Si hay algún problema como burbujas o líquido en las paredes dar un spin.

1. Transferir todo el contenido del tubo de PCR a un pocillo de una placa midi o a un tubo de 1.5 ml.
2. Añadir 52 μl de SPB a cada pocillo, mezclar muy bien por pipeteo.
3. Incubar a temperatura ambiente 5 min.
4. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
5. Transferir 98 μl de **sobrenadante** a un nuevo pocillo o tubo. (Aquí se transfiere el sobrenadante porque contiene fragmentos más pequeños).
6. Añadir 137 μl de SPB a cada pocillo o tubo y mezclar muy bien.
7. Incubar a temperatura ambiente 5 min.
8. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)

9. Descartar todo el **sobrenadante** de cada pocillo.
10. Lavar dos veces como se indica:
 - a. Añadir 200 µl de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante.
11. Usando una pipeta de 20 µl remover todo el EtOH residual.
12. Dejar secar por 5 minutos en la placa magnética a temperatura ambiente.
13. Remover de la placa magnética
14. Añadir 22.5 µl de RSB a cada tubo, mezclar muy bien hasta que el pellet se resuspenda.
15. Transferir mínimo 20 µl de sobrenadante a un nuevo tubo de PCR.
16. Opcional: correr 1 µl de DNA sin diluir en el chip Bioanalyzer de Agilent.
NOTA: mínimo 33 ng/µl es lo esperado.

Amplificación del DNA tagmentado

1. Añadir 5 µl del índice i7 a cada tubo. Cambiar la tapa por una nueva.
2. Añadir 5 µl del índice i5 a cada tubo. Cambiar la tapa por una nueva.
3. Añadir 20 µl de LAM a cada tubo, mezclar muy bien. Dar un spin.
4. Poner en el termociclador con el programa LAM AMP:
 - a. 72°C 3 min
 - b. 98°C 30 seg
 - c. 10 ciclos de:
 - i. 98°C 10 seg
 - ii. 60°C 30 seg
 - iii. 72°C 30 seg
 - d. 72°C 5 min
 - e. Hold a 10°C
 - f. Volúmen final de la reacción: 50 µl.

PUNTO DE PARO

Se pueden guardar las muestras hasta 2 días almacenándolas a 4°C.

Lavado del DNA amplificado

1. Dar un spin al tubo de PCR.
2. Transferir 50 µl de DNA amplificado del tubo de PCR a un nuevo pocillo.
3. Añadir 90 µl de SPB a cada pocillo, mezclar muy bien.
4. Incubar 5 min a temperatura ambiente.
5. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
6. Descartar todo el sobrenadante.
7. Lavar dos veces como se indica:
 - a. Añadir 200 µl de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante.
8. Usando una pipeta de 20 µl remover todo el EtOH residual.

9. Dejar secar por 5 minutos en la placa magnética.
10. Remover de la placa magnética
11. Añadir 17 μ l de RSB a cada tubo, mezclar muy bien hasta que el pellet se resuspenda.
12. Transferir mínimo 15 μ l de sobrenadante a un nuevo tubo de 0.6 ml.
13. Cuantificar la biblioteca usando el Qubit.
14. Opcional: correr 1 μ l de DNA diluido 1:10 en el chip Bioanalyzer de Agilent. Se esperan fragmentos de 200-500 pb.

PUNTO DE PARO

Se pueden guardar las muestras hasta 14 días almacenándolas a -20°C.

Hibridación de las sondas

Si el volumen total de la muestra recuperado en el paso anterior es menor a 30 μ l, completar a este volumen con RSB.

1. Añadir los siguientes reactivos en el orden que se indica a un nuevo pozo de una placa midi.
 - a. Biblioteca de DNA 30 μ l (completado con RSB)
 - b. BLR 10 μ l
 - c. CEX 10 μ l
2. Mezclar muy bien por pipeteo
3. Dar un spin
4. Añadir 125 μ l de SPB a cada pocillo y después mezclar por pipeteo.
5. Incubar a temperatura ambiente por 10 min
6. Dar un spin
7. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
8. Descartar todo el sobrenadante.
9. Lavar dos veces como se indica:
 - a. Añadir 200 μ l de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante.
10. Usando una pipeta de 20 μ l remover todo el EtOH residual.
11. Dejar secar por 10 minutos en la placa magnética.
12. Remover de la placa magnética
13. Añadir 7.7 μ l de EHB1 a cada tubo, mezclar muy bien hasta que el pellet se resuspenda. (Checar que el EHB1 no tenga cristales, de ser así, dar vortex).
14. Incubar a temperatura ambiente por 2 min.
15. Dar un spin.
16. Colocar en la placa magnética hasta que el líquido este claro (2 min)
17. Transferir mínimo 7.5 μ l de sobrenadante a un tubo nuevo de PCR.
18. Añadir 2.5 μ l de EHB2 a cada tubo.
19. Dar un spin
20. Colocar en el termociclador con el programa TRE HYB:
 - a. 95°C 10 min

- b. 58°C 30 min
- c. Volúmen final de la reacción: 10 µl.

Captura de las sondas hibridadas

- Prender el Hybex a 50°C
- Los pellets de las perlas con streptavidina son más difíciles de resuspender, asegurarse de su resuspensión.
- Evitar la generación de burbujas.

Primera unión

1. Dar un spin
2. Transferir el contenido del tubo a un pocillo de una placa midi.
3. Añadir 250 de SMB mezclar muy bien (min 1 min)
4. Incubar a temperatura ambiente 25 min
5. Dar un spin
6. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
7. Descartar el sobrenadante.
8. Remover de la placa magnética

Primer lavado

1. Añadir 200 µl de EEW a cada pocillo y mezclar muy bien hasta que el pellet se resuspenda.
2. Tapar con adhesivo toda la placa. Incubar 30 min a 50°C en el equipo Hybex.
3. Inmediatamente colocar en la placa magnética hasta que el líquido este claro (2 min)
4. Descartar el sobrenadante
5. Remover de la placa magnética
6. Repetir los pasos del 1 al 5 para un total de 2 lavados.

Primera elución

1. Crear un premix en un tubo de 1.5 ml y darle vortex
 - a. EE1 28 µl
 - b. HP3 1.5 µl
2. Añadir 23 µl del premix a cada pocillo y mezclar muy bien por pipeteo.
3. Incubar a temperatura ambiente por 2 minutos.
4. Dar un spin
5. Colocar en la placa magnética hasta que el líquido este claro (2 min)
6. Transferir 21 µl de sobrenadante a un nuevo pocillo.
7. Añadir 4 µl de ET2 y mezclar muy bien por pipeteo.
8. Añadir 5 µl de RSB y mezclar muy bien por pipeteo.
9. Dar un spin

PUNTO DE PARO

Se pueden guardar las muestras hasta 7 días almacenándolas a -20°C.

Realizar la segunda hibridación

1. Añadir los siguientes reactivos en el orden que se indica a un nuevo pozo de una placa midi.
 - a. BLR 10 μ l
 - b. CEX 10 μ l
2. Mezclar muy bien por pipeteo
3. Dar un spin
4. Añadir 125 μ l de SPB y mezclar por pipeteo.
5. Incubar a temperatura ambiente por 10 min
6. Dar un spin
7. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
8. Descartar todo el sobrenadante.
9. Lavar dos veces como se indica:
 - a. Añadir 200 μ l de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante.
10. Usando una pipeta de 20 μ l remover todo el EtOH residual.
11. Dejar secar por 10 minutos en la placa magnética.
12. Remover de la placa magnética
13. Añadir 7.7 μ l de EHB1 a cada tubo, mezclar muy bien hasta que el pellet se resuspenda.
14. Incubar a temperatura ambiente por 2 min.
15. Dar un spin.
16. Colocar en la placa magnética hasta que el líquido este claro (2 min)
17. Transferir mínimo 7.5 μ l de sobrenadante a un tubo nuevo de PCR.
18. Añadir 2.5 μ l de EHB2 a cada tubo.
19. Dar un spin
20. Colocar en el termociclador con el programa TRE HYB:
 - a. 95°C 10 min
 - b. 58°C 30 min
 - c. Volúmen final de la reacción: 10 μ l.

Segunda captura de las sondas hibridadas

- Prender el Hybex a 50°C
- Los pellets de las perlas con streptavidina son más difíciles de resuspender, asegurarse de su resuspensión.
- Evitar la generación de burbujas.

Segunda unión

1. Dar un spin

2. Transferir el contenido del tubo a un pocillo de una placa midi.
3. Añadir 250 de SMB y mezclar muy bien (min 1 min)
4. Incubar a temperatura ambiente 25 min
5. Dar un spin
6. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
7. Descartar el sobrenadante de cada poco.
8. Remover de la placa magnética

Segundo lavado

1. Añadir 200 μ l de EEW a cada pocillo y mezclar muy bien hasta que el pellet se resuspenda.
2. Tapar con adhesivo toda la placa. Incubar 30 min a 50°C en el equipo Hybex.
3. Inmediatamente colocar en la placa magnética hasta que el líquido este claro (2 min)
4. Descartar el sobrenadante
5. Remover de la placa magnética
6. Repetir los pasos del 1 al 5 para un total de 2 lavados.

Segunda elución

1. Crear un premix en un tubo de 1.5 ml y darle vortex
 - a. EE1 28 μ l
 - b. HP3 1.5 μ l
2. Añadir 23 μ l del premix a cada pocillo y mezclar muy bien por pipeteo.
3. Incubar a temperatura ambiente por 2 minutos.
4. Dar un spin
5. Colocar en la placa magnética hasta que el líquido este claro (2 min)
6. Transferir 21 μ l de sobrenadante a un nuevo pocillo.
7. Añadir 4 μ l de ET2 a cada pocillo y mezclar muy bien por pipeteo.
8. Dar un spin

Lavado de la biblioteca

1. Vortex a las SPB hasta que estén homogéneas
2. Añadir 45 μ l de SPB a cada tubo o pocillo, mezclar muy bien por pipeteo.
3. Incubar a temperatura ambiente por 5 min.
4. Dar un spin
5. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
6. Descartar todo el sobrenadante de cada pocillo.
7. Lavar dos veces como se indica:
 - a. Añadir 200 μ l de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante
8. Usando una pipeta de 20 μ l remover todo el EtOH residual.
9. Dejar secar por 5 minutos en la placa magnética.

10. Remover de la placa magnética
11. Añadir 27.5 µl de RSB a cada pocillo y mezclar muy bien por pipeteo
12. Incubar a temperatura ambiente 2 minutos
13. Dar un spin
14. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
15. Transferir mínimo 25 µl de sobrenadante a un tubo de PCR.

PUNTO DE PARO

Se pueden guardar las muestras hasta 7 días almacenándolas a -20°C.

Amplificar la biblioteca enriquecida

1. Añadir 5 µl de PPC a cada tubo
2. Añadir 20 µl de EAM a cada tubo, mezclar por pipeteo
3. Dar un spin
4. Colocar en el termociclador con el programa AMP10
 - a. 98°C 30 seg
 - b. 10 ciclos de
 - i. 98°C 10 seg
 - ii. 60°C 30 seg
 - iii. 72°C 30 seg
 - c. 72°C 5 min
 - d. Hold a 10°C
 - e. Volúmen final de reacción: 50 µl

PUNTO DE PARO

Se pueden guardar las muestras hasta 2 días almacenándolas a 4°C.

Lavado de la biblioteca enriquecida amplificada

1. Dar un spin
2. Dar vortex a las SPB hasta que estén homogéneas
3. Transferir los 50 µl de la PCR a un pocillo de una placa midi
4. Añadir 50 µl de SPB al pocillo
5. Incubar a temperatura ambiente por 5 min
6. Dar un spin
7. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
8. Descartar el sobrenadante
9. Lavar dos veces como se indica
 - a. Añadir 200 µl de EtOH 80% a cada pocillo
 - b. Incubar 30 seg
 - c. Descartar el sobrenadante
10. Usando una pipeta de 20 µl remover todo el etanol residual
11. Dejar secar en la placa magnética por 5 minutos
12. Remover de la placa magnética
13. Añadir 20 µL de RSB a cada pocillo y mezclar muy bien por pipeteo

14. Incubar a temperatura ambiente por 2 minutos.
15. Dar un spin
16. Colocar en la placa magnética hasta que el líquido este claro (2-5 min)
17. Transferir mínimo 20 µl de sobrenadante a 0.6 ml. En esta muestra se contiene la biblioteca enriquecida

PUNTO DE PARO

Se pueden guardar las muestras hasta 7 días almacenándolas a -20°C.

Cuantificación y dilución de las bibliotecas

Cuantificar las bibliotecas de DNA para asegurar una densidad de clusters óptima en la celda de flujo.

1. Cuantificar la biblioteca enriquecida con el equipo Qubit
2. Correr un gel de agarosa al 4% para determinar el tamaño promedio de la biblioteca, cargar 6 µl de muestra si es po
3. Usar la siguiente fórmula para convertir ng/µl a nM

$$\frac{\text{Concentracion en ng/}\mu\text{l}}{\left(660 \frac{\text{g}}{\text{mol}} * \text{tamaño promedio de libreria}\right)} \times 10^6 = \text{concentracion en nM}$$

*Los 660 g/mol son equivalentes al PM de 1 par de bases.

Hacer los cálculos con la formula $C_1V_1=C_2V_2$ para hacer una dilución con RSB de la biblioteca a 4 nM contemplado tomar mínimo 2 µl de la biblioteca.

$$\frac{\text{Concentración en nM} * 2 \mu\text{l de biblioteca}}{4 \text{ nM}} = \text{Vol. final de la dilución a 4 nM.}$$

Dilución 4 nM=

- Vol. De Muestra = 2 µl
- Vol. de RSB= Vol. Final de la dilución a 4 nM - 2 µl.

Al finalizar esta dilución se tendrá la biblioteca a 4 nM, y estará lista para comenzar con las instrucciones del manual del PhiX.

Anexo 2: Secuenciación.

Nota: no mezclar por pipeteo, depositar el DNA o el reactivo e ir subiendo la pipeta, posteriormente mezclar por vortex.

Preparar una dilución fresca de NaOH

1. Combinar los siguientes volúmenes en un tubo de 1.5 ml.
 - a. 800 μ l Agua
 - b. 200 μ l de NaOH 1.0 N
2. Mezclar por vortex.

Desnaturalización y dilución de las bibliotecas

Desnaturalizar la biblioteca 4 nM

1. Combinar los siguientes volúmenes en un tubo de 1.5 ml
 - a. Biblioteca 4 nM 5 μ l
 - b. 0.2 NaOH 5 μ l
2. Vortex y spin
3. Incubar por 5 minutos a temperatura ambiente
4. Añadir 990 μ l de HT1 descongelado. Dar vortex por 30 seg. Y poner en hielo.

Esto resulta en una biblioteca desnaturalizada a **20 pM** (1 ml).

Diluir la biblioteca 20 pM desnaturalizada a 12.7 pM en paralelo con el PhiX.

1. Añadir:
 - a. 381 μ l de muestra 20 pM desnaturalizado
 - b. 219 μ l de HT1
 - c. Dar vortex por 30 seg.
2. Esto resulta en 600 μ l de muestra **12.7 pM**, poner en hielo.

Desnaturalización y dilución de las secuencias PhiX (control)

Diluir el PhiX a 4 nM

1. Combinar los siguientes volúmenes en un tubo de 1.5 ml
 - a. 2 μ l 10 nM PhiX
 - b. 3 μ l 10 mM Tris-HCl pH 8.5 con 0.1 tween 20, o agua o RSB.

Desnaturalizar el PhiX

1. Combinar los siguientes volúmenes en un tubo de 1.5 ml
 - a. 5 μ l de 4 PhiX 4 nM
 - b. 5 μ l de NaOH 0.2 N

2. Dar un vortex
3. Dar un spin
4. Incubar a temperatura ambiente por 5 min

Diluir el PhiX desnaturalizado a 20 pM

1. Añadir:
 - a. 10 μ l de PhiX desnaturalizado
 - b. 990 μ l de HT1
2. Dar vortex 20-30 seg. Poner en hielo. Esto resulta en PhiX desanturalizado a **20 pM**.

Diluir el PhiX 20 pM desnaturalizado a 12.7 pM en paralelo con la muestra

Nota: diluir el PhiX y la muestra a la misma concentración. Los dos a 12.7 pM.

1. Añadir:
 - a. 381 μ l de PhiX 20 pM desnaturalizado
 - b. 219 μ l de HT1
2. Esto resulta en 600 μ l de PhiX **12.7 pM**. Dar vortex.

Combinar la biblioteca de la muestra con el PhiX y secuenciar

1. Retirar 6 μ l de la biblioteca de la muestra 12.7 pM
2. Añadir a la biblioteca de la muestra 12.7 pM, 6 μ l del PhiX 12.7 pM para tener un volumen final de 600 μ l. Mantener en hielo
3. Agitar el cartucho varias veces para mezclar los reactivos
4. Cargar los 600 μ l de biblioteca en el pozo indicado del cartucho
5. Limpiar la celda de flujo con agua y EtOH 80% (el vidrio)
6. Generar la sample sheet y cargarla en el equipo

Anexo 3: Preparación de gel de agarosa

Para la preparación de los geles de agarosa al 1%:

1. Pesar 0.35 g de agarosa Ultra pura (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 3.5 μ l de SYBR Safe (Invitrogene) y se agitar brevemente hasta que se diluya.
4. Vaciar la agarosa en la cámara de electroforesis.

Para la preparación de los geles al 4% se seguir el mismo proceso pero pesar 2 g de agarosa y diluirlos en 50 ml de TBE 1X, usar 6 μ l de SYBR Safe.

Buffer TBE 10X:

1. 108 g de Tris base.
2. 55 g de ácido bórico
3. 40 ml de EDTA 0.5 M pH 8.0

Para preparar el buffer TBE 1X hacer una dilución 1:10 del buffer TBE 10X con agua miliQ.