



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**ANÁLISIS MUTACIONAL Y DE INTERACCIONES
GÉNICAS EN UNA PACIENTE CON AUSENCIA
CONGÉNITA DE ÚTERO**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGA

P R E S E N T A:

ARISBET DÍAZ BALLESTEROS



**DIRECTOR DE TESIS:
DR. EN C. FERNANDO FERNÁNDEZ RAMÍREZ**

2021

Ciudad de Mexico



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

RESUMEN	3
ABREVIATURAS	4
INTRODUCCIÓN	5
DESARROLLO DEL APARATO REPRODUCTOR FEMENINO	5
MALFORMACIONES UTERINAS	9
EL SÍNDROME MAYER-ROKITANSKY-KÜSTER-HAUSER	10
DESARROLLO Y ETIOLOGÍA DEL SÍNDROME MRKH	12
DIAGNÓSTICO Y MANEJO	12
GENES CANDIDATOS EN HUMANOS	14
ANTECEDENTES	15
JUSTIFICACIÓN	18
OBJETIVO	19
OBJETIVOS ESPECÍFICOS	19
HIPÓTESIS	19
MÉTODOS	19
RESULTADOS	27
ANÁLISIS DE CNV	27
ANÁLISIS DE WES	29
PRIORIZACIÓN DE VARIANTES	30
ANÁLISIS DEL IMPACTO ESTRUCTURAL EN PROTEÍNA	34
REDES DE INTERACCIÓN PROTEÍNA-PROTEÍNA.	35
ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL DE REDES PPI	38
DISCUSIÓN	41
CONCLUSIONES	52
LITERATURA CITADA	53
ANEXOS	65

Resumen

El síndrome de Mayer-Rokitansky-Küster-Hauser (MRKH, ORPHA:3109) es un trastorno que se presenta en 1 de cada 4500 mujeres, caracterizado por la ausencia congénita de útero, cérvix y el tercio superior de la vagina, con un cariotipo usualmente sin alteración (46, XX) y demás rasgos sexuales secundarios normales. La etiología del síndrome se desconoce, sin embargo, la existencia de casos familiares supone un componente genético, y los hallazgos a la fecha sugieren un mecanismo complejo que podría incluir un componente poligénico, interacciones génicas y/o epistasis. Se investigó el caso de una paciente con MRKH típico, realizando un análisis mutacional para identificar a escala genómica variación en el número de copias (CNV) y variantes de nucleótido único (SNV) potencialmente patogénicas, mediante microarreglos cromosómicos y secuenciación de exoma completo. Mediante priorización basada en frecuencias poblacionales, predictores bioinformáticos sobre impacto estructural y funcional, conservación filogenética y otras bases de datos, se identificó una variante sin descripción previa, afectando al gen *WNT9B*; por lo que se realizó un análisis adicional sobre su impacto en la estructura y función de la proteína, mediante modelado por homología. Los resultados corroboraron que la mutación podría generar un impacto funcional negativo, sin embargo, la calidad del modelo *in-silico* obtenido es limitada. Los genes afectados por variantes genéticas se priorizaron en categorías de impacto; aquellos en la categoría de alto-moderado impacto, por si solos no sugieren un mecanismo que explique el desarrollo del síndrome. Por esta razón, se realizó un análisis basado en redes de Interacción Proteína-Proteína (PPI), integrando el conjunto de genes afectados en la paciente, así como genes interactores adicionales a partir de la base de datos STRING. Los resultados mostraron la formación de módulos funcionales relacionados con el procesamiento de RNAs, y con procesos de organización durante el desarrollo. Además, se generó una red PPI a partir de genes candidatos de MRKH descritos en literatura, y se integraron los genes de la paciente, encontrando asociación a procesos de formación de estructuras reproductivas. En conclusión, las alteraciones génicas y las interacciones observadas en la paciente pueden intervenir en procesos cruciales durante el desarrollo embrionario, tales como la alteración en la vía de Notch y en la formación del ducto Müllleriano, y podrían estar implicadas en el desarrollo del MRKH. Estos resultados muestran una nueva visión de las interacciones complejas que se

dan en el síndrome MRKH, sin embargo, es necesario validar que estos mismos procesos se observen en más pacientes con dicha enfermedad.

Abreviaturas

ACMG: American College of Medical genetics

AMP: Asociación de Patología Molecular

CGH: Hibridación genómica comparativa

CNV: Variación del número de copias

dbSNP: Base de datos de Polimorfismos de nucleótido único

ESGE: Sociedad Europea de Endoscopia Ginecológica

ESHRE: Sociedad Europea de Embriología y Reproducción Humana

FDR: False Discovery Rate

LOF: Pérdida de función

M-CAP: Puntuación mendeliana de patogenicidad clínicamente aplicable

MD: Conducto de Müller

MLPA: Amplificación de sondas tras ligación múltiple

MRKH: Mayer-Rokitansky-Küster-Hauser

NGS: Secuenciación de nueva generación

pHI: Índice predictivo de haploinsuficiencia

pLI: Índice predictivo de tolerancia a la pérdida de función

PPI: Interacción proteína-proteína

QMQE: Estimación de la calidad del modelo global

QUAL: Calidad del genotipo asignado

SHh: Sonic Hedgehog

SNV: Variantes de nucleótido único

TFR: Tracto reproductor femenino

WD: Conducto de Wolf

WES: Secuenciación de exoma completo

Introducción

Desarrollo del aparato reproductor femenino

En las primeras semanas de gestación ocurren los procesos de determinación del sexo y la diferenciación sexual, los cuales son procesos diferentes; la determinación del sexo consiste en el desarrollo de la gónada sexual (ovarios o testículos) a partir de una gónada bipotencial y en función del complemento cromosómico establecido durante la fertilización; mientras que la diferenciación sexual se refiere al desarrollo embrionario de los órganos que constituyen el aparato reproductor, incluyendo los genitales externos e internos (Arboleda et al., 2015), así como la diferenciación somática y cerebral subsecuente. Cerca de la semana 5-6 de gestación los embriones 46,XX y 46,XY no presentan diferencias ya que ambos portan una estructura denominada cresta urogenital, la cual está conformada de células con el potencial de diferenciarse tanto al sexo masculino como el femenino (Wilhelm et al., 2007). Sin importar el sexo al que el embrión se diferenciará en un futuro durante la etapa de diferenciación sexual los embriones poseen el conducto mesonéfrico y el paramesonéfrico, llamados así ya que estos se extienden a lo largo de los riñones embrionarios o mesonefros (Arboleda et al., 2015; Roly et al., 2018); éstos se conocen también como conducto de Wolf (WD) y conducto de Müller (MD), respectivamente, ya que en 1830 el fisiólogo Johannes Peter Müller describió los conductos paramesonéfricos, indicando además que al desarrollarse daban origen a las trompas de Falopio, útero y la parte superior de la vagina. (Roly et al., 2018).

Dependiendo de las señales recibidas, el primordio gonadal dará origen a un testículo o un ovario dentro de las semanas 8-11 de gestación dependiendo de la presencia del cromosoma Y (Arboleda et al., 2015). Una vez que el sexo gonadal se ha completado, se producen las hormonas necesarias para la siguiente etapa: al formarse un testículo, éste produce andrógenos y Hormona Anti-mülleriana, lo cual induce la regresión de los MD; por el contrario, la ausencia de estas hormonas conduce a la regresión de los WD, y la diferenciación del tracto reproductor femenino (TRF) (Roly et al., 2018). El desarrollo del TRF consiste en una serie de procesos complejos y ordenados que ocurren durante el

desarrollo embrionario y que involucran la diferenciación, migración y fusión de los conductos müllerianos (Cunha et al., 2018).

A la fecha, la mayor parte del conocimiento que se tiene sobre el desarrollo del TRF se ha obtenido mediante el estudio de modelos animales, principalmente ratón (*Mus musculus*), en el cual se ha descrito principalmente a nivel citológico y tisular, mientras que los procesos moleculares subyacentes se conocen en menor medida. La diferenciación gonadal ocurre entre los días embrionarios E10 y E11, y durante el final de esta etapa comienza el desarrollo de los conductos pronéfricos, que darán lugar a los WD y MD (Wilhelm et al., 2007).

Los WD son los primeros en formarse y extenderse a lo largo del mesonefros en dirección caudal. Para el correcto desarrollo del MD se requiere el WD como guía de elongación; esto se ha demostrado en estudios donde se utilizaron ratones en los que se inactivó *Lhx1* específicamente en células del WD, lo que propició una degeneración del mismo y en consecuencia la ausencia del desarrollo del conducto mülleriano (Huang et al., 2014); confirmando que la pérdida de este gen da como resultado un desarrollo incompleto en el conducto de Müller (Orvis & Behringer, 2007).

El proceso de formación de los MD se puede dividir en tres fases (Fig. 1): el inicio o especificación, la invaginación y la extensión (Atsuta & Takahashi, 2016; Roly et al., 2018). Durante la fase de inicio las células de la placa lateral del mesodermo se especifican en el epitelio celómico del polo craneal del mesonefro; estas células se encuentran marcadas por un engrosamiento y además expresan el gen *Lim1* (Atsuta & Takahashi, 2016). Posteriormente, estas células inician la expresión del gen *Wnt4*, lo que induce la fase de invaginación (Orvis & Behringer, 2007), convirtiéndose en células mesoepiteliares, se invaginan y proliferan en dirección caudal entre el mesénquima del epitelio celómico y el conducto de WD (Fig. 1B); al entrar en contacto con el WD, las células proliferan generando el epitelio Mülleriano, que se elonga en dirección craneocaudal. El conducto de Wolff no solo funciona físicamente como una guía, sino también se ha demostrado que también exhibe una señal necesaria para la fase de migración caudal o elongación; esta señal es *Wnt9b*, la cual es expresada por el epitelio del WD. Cuando el conducto Mülleriano se fusiona con el seno urogenital, concluye la fase de elongación (Roly et al., 2018). Para el día embrionario

~ E13.5 ambos conductos, de Wolff y Müller, han terminado su elongación y comienza su diferenciación o regresión, misma que es competentemente visible al E15.5 (Roly et al., 2018).

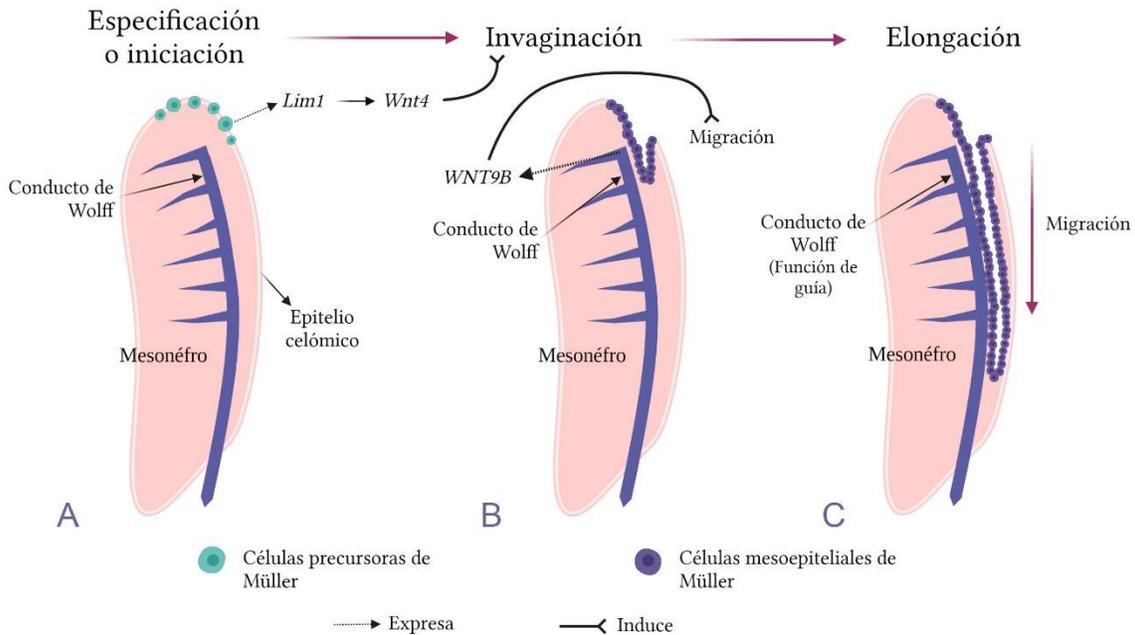


Fig. 1. Esquema de las fases del desarrollo del conducto de Müller. Se muestran las estructuras y los genes involucrados. Creado con BioRender.com

En humanos, posterior a la fase de elongación y cerca de la octava semana de gestación, inicia la diferenciación de la estructural de los MD, donde se unen para formar el canal uterovaginal, el cual más adelante formará el útero y el tercio superior de la vagina (Horii et al., 2018). Al fusionarse los MD el tejido de sus paredes genera un septo en la línea media, el cual comienza a ser reabsorbido en la semana 9 de gestación en la zona caudal y prosigue en dirección craneal, proceso que al terminar conforma la cavidad uterina (Fig. 2A) (Cunha et al., 2018). La zona más caudal donde los MD ya fusionados entran en contacto con el seno urogenital permanece sólida y se fusiona posteriormente con el creciente epitelio del bulbo sinovaginal (Fig. 2B), el cual en la semana 20 se canaliza para formar el resto de la vagina (Cunha et al., 2018; Horii et al., 2018).

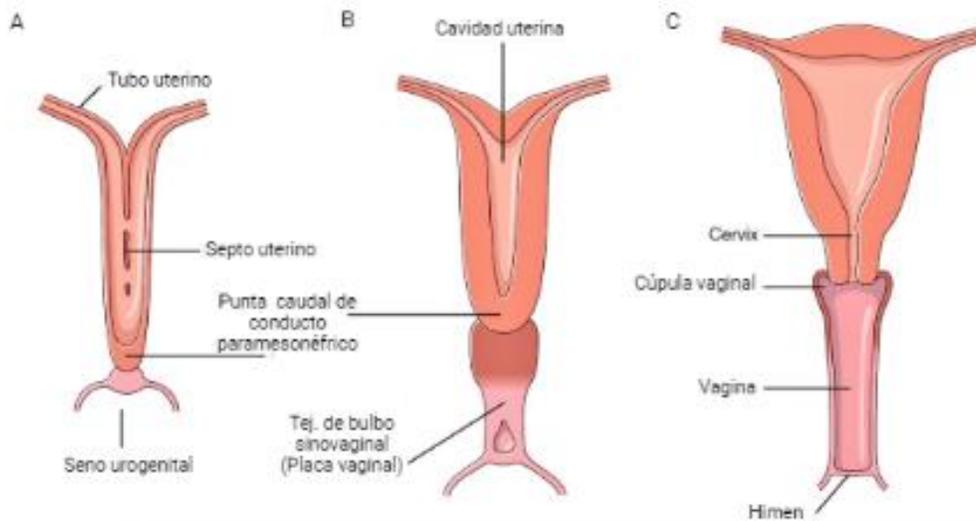


Fig. 2 Regiones anatómicas del útero durante el desarrollo. Tomado y modificado de (Sadler & Langman, 2012). A. Estructura a las 9 semanas, presencia de septo uterino. B. Estructura al final del tercer mes, se observa el tejido de los bulbos sinovaginales. C. Estructura de recién nacida, la cupula vaginal y la porción superior de la vagina se forman por vacuolización del tejido paramesonefrico, y la porción inferior de la vagina se forma por vacuolización de los bulbos sinovaginales.

La familia de genes HOX es importante para el desarrollo de muchas estructuras anatómicas; para la diferenciación y desarrollo adecuado de los MD se requiere un subgrupo de genes llamado *Hoxa* (*Hoxa9*, *Hoxa10*, *Hoxa11* y *Hoxa13*). Los genes *Hoxa* se expresan de forma segmentada y promueven la diferenciación segmentada del MD expresándose en trompas de Falopio, útero, cuello uterino y vagina de forma específica (Roly et al., 2018).

Como se mencionó previamente, el gen *WNT9B* participa en la elongación de los MD. En ratones mutantes *Wnt9b*^{-/-} se expresa un fenotipo similar a MRKH, donde las gónadas parecen normales y hay ausencia de útero y parte superior de la vagina. Además, *Wnt9b* se encarga de regular río arriba la expresión de *Wnt4*, *Fgf8* y *Pax8* (Carroll et al., 2005). Los genes *Wnt4*, *Wnt5a* y *Wnt7a* también están involucrados con la diferenciación de los MD durante el desarrollo. Se ha demostrado que en ratones *Wnt7a*^{-/-} se presentan malformaciones en el tracto reproductivo femenino como oviductos cortos y cuernos uterinos hipoplásicos, además que son infértiles, y presentan un desarrollo postnatal alterado (Dunlap et al., 2011).

Malformaciones uterinas

En cualquiera de los procesos que ocurren durante el desarrollo del TRF, pueden acontecer diversas alteraciones que dan lugar a malformaciones congénitas del aparato reproductor.

Se reconocen tres grandes grupos de alteraciones: defectos de formación, de fusión o de absorción del septo (Fig. 3) (Jacquinet et al., 2016). Estos defectos pueden dar origen a malformaciones que van desde dimorfismo uterino, hasta displasia e incluso aplasia o agenesia (ausencia de la estructura). Sin embargo, cada uno de estos tipos de malformaciones ligadas a la cavidad uterina, puede afectar en menor o mayor grado la capacidad reproductiva de las mujeres que las padecen (Grimbizis et al., 2001).

El diagnóstico de las malformaciones congénitas uterinas es frecuentemente complicado, debido a que algunas de éstas son consideradas variaciones anatómicas normales o asintomáticas y, por lo tanto, son subdiagnosticadas. Sin embargo, se ha determinado que cerca del 5.5% de la población femenina las presenta (Chan et al., 2011). Debido a la falta de consenso en la clasificación de estas malformaciones, la Sociedad Europea de Embriología y Reproducción Humana (ESHRE) y la Sociedad Europea de Endoscopia Ginecológica (ESGE), generaron una clasificación de las malformaciones del TFR según su componente anatómico y su origen embrionario. La clasificación incluye las distintas malformaciones con las siguientes claves: U para las malformaciones del útero, C para las de cérvix y V para las de vagina; éstas se acompañan de un número para indicar su gravedad: 0 indicando un órgano normal, y se incrementa el número dependiendo de la gravedad de la malformación. En el caso del cérvix el máximo de gravedad tiene una clave C4, la cual indica aplasia o displasia, hasta V4 en caso de la vagina donde se indica aplasia y en el caso de útero hasta U6, que indica malformaciones no clasificadas, pasando por U5 que indica aplasia o displasia (Grimbizis et al., 2013).

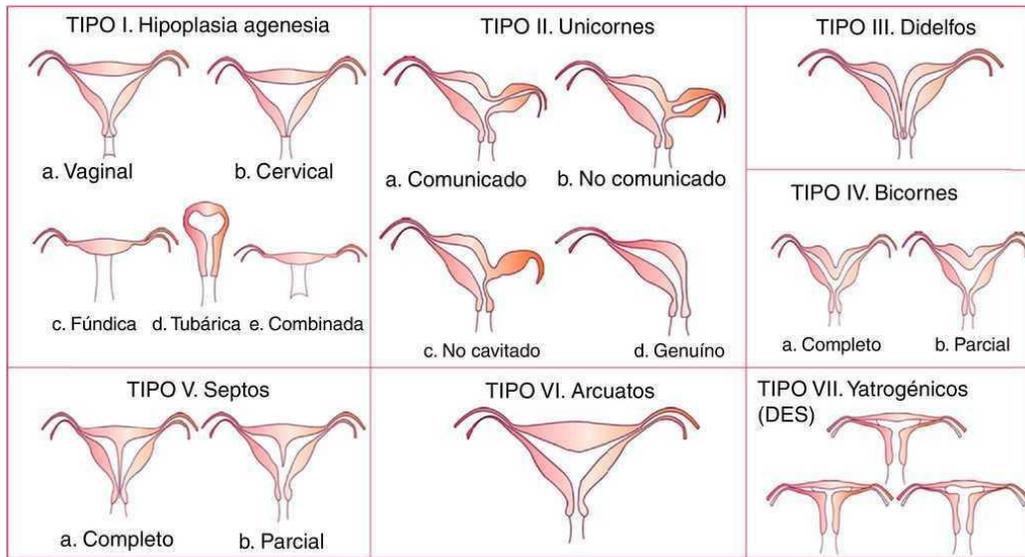


Fig. 3. Tipos de malformaciones de útero. Tomado de (Sociedad Española de Ginecología y Obstetricia, 2014)

Algunas de estas malformaciones fueron descritas inicialmente por el anatomista y fisiólogo Karl Mayer, posterior a la revisión de autopsias, denominó una condición llamada “útero bipartito” que consistía en un tipo de malformación del útero y la vagina obstruida, además, también determinó otros defectos asociados con el sistema renal y cardiaco (Patnaik et al., 2015). Posteriormente, el patólogo Carl Rokitansky estudió a través de autopsias la agenesia mülleriana debido a defectos en el desarrollo, lo cual podía causar formas incompletas y ausencias completas de órganos, como el útero y la vagina (Patnaik et al., 2015). Fue hasta 1910 que Küster, un ginecólogo alemán, describió el caso en una paciente viva, y lo relacionó con deformidades esqueléticas presentes. Por último, el ginecólogo Hauser compiló la información anterior y completó la descripción del síndrome de agenesia mülleriana, añadiendo la diferencia del cariotipo normal (46,XX) a diferencia del síndrome de feminización testicular donde se observa un cariotipo 46,XY (Patnaik et al., 2015). Este desarrollo histórico muestra la descripción de una entidad clínica conocida como el síndrome Mayer-Rokitansky-Küster-Hauser.

El síndrome Mayer-Rokitansky-Küster-Hauser

El síndrome Mayer-Rokitansky-Küster-Hauser (MRKH; OMIM%277000) está considerado como una de las malformaciones congénitas más severas, con clasificación U5C4V4 por la

ESHRE/ESGE. Es la forma más común de agenesia vaginal y se caracteriza por la ausencia congénita de útero, cérvix y el tercio superior de la vagina; en estas mujeres, la función ovárica y el desarrollo de caracteres sexuales secundarios son normales, y usualmente el cariotipo es normal (46,XX) (Susanne Ledig & Wieacker, 2018) (Fig. 4).

El síndrome MRKH puede ser clasificado en dos tipos: el tipo I o típico, el cual presenta las afectaciones urogenitales características, mostrando la aplasia completa del útero en presencia de dos cuernos rudimentarios simétricos unidos por un pliegue peritoneal, y no se relaciona con signos de hiperandrogenismo; el tipo II, forma atípica, asociada o incompleta, se caracteriza por aplasia o hipoplasia uterina, y sobre todo presenta además otras malformaciones, tales como malformaciones en trompas de Falopio, anomalías ováricas, malformaciones renales (agenesia unilateral de riñón, ectopia de uno o ambos riñones, riñones fusionados), esqueléticas (principalmente en cervicales) y en los casos menos frecuentes malformaciones cardíacas, digitales y/o algún grado de discapacidad auditiva.

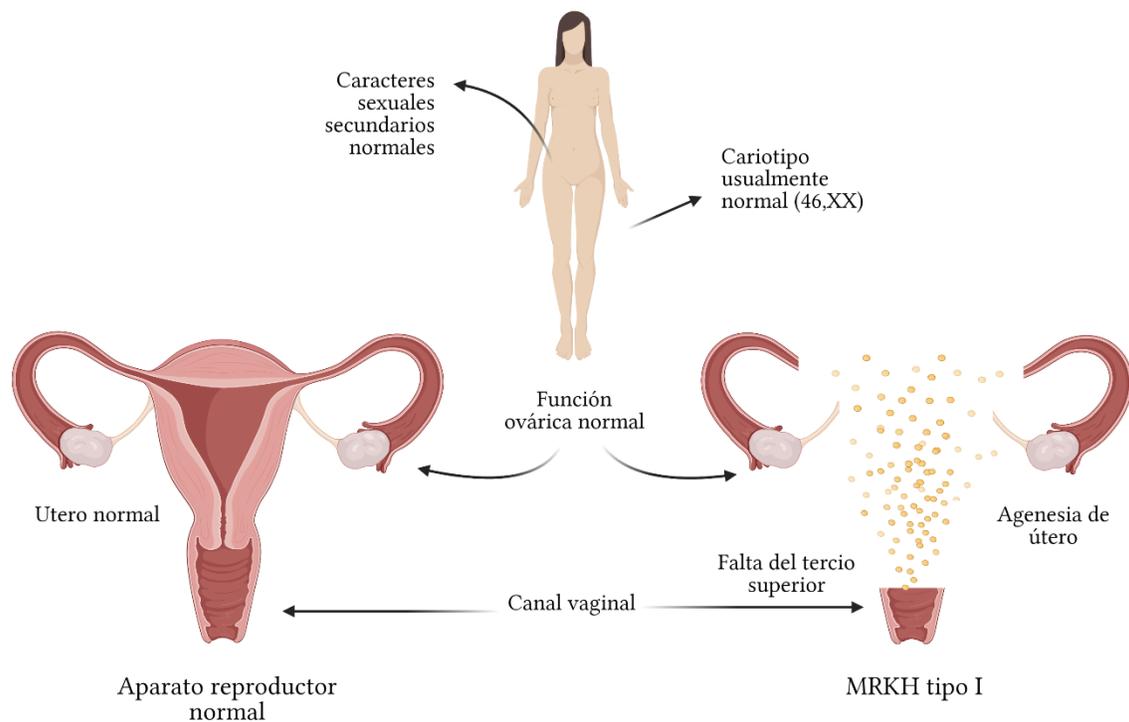


Fig. 4. Esquema del aparato reproductor femenino normal y con MRKH. Creado con BioRender.com

Existen subtipos de este síndrome, por ejemplo, la asociación MURCS (aplasia del conducto Mülleriano, displasia renal, anomalías de los somitas cervicales) presente en el 16% de las pacientes con MRKH tipo II y se puede considerar como un subtipo específico de este (Fontana et al., 2017), de tal manera que, desde el punto de vista clínico, algunos autores lo consideran como una entidad particular, probablemente con etiología relacionada.

La prevalencia de este síndrome es aproximadamente de 1 en 4500 mujeres; algunos reportes indican una mayor prevalencia del tipo I con un 56.5% (M. Herlin et al., 2016), mientras otros presentan mayor prevalencia en el tipo II con un 56% (Strübbe et al., 1994). Sin embargo, la mayoría de los estudios son realizados en cohortes europeas, por lo que la prevalencia en otras poblaciones es desconocida.

Desarrollo y etiología del síndrome MRKH

Se considera que el desarrollo del síndrome MRKH ocurre durante la embriogénesis, y puede atribuirse a una afección inicial que ocurre en el mesodermo intermedio, donde se genera una alteración del blastema de los somitas cervicotorácicos y de los conductos pronéfricos; éstos, posteriormente, inducen la diferenciación de los conductos mesonéfricos y paramesonéfricos, que darán lugar a los conductos de Wolff y Müller, respectivamente (Duncan et al., 1979). La etiología del síndrome MRKH es desconocida; la mayoría de los casos son esporádicos, sin embargo, existen reportes de casos familiares, lo que supone un componente genético. También se han observado discordancias entre gemelos monocigotos, por lo que se clasifica el síndrome MRKH como una enfermedad compuesta, donde pueden interaccionar factores epigenéticos, ambientales y genéticos. La mayoría de los linajes analizados sugiere una herencia autosómica dominante con penetrancia incompleta (M. Herlin et al., 2014).

Diagnóstico y manejo

Debido a que las pacientes presentan genitales externos normales con caracteres sexuales secundarios desarrollados y una función endócrina normal, el síndrome MRKH se

diagnostica generalmente hasta la adolescencia; el signo por el cual las pacientes acuden a revisión es el de amenorrea primaria. El síndrome MRKH constituye así la segunda causa más común de amenorrea primaria (16% de las pacientes), precedido solamente por la disgenesia gonadal (Herlin et al., 2020). Debido a la similitud de las características del síndrome MRKH con algunos otros síndromes como el Síndrome de Insensibilidad a Andrógenos o Atresia vaginal aislada, se debe realizar un diagnóstico diferencial por un ginecólogo pediatra por medio de examen pélvico, una examinación radiológica basada principalmente en técnicas no invasivas que permitan obtener imágenes que revelen la ausencia de estructuras (ecografía transabdominal o resonancia magnética), análisis bioquímicos de niveles hormonales, e incluso un análisis cromosómico (Herlin et al., 2020; Morcel et al., 2007).

Una vez diagnosticadas las pacientes pueden presentar diversos efectos psicológicos como: depresión, baja en la autoestima, dudas de identidad femenina y ansiedad (Song et al., 2020). Además, dichos efectos psicológicos pueden aumentar una vez que se asimilan las consecuencias del síndrome MRKH, ya que si continúan sin intervención médica estas mujeres tendrán dificultades para tener relaciones sexuales por vía vaginal, no menstruarán y con ello no podrán intentar un embarazo, lo cual genera un gran impacto en la mente de una mujer adolescente, no solo en la esfera psicológica sino en general disminuyendo su calidad de vida (Bean et al., 2009). Es por lo que se ha planteado y recomendado un abordaje multidisciplinario, para el manejo y tratamiento de estas pacientes.

El manejo que se da a las pacientes con MRKH incluye tratamientos quirúrgicos y no quirúrgicos. Entre los no quirúrgicos se incluyen el método de Frank y el de Ingran, los cuales consisten en ir adecuando la vagina con dilatadores, los cuales van aumentando de calibre de forma progresiva; las ventajas que presentan este tipo de métodos es que no es necesaria la hospitalización, tienen una buena relación costo/eficiencia, presentan complicaciones mínimas y, si no es efectivo, todavía se puede acceder a los métodos quirúrgicos. Sin embargo, también presentan algunas desventajas, como el tiempo que hay que invertir en ello (de 20 minutos a 2 horas diarias por cerca de 6 meses), o que no permiten recuperar la fertilidad (Londra et al., 2015). Por otra parte, los métodos quirúrgicos incluyen las vaginoplastias, las cuales pueden ser con autoinjertos, es decir, se toma una parte de un tejido autólogo y se injerta en la zona a reconstruir; estos pueden ser de tejido de disección de recto

y vejiga, intestino, peritoneo y de los labios mayores (Herlin et al., 2020). Dichas vaginoplastías también se pueden realizar con ingeniería de tejidos, es decir, se construye una neovagina a partir del material obtenido de una biopsia autóloga vulvar, utilizando biomateriales como soporte para el cultivo (Raya-Rivera et al., 2014). Otra posibilidad quirúrgica que continua en proceso de investigación es el trasplante de útero, el cual permitiría a las pacientes resolver el problema de infertilidad, sin embargo, debido a las complicaciones operatorias y postoperatorias todavía no se incluye este método como el manejo principal del síndrome (Georgopapadakis et al., 2019). Ya que el trasplante de útero aún no es una opción accesible para todas las mujeres, muchas optan por la reproducción asistida o la adopción como medida para poder ejercer la maternidad (Bean et al., 2009). Cabe señalar que, considerando un probable componente genético en el síndrome MRKH, la reproducción asistida puede representar un riesgo para la descendencia de estas pacientes. Sin embargo, aún con estas opciones, hay un gran impacto en la calidad de vida de estas mujeres, sobre todo en países donde socioculturalmente la maternidad es considerada como un aspecto “necesario”, por lo cual además de los tratamientos directos para tratar las malformaciones también se debe incluir el apoyo psicosexual y grupos de apoyo para mujeres con este síndrome (Bean et al., 2009).

Genes candidatos en humanos

Tras al menos dos décadas de estudios moleculares, se han identificado mutaciones en algunos genes, sin embargo, éstas se han observado con muy baja prevalencia, por lo que el síndrome se mantiene sin una etiología clara. Debido a la escasez de estudios de árboles genealógicos completos, ha preponderado el enfoque del estudio de genes candidatos, los cuales se han propuesto por su participación en vías de señalización compartidas por ciertos órganos afectados durante el desarrollo embrionario, o por genes que son conocidos por producir efectos semejantes a los asociados con el MRKH en modelos animales (Fontana et al., 2017). También se encuentran aquellos que regulan el desarrollo, entre ellas las familias de WNT y HOX. La familia de WNT es un grupo de genes que juega un papel importante durante el desarrollo del sistema urogenital de los mamíferos. Dentro de la familia de WNT, se propuso al gen *WNT4*, debido a su función como efector de determinación del

sexo y porque constituye la señal para la invaginación de las células epiteliales celómicas. Sin embargo, las mutaciones en *WNT4* se han identificado en casos de aplasia uterina asociada con hiperandrogenismo; aun así, algunos autores han considerado la mutación de *WNT4* como un elemento de diagnóstico genético del síndrome MRKH. (Backhouse et al., 2019; Layman, 2014). Las mutaciones que se consideran como diagnósticas son p.E226G (Biaison-Lauber et al., 2004), p.A233T (Philibert et al., 2011), sin embargo, con el paso del tiempo se han descrito más mutaciones de este gen ligadas a MRKH. Además de *WNT4*, se han determinado solamente otros dos genes como diagnósticos del MRKH, el gen *HNF1B* y el gen *LHX1*, ambos localizados en el cromosoma 17q12, y con mutaciones encontradas en casos de MRKH, principalmente en forma de delección (Layman, 2014); particularmente, las mutaciones de *HNF1B* se han identificado en pacientes con anomalías müllerianas, quistes renales y diabetes juvenil. Por lo que hasta el momento, el candidato más sólido para los casos de aplasia mülleriana aislada es *LHX1*, sin embargo la frecuencia en la que se observan mutaciones en este gen es baja (<6% para delección, y aún menor para mutaciones puntuales) (Layman, 2014).

Antecedentes

Del año 2010 a la fecha ha aumentado la disponibilidad de tecnologías de alto rendimiento tales como la secuenciación masiva en paralelo o secuenciación de nueva generación (NGS, por sus siglas en inglés) y los microarreglos genómicos. La disminución en el costo de estas metodologías también ha permitido extender su aplicación para buscar las variantes genéticas implicadas en síndromes poligénicos y enfermedades complejas. En el caso de MRKH, y malformaciones müllerianas en general, esto permite un abordaje desde una perspectiva más integral, permitiendo una mayor tasa de identificación de variantes y/o genes candidatos.

Dado que la mayoría de las mujeres afectadas presentan un cariotipo 46,XX aparentemente normal, entre las hipótesis propuestas para explicar la etiología del síndrome MRKH se han considerado las variaciones del número de copias (CNV) submicroscópicas, ubicadas en diferentes y múltiples regiones cromosómicas, tales como 4q34-qter, 8p23.1, 10p14 (Morcel

et al., 2011), 7q31.2, 15q26.3 (McGowan et al., 2015), entre otras. Aunque los estudios de hibridación genómica comparativa (CGH) han proporcionado mucha información acerca de estas microdeleciones o microduplicaciones, sólo se han podido observar en cerca de 35% de las pacientes con el síndrome MRKH (McGowan et al., 2015), por lo que no se consideran determinantes de la enfermedad, y tampoco se han logrado integrar en un mecanismo etiológico que explique por completo esta condición.

Sin embargo, las microdeleciones o microduplicaciones más frecuentes ocurren en las regiones 1q21.1, 17q12 y 22q11.21 (Cheroki et al., 2008; Susanne Ledig et al., 2011; McGowan et al., 2015). Además, estos últimos *loci* abarcan genes previamente identificados en los estudios en modelos murinos, y relacionados con fenotipos similares a los del MRKH. Por ejemplo, se ha mencionado que el *locus* 17q12 contiene el gen *LHX1* (antes nombrado como su homólogo murino *LIM-1*), el cual se ha encontrado como parte de una deleción en pacientes con síndrome MRKH (Ledig et al., 2012) y como se había mencionado anteriormente es necesario para la elongación del MD (Huang et al., 2014). Dentro del *locus* 17q12 también se codifica el gen *HNF1B*, el cual es un factor de transcripción que las mutaciones en este gen dan lugar quistes de riñón, malformaciones del riñón y del sistema genitourinario (Ferrè & Igarashi, 2019).

Durante la última década, se han acumulado estudios que abordan al menos dos distintos mecanismos que pudieran respaldar la identificación de genes candidatos (ya sea por pérdida o ganancia de función): uno orientado a dosis génica y otro a SNVs. Por ejemplo, un miembro de la familia *HOX*, el gen *SHOX*, se encarga de controlar aspectos del desarrollo y del crecimiento; la mutación de éste se ve implicada en distintas enfermedades de defectos en el desarrollo, que van desde efectos graves como el síndrome de Turner, hasta efectos leves como estatura baja idiopática (Binder, 2011). En un estudio Gervasini y colaboradores (2010) por medio de Amplificación de sondas tras ligación múltiple (MLPA) encontraron el gen *SHOX* duplicado de forma parcial en cerca del 16% de la cohorte de MRKH (n=30), dicha duplicación apareció tanto en casos familiares como esporádicos de pacientes con MRKH tipo I; en el caso familiar se observó que la duplicación provenía del padre y que estaba ausente en 2 hermanas sanas. Por otra parte, se ha observado que los ratones *knockout* para

los genes *Hox* presentan anomalías del aparato urogenital, sin embargo, no se han identificado mutaciones claramente patogénicas en estos genes en pacientes con MRKH (Fontana et al., 2017), salvo variantes de significado incierto afectando los genes *HOXA10* y *HOXA11* en al menos 10 individuos con malformaciones uterinas esporádicas, así como mutaciones patogénicas en *HOXA13* asociadas al síndrome Mano-Pie-Genital, el cual tiene sobrelapamiento fenotípico con MRKH (Jacquinet et al., 2016).

Otro estudio mostró por secuenciación completa de exoma que las variaciones del gen *GREBIL* son parte de la etiología del MRKH. Sin embargo, se desconoce el mecanismo en el que éste se relaciona con el desarrollo del sistema genitourinario y también concluyeron que estas variantes no son por completo la causa de este síndrome y que se requiere de otros factores ambientales o genéticos para generar la enfermedad (Jacquinet et al., 2020). Gran cantidad de los genes candidatos se han obtenido a partir de la secuenciación completa del exoma en pacientes con MRKH o por su relación en modelos de desarrollo animales, esta lista incluye: *Amh*, *Amhr2*, *Ctnnb1*, *Dach1*, *Dach2*, *Emx2*, *Hoxa11*, *Hoxa13*, *Lamc1*, *Pbx1*, *Rara*, *Rarb*, *Rarc*, *Wt1*, *DOCK4*, *HNRNPCL1*, *ITIH5*, *LRP10*, *MMRP14*, *OR2T2*, *OR4M2*, *PAX8*, *PDE11A*, *RBM8A*, *TBX6*, y *ZNF816* (Backhouse et al., 2019). De manera general, los hallazgos de SNV no han permitido integrar a la fecha un circuito genético que permita comprender la naturaleza de los procesos de desarrollo que conducen a las malformaciones uterinas, en parte por el conocimiento parcial que se tiene del desarrollo normal del TRF, y también porque los hallazgos genéticos han sido heterogéneos y con muy baja recurrencia general entre las distintas pacientes estudiadas.

Debido a lo anterior, también se han realizado otro tipo de abordajes para entender la etiología del MRKH, orientados en aspectos genéticos funcionales. En uno de tales estudios, se compararon microarreglos de genoma completo de tejido vaginal cultivado *in vitro* de pacientes con MRKH; se encontró un patrón diferente de expresión en múltiples genes en comparación de mujeres clínicamente sanas, en total se encontraron genes regulados a la alza y 133 a la baja, sin embargo, los genes *MUC1*, *HOXB2*, *HOXB5*, *HOXC8*, *JAG1* y *DLL1* se determinaron como genes candidatos por su expresión diferencial y su papel en el desarrollo embrionario humano (Nodale et al., 2014).

Debido a que los genes candidatos, como una entidad individual, no han proporcionado un esclarecimiento de la etiología del síndrome MRKH, otros estudios proponen observar las interacciones biológicas que se perturban por las distintas mutaciones observadas en las pacientes con este síndrome, con la idea de integrar un modelo etiológico. En este sentido, el enfoque de biología de sistemas y el estudio de redes genéticas permite conjuntar las interacciones físicas y funcionales entre las proteínas que participan en un proceso biológico, de manera tal que se puedan integrar relaciones mecánicas y causales entre agentes aparentemente heterogéneos, y poder vislumbrar eventualmente su relación con el desarrollo del TRF y el MRKH. Recientemente, el estudio realizado por Pontecorvi y colaboradores (2021) mostró un análisis de redes, utilizando las proteínas cuyos genes se encontraron alterados en las pacientes con MRKH. Este análisis destacó algunas conexiones biológicas y permitió determinar nuevas interacciones genéticas que probablemente pueden estar involucradas en el desarrollo de esta condición. Mientras que dicho trabajo se limitó al uso de las CNV y análisis de redes, el presente trabajo pretende abordar un panorama más amplio con el uso de las CNV, mutaciones puntuales e integración de redes de PPI.

Justificación

Las pacientes con el síndrome de MRKH se ven afectadas por la falta de las estructuras anatómicas, la dificultad para llevar una vida sexual normal y la incapacidad de procrear, lo cual conlleva a otras afecciones que recaen en la esfera psicológica y emocional. El mecanismo etiológico que subyace al síndrome MRKH se desconoce, y la identificación de distintos genes candidatos no ha permitido explicar la mayoría de los casos.

El presente trabajo tiene la intención de realizar un análisis integrativo y detallado de las variantes genéticas identificadas en una paciente con síndrome MRKH típico, orientado a comprender los genes e interacciones que pueden ocurrir entre sus productos, con el fin de identificar vías biológicas o redes regulatorias que pudieran relacionarse con los mecanismos patogénicos del MRKH. El conocimiento genético de las causas de este síndrome puede ayudar a las pacientes a recibir un diagnóstico oportuno, a médicos tratantes a elegir las mejores terapias y probablemente en un futuro contribuir a disminuir la recurrencia del padecimiento.

Objetivo

Identificar a escala genómica las variantes genéticas potencialmente patogénicas de una paciente con síndrome de MRKH típico, y realizar un análisis integrativo de los hallazgos, utilizando redes biológicas de PPI y categorías ontológicas establecidas.

Objetivos específicos

- A partir de datos de microarreglos genómicos y de secuenciación de exoma completo, identificar y priorizar variantes y genes con potencial patogénico para MRKH.
- Analizar el efecto de las variantes puntuales con mayor impacto en la estructura y función proteínica.
- Determinar la relación física o funcional entre los productos de los genes candidatos establecidos, y las variantes potencialmente patogénicas identificadas en la paciente, mediante el análisis de redes de PPI.
- Examinar en módulos o categorías biológicas la posible relación funcional de los genes candidatos establecidos y las variantes posiblemente patogénicas identificadas en la paciente

Hipótesis

Dado que la evidencia existente apunta a que el síndrome MRKH es una condición con base genética producida por un mecanismo poligénico, y que las variantes genéticas en los genes implicados pueden producir perturbaciones en sus funciones e interacciones, por lo tanto, un abordaje exploratorio e integrativo, basado en técnicas genómicas y análisis de redes PPI, permitirá reconstruir e identificar vías biológicas que puedan estar relacionadas con el desarrollo de este síndrome.

Métodos

Tipo y diseño del estudio

El presente trabajo corresponde a un estudio observacional descriptivo, exploratorio y transversal, centrado en un caso clínico seleccionado.

Población y muestras

Este estudio se enfoca en el análisis de datos genómicos de un caso clínico seleccionado. El caso clínico corresponde a una muestra de estudio conformada por una serie de mujeres adultos diagnosticadas con MRKH, las cuales fueron incluidas durante el periodo 2014-2017 como parte de un proyecto de investigación previo. A partir de esta población de estudio se constituyó un banco de ácidos nucleicos a partir de sangre periférica para su análisis genético.

Aspectos éticos y de bioseguridad

El presente trabajo se enfocó en el análisis computacional de datos genómicos, por lo que constituye una investigación sin riesgos de bioseguridad. Los datos genómicos fueron obtenidos a partir de metodologías experimentales realizadas previamente, utilizando muestras del banco de ácidos nucleicos conformado durante el desarrollo del proyecto titulado “Identificación de genes implicados en la morfogénesis uterina”, con números de registro vigentes DI/11/310A/5/35 (Hospital General de México), y HIM/2017/065SSA1400 (Hospital Infantil de México), así como DGAPA-UNAM IN219912 (concluido, Facultad de Medicina, UNAM). La colecta de muestras para el banco de DNA se realizó mediante consentimiento informado, utilizando los formatos aprobados por las comisiones de Ética e Investigación de las Instituciones participantes.

Criterios de inclusión, exclusión y eliminación

Para la generación del banco de ácidos nucleicos se consideró como criterio de inclusión de caso, a pacientes adultos diagnosticadas con síndrome MRKH de tipo I y con muestras de DNA disponibles, tanto de la paciente como parentales; se excluyeron las muestras de DNA de pacientes con otro tipo de malformaciones y/o diagnóstico (displasias u otras asociaciones clínicas), y se eliminaron aquellos casos cuya muestra de DNA no presentó integridad y concentración adecuada para los procedimientos experimentales, y/o que no se acompañara de muestras de DNA parental disponibles.

Estrategia general de estudio

Se tomó el DNA de una paciente con MRKH seleccionada del banco de ácidos nucleicos, a dichos datos genómicos se les realizaron dos metodologías exploratorias de alto rendimiento: microarreglo genómico (para lo cual también se obtuvo el DNA de la madre) y secuenciación de exoma completo (WES, por sus siglas en inglés). A partir de los datos generados, se inició el análisis computacional planteado en el presente trabajo: se realizó un preprocesamiento (análisis estadístico y control de calidad), anotación biológica de las variantes identificadas, y posteriormente se realizó el análisis de priorización de variantes genéticas. Una vez obtenidos los conjuntos de datos (*datasets*) priorizados, se procedió de forma paralela a un análisis individual sobre el impacto estructural de las variantes puntuales con mayor potencial patogénico, así como a un análisis integrativo usando los distintos *datasets*, basado en redes de interacciones proteína-proteína y categorías ontológicas, con el fin de reconstruir e identificar vías biológicas cuya perturbación pudiera incidir en el desarrollo del síndrome MRKH.

Metodologías de alto rendimiento

Se realizaron microarreglos genómicos con el fin de identificar CNV, es decir, duplicaciones y deleciones submicroscópicas. El modelo de microarreglo utilizado fue Cytoscan HD (Affymetrix Inc.), que contiene 2.7 millones de sondas con cobertura genómica, y una distancia promedio de 1100 pb entre sondas. A partir de 250ng de gDNA de la paciente y de la madre, se efectuó el procedimiento de acuerdo con el protocolo provisto por el fabricante, utilizando la plataforma Affymetrix 3000 7G (Thermo-Fisher Scientific) de la Unidad de Genética del Hospital General de México. Los datos crudos (archivo. CEL) fueron preprocesados en Chromosome Analysis Suite v4.2 bajo parámetros de alta resolución: segmento mínimo detectable: 100 kb y 50 sondas contiguas; 5000 sondas para mosaicismo, y 5000 pb para regiones de alta homocigosidad. Se utilizó un set de referencia provisto en el programa, constituido por 380 controles (individuos clínicamente sanos) y el mapeo se realizó según el ensamble [GRCh37] del genoma humano, usando las anotaciones NetAffx Build 20200828 (hg19).

La secuenciación de exoma completo se realizó con la finalidad de determinar la presencia de variantes puntuales (SNV o SNP), así como variantes génicas (*indels*). La metodología fue realizada a partir de una muestra de 50 ng de gDNA de la paciente seleccionada, la cual

fue enviada a un servicio de secuenciación en paralelo (Macrogen, Korea del Sur). Se utilizó la plataforma NovaSeq 6000 (Illumina Inc.), y el kit de enriquecimiento SureSelect v6 (Agilent). Los datos crudos (archivo .FASTq) fueron pre-procesados y alineados usando el ensamble [GRCh37] del genoma humano; las variantes identificadas (archivo .VCF) fueron extensivamente anotadas a través de un servicio externo, y desplegadas para su posterior análisis usando la herramienta Bplatform (Bitgenia, Argentina).

Análisis computacional de datos de microarreglos

El análisis se realizó en Chromosome Analysis Suite v4.2. Se anotaron y priorizaron las microdeleciones y microduplicaciones halladas en el genoma de la paciente, de acuerdo con la frecuencia poblacional (Database of Genomic Variants (MacDonald et al., 2014)) y características de su contenido génico, tales como predicción de sensibilidad a dosis (base de datos Decipher), perfil de expresión (base de datos GTEx; (Lonsdale et al., 2013), así como anotaciones clínicas (base de datos OMIM). Se priorizaron las deleciones y duplicaciones de baja frecuencia poblacional (<1%), cuyo contenido génico resultara relevante, según los criterios mencionados.

Análisis computacional de datos de WES

Las anotaciones generadas para las variantes genéticas identificadas, incluyeron los siguientes datos: profundidad, GQ y QUAL (calidad del genotipo asignado), posición (genómica, en transcrito y en proteína); cambio de nucleótido; tipo e impacto del cambio esperado en la codificación de la proteína; interpretación del impacto clínico de la variante según la American College of Medical genetics (ACMG) y la Asociación de Patología Molecular (AMP) (Richards et al., 2015), frecuencia poblacional en bases de datos genómicos (gnomAD, TGP), registro en la base de datos dbSNP, anotaciones clínicas (ClinVar), conservación filogenética del residuo, así como valores obtenidos en distintos predictores de impacto funcional, incluyendo Polyphen-2, CADD, M-Cap, Mutation Taster, SIFT, entre otros. Utilizando esta información, se realizó un análisis bioinformático orientado a priorizar las variantes en categorías de patogenicidad, mediante un algoritmo de filtrado de datos (Fig. 5). De manera paralela se realizaron dos abordajes de análisis: uno centrado en

genes asociados con fenotipos urogenitales (Tablas 1 y 2), y otro a manera de análisis abierto, incluyendo todas las variantes genéticas halladas en todos los genes estructurales.

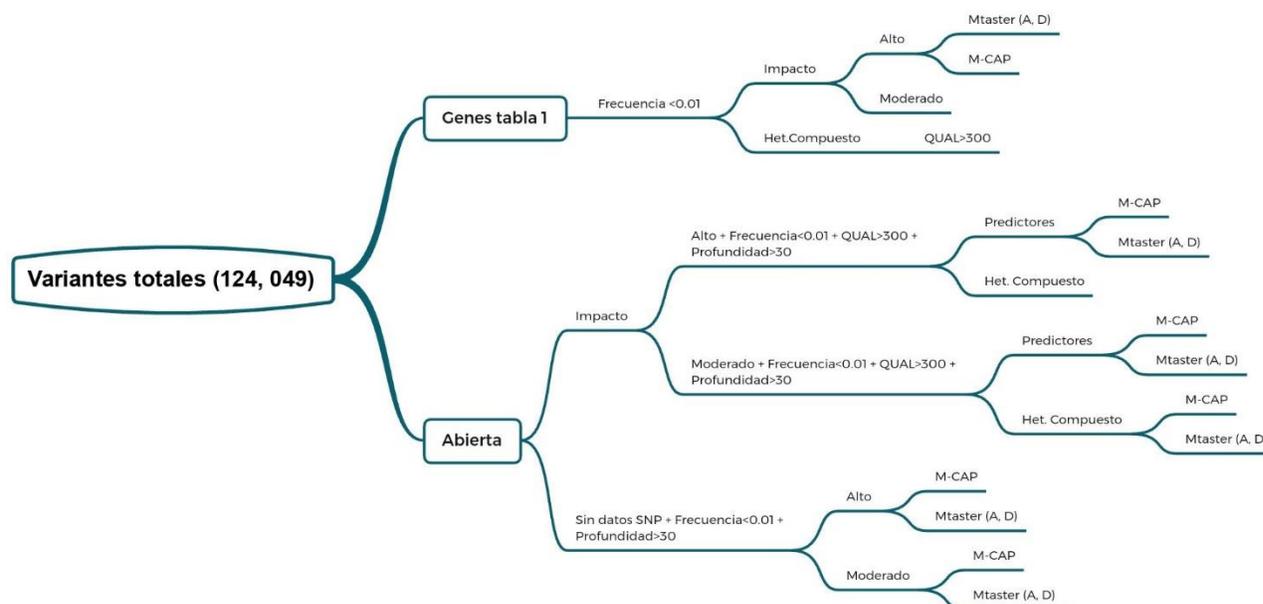


Fig. 5. Algoritmo de priorización de variantes genéticas identificadas mediante WES. *Het.Compuesto*: Modelo de enfermedad autosómico recesivo hetero compuesto.

Para el primer abordaje se integraron dos listas de genes relacionados con MRKH, a partir de literatura y bases de datos: una de ellas de genes directamente asociados con MRKH (Anexo, Tabla 1A) y otra de genes asociados con patologías y síndromes que presentan solapamiento fenotípico con las características clínicas de MRKH (Anexo, Tabla 1B). Para la integración de las listas de genes candidatos se incluyeron los genes propuestos en las revisiones de Backhouse y colaboradores (2019), Herlin (2020) y Fontana (2017).

Los parámetros de cada criterio de priorización fueron como se describe a continuación:

- a. Profundidad mínima de lectura >30 y calidad del genotipo $QUAL>300$.
- b. Frecuencia poblacional de las variantes, utilizando tres bases de datos: gnomAD Exome, gnomAD Genome y 1000G. Se filtraron variantes de alta frecuencia poblacional, quedando aquellas consideradas de baja frecuencia (<0.01).
- c. Priorización por consecuencia de la variante en la codificación de la proteína, donde se consideraron como impacto alto:

- Ganancia de codón de paro o sin sentido (*nonsense*): Variante de secuencia por la que se cambia al menos una base de un codón, lo que da como resultado un codón de paro prematuro, lo que lleva a una proteína truncada.
- Cambio del marco de lectura (*frameshift*): Variante de inserción o deleción que provoca el corrimiento del marco de lectura, porque el número de nucleótidos insertados o eliminados no es múltiplo de tres.
- Variante en el sitio aceptor de maduración (*splicing*): Una variante de que cambia la región de 2 bases en el extremo 3' de un intrón o base aceptor.
- Variante en el sitio donador de splicing: Una variante que cambia la región de 2 bases en el extremo 5' principal de un intrón.
- Variante de interacción estructural: Una variante que impacta en las interacciones internas de la estructura polipeptídica resultante.

De impacto moderado se consideraron:

- Variantes con sentido equivocado (*missense*): Variante que da como resultado una secuencia de nucleótidos diferente, la cual puede cambiar el aminoácido resultante sin alterar el marco de lectura.
- d. Variantes no registradas en la base de datos de Polimorfismos de nucleótido único (dbSNP).
- e. Valoración de impacto en la estructura y función proteínica mediante predictores y metapredictores computacionales.
- Mutation Taster: Meta-predicor que permite evaluar el potencial causante de enfermedad de variantes de secuencia, integrando un valor predictivo con base a la conservación evolutiva, cambios en sitios de “splicing”, pérdidas de función de proteínas y cambios que pueden afectar la abundancia del mRNA (Dong et al., 2015); clasifica las variantes como polimorfismos, patogénicas o patogénicas-automáticas, y asigna un valor de probabilidad a clasificación efectuada.
 - Puntuación mendeliana de patogenicidad clínicamente aplicable (M-CAP): Es un meta-predicor de variantes (sustituciones) que remodela un algoritmo de clasificación basado en valores de predictores simples como SIFT, Polyphen-2 y

CADD, integrando un valor predictivo entre 0 y 1; se utilizó un umbral de >0.025 para separar las variantes patogénicas (Jagadeesh et al., 2016), considerando que aquellas con menor valor se clasifican como benignas con una confianza de 95%.

- SIFT: Se basa en la homología de la secuencia para predecir si una variante de aminoácido alterará la función de la proteína, y con ello su fenotipo (Ng & Henikoff, 2003) Se utilizó el criterio de >0.95 para determinar un efecto deletéreo.
- CADD v1.4: Predicción funcional de un SNP. Esta herramienta compara los efectos de alelos fijos o casi fijos de 14.7 millones de variantes de alta frecuencia en humanos contra variantes simuladas, con lo cual genera un índice de nocividad, es decir, la correlación entre la funcionalidad y la patogenicidad que tiene dicha variante (Kircher et al., 2014). Se consideraron las variantes con un score >11 como posiblemente Deletérea, y >14 como deletérea.

f. Evaluación de la conservación filogenética de la base.

- GERP++: Score de conservación que cuantifica los déficits de sustitución en alineamientos múltiples de genes ortólogos de 35 mamíferos (Dong et al., 2015). El umbral de >4.4 se tomó como conservación de la base/posición, siendo el índice máximo de conservación de 6.17.
- PhyloP: Índice de conservación de DNA basado en alineamientos de 146 vertebrados, con valores de -14 para los menos conservados hasta +6 los más conservados. El valor >1.6 se utilizó como umbral para determinar base/posición conservada.
- PhastCons: Índice que representa la probabilidad de selección negativa de un nucleótido, de acuerdo con alineamientos locales y de regiones vecinas (Basu et al., 2013); genera valores en el rango de 0 a 1, donde 1 indica una alta conservación.

Además de las anotaciones previamente descritas, se realizaron tres anotaciones adicionales a las variantes previamente priorizadas, incluyendo el índice predictivo de haploinsuficiencia (pHI) y el índice predictivo de tolerancia a la pérdida de función (pLI), así como el perfil de expresión cualitativo. Los pHI y pLI fueron tomados de base de datos Decipher (Firth et al., 2009), la cual tiene una orientación clínica. Por un lado, el pLI indica la probabilidad de que

un gen sea intolerante a las mutaciones que producen pérdida de función (LOF); mientras que el pHI indica la probabilidad de que un gen sea haploinsuficiente (Firth et al., 2009), es decir, que requiere de las dos copias genómicas para producir el fenotipo silvestre. El perfil de expresión se determinó de forma cualitativa utilizando la base de datos GTEx (Lonsdale et al., 2013), considerando como un patrón positivo cuando la expresión del gen ocurre preponderantemente en tejidos relacionados con MKRH (útero, vagina, endometrio, miometrio, trompas de Falopio, etc.); patrón compatible cuando la expresión génica incluye dichos tejidos y otros relacionados con MRKH tipo II, y patrón negativo cuando no hay expresión génica en los mismos o ésta es residual (<5 transcritos por millón).

Análisis proteico estructural

Una vez priorizadas las variantes según los criterios previamente descritos, para aquellas con mayor rango se realizó un modelaje por homología tanto de la proteína silvestre como la versión mutante, utilizando el servidor Swiss-Model (Waterhouse et al., 2018), con la finalidad de evaluar o corroborar las posibles perturbaciones en el plegamiento o función de la proteína afectada.

Análisis de redes de interacción proteína-proteína (PPI)

Con el fin de realizar un análisis integrativo de las variantes genéticas identificadas, así como explorar un posible mecanismo poligénico para esta condición, se reconstruyeron redes de interacción funcional o física entre proteínas, utilizando la base de datos STRING (Szklarczyk et al., 2019). Los parámetros considerados para la reconstrucción de redes de interacciones fueron: a) Fuentes de datos: experimentos, bases de datos, co-expresión y co-ocurrencia. b) Confianza: Media (0.5). c) Número máximo de interactores: 10, en primer grado o primera capa. Se generaron varios modelos de redes: una exploratoria construida con los genes identificados en la paciente (a partir de las variantes potencialmente patogénicas, priorizadas según el algoritmo descrito previamente, Anexo 2). Posteriormente, se generó otra red que incluye los genes previamente identificados en literatura como candidatos y/o

asociados a MRKH (Anexo, Tabla 1A). Finalmente, se integró una red utilizando los datos de las dos redes previas; para cada caso, se analizaron los resultados obtenidos, incluyendo el número total de módulos obtenidos, número de nodos que constituyen cada módulo, así como número de nodos sin integrar; posteriormente se realizó minería de datos mediante el incremento de interactores, reevaluando los resultados obtenidos.

Análisis enriquecimiento funcional de redes PPI

Utilizando las herramientas provistas en STRING, se realizaron análisis de enriquecimiento funcional en cada una de las redes PPI construidas, basándose en categorías de procesos biológicos de Gene Ontology (Carbon et al., 2021).

Resultados

Análisis de CNV

Los datos de microarreglos permitieron la identificación de las siguientes microduplicaciones: 17q21.31, 14q32.33, 22q11.22 y 22q13.32q13.33, y dos microdeleciones en Xq24 y Xq23 (Tabla 1).

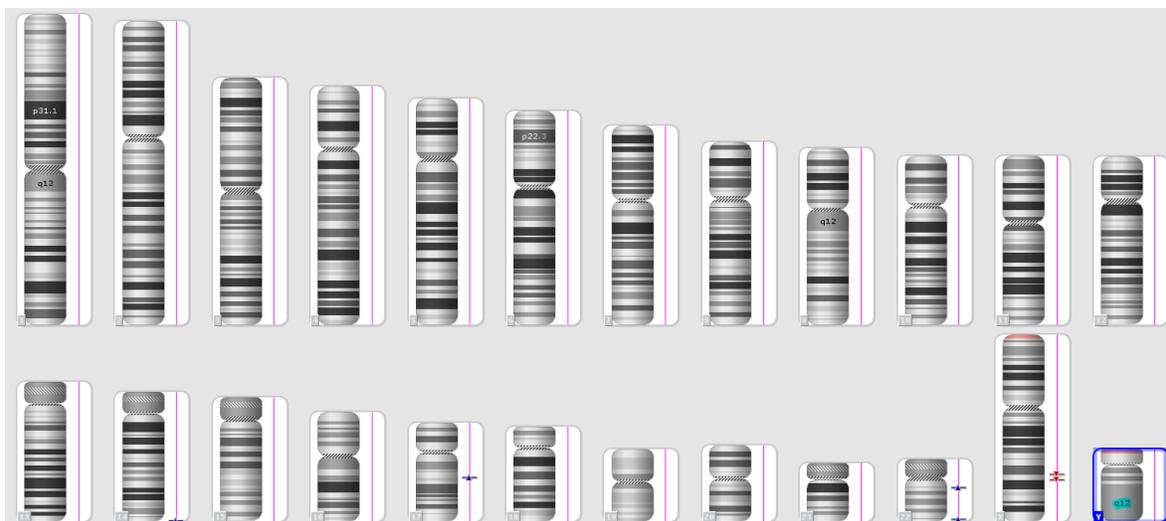


Fig. 6. Cariotipo molecular de la paciente en estudio. Los puntos azules indican ganancia de material genético, en rojo se muestran las pérdidas.

Número de copias	Tipo	Cromosoma	Cyto band	Tamaño (kbp)	Número de genes	Genes	Clasificación clínica	Nomenclatura microarreglo	Anotaciones del segmento
3.00	Ganancia	17	q21.31	572	8	KANSLI, KANSLI-ASI, ARLI7B, LRRC37A, ARL17A, NSFPI, LRRC37A2, NSF	Benigna	arr[GRCh37]17q21.31(44212824_44784639)x3	Origen materno SCNVP (Frec.12.8%)
1.00	Pérdida	X	q24	119	13	CT47B1, CT47A1, CT47A10, CT47A11, CT47A12, CT47A2, CT47A3, CT47A4, CT47A5, CT47A8, CT47A9, CT47A6, CT47A7	Probable Benigna	arr[GRCh37]Xq24(120002999_120122287)x1	SCNVP CN id gssvL137268 (Frec.1.8%)
3.00	Ganancia	14	q32.33	284	6	MIR4507, MIR4538, MIR4537, MIR4539, FAM30A, ADAM6	Significado incierto	arr[GRCh37]14q32.33(106246289_106530460)x3	Origen materno SCNVP CN id gssvG12033 (Frec.0.09%)
3.00	Ganancia	22	q11.22	134	3	MIR650, MIR5571, IGLL5	Significado incierto	arr[GRCh37]22q11.22(23124498_23258369)x3	Origen materno. SCNVP (Frec.0.02%, 3.64%)
3.00	Ganancia	22	q13.32 - q13.33	608	2	C22orf34, MIR3667	Significado incierto	arr[GRCh37]22q13.32q13.33(49396259_5000399)x3	SCNVP CN id nsv4286863 (Frec.<0.00001%)
1.00	Pérdida	X	q23	418	1	LOC100126447	Significado incierto	arr[GRCh37]Xq23(115905893_116323691)x1	SCNVP CN id nsv4039570 /// (Frec.<0.00001%) SCNVP CN id CN id nsv4051944 (Frec.<0.00001%)

Tabla 1. Tabla de segmentos con duplicaciones o deleciones encontrados en la paciente. Resaltado en negrita se muestran los genes codificantes de proteínas, los no resaltados pueden ser genes de RNA o pseudogenes. SCNVP: Sobreapamiento con CNV poblacionales.

A partir de estos datos, se realizó la priorización de CNV mediante el análisis de origen parental, así como del contenido génico. Los resultados indicaron la presencia de 2 deleciones y 4 duplicaciones, tres de las cuales también las presentaba la madre, y dos de ellas presentan una frecuencia poblacional alta. Las CNV de significancia incierta contienen genes que mayormente participan en el procesamiento de distintos tipos de RNA, no se encuentran directamente relacionados a padecimientos ni con la expresión específica en regiones anatómicas relacionadas con el MRKH (Lonsdale et al., 2013). Aunque estos datos no muestran una relación directa, se incluyeron los genes de los CNV de significancia incierta

y que no fueran de origen materno a la red de IPP, para no descartar su posible participación en la generación de este síndrome, por lo que se incluyeron 6 genes a la red (*MIR650*, *MIR5571*, *IGLL5*, *C22orf34*, *MIR3667*, *LOC100126447*).

Análisis de WES

Como parte de la secuenciación de exoma completo se obtuvieron los estadísticos de los datos crudos, lo cual indica que se leyeron un total de 6,851,127,002 pares de bases, a partir de un total de 45,371,702 lecturas. Se determinó un contenido GC de 51.55% y un contenido AT de 48.45%. El indicador de calidad Q20 se alcanzó en el 98.12% de las lecturas, mientras el indicador Q30 se alcanzó en el 94.72% de ellas, por lo que se determinó que los datos tenían la calidad necesaria para su análisis. El indicador Q indica una relación logarítmica de la probabilidad de que la asignación de una base sea incorrecta, por lo que el indicador Q10 indica que 1 en 10 lecturas es probablemente incorrecta, Q20 1 en 100, Q30 1 en 1000 (Ewing & Green, 1998);

El alineamiento y llamado de variantes arrojó los resultados mostrados en la tabla 1, donde se muestra que, de 124,049 variantes el 99.697% estaban ya reportadas anteriormente (Fig. 7). La profundidad de lectura en promedio fue de 47 lecturas, lo que reafirmó la calidad de los datos (Fig. 7).

Kit	SureSelect All Exon V6
Referencia	GRCh37
#Variantes	124049
#Variantes conocidas	123673
dbSNP ratio	1
Het/Hom ratio	1.57
TI/TV ratio	2.33
Missense/Silent	-
Profundidad prom. variantes	47
#Genes	21900
#SNP	110247
#INS	6316
#DEL	7165



Fig. 7 Tabla de estadísticos y gráficos de resultados

Priorización de variantes

Partiendo de 124,049 variantes identificadas en la paciente, se aplicó el algoritmo de priorización, obteniendo los resultados mostrados en la figura 8. Se identificaron un total de 101 variantes de impacto moderado a grave (Anexo 2, Lista de Genes). De estas 101 variantes, 17 mostraban mayor impacto estructural en las proteínas, por lo que se analizaron de forma más profunda. De estas 17 variantes (Tabla 2), se encontró que 8 se expresan de forma diferencial en los tejidos relacionados con el MRKH (útero, vagina, endometrio, miometrio, trompas de Falopio, etc). Además, tres de ellos son sensibles a dosis (considerados como haploinsuficientes) y a mutaciones de pérdida de función (*ESRI*, *SALL1*, *JAG1*), uno de ellos intolerante a la pérdida de función (*PDGFC*) y dos más se consideran solo haploinsuficientes (*DYNC2H1*, *RPGRIPL*).

Gen	Variante Pos. Genómica (Cr:Posición)	Variante proteína y mRNA	Efecto	Patrón de expresión (GTEX)	pLI (Decipher)	pHI (Decipher)
<u>WNT9B</u>	G > C 17:44952732	p.Lys200Asn c.600G>C	Var. de sentido equivocado	Negativo	0.54	34.84
<u>FANCC</u>	A > G 9:97912223	p.Val223Ala c.668T>C	Var. de sentido equivocado	si (II)	0	43.39
<u>ESRI</u>	C > A 6:152129484	p.Pro146Gln c.437C>A	Var. de sentido equivocado	si (I)	1*	0.06*
<u>DYNC2H1</u>	G > A 11:103024105	p.Arg1057His c.3170G>A	Var. de sentido equivocado	si (I)	0	0.57*
<u>SALL1</u>	G > A 16:51175096	p.Ala346Val c.1037C>T	Var. de sentido equivocado	sí (II)	1*	5.62*
<u>RPGRIPL</u>	C > G 16:53653005	p.Ala1183Gly 3548C>G	Var. de sentido equivocado	si (II)	0	9.18*
<u>PDGFC</u>	C->G 4:55139727	p.Thr463Ser c.1388C>G	Var. de sentido equivocado	si (I)	0.99*	27.48
<u>TTC7A</u>	C > T 2:47300917	p.Thr835Met c.2504C>T	Var. de sentido equivocado	Negativo	0	45.91
<u>MAP3K9</u>	G > C 14:71216698	p.Leu368Val c.1102C>G	Var. de sentido equivocado	Negativo	0.07	37.92
<u>JAG1</u>	G > A	c.588C>T	Variante de interacción estructural	si (II)	1*	1.06*
<u>PROZ</u>	C > T 13:113826030	p.Gln294* c.880C>T	Var. sin sentido	Negativo	0	89.66
<u>USP54</u>	G > A 10:75264736	p.Arg1395* c.4183C>T	Var. sin sentido	Negativo	0	25.07
<u>RCVRN</u>	C > A 17:9801504	p.Glu171* c.511G>T	Var. sin sentido	Negativo	0	27.95
<u>GPR113</u>	G > A 2:26540944	p.Arg76* c.226C>T	Var. sin sentido	Negativo	0	-
<u>TUFT1</u>	C > T 1:151554166	p.Arg405* c.1213C>T	Var. sin sentido	si (II)	0	48.72
<u>RITAI</u>	C > T 12:113629581	p.Arg281* c.841C>T	Var. sin sentido	Negativo	0	-
<u>MYH1</u>	C > T 17:10417137	c.741+1G>A	Var. donador splicing	Negativo	0	15.1

Tabla 2. Anotación adicional de mutaciones identificadas como de más alto impacto. Patrón de expresión: I. Indica patrón positivo, cuando la expresión del gen ocurre preponderantemente en tejidos relacionados con MKRH (útero,

vagina, endometrio, miometrio, trompas de Falopio, etc.); *II* Indica patrón compatible, cuando la expresión génica incluye dichos tejidos y otros relacionados con MRKH tipo I; *Negativo*: cuando no hay expresión génica en los tejidos relacionada al MRKH o ésta es residual. El color muestra la clasificación del gen: amarillo: gen intolerante a la pérdida de función, azul: gen haploinsuficiente, verde: gene intolerante a pérdida de función y haploinsuficiente. *pLI*: índice predictivo de tolerancia a la pérdida de, *pHI*:índice predictivo de haploinsuficiencia. Los genes subrayados se encuentran descritos como candidatos (Anexo, Tabla 1) *Niveles considerados positivos.

Análisis del impacto estructural en proteína

A partir del análisis de WES, se identificó la variante c.317T>C (rs4968281 en la base de datos dbSNP) que produce alteraciones en la proteína WNT9B, específicamente en la posición 106, generando un cambio de aminoácido Metionina p.(Met106Thr). Además, se identificó otra variante c.600G>C p.(Lys200Asn) la cual no tiene registro en la base de datos dbSNP; ésta es una transversión de G>C, generando un cambio de aminoácido de Lisina a Asparagina en la posición 200 (Fig. 9). Con el fin de mapear esta mutación en la estructura tridimensional de la proteína WNT9B y evaluar su impacto funcional, se realizó modelado por homología mediante la herramienta *online* Swiss-Model (Waterhouse et al., 2018); para esto se utilizó el modelo cristalográfico reportado para la proteína *Wnt3*, con la cual *WNT9B* comparte 34% de similitud en secuencia. Se reportó una Estimación de la calidad del modelo global (QMQE) de 0.59 y una calidad por residuo de 0.63 ± 0.05 , esto para ambos modelos (silvestre y mutante). De ambos se generó un archivo .PDB, el cual permitió explorar la estructura tridimensional de las proteínas y analizar el efecto que se podría tener debido a la zona donde se encuentran las variantes (Figura 9).

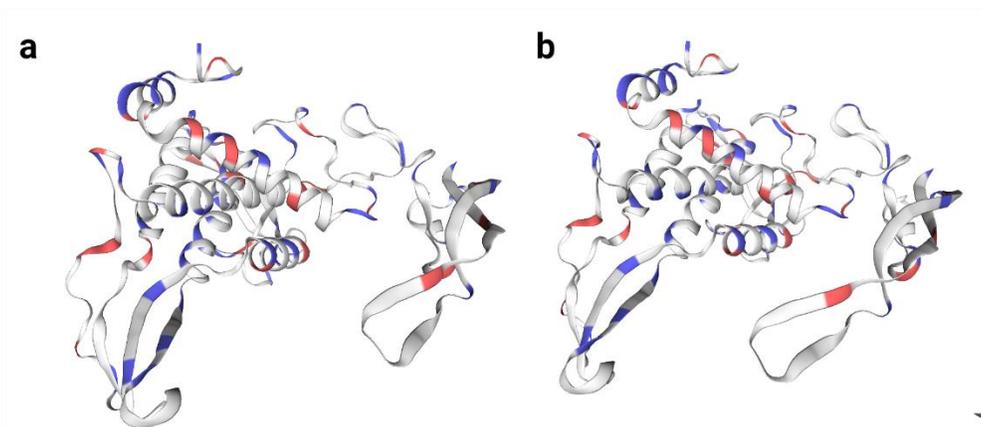


Fig. 9. Modelado por homología de proteína WNT9B. Se utilizó la herramienta Swiss-Model y el modelo cristalográfico de la proteína homóloga *Wnt3*. a. Vista general de proteína silvestre. b. Vista general de proteína WNT9BK200N de paciente MRKH.

Redes de interacción proteína-proteína.

Para la generación de la red de PPI se utilizó la base de datos de STRING 11.0; <http://string-db.org/>) (Szklarczyk et al., 2019). Esta herramienta muestra asociaciones directas (físicas) e indirectas (funcionales) entre proteínas, lo que permite relacionar diversos conjuntos de datos basados en el contexto genómico, coexpresión, experimentos de alto rendimiento y conocimientos previos. Se introdujeron como datos de entrada 98 genes identificados en la paciente con variantes potencialmente patogénicas (Anexo. Tabla 2, Lista de genes), 6 genes

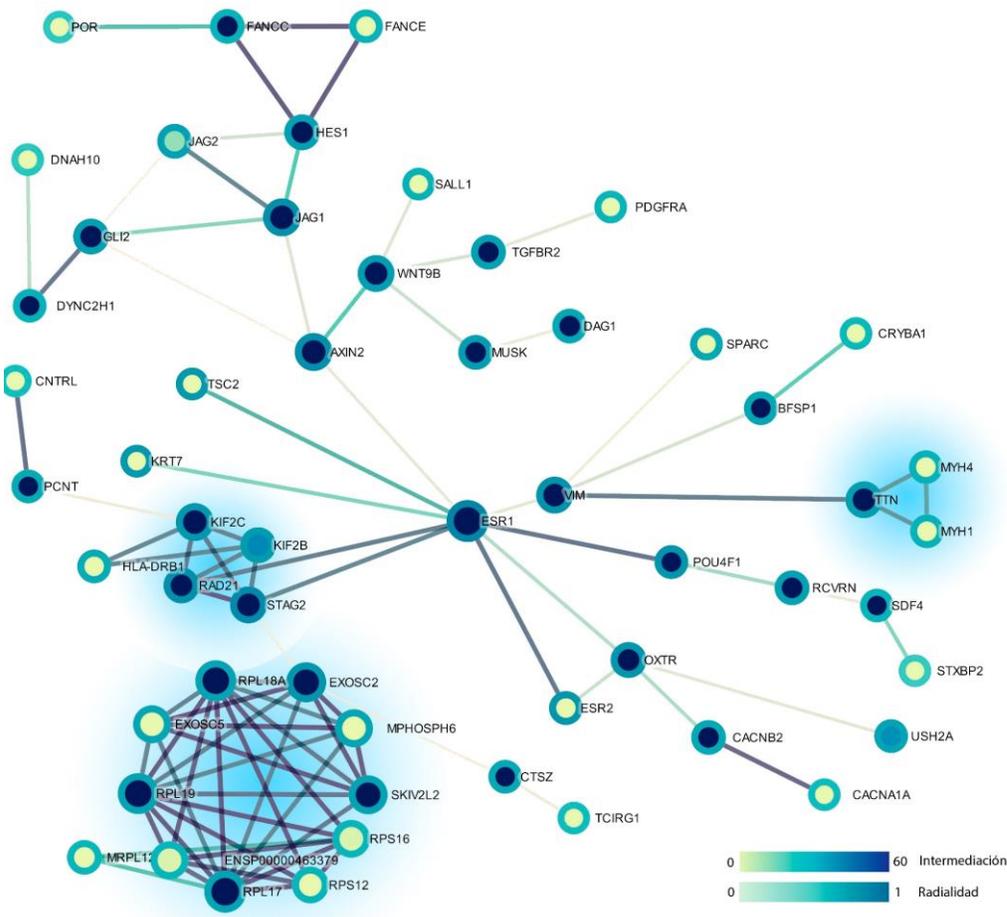


Fig. 10. Red de interacciones proteicas de la paciente, se muestran las interacciones proteína-proteína codificadas por los genes con mutaciones identificadas como moderadas-graves en la paciente (Anexo, tabla 2). *Solo se muestra la red mayor, los nodos que no forman parte de la red se encuentran ocultos. Las zonas sombreadas muestran los clústeres funcionales generados identificados por M-CAP. El tamaño de los nodos indica el parámetro de Grado, donde en este caso indica un valor de 10.

con mayor potencial patogénico pertenecientes a las CNV, y se añadieron 10 interactores con el fin de ampliar la red y

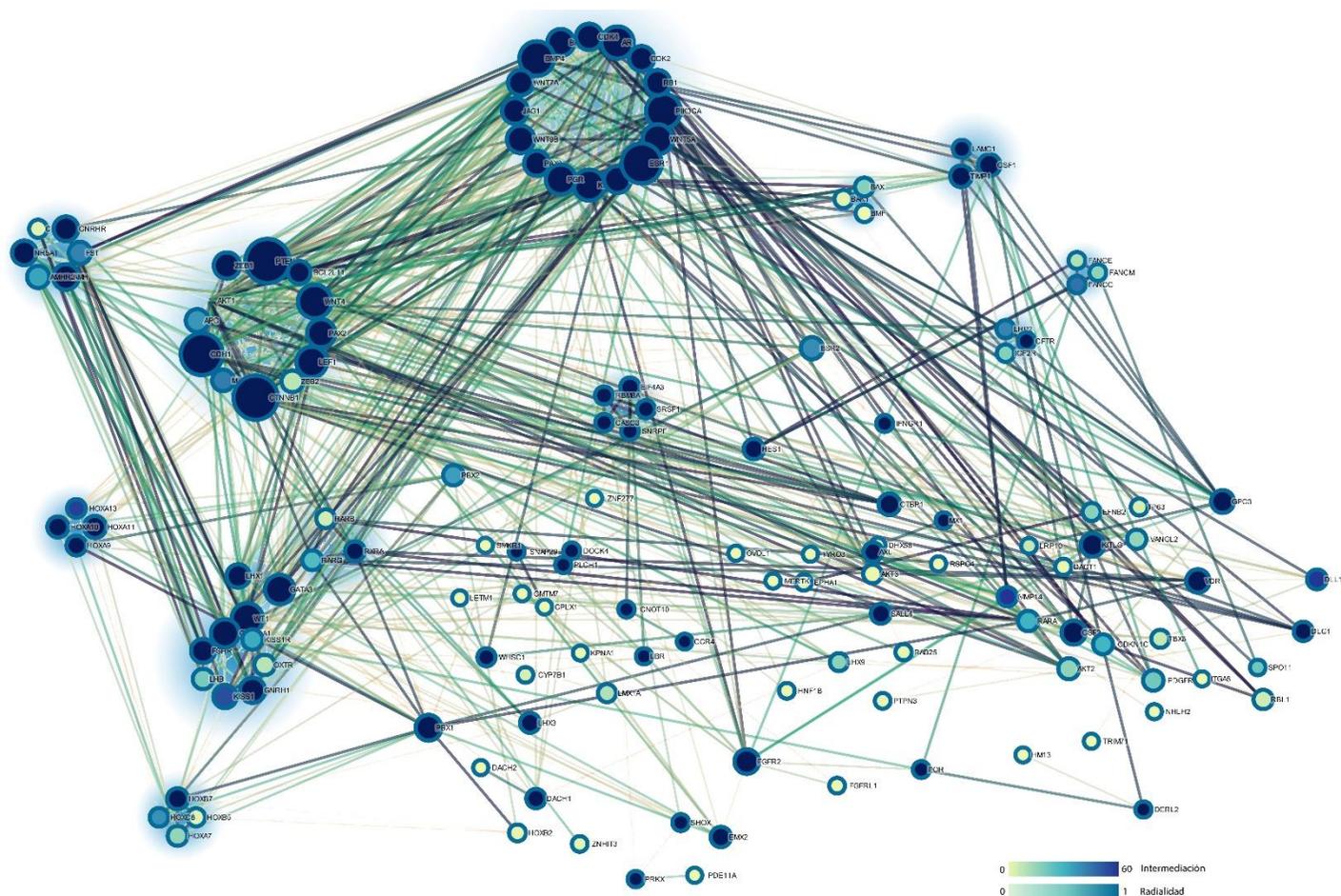


Fig. 11 Red de interacciones proteicas, se muestran las interacciones proteína-proteína codificadas por los genes candidatos de MRKH (lista en Anexo, tabla 1A). *Solo se muestra la red mayor, los nodos que no forman parte de la red se encuentran ocultos. Las zonas sombreadas muestran los clústeres funcionales generados por M-CAP. El tamaño de los nodos indica el parámetro de Grado, donde en este caso indica un valor de 60.

enriquecer su conectividad. La red generada en STRING fue posteriormente analizada en el programa Cytoscape (Paul Shannon et al., 2003) con la ayuda de los plugins stringApp, CentiScaPe 2.2 y M-CAP. Con CentiScaPe se obtuvieron los parámetros de grado, centralidad, intermediación y estrés. Esto generó una red con 114 nodos, de los cuales 57 de éstos se encontraban conectados. Se generaron un total de 2 grupos: uno de 2 nodos, y uno de 55 nodos (Fig. 10).

De manera paralela, se construyó una red PPI a partir de la lista de genes candidatos de MRKH, la cual está conformada por 153 elementos (Anexo 1, Tabla 1A), los cuales se

incluyeron junto con 10 interactores adicionales para enriquecer la red y poder observar de mejor manera las interacciones entre ellas. Por esta razón se obtuvo una red que incluye un total de 163 nodos, de los cuales 150 se encuentran interconectados. La red generada se conforma por 3 grupos que se encuentran distribuidos de la siguiente manera: un grupo de 2 nodos, uno de 4 y un grupo mayor de 143 nodos (Fig. 11).

Al unir los datos de ambas redes se obtuvo una red que muestra en conjunto los genes con mutaciones de la paciente más los genes candidatos de MRKH (Fig. 12), constituida por 256 nodos, de los cuales 207 se encontraron interconectados. En esta red se conformaron 6 grupos, cuatro de ellos de 2 nodos cada uno, uno de 4 nodos y el mayor de 195 nodos.

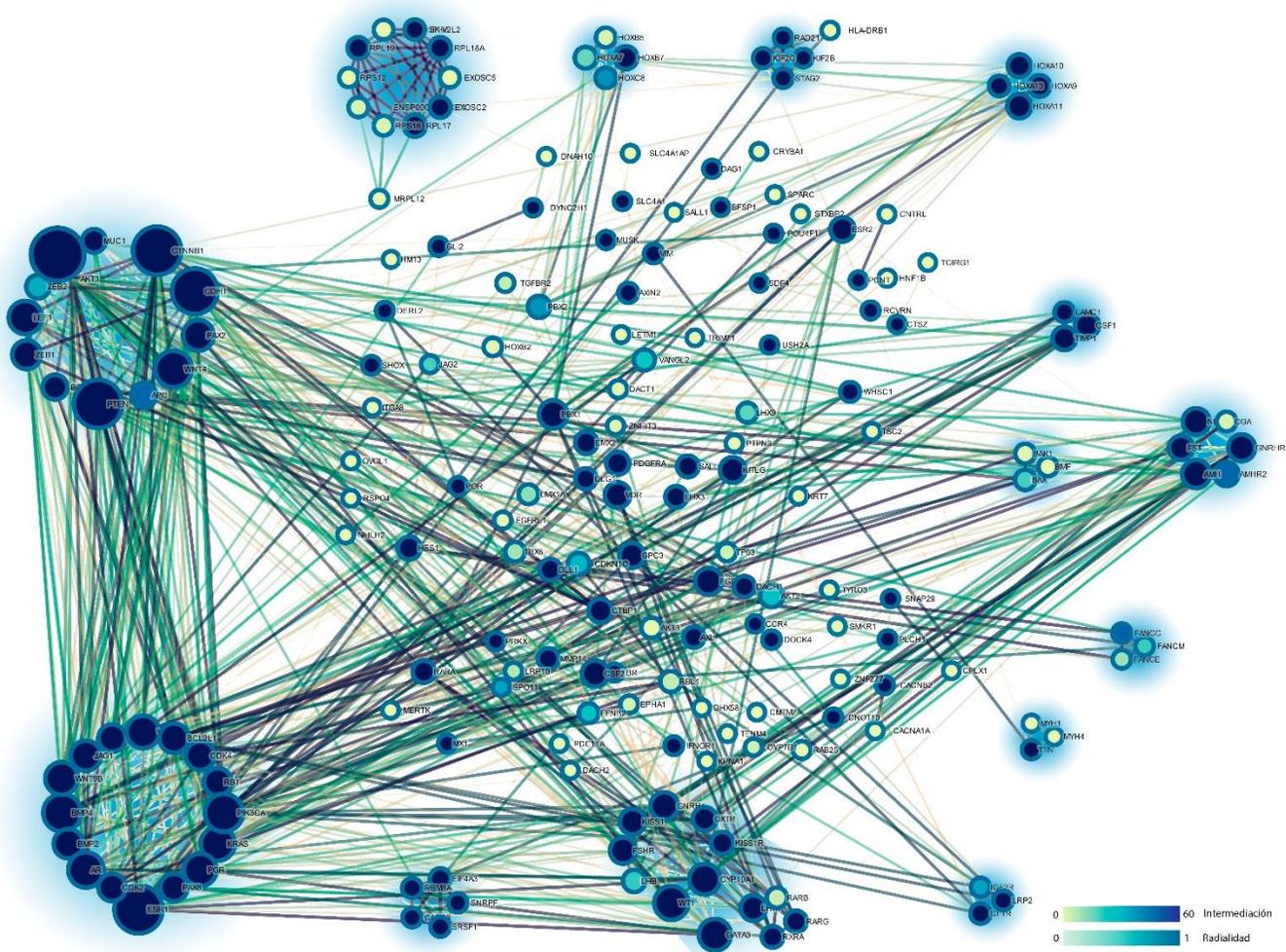


Fig. 12 Red de interacciones proteicas, se muestran las interacciones proteína-proteína codificadas por los genes candidatos de MRKH (lista en Anexo, tabla IA). *Solo se muestra la red mayor, los nodos que no forman parte de la red se encuentran ocultos. Las zonas sombreadas muestran los clústeres/clústeres funcionales generados por M-CAP. El tamaño de los nodos indica el parámetro de Grado, donde en este caso indica un valor de 60.

Análisis de enriquecimiento funcional de redes PPI

El enriquecimiento funcional con Gene Ontology de la red de la paciente arrojó en primer lugar una mayor representación de genes relacionados con el desarrollo glandular con un False Discovery Rate (FDR) de 0.00046 (Tabla 3). El enriquecimiento funcional de los genes candidatos mostró una mayor representación de procesos del desarrollo implicados en la reproducción (FDR= 1.71E-39), es decir, procesos del desarrollo que están relacionados a la capacidad del organismo para generar descendencia (Tabla 4). Por último, el conjunto de genes candidatos más los de la paciente, mostró resultados muy semejantes a la de los genes candidatos, sin embargo, se identificó en cuarto lugar el proceso de morfogénesis del tubo, así como un aumento en el número de genes implicados en ambas redes, pues se complementa con el número de genes con los proporcionados por la paciente y con ello aumenta la fuerza de la relación. Pasando de 58 a 63 en “Proceso del desarrollo implicados en la reproducción”, 59 a 72 en “Morfogénesis de órganos animales” y de 50 a 54 en “Desarrollo de la estructura reproductiva” (Tabla 5).

Proceso celular	Genes presentes	Fuerza	false discovery rate	Proteínas dentro de la red
Desarrollo glandular	13 de 395	0.75	0.00046	RPL19, FOXN1, HES1, SALL1, RPS16, PDGFRA, OXTR, PLAG1, TGFBR2, GLI2, ESR1, DDR1, DAG1
Morfogénesis del epitelio	13 de 414	0.73	0.00046	CTSZ, TSC2, FOXN1, HES1, SALL1, JAG1, WNT9B, JAG2, TGFBR2, GLI2, ESR1, DDR1, DAG1
Desarrollo de órganos sensoriales	15 de 515	0.7	0.00046	CRYBA1, SPARC, HES1, SALL1, JAG1, PDGFRA, WNT9B, USH2A, JAG2, TGFBR2, BFSP1, DSCAM, GLI2, DDR1, VIM
Desarrollo del corazón	14 de 485	0.69	0.00046	TSC2, SPARC, HES1, SALL1, JAG1, PDGFRA, TENM4, AXIN2, OXTR, TGFBR2, POU4F1, DYNC2H1, GLI2, TTN
Desarrollo de tejidos	27 de 1626	0.45	0.00046	CTSZ, TSC2, FOXN1, HES1, SALL1, JAG1, PDGFRA, TENM4, WNT9B, AXIN2, USH2A, JAG2, KRT7, HYAL3, PTPRS, TGFBR2, FLG, TUFT1, PHGDH, POU4F1, BFSP1, GLI2, ESR1, DDR1, DAG1, VIM, TTN

Tabla 3. Enriquecimiento funcional de Gene Ontology obtenido a partir de la red resultante de la paciente. *Solo se muestran los 5 primeros resultados, ordenados por el menor FDR.

Proceso celular	Genes presentes	Fuerza	false discovery rate	Proteínas dentro de la red
Proceso del desarrollo implicados en la reproducción	58 de 622	1.05	1.71E-39	CFTR, HOXA11, LHB, AMH, KITLG, HES1, BMP4, RARA, PDGFRA, AMHR2, TYRO3, LRP2, WNT5A, TP63, LEF1, GNRH1, HOXA10, WNT7A, FANCC, WNT9B, WNT4, BAX, MERTK, CSF2, AXL, BCL2L1, MMP14, CYP7B1, PGR, CSF1, WT1, CTNNB1, DLG1, LHX9, SPO11, LHX3, PTEN, DACH2, SGPL1, NR5A1, BAK1, AR, GATA3, LHFPL2, BCL2L11, CYP19A1, FSHR, RARG, ESR1, PBX1, FGFR2, RXRA, VDR, AKT1, GREB1L, LHX1, HNF1B, DACH1
Desarrollo de la estructura reproductiva	50 de 405	1.17	6.09E-39	HOXA11, LHB, AMH, KITLG, HES1, BMP4, RARA, PDGFRA, AMHR2, TYRO3, LRP2, WNT5A, TP63, LEF1, HOXA10, WNT7A, WNT9B, WNT4, BAX, MERTK, CSF2, AXL, BCL2L1, MMP14, CYP7B1, PGR, WT1, CTNNB1, DLG1, LHX9, SPO11, LHX3, PTEN, SGPL1, NR5A1, BAK1, AR, GATA3, BCL2L11, CYP19A1, FSHR, RARG, ESR1, FGFR2, RXRA, VDR, AKT1, GREB1L, LHX1, HNF1B
Diferenciación sexual	41 de 252	1.29	9.07E-36	HOXA11, LHB, AMH, KITLG, RARA, PDGFRA, AMHR2, TYRO3, LRP2, WNT5A, TP63, HOXA10, WNT7A, WNT9B, WNT4, BAX, MERTK, AXL, BCL2L1, MMP14, PGR, WT1, CTNNB1, LHX9, SPO11, DACH2, SGPL1, NR5A1, BAK1, AR, GATA3, LHFPL2, BCL2L11, CYP19A1, FSHR, ESR1, PBX1, GREB1L, LHX1, HNF1B, DACH1
Morfogénesis de órganos animales	59 de 865	0.91	7.71E-34	HOXA11, HOXC8, HES1, HOXB5, HOXB7, HOXA7, BMP4, RARA, JAG1, KRAS, PDGFRA, PRKX, PAX8, LRP2, AKT3, WNT5A, TP63, LEF1, EIF4A3, WNT7A, FANCC, WNT9B, WNT4, BAX, MMP14, CYP7B1, PGR, CSF1, WT1, HOXB2, RARB, CTNNB1, DLG1, ZEB1, DLL1, LHX9, VANGL2, PAX2, LHX3, PTEN, SGPL1, BAK1, AR, ITGA8, BMP2, GATA3, WHSC1, BCL2L11, GPC3, FGFR1, RARG, ESR1, PBX1, FGFR2, RXRA, VDR, GREB1L, LHX1, HNF1B
Desarrollo del sistema urogenital	41 de 299	1.22	3.08E-33	HOXA11, AMH, HES1, BMP4, RARA, JAG1, PDGFRA, PRKX, PAX8, WNT5A, TP63, WNT9B, WNT4, BAX, CYP7B1, WT1, RARB, OVOL1, CTNNB1, DLG1, DLL1, VANGL2, PAX2, PTEN, SGPL1, AR, ITGA8, BMP2, GATA3, BCL2L11,

				GPC3, CYP19A1, RARG, ESR1, PBX1, FGFR2, RXRA, EMX2, GREB1L, LHX1, HNF1B
--	--	--	--	---

Tabla 4. Enriquecimiento funcional de Gene Ontology obtenido a partir de la red de genes candidatos. *Solo se muestran los 5 primeros resultados, ordenados por el menor FDR

Proceso celular	Genes presentes	Fuerza	false discovery rate	Proteínas dentro de la red
Proceso de desarrollo implicado en la reproducción	63 de 622	0.89	1.97E-32	CFTR, HOXA11, LHB, AMH, KITLG, HES1, BMP4, SALL1, RARA, PDGFRA, AMHR2, TYRO3, LRP2, WNT5A, TP63, LEF1, GNRH1, HOXA10, WNT7A, FANCC, WNT9B, WNT4, BAX, MERTK, CSF2, AXL, BCL2L1, PDILT, MMP14, CYP7B1, PGR, PLAG1, CSF1, WT1, HYAL3, CTNNB1, DLG1, LHX9, SPO11, LHX3, PTEN, DACH2, SGPL1, NR5A1, BAK1, AR, GATA3, LHFPL2, BCL2L11, CYP19A1, FSHR, RARG, GLI2, ESR1, PBX1, FGFR2, RXRA, VDR, AKT1, GREB1L, LHX1, HNF1B, DACH1
Morfogénesis de órganos animales	72 de 865	0.8	1.97E-32	HOXA11, HOXC8, CTSZ, FOXN1, HES1, HOXB5, HOXB7, HOXA7, BMP4, SALL1, RARA, JAG1, KRAS, PDGFRA, PRKX, PAX8, LRP2, AKT3, WNT5A, TP63, LEF1, WNT7A, FANCC, WNT9B, WNT4, BAX, AXIN2, MMP14, CYP7B1, TCAP, PGR, PLAG1, CSF1, JAG2, WT1, HOXB2, RARB, CTNNB1, DLG1, TGFB2, ZEB1, DLL1, LHX9, VANGL2, TUFT1, PAX2, LHX3, PTEN, SGPL1, BAK1, AR, POU4F1, ITGA8, BMP2, GATA3, WHSC1, BCL2L11, GPC3, FGFR1, DSCAM, RARG, GLI2, ESR1, PBX1, FGFR2, RXRA, DDR1, DAG1, VDR, GREB1L, LHX1, HNF1B
Desarrollo de la estructura reproductiva	54 de 405	1.01	1.97E-32	HOXA11, LHB, AMH, KITLG, HES1, BMP4, SALL1, RARA, PDGFRA, AMHR2, TYRO3, LRP2, WNT5A, TP63, LEF1, HOXA10, WNT7A, WNT9B, WNT4, BAX, MERTK, CSF2, AXL, BCL2L1, MMP14, CYP7B1, PGR, PLAG1, WT1, HYAL3, CTNNB1, DLG1, LHX9, SPO11, LHX3, PTEN, SGPL1, NR5A1, BAK1, AR, GATA3, BCL2L11, CYP19A1, FSHR, RARG, GLI2, ESR1, FGFR2, RXRA, VDR, AKT1, GREB1L, LHX1, HNF1B
Morfogénesis del tubo	60 de 615	0.87	2.84E-30	HOXA11, SALL4, CTSZ, TSC2, FOXN1, HES1, HOXA7, EFNB2, BMP4, SALL1, RARA, JAG1, KRAS, PDGFRA, PRKX, PAX8, LRP2, PIK3CA, WNT5A, TP63, LEF1, EPHA1, WNT7A, WNT9B, WNT4, BAX, MMP14, TCAP, PGR, CSF1, WT1,

				DACT1, CTNNB1, DLG1, TGFBR2, ZEB1, DLL1, VANGL2, PAX2, PTEN, SGPL1, BAK1, AR, BMP2, GATA3, TRIM71, BCL2L11, GPC3, RARG, GLI2, ESR1, PBX1, FGFR2, DDR1, DAG1, VDR, ZEB2, GREB1L, LHX1, HNF1B
Desarrollo de órganos animales	120 de 2926	0.5	3.57E-30	HOXA11, HOXC8, SALL4, CTSZ, TIMP1, TSC2, LHB, AMH, CRYBA1, RPL19, FOXN1, KITLG, SPARC, HES1, HOXB5, HOXB7, HOXA7, BMP4, SALL1, RPS16, RARA, JAG1, KRAS, FST, PDGFRA, AMHR2, CDK4, PRKX, PAX8, TYRO3, LRP2, AKT3, PIK3CA, WNT5A, TP63, LEF1, RB1, TENM4, HOXA10, WNT7A, FANCC, WNT9B, WNT4, BAX, MERTK, CSF2, AXL, BCL2L1, AXIN2, USH2A, MMP14, CYP7B1, TCAP, OXTR, PGR, PLAG1, CSF1, JAG2, KRT7, WT1, HOXB2, RARB, CMTM7, HYAL3, OVOL1, LMX1A, CTNNB1, DLG1, IGF2R, PTPRS, TGFBR2, LRP10, SDF4, ZEB1, DLL1, LHX9, VANGL2, FLG, TUFT1, PHGDH, PAX2, SPO11, LHX3, PTEN, SGPL1, NR5A1, BAK1, PTPN3, AR, SLC38A10, POU4F1, BFSP1, ITGA8, BMP2, GATA3, WHSC1, AKT2, BCL2L11, GPC3, TTC7A, CYP19A1, DYNC2H1, FGFR1, DSCAM, FSHR, RARG, GLI2, ESR1, PBX1, FGFR2, RXRA, DDR1, DAG1, VDR, EMX2, AKT1, ZEB2, GREB1L, LHX1, HNF1B

Tabla 5. Enriquecimiento funcional de Gene Ontology obtenido a partir de la red conjunta de genes de la paciente mas genes candidatos. *Solo se muestran los 5 primeros resultados, ordenados por el menor FDR.

Discusión

Los análisis genéticos y la presencia de casos familiares de MRKH apuntan a que este síndrome es causado por un mecanismo poligénico, sin embargo, aún no se logra dilucidar una etiología genética clara. Dentro de los mecanismos propuestos se han identificado alteraciones en las CNV, algunas de estas alteraciones se encuentran reportadas más frecuentemente en pacientes como las asociadas con la enfermedad, tales como las duplicaciones en 1q21.1 (M. J. Chen et al., 2015; Cheroki et al., 2008; Susanne Ledig et al., 2011; McGowan et al., 2015) y Xpter-p22.32 (Morcel et al., 2012), o deleciones en 4q34-qter, 8p23.1, 10p14, 16p11.2, 17q12, 22q11.21 y Xq21.31 (Cheroki et al., 2008; Gatti et al., 2018; Morcel et al., 2011; Nik-Zainal et al., 2011). Aunque muchos de estos sitios

cromosómicos contienen genes que podrían considerarse como candidatos, no se ha observado una causalidad directa. Sin embargo, algunos estudios han buscado SNP concomitantes con la CNV; entre éstos, el de Sandbacka y colaboradores (2013), buscaron SNP dentro de los genes que eran parte de las CNV más comunes en pacientes con MRKH, por lo que determinaron que en su cohorte (n=112) se encontraba de forma simultánea la deleción de 16p11.2 y una variante del gen *TBX6* con una frecuencia del 3.6%. Por otra parte, Ledig y colaboradores (2011) encontraron de forma recurrente la deleción en la región 17q12 y en menor frecuencia la mutación puntual de *LHX1*; por lo que sugieren que podría ser una enfermedad de herencia oligogénica. Ya sea de forma conjunta o por separado, los estudios de CNV no han logrado generar un modelo etiológico, sobre todo porque las alteraciones no se presentan en gran parte de las mujeres con MRKH y además hasta el momento no se ha podido analizar la herencia directamente, ya que las mujeres con MRKH pocas veces logran tener hijos lo cual no permite observar si las alteraciones en las madres permanecen en los hijos, y el impacto en las generaciones subsecuentes.

Otros estudios han intentado generar una hipótesis del desarrollo de la enfermedad con la presencia de SNP. La mayoría de los trabajos iniciales por SNPs mantuvieron la idea de un gen candidato con un enfoque monogénico u oligogénico. Como es el caso del estudio presentado por Waschke y colaboradores (2016) donde se propone el gen *WNT9B* como gen candidato, además de su asociación con mutaciones en los genes *LHX1* y *TBX6*, por lo que se propone un modelo de enfermedad digénico, sin embargo, estas mutaciones solo se encontraron presentes en el 4.6% de las pacientes. Este tipo de enfoque obtuvo relaciones causales poco frecuentes y genes candidatos que poca relación terminaban guardando con la enfermedad (Williams et al., 2017).

Debido a esto, otro nuevo enfoque en la búsqueda de la etiología del MRKH surge al conjuntar la contribución de CNV con las redes de PPI. El estudio desarrollado por Pontecorvi y colaboradores (2021) muestra alteraciones en la CNV en el 25% de las pacientes, e integran los genes que se encuentran en estas duplicaciones o deleciones en una red de interacción; dentro de la red de interacción se encontró una relación de dichos genes con el desarrollo anatómico estructural. Hasta ahora este tipo de enfoque es el que más ha

permitido vislumbrar una posible relación causal del síndrome MRKH, sin embargo, consideramos que no solo es importante tomar en cuenta los cambios de dosis producidos en las CNV, sino también integrar a este tipo de modelos las mutaciones puntuales que pudieran estar alterando las interacciones proteína-proteína, contribuyendo de un modo sutil o poco cuantificable; lo cual permite integrar un modelo etiológico más completo

Dentro de los datos de WES se obtuvo una profundidad promedio de 47 lecturas, lo cual indica que la secuenciación se encuentra en un buen parámetro ya que se espera entre 40 y 50 lecturas por base para considerarlo como una lectura de calidad (Eastman & Yuan, 2015). El indicador Q30 se alcanzó en el 94.72% de las lecturas y el Q20 en el 29.12%, es decir, el 94.72% de los datos tiene una probabilidad del 99.9% de ser correcta y el 29.12% un 99.99%. De las 124,049 variantes encontradas en la paciente, se encontró que cerca del 1% tenía un impacto potencial alto en la proteína y un 10% un impacto moderado. Para priorizar las variantes en la paciente se consideraron distintas características. Entre ellos la frecuencia de estas variaciones en la población en general, esto debido a que entre más baja es la distribución en la frecuencia de un polimorfismo en un gen refleja una mayor presión de selección (Desai & Plotkin, 2008). Ya que si bien los polimorfismos génicos surgen a través de las mutaciones, cuando estos generan un impacto mayor a rasgos funcionalmente importantes, tienden a eliminarse rápidamente por fuerzas de selección (Lande, 1976). Además, se utilizaron puntuaciones de conservación filogenética, pues con ello se trata de observar si el locus se preserva tanto a nivel intraespecífico como interespecífico. Ya que se ha observado que los polimorfismos de un solo nucleótido (SNP) no se presentan en todas las regiones por igual, los sitios más conservados son los que son cruciales para codificar regiones funcionales, por tanto soportan menos los cambios (Castle, 2011). Por último, se utilizaron predictores de daño en la proteína, los cuales consideran diversos factores como: la conservación filogenética del residuo, afectación de sitios de *splicing*, daño estructural de la proteína, y varios más. Gracias a este algoritmo de filtrado se logró acotar de 124,049 a 101 variantes, de las cuales, más del 99% se encuentran categorizadas por la ACMG como de “significancia incierta” o VUS por sus siglas en inglés (Variant of Uncertain Significance). Por dicha razón se prosiguió a hacer un análisis más exhaustivo de las variantes de mayor prioridad. De estas 101 variantes, 17 mostraban predicción de mayor impacto estructural en

las proteínas que codifican (debido a las puntuaciones de los predictores) y una relación con el desarrollo del aparato genitourinario. La tabla 2 muestra el tipo de mutación al que corresponde cada una de las variantes, 9 de estas variantes generan un codón de paro en un sitio equivocado, dando lugar a una proteína trunca y/o no funcional. Este tipo de mutaciones pueden tener consecuencias muy graves y muchas de estas están asociadas a enfermedades genéticamente heredadas y a distintos tipos de cáncer (Catherine L Jopling, 2014). Una mutación generó un cambio en un sitio donador de *splicing*; este tipo de cambios han tomado importancia en los años recientes pues aunque pueden ocurrir tanto en los intrones como en los exones puede generar daños en el transcrito y por ende en la proteína; ya que estas mutaciones normalmente generan errores al reconocer los sitios de *splicing*, lo que puede generar una alteración al momento de remover los intrones, retirar parte de los exones y con ello mover el marco de lectura, por lo que en la traducción se generan proteínas defectuosas (Anna & Monika, 2018). Otro parámetro utilizado para analizar el impacto de estas variantes fue evaluar la probabilidad de que estas pertenezcan a genes intolerantes a LOF, es decir, que sean genes que no toleran las mutaciones pudiendo alterar cualquier elemento genético, lo cual incluye las proteínas. La intolerancia a la pérdida de función se incluyó debido a que variaciones cerca del en estos genes se han visto relacionadas con enfermedades hereditarias (MacArthur & Tyler-Smith, 2010). A su vez también se verificó la probabilidad de que estos genes fueran haploinsuficientes, esto debido a que estos genes no toleran la mutación por LOF o delección de alguno de los alelos cuando estos pertenecen a un organismo diploide, porque un solo alelo no es capaz de producir y fenotipo normal. Además, este tipo de genes se encuentra relacionado a enfermedades de herencia dominante (Johnson et al., 2019).

Análisis orientado a genes candidatos

Dentro de las mutaciones de la paciente se encontró una en el gen *DYNC2H* p.Arg1057His, además que se identificó este gen como sensible a la pérdida de función. Dicho gen codifica para la cadena pesada de dineína, proteína que participa dentro del proceso de transporte intraflagelar ciliar; el cual es un proceso ampliamente regulado, conservado y necesario para la señalización durante los procesos de desarrollo normal (Schmidts et al., 2013). La vía de Sonic Hedgehog (SHh) es una vía de señalización de morfógenos muy conservada y que está relacionada en la organogénesis de muchos metazoarios. Se ha observado que las mutaciones

que afectan la función del gen *DYNC2H1* pueden conducir a la regulación a la baja de la vía SHh, causando una alteración importante durante el desarrollo, llegando a fenotipos de ausencia total de neuronas motoras en la región cervical, sin embargo, la señalización de SHh actúa de forma diferente a lo largo del axis embrionario de una forma aun no comprendida (Legué & Liem, 2020). La vía de SHh normalmente se relaciona al desarrollo del sistema neural, sin embargo, se ha visto implicado en la diferenciación sexual, donde juega un papel crítico en el desarrollo embrionario del tracto reproductivo y el desarrollo de genitales externos; esto por un mecanismo de comunicación entre los epitelios y las células mesenquimales (Franco & Yao, 2012). Además, se ha visto que la activación de esta vía por SHh es indispensable para el desarrollo uterino embrionario y adulto, ya que se ha observado que los WD son necesarios para la activación de SHh, la cual se requiere para el desarrollo de los oviductos, el cuerno y el cuello uterinos (Haraguchi et al., 2021). La vía de SHh tiene 3 ligandos distintos, donde el más estudiado es el de SHh; al no encontrarse la presencia de estos ligandos la vía se encuentra reprimida por el receptor PTCH1 o PTCH2, y una vez activada, se produce la activación de los factores de transcripción: *GLI1*, *GLI2*, *GLI3* (Franco & Yao, 2012). Entre éstos, el gen *GLI2* presentó dos variantes de importancia (Anexo, Tabla 2), lo que podría afectar también los procesos normales de esta vía. *Gli2* actúa principalmente como activador de la transcripción, sin embargo, también puede actuar como represor aunque en menor medida (Pan et al., 2015). A pesar de que *Gli2* no tiene un papel bien definido durante el proceso de sexualización, se ha observado que es indispensable para algunos procesos como la diferenciación sexual de órganos sexuales masculinos externos (Miyagawa et al., 2011). La alteración de la vía de SHh presentaría congruencia fenotípica con el síndrome MRKH, ya que se ha reportado que al verse alterada la vía de SHh no se observan cambios en el desarrollo ovárico, debido a que esta vía se debe mantener inactiva en el ovario fetal, sin embargo, hay indicios de que podría afectar el desarrollo de las células pre-foliculares (Franco & Yao, 2012).

Las mutaciones de pérdida de función del gen *JAG1* están relacionadas con el síndrome de Alagille ya que se presentan en cerca del 70% de los casos. Este es un síndrome caracterizado por anomalías en el desarrollo que afecta múltiples órganos, incluyendo el corazón, hígado, ojos, cara, riñones y páncreas (Zhang et al., 2018). La relación que guarda la proteína

Jag1 y los defectos en el desarrollo se debe posiblemente a que forma parte de la vía de señalización de Notch como uno de cinco ligandos; la vía de Notch se encuentra altamente conservada y controla procesos de diferenciación en muchos tejidos por señales de contacto célula-célula (Hunkapiller et al., 2011). La vía de Notch forma parte de la regulación del desarrollo del cérvix, ovarios y tejido epitelial uterino (Orzechowska et al., 2020). Aunque se ha visto que una desregulación en la vía de Notch tiene efectos sobre el desarrollo del tracto uterino, solo se observan defectos en el oviducto, vascularización del sistema (Ferguson et al., 2016), células estromales uterinas y en el ovario (Orzechowska et al., 2020). Lo que nos lleva a sugerir que esta vía no se ve implicada al menos de forma directa en el desarrollo del síndrome de MRKH. Sin embargo, la vía de Notch tiene una comunicación directa con la vía de WNT/b-catenina (Ferguson et al., 2016).

Entre de los genes de mayor importancia mostrados en la Tabla 1, se encuentra *WNT9B*, un miembro de la familia WNT, que es esencial para el alargamiento de los MD y con ello el desarrollo del tracto reproductor femenino (Carroll et al., 2005; Waschke et al., 2016), además de ser un factor en el desarrollo de riñones, pues las células progenitoras de los riñones necesitan esta la señal de *Wnt9b* para su correcto desarrollo y diferenciación (Karner et al., 2011). En conjunto, la señal de *Wnt9b* se considera que regula la transición del mesénquima metanéfrico a epitelio de los riñones adultos, y regula diferentes procesos de diferenciación del WD, MD y los túbulos mesonéfricos (Carroll et al., 2005). Por dichas razones Tang y colaboradores (2014) propusieron las mutaciones del gen *WNT9B* como causantes de las anomalías del conducto mülleriano, entre ellas el MRKH; sin embargo, en cerca de las 550 pacientes sometidas al estudio no encontraron que estos SNPs fueran lo daño no genera señales de alerta s responsables de estas anomalías. Ya específicamente hablado del síndrome de MRKH gracias a la secuenciación de exoma en una cohorte de 442 mujeres de China, se han observado en *WNT9B* variantes causantes de LOF, sin observar estas variantes en la población control, por lo que se ha categorizado a este gen como uno de los más recientes genes candidatos (N. Chen et al., 2021). Experimentalmente se ha observado que el doble *knock out* presenta un fenotipo similar al del MRKH (Waschke et al., 2016). Un estudio realizado por Ravel y colaboradores (2009), al momento de buscar mutaciones de *WNT9B* en mujeres con MRKH concluye que el SNP (*WNT9B* c.317T>C) mayormente encontrado

en su cohorte no puede ser considerado como un agente causal pues está presente en más del 30% de la población caucásica. Otro estudio concluye que hay una asociación entre una mutación sinónima de *WNT9B* (c.399G>T) y la presencia de MRKH y además la interacción aditiva que fomentaba un mayor riesgo de padecer MRKH cuando se encontraba una variante de *PBX1*; además, documentan efectos de epistasia al conjuntar los efectos de los genes *AMH*, *PBXQ*, *WNT7A* y *WNT9B* (Ma et al., 2015). Si bien la evidencia no es concluyente, la integración a un modelo oligogénico parece incrementar la certeza de que el gen *WNT9B* participa en la etiología del síndrome, por lo que el análisis de las redes funcionales podría dar una mejor idea de su contribución en el desarrollo del MRKH.

El análisis estructural de las proteínas *WNT9B* silvestre y mutante mostró un QMQE de 0.59 y uno de calidad global por residuo de 0.63 ± 0.05 , ambos puntajes arrojan valores entre 0 y 1, por lo que un valor de 0.59 y 0.63 es aceptable para hacer aproximaciones, sin embargo, el modelo tiene limitaciones. En la figura 13 se observa que la proteína silvestre presenta una metionina en el sitio 106, el cual cambia a treonina en la proteína de la paciente, esta mutación según los predictores computacionales no genera daño estructural (MutationTaster: polimorfismo, CADD:Benigno, SIFT:Tolerable). El cambio de metionina a treonina podría generar la aparición de un dipolo que la metionina es un aminoácido hidrofóbico y la treonina un aminoácido polar sin carga, pudiendo generar una interferencia estérica (Biro, 2006). Sin embargo, se ha mostrado que las sustituciones específicas de metionina dentro de una estructura beta-plegada por algún otro aminoácido, genera cambios neutros o lo más parecido a una proteína *wild-type*, por lo que se espera que la mutación p.Met106Thr no genere un gran impacto en la proteína (Gray et al., 2017). Por otra parte la mutación c.600G>C p.(Lys200Asn) cambia la lisina que es un aminoácido polar con carga por una Asparagina que es un aminoácido polar sin carga (Bruce et al., 2002), lo anterior cobra importancia ya que el modelo tridimensional muestra que este residuo forma parte de una alfa-hélice, la cual se encuentra cargada de forma negativa en una de sus caras (superficie expuesta); la homogeneidad de cargas de esta superficie se encontraría interrumpida en la proteína mutante. Los predictores computacionales de impacto en la proteína mutante muestran concordancia en relación con un impacto dañino (MutationTaster: A “Patogénico”, M-CAP: Posiblemente patogénico, CADD:Deletéreo, SIFT:Damaging). Se ha observado la

intolerancia a la sustitución de aminoácidos según la estructura secundaria en la que participan, y se ha observado que la sustitución de la Lisina es la cuarta sustitución que más impacto genera dentro de las alfa-hélices (Gray et al., 2017). Además, esta zona con carga podría ser una zona de interacción o por lo que la pérdida de la carga podría generar un impacto alto en la función de la proteína.

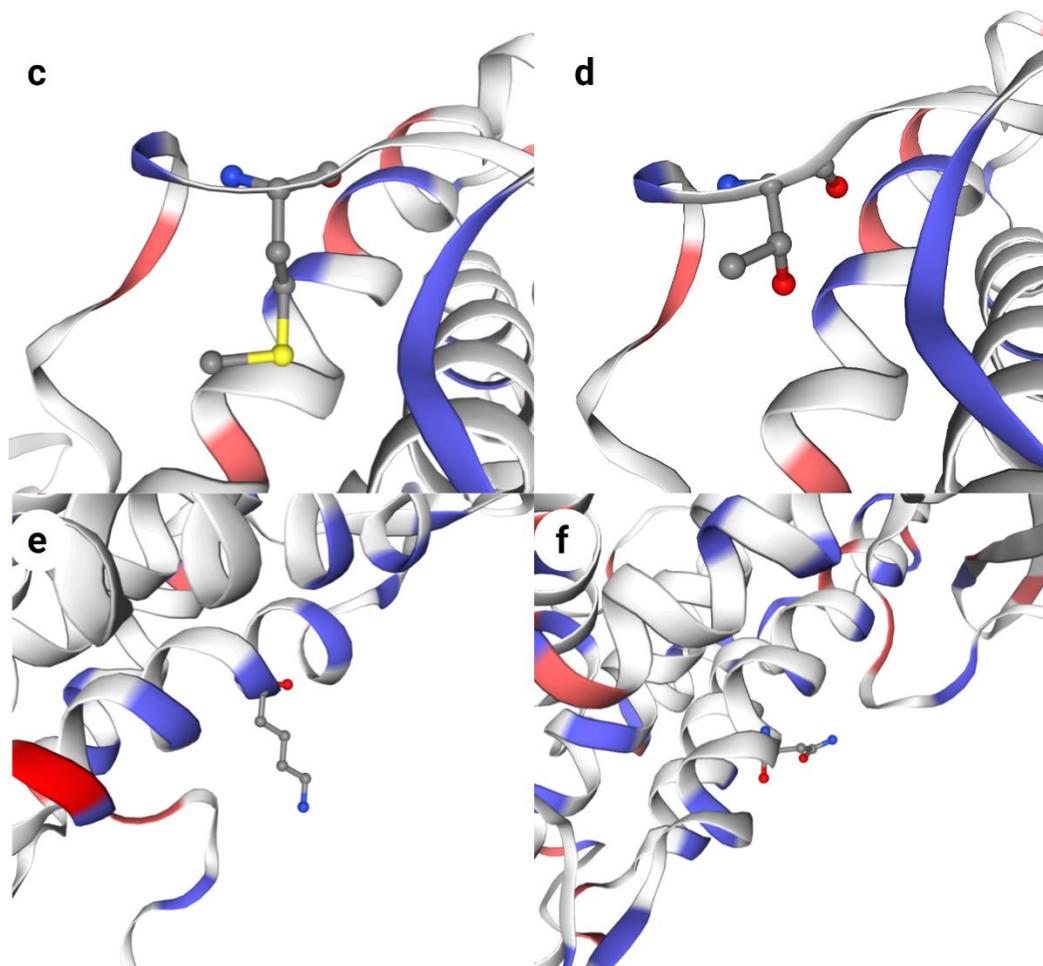


Fig. 13. Acercamiento estructural de proteína WNT9B. c. Acercamiento que muestra la estructura de Metionina 106 (wild-type). d. Acercamiento y estructura de la mutación por Lisina en la posición 106. e. Acercamiento y estructura de Lisina 200 (polar con carga) wild-type. f. Acercamiento y estructura de la mutación por Asparagina 200 (polar sin carga). En rojo se muestran los aminoácidos con carga negativa, en azul aminoácidos con carga positiva.

El análisis de las PPI mostró que la red de la paciente tiene un valor de p de 0.023, lo cual indica que el número de interacciones en la red es significativo, es decir, que al menos las proteínas que la integran se encuentran biológicamente conectadas. Si bien la red de la paciente muestra menos grupos o *clústeres* (Fig. 10), las proteínas que los conforman o bien

el *clúster* completo se integran a distintos *clústeres* cuando se genera la red conjunta de la paciente más genes candidatos (Fig. 13); al adicionar los genes de la paciente, la heterogeneidad de la red aumenta 0.125; la heterogeneidad de la red está directamente ligada a su topología, donde entre más heterogénea es, más se asemeja a una red libre de escala (Ashtiani et al., 2018); éste es un tipo de red que tiene algunos elementos muy conectados y otros no tan conectados, pero no son uniones aleatorias y por tanto son consideradas como redes naturales (Scardoni & Laudanna, 2012). Los clústeres de mayor conectividad encontrados en la red de datos de la paciente no muestran relación directa con el MRKH, ya que estos intervienen en procesos como la maduración del 5.8 rRNA y la segregación cromosómica nuclear. Sin embargo, estos clústeres se integran en la red de candidatos (Fig. 12) lo que podría indicar una función de unión de la red.

La herramienta CentiScaPe nos permitió identificar las proteínas con mayor fuerza dentro de la red, ocupando medidas de centralidad. Las medidas de centralidad que se tomaron en cuenta son: grado, excentricidad, estrés e intermediación. El grado indica el número de nodos a los que se encuentra conectado directamente un nodo específico; biológicamente indica la capacidad regulatoria de la proteína, en redes de señalización las proteínas con un grado alto son las que juegan como “centros” de regulación (Scardoni & Laudanna, 2012). La excentricidad matemáticamente indica la distancia más corta para llegar de un nodo de la red al nodo más distante, en términos biológicos indica la facilidad con la que una proteína puede influir funcionalmente a otra proteína de la red, es decir, que fácilmente influye en la regulación de otras proteínas, por lo que una excentricidad alta indica una regulación más compleja y precisa por muchas proteínas o que, por el contrario, muchas proteínas pueden influir sobre ella (Scardoni & Laudanna, 2012). El estrés biológicamente indica qué tan importante es una proteína que funciona como “puente” para comunicar nodos, es decir, entre más alto el valor del estrés la proteína es más relevante para comunicar moléculas reguladoras (Scardoni & Laudanna, 2012). La intermediación indica la capacidad de una proteína para mantener juntas otras proteínas comunicantes, es decir, indica su capacidad como proteína organizadora y en módulos de señalización su capacidad de señalización (Scardoni & Laudanna, 2012). Como se muestra en la figura 14, la proteína con mayor grado es el ESR1, además ésta también obtiene los valores más altos de estrés e intermediación, por lo que podemos apreciar que dentro de las proteínas mutadas de la paciente ésta juega un papel muy

importante como proteína comunicante, además, al conjuntar la redes (paciente + genes candidatos) las medidas de centralidad aumentan, reforzando su papel como elemento regulador de la red, sin embargo, si la mutación altera su función puede no funcionar adecuadamente como ese centro regulatorio. Además, un estudio realizado por Rall y colaboradores (2011) demostró cambios epigenéticos en el gen del *ESR1* y una sobreexpresión de estos receptores en tejido uterino rudimentario en pacientes con MRKH. *STAG2*, identificado en la red de la paciente, también parece un centro organizador, pues obtiene el segundo lugar en la red de la paciente con los valores de centralidad además que también se observa como un centro organizador de los genes candidatos, por lo que sería importante para mantener una buena señalización dentro de los genes que están relacionados con el MRKH. *STAG2* es una de las cuatro subunidades de la cohesina, un complejo que mantiene unidas a las cromátidas hermanas durante la división celular; en ratones *Stag2*^{-/-} se presenta un retraso en el desarrollo de las somitas durante E9.5 y es letal al E10.5 (De Koninck et al., 2020). En la base de datos *Human Phenotype Ontology Project*, *STAG2* está relacionado a anomalías del sistema genitourinario (Köhler et al., 2021), sin embargo, no existe mayor evidencia de esta relación.

Por otra parte, la proteína *EXOSC2* presenta el tercer lugar en las medidas de centralidad y se integra a la red de candidatos como un centro organizador. La deficiencia de *Exosc2* se ha observado que produce problemas en el desarrollo, aunque estos están principalmente relacionados con microcefalia y pérdida de neuronas motoras (Yatsuka et al., 2020). Este gen es ampliamente conservado desde levaduras hasta humanos, su función se centra en procesar la región 3' RNA y degradar diversos tipos de RNA (Yatsuka et al., 2020). Aunque quizá la relevancia directa de *EXOSC2* y *STAG2* no es muy clara, la red de la paciente muestra que ambas proteínas funcionan como “puente” o comunicación entre dos *clústeres* funcionales (Fig. 10), por lo que su función adecuada es muy importante, ya que aunque las redes soportan pequeñas perturbaciones aleatorias, sin embargo una variación en ese tipo de zonas frecuentemente conduce a una perturbación en todo el sistema (Pavlopoulos et al., 2011).

Paciente					Genes candidatos					Paciente + Genes candidatos				
Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad
ESR1	10	2349.333333	4248	6.448275862	AKT1	60	3564.018237	19698	7.23943662	AKT1	60	4760.12331	28516	5.611320755
STAG2	5	1226	2494	6.103448276	CTNNB1	50	1741.22435	10894	7.133802817	PTEN	50	3817.463309	23900	5.573584906
EXOSC2	8	989	1860	5.534482759	PTEN	50	2553.093185	16000	7.154929577	CTNNB1	50	2335.848579	15704	5.554716981
VIM	5	872.3333333	1334	5.862068966	CDH1	46	1537.487118	11110	7.077464789	ESR1	50	12157.99919	72784	5.649056604
JAG1	6	715.3333333	1290	5.896551724	ESR1	43	970.4299986	7994	7.049295775	CDH1	46	2088.33079	15728	5.524528302
Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad
ESR1	10	2349.333333	4248	6.448275862	AKT1	60	3564.018237	19698	7.23943662	ESR1	50	12157.99919	72784	5.649056604
STAG2	5	1226	2494	6.103448276	PTEN	50	2553.093185	16000	7.154929577	STAG2	5	5230	33486	5.075471698
EXOSC2	8	989	1860	5.534482759	CTNNB1	50	1741.22435	10894	7.133802817	AKT1	60	4760.12331	28516	5.611320755
VIM	5	872.3333333	1334	5.862068966	CDH1	46	1537.487118	11110	7.077464789	EXOSC2	8	3981	23316	4.452830189
JAG1	6	715.3333333	1290	5.896551724	BMP4	34	1168.363368	8344	6.992957746	PTEN	50	3817.463309	23900	5.573584906
					PIK3CA	34	1067.513143	7056	6.964788732					
					ESR1	43	970.4299986	7994	7.049295775					
Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad	Nombre	Grado	Intermediación	Estrés	Radialidad
ESR1	10	2349.333333	4248	6.448275862	AKT1	60	3564.018237	19698	7.23943662	ESR1	50	12157.99919	72784	5.649056604
STAG2	5	1226	2494	6.103448276	PTEN	50	2553.093185	16000	7.154929577	STAG2	5	5230	33486	5.075471698
EXOSC2	8	989	1860	5.534482759	CDH1	46	1537.487118	11110	7.077464789	AKT1	60	4760.12331	28516	5.611320755
VIM	5	872.3333333	1334	5.862068966	CTNNB1	50	1741.22435	10894	7.133802817	PTEN	50	3817.463309	23900	5.573584906
JAG1	6	715.3333333	1290	5.896551724	BMP4	34	1168.363368	8344	6.992957746	EXOSC2	8	3981	23316	4.452830189
					ESR1	43	970.4299986	7994	7.049295775					

Fig. 14. Medidas de centralidad de las redes de PPI de la paciente, de los genes candidatos y la conjunción de ambas redes. Solo se muestran las proteínas que mostraron las medidas de centralidad más altas. La parte superior indica las proteínas de mayor grado en cada red, el compendio del centro las más altas en centralidad y por último las de mayor estrés.

El enriquecimiento funcional de la red de la paciente demostró que la mayoría de los genes encontrados presentan una relación con procesos del desarrollo, sin embargo, no todos estos están directamente relacionados con el desarrollo del sistema genitourinario (Tabla 1a). Es importante aclarar que las bases de datos empleadas para el análisis de enriquecimiento de categorías biológicas, no necesariamente cuentan con la información que permita asociar un grupo de genes a un proceso biológico particular, sobre todo si dicho proceso se desconoce o no ha sido cabalmente descrito, como es el caso del desarrollo del TRF. El desarrollo de tejidos (FDR=0.00046) engloba por sí mismo distintos procesos que podrían o no estar relacionados al MRKH. La morfogénesis del epitelio (FDR=0.00046) también se encuentra dentro de los procesos que podrían resultar afectados debido a las mutaciones encontradas en la paciente; la transformación (diferenciación celular) y el desarrollo del epitelio son imprescindibles en la transformación del MD a distintos órganos con múltiples características epiteliales (Kurita & Cunha, 2001). El enriquecimiento funcional obtenido de los genes candidatos aumenta el número de genes asociados a los procesos implicados en la reproducción y el desarrollo de la estructura reproductiva, lo que indica que, aunque es el proceso con mayor significancia, las mutaciones de la paciente podrían interferir en estos procesos. Estos resultados corroboran lo encontrado por Pontecorvi y colaboradores (2021), los cuales revelan un enriquecimiento de elementos funcionales para el desarrollo de las estructuras anatómicas relacionadas al MRKH.

La inclusión de más pacientes podría generar redes más robustas para así evitar la necesidad de incluir genes candidatos, ya que estos pueden generar un sesgo por la forma en la que algunos de ellos se han integrado a la literatura (ya que algunos candidatos surgieron según la función durante el desarrollo), es decir, solo incluir datos de pacientes y observar directamente las interacciones de las mutaciones en redes de PPI. Además, es necesaria la corroboración de las mutaciones de mayor impacto mediante una técnica alternativa, ya que, aunque las lecturas de la secuenciación tienen buenos valores de calidad, el estándar de oro para variantes puntuales de interés clínico es la secuenciación de tipo Sanger. Asimismo, es necesaria la validación de estas mutaciones de manera experimental (pues en series amplias de pacientes no se ha logrado), con el fin de proporcionar mayor grado de evidencia sobre su función y asociaciones clínicas.

Conclusiones

A partir del análisis mutacional a escala genómica, se identificaron 103 SNV con potencial patogénico y 4 genes que forman parte de CNVs. De estas variantes, 17 tienen la mayor relevancia debido a su potencial patogénico, evaluado mediante diversos criterios como frecuencia poblacional, predicción computacional de impacto en la estructura y función proteínica, entre otras. El análisis de estas variantes apunta a la perturbación de algunas vías de señalización como SHh y Notch, sin embargo, sería necesaria la inclusión de los datos de más pacientes para poder observar si esto constituye un patrón recurrente.

Se identificó una SNV sin previo reporte en *WNT9B*, la cual se considera patogénica con base en predicción computacional. Por lo que se realizó el modelado *in silico* de la proteína utilizando un método por homología; los resultados de esta evaluación estructural son consistentes con los predictores de impacto, aunque las estructuras tridimensionales obtenidas presentan limitaciones de calidad intrínsecas al método de modelado.

El análisis basado en redes PPI, permitió identificar tres genes (*ESR1*, *STAG2* y *EXOSC2*) que constituyen centros organizadores y reguladores, los cuales están relacionados también con problemas del desarrollo. La red de PPI de la paciente mostró la formación de *clústeres* funcionales, sin embargo, éstos no tenían relación a proceso del desarrollo, sino están relacionados principalmente con el procesamiento de diferentes tipos de RNA.

La relación funcional de la red de la paciente se concentró en procesos relacionados al desarrollo, sin embargo, la “morfogénesis del epitelio” es el único proceso directamente ligado al desarrollo de las estructuras anatómicas relacionadas con el TRF, y cuya disfunción permite asociarlo al MRKH.

Literatura citada

- Anna, A., & Monika, G. (2018). Splicing mutations in human genetic disorders: examples, detection, and confirmation. En *Journal of Applied Genetics* (Vol. 59, Número 3, pp. 253–268). <https://doi.org/10.1007/s13353-018-0444-7>
- Arboleda, V. A., Quigley, C. A., & Vilain, E. (2015). Genetic Basis of Gonadal and Genital Development. En *Endocrinology: Adult and Pediatric* (Seventh Ed, Vols. 2–2). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-18907-1.00118-9>
- Ashtiani, M., Salehzadeh-Yazdi, A., Razaghi-Moghadam, Z., Hennig, H., Wolkenhauer, O., Mirzaie, M., & Jafari, M. (2018). A systematic survey of centrality measures for protein-protein interaction networks. *BMC Systems Biology*, *12*(1), 1–17. <https://doi.org/10.1186/s12918-018-0598-2>
- Backhouse, B., Hanna, C., Robevska, G., Van Den Bergen, J., Pelosi, E., Simons, C., Koopman, P., Juniarto, A. Z., Grover, S., Faradz, S., Sinclair, A., Ayers, K., & Tan, T. Y. (2019). Identification of Candidate Genes for Mayer-Rokitansky-Küster-Hauser Syndrome Using Genomic Approaches. *Sexual Development*, *13*(1), 26–34. <https://doi.org/10.1159/000494896>
- Basu, S., Müller, F., & Sanges, R. (2013). Examples of sequence conservation analyses capture a subset of mouse long non-coding RNAs sharing homology with fish conserved genomic elements. *BMC Bioinformatics*, *14*(7), S14. <https://doi.org/10.1186/1471-2105-14-S7-S14>
- Bean, E. J., Mazur, T., & Robinson, A. D. (2009). Mayer-Rokitansky-Küster-Hauser Syndrome: Sexuality, Psychological Effects, and Quality of Life. *Journal of Pediatric and Adolescent Gynecology*, *22*(6), 339–346. <https://doi.org/10.1016/j.jpog.2008.11.006>
- Biason-Lauber, A., Konrad, D., Navratil, F., & Schoenle, E. J. (2004). A WNT4 Mutation

- Associated with Müllerian-Duct Regression and Virilization in a 46,XX Woman . *New England Journal of Medicine*, 351(8), 792–798. <https://doi.org/10.1056/nejmoa040533>
- Binder, G. (2011). Short stature due to SHOX deficiency: Genotype, phenotype, and therapy. *Hormone Research in Paediatrics*, 75(2), 81–89. <https://doi.org/10.1159/000324105>
- Biro, J. C. (2006). Amino acid size, charge, hydrophathy indices and matrices for protein structure analysis. *Theoretical Biology & Medical Modelling*, 3, 15. <https://doi.org/10.1186/1742-4682-3-15>
- Bruce, A., Johnson, A., Lewis, J., Raff, M., Keith, R., & Walter, P. (2002). *Molecular Biology of the Cell* (4a ed.). Garland Science.
- Carbon, S., Douglass, E., Good, B. M., Unni, D. R., Harris, N. L., Mungall, C. J., Basu, S., Chisholm, R. L., Dodson, R. J., Hartline, E., Fey, P., Thomas, P. D., Albou, L. P., Ebert, D., Kesling, M. J., Mi, H., Muruganujan, A., Huang, X., Mushayahama, T., ... Elser, J. (2021). The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Research*, 49(D1), D325–D334. <https://doi.org/10.1093/nar/gkaa1113>
- Carroll, T. J., Park, J. S., Hayashi, S., Majumdar, A., & McMahon, A. P. (2005). Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. *Developmental Cell*, 9(2), 283–292. <https://doi.org/10.1016/j.devcel.2005.05.016>
- Castle, J. C. (2011). SNPs occur in regions with less genomic sequence conservation. *PLoS ONE*, 6(6). <https://doi.org/10.1371/journal.pone.0020660>
- Catherine L Jopling. (2014). Stop that nonsense! *eLife*, 3, 1–3. <https://doi.org/10.7554/eLife.04300>
- Chan, Y. Y., Jayaprakasan, K., Zamora, J., Thornton, J. G., Raine-Fenning, N., & Coomarasamy, A. (2011). The prevalence of congenital uterine anomalies in unselected and high-risk populations: A systematic review. *Human Reproduction Update*, 17(6), 761–771. <https://doi.org/10.1093/humupd/dmr028>
- Chen, M. J., Wei, S. Y., Yang, W. S., Wu, T. T., Li, H. Y., Ho, H. N., Yang, Y. S., & Chen, P. L. (2015). Concurrent exome-targeted next-generation sequencing and single nucleotide polymorphism array to identify the causative genetic aberrations of isolated Mayer-Rokitansky-Küster-Hauser syndrome. *Human Reproduction*, 30(7), 1732–1742. <https://doi.org/10.1093/humrep/dev095>

- Chen, N., Zhao, S., Jolly, A., Wang, L., Pan, H., Yuan, J., Chen, S., Koch, A., Ma, C., Tian, W., Jia, Z., Kang, J., Zhao, L., Qin, C., Fan, X., Rall, K., Coban-Akdemir, Z., Chen, Z., Jhangiani, S., ... Zhu, L. (2021). Perturbations of genes essential for Müllerian duct and Wölffian duct development in Mayer-Rokitansky-Küster-Hauser syndrome. *American Journal of Human Genetics*, *108*(2), 337–345. <https://doi.org/10.1016/j.ajhg.2020.12.014>
- Cheroki, C., Krepischi-Santos, A. C. V., Szuhai, K., Brenner, V., Kim, C. A. E., Otto, P. A., & Rosenberg, C. (2008). Genomic imbalances associated with müllerian aplasia. *Journal of Medical Genetics*, *45*(4), 228–232. <https://doi.org/10.1136/jmg.2007.051839>
- Cunha, G. R., Robboy, S. J., Kurita, T., Isaacson, D., Shen, J., Cao, M., & Baskin, L. S. (2018). Development of the human female reproductive tract. *Differentiation*, *103*, 46–65. <https://doi.org/10.1016/j.diff.2018.09.001>
- De Koninck, M., Lapi, E., Badía-Careaga, C., Cossío, I., Giménez-Llorente, D., Rodríguez-Corsino, M., Andrada, E., Hidalgo, A., Manzanares, M., Real, F. X., & Losada, A. (2020). Essential Roles of Cohesin STAG2 in Mouse Embryonic Development and Adult Tissue Homeostasis. *Cell Reports*, *32*(6). <https://doi.org/10.1016/j.celrep.2020.108014>
- Desai, M. M., & Plotkin, J. B. (2008). The polymorphism frequency spectrum of finitely many sites under selection. *Genetics*, *180*(4), 2175–2191. <https://doi.org/10.1534/genetics.108.087361>
- Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., & Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human Molecular Genetics*, *24*(8), 2125–2137. <https://doi.org/10.1093/hmg/ddu733>
- Duncan, P. A., Shapiro, L. R., Stangel, J. J., Klein, R. M., & Addonizio, J. C. (1979). The MURCS association: Müllerian duct aplasia, renal aplasia, and cervicothoracic somite dysplasia. *The Journal of Pediatrics*, *95*(3), 399–402. [https://doi.org/10.1016/s0022-3476\(79\)80514-4](https://doi.org/10.1016/s0022-3476(79)80514-4)
- Dunlap, K. A., Filant, J., Hayashi, K., Rucker III, E. B., Song, G., Deng, J. M., Behringer, R. R., DeMayo, F. J., Lydon, J., Jeong, J.-W., & Spencer, T. E. (2011). Postnatal Deletion of *Wnt7a* Inhibits Uterine Gland Morphogenesis and Compromises Adult

- Fertility in Mice¹. *Biology of Reproduction*, 85(2), 386–396.
<https://doi.org/10.1095/biolreprod.111.091769>
- Eastman, A., & Yuan, Z.-C. (2015). Development and validation of an rDNA operon based primer walking strategy applicable to de novo bacterial genome finishing. *Frontiers in Microbiology*, 5, 769. <https://doi.org/10.3389/fmicb.2014.00769>
- Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research*, 8(3), 186–194.
- Ferguson, L., Kaftanovskaya, E. M., Manresa, C., Barbara, A. M., Poppiti, R. J., Tan, Y., & AgoulNIK, A. I. (2016). Constitutive Notch Signaling Causes Abnormal Development of the Oviducts, Abnormal Angiogenesis, and Cyst Formation in Mouse Female Reproductive Tract¹. *Biology of Reproduction*, 94(3), 1-12,67.
<https://doi.org/10.1095/biolreprod.115.134569>
- Ferrè, S., & Igarashi, P. (2019). New insights into the role of HNF-1 β in kidney (patho)physiology. *Pediatric Nephrology*, 34(8), 1325–1335.
<https://doi.org/10.1007/s00467-018-3990-7>
- Firth, H. V, Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., Vooren, S. Van, Moreau, Y., Pettett, R. M., & Carter, N. P. (2009). DECIPHER: *D*atabases of *C*hromosomal *I*mbalance and *P*henotype in *H*umans Using *E*nsembl *R*esources. *The American Journal of Human Genetics*, 84(4), 524–533.
<https://doi.org/10.1016/j.ajhg.2009.03.010>
- Fontana, L., Gentilin, B., Fedele, L., Gervasini, C., & Miozzo, M. (2017). Genetics of Mayer–Rokitansky–Küster–Hauser (MRKH) syndrome. *Clinical Genetics*, 91(2), 233–246.
<https://doi.org/10.1111/cge.12883>
- Franco, H. L., & Yao, H. H. (2012). Pathway in Mammalian Sexual Differentiation. *Chromosome Res.*, 20(1), 247–258. <https://doi.org/10.1007/s10577-011-9254-z>.Sex
- Gatti, M., Tolva, G., Bergamaschi, S., Giavoli, C., Esposito, S., Marchisio, P., & Milani, D. (2018). Mayer–Rokitansky–Küster–Hauser Syndrome and 16p11.2 Recurrent Microdeletion: A Case Report and Review of the Literature. *Journal of Pediatric and Adolescent Gynecology*, 31(5), 533–535. <https://doi.org/10.1016/j.jpag.2018.04.003>
- Georgopapadakos, N., Manoli, A., Passia, G., Skandalakis, P. N., & Filipou, D. (2019).

- Uterus Transplantation as a Therapy Method in Mayer-Rokitansky-Küster-Hauser Syndrome. *Cureus*, *11*(12), 1–8. <https://doi.org/10.7759/cureus.6333>
- Gervasini, C., Grati, F. R., Lalatta, F., Tabano, S., Gentilin, B., Colapietro, P., De Toffol, S., Frontino, G., Motta, F., Maitz, S., Bernardini, L., Dallapiccola, B., Fedele, L., Larizza, L., & Miozzo, M. (2010). SHOX duplications found in some cases with type I Mayer-Rokitansky-Küster-Hauser syndrome. *Genetics in Medicine*, *12*(10), 634–640. <https://doi.org/10.1097/GIM.0b013e3181ed6185>
- Gray, V. E., Hause, R. J., & Fowler, D. M. (2017). Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. *Genetics*, *207*(1), 53–61. <https://doi.org/10.1534/genetics.117.300064>
- Grimbizis, G. F., Camus, M., Tarlatzis, B. C., Bontis, J. N., & Devroey, P. (2001). Clinical implications of uterine malformations and hysteroscopic treatment results. *Human Reproduction Update*, *7*(2), 161–174. <https://doi.org/10.1093/humupd/7.2.161>
- Grimbizis, G. F., Gordts, S., Di Spiezio Sardo, A., Brucker, S., De Angelis, C., Gergolet, M., Li, T. C., Tanos, V., Brölmann, H., Gianaroli, L., & Campo, R. (2013). The ESHRE/ESGE consensus on the classification of female genital tract congenital anomalies. *Human Reproduction*, *28*(8), 2032–2044. <https://doi.org/10.1093/humrep/det098>
- Haraguchi, R., Yamada, G., Murashima, A., Matsumaru, D., Kitazawa, R., & Kitazawa, S. (2021). New insights into development of female reproductive tract-hedgehog-signal response in wolffian tissues directly contributes to uterus development. En *International Journal of Molecular Sciences* (Vol. 22, Número 3, pp. 1–13). <https://doi.org/10.3390/ijms22031211>
- Herlin, M., Bjørn, A.-M. B., Rasmussen, M., Trolle, B., & Petersen, M. B. (2016). Prevalence and patient characteristics of Mayer-Rokitansky-Küster-Hauser syndrome: a nationwide registry-based study. *Human Reproduction*, *31*(10), 2384–2390. <https://doi.org/10.1093/humrep/dew220>
- Herlin, M., Højland, A. T., & Petersen, M. B. (2014). Familial occurrence of Mayer-Rokitansky-Küster-Hauser syndrome: a case report and review of the literature. *American Journal of Medical Genetics. Part A*, *164A*(9), 2276–2286. <https://doi.org/10.1002/ajmg.a.36652>

- Herlin, M. K., Petersen, M. B., & Brännström, M. (2020). Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome: a comprehensive update. *Orphanet Journal of Rare Diseases*, *15*(1), 214. <https://doi.org/10.1186/s13023-020-01491-9>
- Horii, M., Boyd, T. K., Quade, B. J., Crum, C. P., & Parast, M. M. (2018). Chapter 1 - Female Genital Tract Development and Disorders of Childhood. En *Diagnostic Gynecologic and Obstetric Pathology* (Third Edit). Elsevier Inc. <https://doi.org/10.1016/B978-0-323-44732-4.00001-7>
- Huang, C. C., Orvis, G. D., Kwan, K. M., & Behringer, R. R. (2014). Lhx1 is required in Müllerian duct epithelium for uterine development. *Developmental Biology*, *389*(2), 124–136. <https://doi.org/10.1016/j.ydbio.2014.01.025>
- Hunkapiller, N. M., Gasperowicz, M., Kapidzic, M., Plaks, V., Maltepe, E., Kitajewski, J., Cross, J. C., & Fisher, S. J. (2011). A role for Notch signaling in trophoblast endovascular invasion and in the pathogenesis of pre-eclampsia. *Development (Cambridge, England)*, *138*(14), 2987–2998. <https://doi.org/10.1242/dev.066589>
- Jacquinet, A., Boujemla, B., Fasquelle, C., Thiry, J., Josse, C., Lumaka, A., Brischoux-Boucher, E., Dubourg, C., David, V., Pasquier, L., Lehman, A., Morcel, K., Guerrier, D., & Bours, V. (2020). GREB1L variants in familial and sporadic hereditary urogenital adysplasia and Mayer-Rokitansky-Kuster-Hauser syndrome. *Clinical Genetics*, *98*(2), 126–137. <https://doi.org/10.1111/cge.13769>
- Jacquinet, A., Millar, D., & Lehman, A. (2016). Etiologies of uterine malformations. *American Journal of Medical Genetics, Part A*, *170*(8), 2141–2172. <https://doi.org/10.1002/ajmg.a.37775>
- Jagadeesh, K. A., Wenger, A. M., Berger, M. J., Guturu, H., Stenson, P. D., Cooper, D. N., Bernstein, J. A., & Bejerano, G. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nature Genetics*, *48*(12), 1581–1586. <https://doi.org/10.1038/ng.3703>
- Johnson, A. F., Nguyen, H. T., & Veitia, R. A. (2019). Causes and effects of haploinsufficiency. *Biological Reviews*, *94*(5), 1774–1785. <https://doi.org/10.1111/brv.12527>
- Karner, C. M., Das, A., Ma, Z., Self, M., Chen, C., Lum, L., Oliver, G., & Carroll, T. J. (2011). Canonical Wnt9b signaling balances progenitor cell expansion and

- differentiation during kidney development. *Development*, 138(7), 1247–1257.
<https://doi.org/10.1242/dev.057646>
- Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, 46(3), 310–315. <https://doi.org/10.1038/ng.2892>
- Köhler, S., Gargano, M., Matentzoglou, N., Carmody, L. C., Lewis-Smith, D., Vasilevsky, N. A., Danis, D., Balagura, G., Baynam, G., Brower, A. M., Callahan, T. J., Chute, C. G., Est, J. L., Galer, P. D., Ganesan, S., Griese, M., Haimel, M., Pazmandi, J., Hanauer, M., ... Robinson, P. N. (2021). The Human Phenotype Ontology in 2021. *Nucleic Acids Research*, 49(D1), D1207–D1217. <https://doi.org/10.1093/nar/gkaa1043>
- Kurita, T., & Cunha, G. R. (2001). Roles of p63 in differentiation of Müllerian duct epithelial cells. *Annals of the New York Academy of Sciences*, 948, 9–12.
<https://doi.org/10.1111/j.1749-6632.2001.tb03982.x>
- Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetics Research*, 26, 221–235.
<https://doi.org/10.1017/S0016672300016037>
- Layman, L. C. (2014). The genetics of Mullerian aplasia. *Expert Review of Endocrinology & Metabolism*, 9(4), 411–419. <https://doi.org/10.1586/17446651.2014.914433>
- Ledig, S., Brucker, S., Barresi, G., Schomburg, J., Rall, K., & Wieacker, P. (2012). Frame shift mutation of LHX1 is associated with Mayer-Rokitansky- Küster-Hauser (MRKH) syndrome. *Human Reproduction*, 27(9), 2872–2875.
<https://doi.org/10.1093/humrep/des206>
- Ledig, Susanne, Schippert, C., Strick, R., Beckmann, M. W., Oppelt, P. G., & Wieacker, P. (2011). Recurrent aberrations identified by array-CGH in patients with Mayer-Rokitansky-Küster-Hauser syndrome. *Fertility and Sterility*, 95(5), 1589–1594.
<https://doi.org/10.1016/j.fertnstert.2010.07.1062>
- Ledig, Susanne, & Wieacker, P. (2018). Klinische und genetische Aspekte des Mayer-Rokitansky-Küster-Hauser Syndroms. *Medizinische Genetik*, 30(1), 3–11.
<https://doi.org/10.1007/s11825-018-0173-7>
- Legué, E., & Liem, K. F. (2020). Mutations in Ciliary Trafficking Genes affect Sonic Hedgehog-dependent Neural Tube Patterning Differentially along the Anterior–

Posterior Axis. *Neuroscience*, 450, 3–14.
<https://doi.org/10.1016/j.neuroscience.2020.07.015>

Londra, L., Chuong, F. S., & Kolp, L. (2015). Mayer-Rokitansky-Kuster-Hauser syndrome: a review. *International Journal of Women's Health*, 7, 865–870.
<https://doi.org/10.2147/IJWH.S75637>

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., ... Moore, H. F. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45(6), 580–585.
<https://doi.org/10.1038/ng.2653>

Ma, W., Li, Y., Wang, M., Li, H., Su, T., Li, Y., & Wang, S. (2015). Associations of polymorphisms in WNT9B and PBX1 with Mayer-Rokitansky-Küster-Hauser syndrome in Chinese Han. *PLoS ONE*, 10(6), 1–15.
<https://doi.org/10.1371/journal.pone.0130202>

MacArthur, D. G., & Tyler-Smith, C. (2010). Loss-of-function variants in the genomes of healthy humans. *Human Molecular Genetics*, 19(R2), R125–R130.
<https://doi.org/10.1093/hmg/ddq365>

MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014). The Database of Genomic Variants: A curated collection of structural variation in the human genome. *Nucleic Acids Research*, 42(D1), 986–992. <https://doi.org/10.1093/nar/gkt958>

McGowan, R., Tydeman, G., Shapiro, D., Craig, T., Morrison, N., Logan, S., Balen, A. H., Ahmed, S. F., Deeny, M., Tolmie, J., & Tobias, E. S. (2015). DNA copy number variations are important in the complex genetic architecture of müllerian disorders. *Fertility and Sterility*, 103(4), 1021-1030.e1.
<https://doi.org/10.1016/j.fertnstert.2015.01.008>

Miyagawa, S., Matsumaru, D., Murashima, A., Omori, A., Satoh, Y., Haraguchi, R., Motoyama, J., Iguchi, T., Nakagata, N., Hui, C.-C., & Yamada, G. (2011). The role of sonic hedgehog-Gli2 pathway in the masculinization of external genitalia. *Endocrinology*, 152(7), 2894–2903. <https://doi.org/10.1210/en.2011-0263>

Morcel, K., Camborieux, L., & Guerrier, D. (2007). Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome. *Orphanet Journal of Rare Diseases*, 2(1), 1–9.

<https://doi.org/10.1186/1750-1172-2-13>

- Morcel, K., Dallapiccola, B., Pasquier, L., Watrin, T., Bernardini, L., & Guerrier, D. (2012). Clinical utility gene card for: Mayer-Rokitansky-Küster-Hauser syndrome. *European Journal of Human Genetics*, 20(2), 246. <https://doi.org/10.1038/ejhg.2011.158>
- Morcel, K., Watrin, T., Pasquier, L., Rochard, L., Le Caignec, C., Dubourg, C., Loget, P., Paniel, B. J., Odent, S., David, V., Pellerin, I., Bendavid, C., & Guerrier, D. (2011). Utero-vaginal aplasia (Mayer-Rokitansky-Küster-Hauser syndrome) associated with deletions in known DiGeorge or DiGeorge-like loci. *Orphanet Journal of Rare Diseases*, 6(1), 1–10. <https://doi.org/10.1186/1750-1172-6-9>
- Ng, P. C., & Henikoff, S. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31(13), 3812–3814. <https://doi.org/10.1093/nar/gkg509>
- Nik-Zainal, S., Strick, R., Storer, M., Huang, N., Rad, R., Willatt, L., Fitzgerald, T., Martin, V., Sandford, R., Carter, N. P., Janecke, A. R., Renner, S. P., Oppelt, P. G., Oppelt, P., Schulze, C., Brucker, S., Hurler, M., Beckmann, M. W., Strissel, P. L., & Shaw-Smith, C. (2011). High incidence of recurrent copy number variants in patients with isolated and syndromic Müllerian aplasia. *Journal of Medical Genetics*, 48(3), 197–204. <https://doi.org/10.1136/jmg.2010.082412>
- Nodale, C., Ceccarelli, S., Giuliano, M., Cammarota, M., D'Amici, S., Vescarelli, E., Maffucci, D., Bellati, F., Panici, P. B., Romano, F., Angeloni, A., & Marchese, C. (2014). Gene expression profile of patients with mayer- Rokitansky-küster- hauser syndrome: New insights into the potential role of developmental pathways. *PLoS ONE*, 9(3). <https://doi.org/10.1371/journal.pone.0091010>
- Orvis, G. D., & Behringer, R. R. (2007). Cellular mechanisms of Müllerian duct formation in the mouse. *Developmental Biology*, 306(2), 493–504. <https://doi.org/10.1016/j.ydbio.2007.03.027>
- Orzechowska, M., Anusewicz, D., & Bednarek, A. K. (2020). Functional Gene Expression Differentiation of the Notch Signaling Pathway in Female Reproductive Tract Tissues— A Comprehensive Review With Analysis . En *Frontiers in Cell and Developmental Biology* (Vol. 8, p. 1469). <https://www.frontiersin.org/article/10.3389/fcell.2020.592616>

- Pan, Y. B., Gong, Y., Ruan, H. F., Pan, L. Y., Wu, X. K., Tang, C., Wang, C. J., Zhu, H. B., Zhang, Z. M., Tang, L. F., Zou, C. C., Wang, H. B., & Wu, X. M. (2015). Sonic hedgehog through Gli2 and Gli3 is required for the proper development of placental labyrinth. *Cell Death & Disease*, *6*(2), e1653–e1653. <https://doi.org/10.1038/cddis.2015.28>
- Patnaik, S. S., Brazile, B., Dandolu, V., Ryan, P. L., & Liao, J. (2015). Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome: A historical perspective. *Gene*, *555*(1), 33–40. <https://doi.org/10.1016/j.gene.2014.09.045>
- Paul Shannon, 1, Andrew Markiel, 1, Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage, 2, Nada Amin, 2, Benno Schwikowski, 1, 5 and Trey Ideker^{2, 3, 4, 5}, 山本隆久, 豊田直平, 深瀬吉邦, & 大森敏行. (2003). Cytoscape: A Software Environment for Integrated Models. *Genome Research*, *13*(22), 426. <https://doi.org/10.1101/gr.1239303.metabolite>
- Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., Schneider, R., & Bagos, P. G. (2011). Using graph theory to analyze biological networks. *BioData Mining*, *4*(1), 1–27. <https://doi.org/10.1186/1756-0381-4-10>
- Philibert, P., Biason-Lauber, A., Gueorguieva, I., Stuckens, C., Pienkowski, C., Lebon-Labich, B., Paris, F., & Sultan, C. (2011). Molecular analysis of WNT4 gene in four adolescent girls with mullerian duct abnormality and hyperandrogenism (atypical Mayer-Rokitansky-Küster-Hauser syndrome). *Fertility and Sterility*, *95*(8), 2683–2686. <https://doi.org/10.1016/j.fertnstert.2011.01.152>
- Pontecorvi, P., Bernardini, L., Capalbo, A., Ceccarelli, S., Megiorni, F., Vescarelli, E., Bottillo, I., Preziosi, N., Fabbretti, M., Perniola, G., Benedetti Panici, P., Pizzuti, A., Grammatico, P., & Marchese, C. (2021). Protein–protein interaction network analysis applied to DNA copy number profiling suggests new perspectives on the aetiology of Mayer–Rokitansky–Küster–Hauser syndrome. *Scientific Reports*, *11*(1), 1–11. <https://doi.org/10.1038/s41598-020-79827-5>
- Rall, K., Barresi, G., Walter, M., Poths, S., Haebig, K., Schaeferhoff, K., Schoenfish, B., Riess, O., Wallwiener, D., Bonin, M., & Brucker, S. (2011). A combination of transcriptome and methylation analyses reveals embryologically-relevant candidate genes in MRKH patients. *Orphanet Journal of Rare Diseases*, *6*(1), 32.

<https://doi.org/10.1186/1750-1172-6-32>

- Ravel, C., Lorenço, D., Dessolle, L., Mandelbaum, J., McElreavey, K., Darai, E., & Siffroi, J. P. (2009). Mutational analysis of the WNT gene family in women with Mayer-Rokitansky-Kuster-Hauser syndrome. *Fertility and Sterility*, *91*(4 SUPPL.), 1604–1607. <https://doi.org/10.1016/j.fertnstert.2008.12.006>
- Raya-Rivera, A. M., Esquiliano, D., Fierro-Pastrana, R., López-Bayghen, E., Valencia, P., Ordorica-Flores, R., Soker, S., Yoo, J. J., & Atala, A. (2014). Tissue-engineered autologous vaginal organs in patients: A pilot cohort study. *The Lancet*, *384*(9940), 329–336. [https://doi.org/10.1016/S0140-6736\(14\)60542-0](https://doi.org/10.1016/S0140-6736(14)60542-0)
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. <https://doi.org/10.1038/gim.2015.30>
- Roly, Z. Y., Backhouse, B., Cutting, A., Tan, T. Y., Sinclair, A. H., Ayers, K. L., Major, A. T., & Smith, C. A. (2018). The cell biology and molecular genetics of Müllerian duct development. *Wiley Interdisciplinary Reviews: Developmental Biology*, *7*(3), 1–13. <https://doi.org/10.1002/wdev.310>
- Sadler, T., & Langman, J. (2012). *Langman's Medical Embryology* (12 edition). Wolters Kluwer Health/Lippincott Williams & Wilkins. <https://doi.org/10.4159/harvard.9780674865853.c10>
- Sandbacka, M., Laivuori, H., Freitas, É., Halttunen, M., Jokimaa, V., Morin-Papunen, L., Rosenberg, C., & Aittomäki, K. (2013). TBX6, LHX1 and copy number variations in the complex genetics of Müllerian aplasia. *Orphanet Journal of Rare Diseases*, *8*(1), 1–13. <https://doi.org/10.1186/1750-1172-8-125>
- Scardoni, G., & Laudanna, C. (2012). Centralities Based Analysis of Complex Networks. *New Frontiers in Graph Theory, March 2012*. <https://doi.org/10.5772/35846>
- Schmidts, M., Arts, H. H., Bongers, E. M. H. F., Yap, Z., Oud, M. M., Antony, D., Duijkers, L., Emes, R. D., Stalker, J., Yntema, J.-B. L., Plagnol, V., Hoischen, A., Gilissen, C., Forsythe, E., Lausch, E., Veltman, J. A., Roeleveld, N., Superti-Furga, A., Kutkowska-

- Kazmierczak, A., ... Mitchison, H. M. (2013). Exome sequencing identifies DYNC2H1 mutations as a common cause of asphyxiating thoracic dystrophy (Jeune syndrome) without major polydactyly, renal or retinal involvement. *Journal of Medical Genetics*, 50(5), 309 LP – 323. <https://doi.org/10.1136/jmedgenet-2012-101284>
- Sociedad Española de Ginecología y Obstetricia (organización). (2014). Anomalías congénitas del útero. *Progresos de Obstetricia y Ginecología*, 57(4), 191–200.
- Song, S., Chen, N., Duan, Y. P., Kang, J., Deng, S., Pan, H. X., & Zhu, L. (2020). Anxiety symptoms in patients with Mayer-Rokitansky-Küster-Hauser syndrome: A cross-sectional study. *Chinese Medical Journal*, 133(4), 388–394. <https://doi.org/10.1097/CM9.0000000000000648>
- Strübbe, E. H., Cremers, C. W., Willemsen, W. N., Rolland, R., & Thijn, C. J. (1994). The Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome without and with associated features: two separate entities? *Clinical Dysmorphology*, 3(3), 192–199. <http://europepmc.org/abstract/MED/7981853>
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47(D1), D607–D613. <https://doi.org/10.1093/nar/gky1131>
- Tang, R., Dang, Y., Qin, Y., Zou, S., Li, G., Wang, Y., & Chen, Z. J. (2014). WNT9B in 542 Chinese women with Müllerian duct abnormalities: Mutation analysis. *Reproductive BioMedicine Online*, 28(4), 503–507. <https://doi.org/10.1016/j.rbmo.2013.11.011>
- Waschk, D. E. J., Tewes, A. C., Römer, T., Hucke, J., Kapczuk, K., Schippert, C., Hillemanns, P., Wieacker, P., & Ledig, S. (2016). Mutations in WNT9B are associated with Mayer-Rokitansky-Küster-Hauser syndrome. *Clinical Genetics*, 89(5), 590–596. <https://doi.org/10.1111/cge.12701>
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>

- Wilhelm, D., Palmer, S., & Koopman, P. (2007). Sex determination and gonadal development in mammals. *Physiological Reviews*, 87(1), 1–28. <https://doi.org/10.1152/physrev.00009.2006>
- Williams, L. S., Eksi, D. D., Ph, D., Shen, Y., Ph, D., Lossie, A. C., Ph, D., Chorich, L. P., Sullivan, M. E., Iii, J. A. P., Kim, H., Ph, D., Alper, O. M., Ph, D., & Layman, L. C. (2017). Genetic Analysis of Mayer-Rokitansky-Kuster-Hauser Syndrome (MRKH) Through Ascertainment of a Large Cohort of Families. *Fertil Steril*, 108(1), 145–151. <https://doi.org/10.1016/j.fertnstert.2017.05.017>.Genetic
- Yatsuka, H., Hada, K., Shiraishi, H., Umeda, R., Morisaki, I., Urushibata, H., Shimizu, N., Sebastian, W. A., Hikida, T., Ishitani, T., Hanada, R., Shimada, T., Kimoto, K., Kubota, T., & Hanada, T. (2020). Exosc2 deficiency leads to developmental disorders by causing a nucleotide pool imbalance in zebrafish. *Biochemical and Biophysical Research Communications*, 533(4), 1470–1476. <https://doi.org/10.1016/j.bbrc.2020.10.044>
- Zhang, E., Xu, Y., Yu, Y., Chen, S., Yu, Y., & Sun, K. (2018). JAG1 loss-of-function mutations contributed to Alagille syndrome in two Chinese families. *Mol Med Rep*, 18(2), 2356–2364. <https://doi.org/10.3892/mmr.2018.9217>

Anexos

Anexo Tabla 1. Genes asociados a MRKH. A. Genes propuestos como candidatos etiológicos de MRKH. B. Genes involucrados en la generación de patologías relacionadas o con solapamiento fenotípico de MRKH.

Genes candidatos MRKH (A)	Genes relacionados con patologías asociadas a MRKH (B)
RB1,OXTR,CSF2,AMHR2,CSF1,DACH2,HOXA13,CCDC28B,FGFRL1,HOXA11,HOXA10,HOXA9,DACH1,AKT2,SALL4,AKT3,DHX58,AKT1,HOXA7,TP63,KPNA1,PDGFRA,FST,WNT5A,HNF1B,OVO L1,AR,OR2T2,POR,LMX1A,TYRO3,PGR,EPHA1,HOXB7,CFTR,ZNF277,RBM8A,CTBP1,LHB,MYCBP2,GATA3,CPLX1,CYP19A1,EFNB2,NHLH2,SGPL1,BCL2L11,ZNHIT3,BMF,CGA,MORC2,WNT4,BBS1,VDR,PBX2,CNOT10,FANCC,ESR	KIF14,SMC3,HOXA12,GRIP1,NIPBL,SALL1,TCTN2,TMEM107,BBS9,BBS7,PITX2,BBS5,CEP290,CC2D2A,BBS4,TP63,MNX1,DYNC2H1,MKKS,PAX3,PAX6,RPGRIP1L,ESCO2,SMC1A,LZTFL1,POR,HOXB9,BBS10,CYP26C1,BBS12,TMEM216,SPECC1L,WDPCP,CCNQ,FREM2,ARL6,TWIST2,B3GLCT,UBR1,HOXB13,HDAC8,HOXB12,HOXB11,HOXB10,TCOF1,RAD21,TMEM67,DVL1,NPHP3,DVL3,FLNA,WNT3,STRA6,BBS2,BBS1,BB

<p>1,ESR2,IGF2R,PBX1,BMP4,DLG1,BMP2 ,VANGL2,LHX3,ZNF816,LHX1,GNRH1, PLCH1,LRP10,FGFR2,LHX9,BCL2L1,C DKN1C,TRIM71,DOCK4,HNRNPCL1,R ARG,SPO11,PTEN,OR4M2,LAMC1,GN RHR,TIMP1,AMH,CCR4,LBR,SNAP29,R SPO4,DACT1,EMX2,IFNGR1,SMKR1,W NT9B,VPS13A,MSL3,CYP7B1,NAALAD L2,PAX2,NR5A1,MMP14,ZEB2,RBL1,Z EB1,NAV3,PAX8,PIK3CA,FSHR,ITGA8, RARA,RARB,ITIH5,GALT,HM13,LEF1, GREB1L,DERL2,PRKX,LRP2,TMEM163 ,CENPB,PDE11A,RXRA,RAB25,NSD2,G PC3,KISS1R,SHOX,BAK1,SH2B1,CMT M7,MX1,WNT7A,LHFPL2,TBX6,MERT K,KISS1,LETM1,KITLG,APC,AXL,WT1, BAX,CTNNB1,KRAS,PTPN3,WHCR,M UC1,JAG1,HOXB2,HOXC8,DLL1,HOXB 5</p>	<p>IP1,TBX3,TMEM231,TTC8,FRAS1,MKS 1,IFT27,B9D1,B9D2,DHCR7,BMPR1B, TRIM32</p>
---	---

Anexo 2. Lista de genes con mutaciones consideradas como de impacto moderado-grave encontradas en la paciente.

Mutaciones moderadas-altas paciente
<p>KANSL1, AIFM1 , PLAG1, ADCY6, ARHGEF17, AXIN2, BFSP1, C10orf90 , CACNA1A , CACNB2 , CFTR , CFTR, CNGA4, CNTRL, CRYBA1 , CTSZ, DAG1, DDR1, DNAH10 , DNAJB13 , DSCAM, DYNC2H1 , ECE2, EFCC1, EIF4E3, ESR1 , ESR2, EXOSC2, FANCC , FAT3, FILIP1, FLG , FOXN1, FRYL, GLI2 , GPR113 , HGSNAT, HHATL, HLA-DRB1 , HYAL3, IGDCC4, JAG1, JAG2, KIF2B, KIF2C, KRT7, MAP3K9, MN1, MRPL12, MUSK , MYH4, NDOR1, NUDT51, PANK1, PCNT, PCNX, PDGFRA , PDILT, PHGDH, PKHD1L1, PLXNB3, POU4F1, PROZ , PTPRS, RAD21 , RBM6, RCE1, RCVRN , RDH14, RHPN1, RITA1 , RPL17, RPL17-C18orf32, SALL1 , SDF4, SHPK, SLC17A9, SLC38A10, SLC4A1, SLC4A1AP, SPARC, STK16, STXBP2, TCIRG1, TENM4, TENM4, TGFBR2, TOP3B, TSC2, TTC7A, TTC7A, TTN*, TUFT1 , USH2A , USP54 , USP6, VARS2, VIM*, WNT9B, WNT9B, XKR4, YY1AP1, ZBTB10, ZFP69B, ZNF574.</p>