



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

**“PROGRAMA DE MAESTRÍA Y DOCTORADO EN
CIENCIAS MÉDICAS, ODONTOLÓGICAS Y DE LA
SALUD”**

**FACULTAD DE MEDICINA
HOSPITAL INFANTIL DE MÉXICO “FEDERICO
GÓMEZ”**

**“BÚSQUEDA DE GENES CANDIDATOS PARA
CONEXIÓN ANÓMALA TOTAL DE VENAS
PULMONARES VARIEDAD MENDELIANA,
MEDIANTE MAPEO DE AUTOCIGOSIDAD”**

TESIS

Para optar por el grado de:

MAESTRÍA EN CIENCIAS MÉDICAS

Presenta:

MARYANGEL PEREA CABRERA

Tutora:

DRA. ROCÍO SÁNCHEZ URBINA

**Programa de maestría y Doctorado en Ciencias Médicas,
Hospital Infantil de México “Federico Gómez”**

Ciudad Universitaria, Cd. Mx. Octubre de 2016



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

INTEGRANTES DEL JURADO

DRA. MARIA DE LA LUZ ARENAS SORDO

PRESIDENTE

DR. JAVIER TADEO GRANADOS RIVERON

SECRETARIO

DRA. ROCIO SANCHEZ URBINA

VOCAL

DRA. JOSEFINA YOALY SANCHEZ LOPEZ

VOCAL

DR. LUIS FELIPE MONTAÑO ESTRADA

VOCAL

DECLARATORIA

Durante los estudios de maestría y la realización de este trabajo de tesis, se contó con el apoyo de la beca CONACYT para posgrado y de Fondos Federales obtenidos a través de la convocatoria de la dirección de investigación del Hospital Infantil de México Federico Gómez.

DEDICATORIA

A mis padres, Ángeles y Saúl, quienes han sido mi apoyo durante toda la vida.

A mis hermanos, Israel, Edgar y Moisés, que me enseñaron a pelear, pero también a reconciliarme.

A mi tutora Rocío, quien ha marcado mi vida al mostrarme las maravillas de la investigación.

A mis compañeros que ahora también son mis amigos y me acompañaron en este camino académico y en la vida.

A Noé, que me hace creer que todo en el futuro es posible.

A mis maestros, no solo los oficiales, sino también a aquellos que me mostraron cosas nuevas en la vida, sin importar si son buenas o malas.

A mis pacientes que son la motivación más importante para seguir buscando respuestas.

Gracias a todos, porque es por ustedes que me he convertido en la persona y profesionalista que soy.

ÍNDICE

| | |
|----------------------------------|----|
| RESUMEN | 7 |
| MARCO TEÓRICO:..... | 10 |
| ANTECEDENTES: | 16 |
| CASO FAMILIAR: | 19 |
| JUSTIFICACIÓN. | 21 |
| PLANTEAMIENTO DEL PROBLEMA. | 22 |
| PREGUNTA DE INVESTIGACIÓN..... | 22 |
| OBJETIVO GENERAL..... | 22 |
| MATERIAL Y MÉTODOS: | 23 |
| ANÁLISIS DEL MICROARREGLO..... | 40 |
| RESULTADOS Y DISCUSIÓN..... | 42 |
| ANÁLISIS DE LA SECUENCIA..... | 48 |
| DISCUSIÓN Y CONCLUSIONES. | 61 |
| PERSPECTIVAS. | 63 |
| BIBLIOGRAFIA:..... | 67 |

RESUMEN

ANTECEDENTES: En México las malformaciones congénitas fueron la 2da causa de muerte en 2010, el 22% de las muertes infantiles ocurrieron en niños menores de un año y la principal causa fueron las cardiopatías congénitas, que aportaron 2, 287 muertes en ese año. La conexión anómala total de venas pulmonares (CATVP) es la Cardiopatía Congénita (CC) cianótica más frecuente en la etapa de lactantes en el Hospital Infantil de México “Federico Gómez” (HIMFG) y a nivel mundial tiene una frecuencia de 1.5% de entre todas las CC. En las CC intervienen factores genéticos presentando heterogeneidad genética y alélica para el caso de la CATVP no sindrómica se han encontrado mutaciones en los genes *TAPVR1*, *ANKRD1*, *ZIC3* y *PDGFRA*, en todos los casos, restringidas a familias individuales, por lo que las causas genéticas no se conocen bien. Los estudios de modelos familiares con patrones de herencia mendeliana son de gran utilidad para determinar la alteración de un solo gen como un indicativo de etiología. En este estudio se presenta una familia con antecedente de consanguinidad (tío-sobrina) y dos hijos afectados con CATVP (concordantes), para mapeo de autocigosidad, ya que la región de ADN genómico que alberga las mutaciones causantes de enfermedad segrega con la enfermedad en la genealogía analizada, lo cual permitirá identificar genes candidatos asociados al desarrollo de la CATVP.

JUSTIFICACIÓN: La CATVP se presenta muy frecuentemente en nuestro medio y es causa de tratamiento médico y quirúrgico, sin embargo, no se conocen las causas genéticas, con la realización de este trabajo se podrán encontrar mutaciones asociadas a su desarrollo y se propondrá el estudio genético de las familias mexicanas con afectados de CATVP, contribuyendo potencialmente a dar un asesoramiento genético adecuado para la toma de decisiones sobre su descendencia futura y en un largo plazo disminuir la incidencia.

OBJETIVO GENERAL: Identificar las regiones de homocigosidad y el gen o genes candidatos en una familia de pacientes con CATVP del HIMFG.

MATERIAL Y METODOS: Previo consentimiento informado, se tomó una muestra de sangre periférica de 5ml al hijo afectado vivo, con la que se realizó extracción de ADN genómico. Para el mapeo de autocigosidad se optó por un estudio de microarreglo, y se utilizó un chip Affymetrix de 250K Snp y la plataforma de Affymetrix® microarray. Los resultados se analizaron con el programa *Homozygosity Mapper* para identificar las regiones homocigotas, considerando el parámetro de pérdida de heterocigosidad (LOH por sus siglas en inglés) que limita regiones basadas en el número de SNPs homocigotos contiguos, la distancia entre los mismos y el índice de heterocigosidad. El criterio empleado para determinar el alto grado de pérdida de heterocigosidad para primer grado de consanguinidad

fue un LOH mínimo de 4 MB en microarreglos de 10K), lo cual sugiere que la región es verdaderamente homocigota por descendencia y derivada del mismo ancestro en común.

ANÁLISIS: Una vez identificada la región, se caracterizó la secuencia, con los recursos bioinformáticos Homozygosity Mapper, gene destiller, UCSC genome browser, Genome, Gene, dbSNP, Ensembl, Gene Data Viewer y NCBI Map Viewer. La determinación de las funciones de la secuencia y genes contenidos en ella se realizó por búsqueda bibliográfica en las bases de datos Pubmed, OMIM, Phenotype-genotype integrator (Phe-GenI) y Scielo para poder establecer su probable relación con la etiología de la CATVP.

RESULTADOS: Se encontró una región homocigota en el cromosoma 4q13.1 de 1.635 Kb aproximadamente, que incluye 17 secuencias codificantes y reguladoras con biotipos diferentes, las cuales incluyen un gen codificante (*EPHA5*), tres pseudogenes (*RPL6P10*, *RPS23P3* y *LOC728048*), un RNA antisentido (*EPHA5-AS1*), cuatro RNA largos no codificantes (*RP11-807H7.2*, *RP11-51P8.1*, *RP11-25H12.1*, *RP11-19C24.1*), dos pseudogenes de RNA pequeños nucleares, un microRNA (*miR1269A*) y cinco secuencias no caracterizadas.

El gen *EPHA5* tiene un gen homólogo *EPHB2*, el cual ha sido asociado con la diferenciación arteria-vena temprana de los vasos sanguíneos embrionarios.

MARCO TEÓRICO:

La conexión anómala total de venas pulmonares (CATVP) se define como la alteración estructural en que las venas pulmonares no tienen una conexión con el atrio izquierdo, condicionando que las venas pulmonares se conecten directamente a los sistemas de venas o drenen en el atrio derecho. De acuerdo a la clasificación de Darling, la CATVP se divide por el sitio de desembocadura de las venas pulmonares¹ (ver Tabla 1), siendo la variante más frecuente la supracardiaca, abarcando aproximadamente el 50% de los casos^{2, 3}.

| Clasificación de Darling | |
|--------------------------|---|
| Tipo I | Nivel supracardiaco (vena cava superior, cayado de la vena ácigos, vena innominada y vena cava superior izquierda persistente). |
| Tipo II | Nivel cardíaco (seno coronario o directamente al atrio derecho, o ambos). |
| Tipo III | Nivel infracardiaco (vena porta, venas hepáticas, <i>ductus venosus</i> , venas gástricas o vena cava inferior) |
| Tipo IV | Tipo mixto (dos o más de los tres niveles previos) |

Tabla 1. Clasificación de Darling para Conexión Anómala Total de Venas Pulmonares¹

La frecuencia reportada de esta patología en series de pacientes con cardiopatía congénita (CC) a nivel internacional se encuentra entre 1.5% al 3.6%⁴, con una incidencia aproximada de 2.5 por cada 100,000 recién nacidos vivos en población no hispana y de 19.8 por cada 100,000 nacidos vivos hispanos⁵, principalmente en varones³. En contraste, en nuestro país la CATVP es una de las CC complejas más frecuentes, situándose en el segundo lugar en frecuencia, únicamente superada por la Tetralogía de Fallot⁶. La CATVP ocupó el cuarto lugar dentro de los procedimientos quirúrgicos para corrección de CC realizados en México en el periodo 2011-2012⁶, no obstante, las 3 primeras cardiopatías intervenidas corresponden a formas simples de CC (comunicación interventricular, persistencia de conducto arterioso y fístula sistémico pulmonar). En el Hospital Infantil de México “Federico Gómez” se reciben aproximadamente 40 casos nuevos anuales, la mayoría de estos son de presentación aislada (68%), con predominancia en el sexo masculino(61%), lo cuales concordante con la literatura en nuestro país³.

DESARROLLO EMBRIOLÓGICO DE LAS VENAS PULMONARES.

El sistema vascular se desarrolla principalmente a partir de las células endoteliales derivadas del mesodermo, que se diferencia a partir de angioblastos, mismos que migran y se agregan para formar “cordones”, que comienzan a formar vasos de distintos calibres mediante la diferenciación de células musculares lisas y pericitos⁷, a través de la expresión de factores de crecimiento de fibroblastos (FGF), integrinas, colágeno, cadherinas vasculares endoteliales y factores de crecimiento vascular endotelial, entre otras⁷. Las venas pulmonares comienzan a desarrollarse en la cuarta semana de gestación (25-27 días de vida intrauterina), con la formación de una evaginación endotelial en la porción dorsal izquierda de la aurícula común, que forma la Vena Pulmonar Común (VPC), en la sexta semana de gestación la VPC se une al plexo pulmonar. Esta absorción genera la división de la vena común en dos ramas (izquierda y derecha), que se dividirán para generar 4 ramas alrededor de la semana 8 de gestación⁸.

Se ha descrito que la CATVP se origina por la falla en el desarrollo de la VPC³. Actualmente se consideran tres posibles mecanismos para explicar su origen embriológico, basándose en estudios patológicos: agenesia, involución y atresia, siendo esta última la causa menos frecuente⁹.

BÚSQUEDA DE GENES CANDIDATOS EN PADECIMIENTOS CON HERENCIA AUTOSÓMICA RECESIVA.

Para la búsqueda de genes candidatos en enfermedades con herencia mendeliana es importante definir inicialmente el patrón de herencia. En familias con uniones consanguíneas y herencia horizontal se considera la herencia autosómica recesiva. Para este patrón el abordaje actual más aceptado es el análisis de autocigosidad^{10,11}, el cual se basa en que dos segmentos cromosómicos provienen de un ancestro común y que son regiones de Homocigosidad alta o pérdida de heterocigosidad (ROH por sus siglas en inglés). Esto ocurre con una alta probabilidad en los hijos de parejas que están estrechamente relacionados (consanguinidad)^{12,13}. Se estima que las uniones consanguíneas entre familiares en segundo grado, o más cercanas, tienen una frecuencia del 10.4% a nivel mundial. La consanguinidad aumenta de manera significativa el número y tamaño de ROH igual o mayor a 10Mb, teniendo en una unión entre primos hermanos, un riesgo de 1.7 a 2.8% de tener hijos con malformaciones congénitas¹⁴.

El análisis molecular de los ROH o secuencias de ADN que comparten los familiares y que sirven para llegar a un locus asociado a la enfermedad se ha desarrollado en los últimos años, existiendo métodos novedosos que sustituyen al análisis de ligamiento que se realizaba en años pasados. Un ejemplo son los microarreglos de polimorfismos de un solo nucleótido (SNP), los cuales pueden detectar ROH a través del genoma y buscar su asociación con la

enfermedad, teniendo ventajas significativas en los casos de consanguinidad al no requerir que múltiples miembros de la familia estén afectadas para su estudio anexar. Para el análisis de ROH existen tres plataformas, entre ellas, Copy Number Analyzer for GeneChip arrays Version 3.011™, Affymetrix Genotyping Console y Homozygosity Mapper, siendo esta última la que posee mayor sensibilidad y especificidad en la determinación, ofreciendo adicionalmente el poder visualizar gráficamente las regiones candidatas en cada uno de los cromosomas anexar¹⁵.

El análisis de LOH, ha permitido desde el año 2005 la identificación exitosa de diversos loci causantes de patologías monogénicas en familias consanguinias, entre ellas el síndrome de Carpenter, Crisponi, Cutis Laxa, discromatosis universal hereditaria, entre otras¹¹.

Otro método útil para identificar mutaciones causantes de enfermedades de origen mendeliano, es el de secuenciación de siguiente generación (NGS por sus siglas en inglés)¹⁶. Las principales tecnologías asociadas a la NGS son la secuenciación exomas y la secuenciación de todo el genoma. La secuenciación del exoma identifica aproximadamente 20,000 variantes de un solo nucleótido (SNVs por sus siglas en inglés)¹⁶. La gran mayoría de los SNVs son polimorfismos que se encuentran en la población general. Sin embargo, usando las bases de datos (por ejemplo dbSNP), pueden ser

excluidas del análisis, aun quedando un gran número de posibles variantes para investigar las relacionadas con el origen genético de la enfermedad ¹⁶.

La ventaja importante para identificación de genes candidatos con los microarreglos de SNPs y la SNG, es la capacidad para obtener resultados en un pequeño número de afectados que pueden encontrarse en familias nucleares ^{13 16}. Hace tres décadas para lograr la identificación de genes responsables de enfermedades mendelianas se realizaba un análisis de ligamiento en grandes familias, como es el caso de la familia de Utah-Idaho en el que se identificó el gen *TAPVRI* para CATVP ¹⁷.

ANTECEDENTES:

Si bien se ha descrito la presencia de CATVP en asociación con diferentes síndromes genéticos bien definidos¹⁸, en la mayor parte de los pacientes se presenta de forma aislada, es decir, la única característica clínica es la CC. En estos casos la etiología no se conoce con certeza, aunque se ha llegado a asociar con la exposición materna a solventes orgánicos¹⁸ y además se ha sospechado una etiología genética, esto último debido que se han descrito algunas mutaciones en genes como *TAPVR1*, *ANKRD1*, *ZIC3* y *PDGFRA* que se asocian a la presencia de CATVP en familias aisladas^{19,20,21,22}, mostrando patrones de herencia autosómica dominante y ligada al cromosoma X, con penetrancia incompleta y expresividad variable^{23 24}; también se han descrito dos casos con alteraciones cromosómicas y CATVP, uno de ellos con una translocación 10;21 t(10;21)(q23.1;q11.2)^{25 20}, y el otro con un anillo 12p²⁶.

La herencia autosómica recesiva de la CATVP se ha propuesto debido a la presencia de la CC en familias en las que solo hay afectados en hermandades, con antecedentes de consanguinidad²⁷, sin embargo, hasta el momento no ha sido descrito un gen que se asocie a la misma^{24 28}.

GENES ASOCIADOS A LA CATVP:

TAPVRI. (Total Anomalous Pulmonary Venous Return 1)

Descrito en 1994, se identificó mediante análisis de ligamiento en una familia con 14 afectados con CATVP aislada (familia Utah-Idaho)²⁴. El gen está localizado en 4q12^{22,17}, en las coordenadas genómicas 4:51,800,000-58,500,000, en el genoma de referencia GRCh38, y cuenta con una forma de herencia autosómica dominante y penetrancia del 40%. *TAPVRI*, se ha caracterizado por presentar un efecto fundador, que se demostró debido a que tanto la familia Utah-Idaho descrita originalmente, como las demás familias identificadas en el área de forma posterior por Ward, *et al.*, contaban con una ancestría rastreable a Escocia¹⁹.

ANKRD1 (Ankyrin Repeat Domain-Containing Protein 1)

Este gen, pertenece a una familia génica altamente conservada conocida como proteínas de repetición de ankirina muscular (MARP por sus siglas en inglés), se localiza en la región 10q23.31, se expresa en el endotelio de los vasos sanguíneos, siendo un participante crucial para la génesis arterial²⁹. Clínicamente se ha asociado con CATVP y aterosclerosis aórtica e iliaca, con una forma de herencia autosómica dominante³⁰. En el año 2008 se determinó que un paciente reportado

previamente con CATVP aislada y una translocación 10;21 t(10;21)(q23.1;q11.2), presentaba pérdida de función de este gen^{20, 25}.

PDGFRA (Platelet-Derived Growth Factor Receptor, Alpha).

Gen localizado en 4q12, en las coordenadas génicas 4:54,229,096-54,298,244, codifica para un receptor de factor de crecimiento derivado de plaquetas alfa, que requiere de la activación del gen *PAX1*, y se asocia generalmente a cáncer de células estromales y a la creación de diferentes proteínas de fusión como lo es la forma atípica de la proteína BCR/ABL en la leucemia mieloide crónica³¹. Mediante modelos animales de pollo y ratón, se observó la asociación de este gen a CATVP ya que mutaciones con pérdida de función generaban la CC. Utilizando estos modelos se determinó una herencia autosómica dominante, expresividad variable y una penetrancia del 7%²².

ZIC3 (Zinc Finger Protein of Cerebellum 3)

El gen ZIC3 se localiza en Xq26.3, codifica para un factor de transcripción tipo dedo de zinc y contribuye a establecer la lateralidad corporal³². ZIC3 presenta una herencia ligada al cromosoma X y se ha asociado con una expresividad variable importante y ha sido identificado en CC complejas aisladas (Transposición de grandes arterias y CATVP) y en formas sindrómicas como asociación VACTERL y *situs inversus* ligado al cromosoma X²¹

CASO FAMILIAR:

En diciembre del año 2014 se ingresó al servicio de cardiología del Hospital Infantil de México “Federico Gómez” a un paciente de un año de edad con diagnóstico de CATVP a vena superior (caso índice) (Figura 1, Individuo V-4). Por sus antecedentes familiares se reportó un hermano mayor de 3 años fallecido que es concordante para la cardiopatía (Figura 1, individuo V-5) y el sitio de conexión. En un interrogatorio posterior se encontró consanguinidad de los padres (tío-sobrina) (Figura 1, Individuos IV-1 y III-3). El antecedente de consanguinidad y la herencia horizontal observada en la genealogía sugiere una herencia autosómica recesiva (ver **¡Error! No se encuentra el origen de la referencia.**). Por lo anterior, la identificación de los genes candidatos en la familia permitiría una búsqueda dirigida de mutaciones en otros pacientes con CATVP aislada en nuestra población, ya que es común que en familias con un solo afectado la herencia autosómica recesiva se confunda con presentación esporádica de la patología.

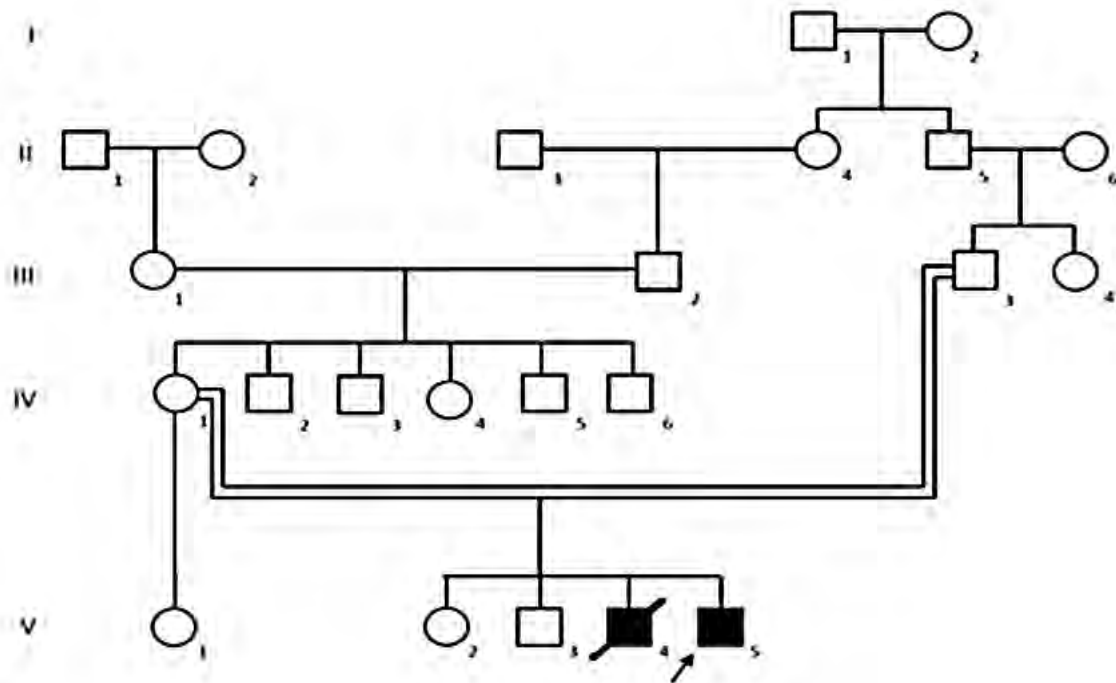


Figura 1. Genealogía de la familia con CATVP. Se observa una unión consanguínea entre los progenitores (individuos III, 3 y IV, 1) y dos hijos afectados con CATVP (V, 4; V, 5).

JUSTIFICACIÓN.

En México las malformaciones congénitas fueron la segunda causa de muerte en 2010, el 22% de las muertes infantiles ocurrieron en niños menores de un año. Dentro de esta categoría, la principal causa son las cardiopatías congénitas, que aportaron 2, 287 muertes en el mismo año. La CATVP se presenta muy frecuentemente en nuestro medio y es causa de tratamiento médico y quirúrgico. Sin embargo, no se conocen aún las causas genéticas. Al encontrar mutaciones asociadas a su desarrollo, podremos proponer el estudio genético de las familias mexicanas con afectados de CATVP y proporcionarles asesoramiento genético para la toma de decisiones sobre su descendencia futura y a largo plazo disminuir la incidencia o favorecer un diagnóstico temprano en aquellas familias con alto riesgo.

PLANTEAMIENTO DEL PROBLEMA.

La evaluación de los pacientes con CATVP debe contemplar el estudio familiar con el fin de determinar los riesgos de recurrencia que pueden seguir patrones de herencia mendeliana. La herencia autosómica recesiva observada en la familia con CATVP, propone la alteración de un gen causante de la enfermedad.

PREGUNTA DE INVESTIGACIÓN.

¿Existirán secuencias provenientes de un ancestro en común que segreguen con la enfermedad y que estén relacionadas con la CATVP en los afectados?

OBJETIVO GENERAL.

Identificar las regiones de homocigosidad y genes candidatos en una familia de pacientes con CATVP del HIMFG.

MATERIAL Y MÉTODOS:

TIPO DE ESTUDIO: transversal descriptivo.

Previo a la realización del estudio, el proyecto se sometió a aprobación por parte de los Comités de Investigación, Ética y Bioseguridad del Hospital Infantil de México “Federico Gómez” (Anexo 1).

Previo firma de consentimiento informado por parte de ambos padres (Anexo 2) se realizó punción venosa del caso índice (individuo V; 5 de la genealogía), y se obtuvieron 3 ml de sangre periférica en un tubo con anticoagulante EDTA para extracción del ADN. Adicionalmente, se volvió a realizar la historia clínica, descartando exposición a organofosforados durante la gestación y se confirmó la genealogía (ver **¡Error! No se encuentra el origen de la referencia.**).

Para la extracción de ADN genómico se utilizó el kit Quiagen® y el equipo automatizado Quiasinphony, distribuido por la misma compañía y que está disponible en el Laboratorio de Cardiopatías Congénitas de este hospital. El ADN obtenido se purificó utilizando el siguiente protocolo:

PURIFICACIÓN DE ADN:

Se colocaron 4 µg de ADN genómico en un volumen total de 20 µl. Se añadió 1 µl de glucógeno a 10 mg/ml, 5 µl NaCl 1 M y 40 µl de etanol al 100%. El ADN se precipitó 30 min a -20°C, y se centrifugó 20 min a 13,000 rpm a 4°C. El botón de ADN se secó y resuspendió en buffer TE (Tris pH 8.0 10 mM, EDTA 0.1 mM).

Posterior a este proceso se cuantificó el ADN en 65.5 ng/µl y se estimó con una pureza por relación 260/280 con un cociente de 1.16/2.28.

El ADN obtenido se mantuvo almacenado hasta su procesamiento en un ultra-congelador a -70°C.

MAPEO DE AUTOCIGOSIDAD:

Se realizó en el Instituto de Oftalmología F.A.P. Conde de Valenciana con la plataforma de Affymetrix® microarray, utilizándose un chip de microarreglos tipo Affymetrix 250k Snp®, siguiendo el protocolo establecido:

1. Dilución del ADN

a) La muestra se diluyó hasta obtener una concentración de 60 ng/µl, con volumen final de 15 µl.

2. Digestión STY (PRE-PCR):

a) En una campana de flujo laminar, se diluyó el ADN genómico obtenido en el paso 1 con Buffer TE, para obtener una concentración de 250 ng en un volumen final de 5 μ L., añadiendo las siguientes cantidades:

| | |
|----------------------|--------------|
| ADN GENÓMICO | 3.81 μ L |
| BUFFER TE | 1.19 μ L |
| VOLUMEN FINAL | 5 μ L |

3. Preparación del Nsp master mix digestión (Affymetrix®).

a) Realizado en la campana de flujo laminar y manteniendo la mezcla en hielo, se mezclaron los siguientes componentes, incluidos en el Kit del proveedor, para 2 tubos de PCR:

| Nsp master mix digestión (2 muestras) | μL |
|--|--------------------------|
| H2O | 26.57 |
| 10X Nsp buffer (NE Buffer 2) | 4.6 |
| BSA (100X) | 0.46 |
| Nsp I (10 U/μL) | 2.3 |
| Total | 33.93 |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

b) Bajo las mismas condiciones de temperatura y en campana de flujo se colocaron 14.75 μl de la mezcla del NSP master mix en dos tubos para PCR y se agregó en uno de ellos 5 μl del ADN obtenido en el paso 2, etiquetándose como Muestra 1. En el segundo tubo se agregaron 5 μl de ADN control (incluido en el kit del microarreglo), etiquetándose como Control. En ambos tubos se obtuvo un volumen final de 19.75 μl .

c) La mezcla anterior se mezcló vigorosamente por 3 segundos y se centrifugó por 30 más previo a la colocación en el termociclador para su digestión.

d) Se colocaron ambos tubos (obtenidos en el paso 3), en un termociclador, asegurándose de que se encuentre en la temperatura de iniciación. Se programó con las siguientes condiciones:

| TEMPERATURA | TIEMPO (minutos) |
|-------------|------------------|
| 37°C | 120 |
| 65°C | 20 |
| 4°C | Indefinido |

e) Se congelaron ambas muestras a -20°C para continuar con el proceso de ligación al día siguiente.

4. Ligación Nsp (Affymetrix®):

Se realizó en campana de flujo laminar y hielo.

a) Se preparó una mezcla para dos muestras, con los siguientes componentes:

| Nsp ligation master mix (2 MUESTRAS) | μl |
|---|-----------|
| T4 Ligase Buffer (10x) | 5.75 |
| Adaptor Nsp I (50μl) | 1.73 |
| T4 ADN Ligase (400 U/μl) | 4.6 |
| Volumen final | 12.08 |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar sesgos de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

b) Posterior a esto se agregaron 5.25 μl de Nsp ligation master mix a cada uno de los 2 tubos digeridos (paso 3), para obtener un volumen final de 25 μl.

c) La mezcla anterior se mezcló vigorosamente por 3 segundos y se centrifugó por 30 más previo a la colocación en el termociclador para su ligación.

d) Se colocaron ambos tubos en un termociclador, mismo que se programó con las siguientes condiciones, asegurándonos de que se encontraba a temperatura de iniciación:

| TEMPERATURA | TIEMPO (minutos) |
|--------------------|-------------------------|
| 16°C | 180 |
| 70°C | 20 |
| 4°C | Indefinido |

e) Se diluyó el ADN ligado con Sty y Nsp con 75 μ l de H₂O, para un volumen final de 100 μ l.

5. Nsp PCR (Affymetrix®):

a) En campana de flujo laminar y manteniendo los componentes en hielo, se preparó el Nsp PCR Master Mix por triplicado para cada muestra (6 tubos) con las siguientes condiciones:

| Nsp PCR Master Mix (6 muestras) | μl |
|--|--------------------------|
| H₂O | 363.4 |
| Titanium Taq PCR Buffer (10x) | 92.0 |
| G-C Melt (5M) | 184.0 |
| dNTP (2.5 mM) | 128.8 |
| PCR Primer 002 (100 mM) | 41.4 |
| Titanium Taq ADN Polymerase (50X) | 18.4 |
| Total | 828.0 |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

b) Se transfirieron 10 μ l del producto ligado (paso 4) de la Muestra 1 a cada uno de los tres tubos y 10 μ l del producto ligado (paso 4) del Control a cada uno de los tres tubos.

c) Se agregaron 90 μ l del Nsp PCR Master Mix a cada uno de los tubos, para un volumen final de 100 μ l.

d) La mezcla anterior se mezcló vigorosamente por 3 segundos y se centrifugó por 30 más previo a la colocación en el termociclador para su amplificación, utilizando el sistema de PCR 9700 (bloque de oro platinado), con el siguiente programa:

| PCR Program for GenAmp™ PCR System 9700 (volumen 100µL) | | |
|---|--------|--------|
| TEMPERATURA | TIEMPO | CICLOS |
| 94°C | 3 min | 1x |
| 94°C | 30 seg | 30 x |
| 60°C | 45 seg | |
| 68°C | 15 seg | |
| 68°C | 7 min | |
| 4°C | ∞ | 1x |

*min= minutos; seg= segundos.

e) Se preparó un gel de agarosa al 2% y se cargó el producto de PCR, un tubo de la muestra y uno de control con 3 µL de buffer de carga, así como marcador de peso molecular, esperando fragmentos de 200 a 1100 pb para considerar una PCR adecuada. En ambos casos se determinó un peso molecular de 800 pb, por lo que se consideró al producto apto para continuar con el proceso.

6. Purificación y cuantificación del producto de PCR.

a) Se homogenizaron los productos de PCR restantes (2 tubos de la muestra y 2 tubos de control).

b) Se transfirió el contenido de ambos tubos a uno solo para así obtener 2 tubos de PCR finales (un control y uno de muestra con volumen final de 200 µl).

c) Se agregaron 12 μl de EDTA 0.5 M en cada tubo de PCR, y se mezcló con pipeta 10 veces.

d) Se dejaron incubar ambos tubos a temperatura ambiente durante 10 minutos, para asegurarse de obtener un producto transparente.

e) Se preparó el Master Mix de Purificación para 2 tubos con los siguientes componentes:

| Master Mix de Purificación | μl |
|----------------------------|----------------------|
| NH4OAC (7.5 M) | 460.0 μL |
| Isopropanol | 1610.0 μL |
| TOTAL | 2070.0 μL |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

f) Se transfirieron 900 μl del Master Mix de Purificación a cada tubo con el producto de PCR, homogenizando con pipeta 5 veces.

g) Se incubaron ambos tubos a temperatura ambiente por 30 minutos, y posteriormente se centrifugó a 2,250 RFC (g), a 4^oC por 30 minutos.

h) Se decantó el sobrenadante por inversión sin disgregar el botón.

- i)** Para garantizar el secado, se dejaron escurrir los tubos hacia abajo por 2 minutos.
- j)** Se realizó el primer lavado, adicionando 1.6 ml de etanol al 75% sobre el botón de cada tubo sin disgregar el botón.
- k)** Centrifugamos a 2,250 RCF (g) por 5 minutos a 4⁰C.
- l)** Se decantó el sobrenadante por inversión sin disgregar el botón y nuevamente se permitió escurrir mecánicamente por 2 minutos.
- m)** Se realizó el segundo lavado, adicionando 1.6 ml de etanol al 75% sobre el botón de cada tubo sin disgregar el botón.
- n)** Centrifugamos a 2,250 RCF (g) por 5 minutos a 4⁰C.
- o)** Se decantó el sobrenadante por inversión sin disgregar el botón y nuevamente se permitió escurrir mecánicamente por 2 minutos.
- p)** Una vez comprobado que el tubo se encontraba seco, se procedió a colocar los tubos en el termoblock 2 minutos a 37⁰C.
- q)** Se disolvió el botón adicionando 55 µl de buffer de elución (Quiagen EB).
- r)** Se agitó en el termoblock a 25⁰C y 500 rpm, por un lapso de 30 minutos (se verificó que el botón se hubiera disuelto completamente).

s) Se diluyeron 2 μL del producto de PCR en 198 μL de H_2O , obteniendo un volumen final de 200 μl .

t) Se cuantificó el producto purificado y se calcularon las diluciones con buffer EB, para obtener 90 ng de producto purificado en 45 μl , obteniendo los siguientes cálculos:

| Muestra | Concentración de ADN | Relación 260/280 | Volumen del ADN purificado | Volumen de Buffer EB | Volumen de reacción total |
|-----------|--------------------------|------------------|----------------------------|----------------------|---------------------------|
| Muestra 1 | 3357.8ng/ μl | 1.99/2.27 | 26.8 μl | 18.2 μl | 45 μl |
| Control | 3572.0 ng/ μl | 1.99/2.29 | 25.2 μl | 19.8 μl | 45 μl |

u) Se congelaron ambas muestras a -20°C para continuar con su fragmentación posteriormente.

7. Fragmentación.

a) En hielo y en el laboratorio principal, se mezcló el buffer de fragmentación y el agua 10 minutos antes de la preparación del Master Mix de fragmentación (**Affymetrix®**), mismo para el que se utilizaron los siguientes elementos:

| Master Mix de Fragmentación | Volumen |
|------------------------------------|----------------|
| H₂O | 77.4 µl |
| 10X FRAGMENTATION BUFFER | 9 µl |
| FRAGMENTATION REAGENT | 3.6 µl |
| TOTAL | 90 µL |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

- b)** Se agregaron a los productos purificados de PCR (paso 6), 5 µl de buffer de fragmentación 10X.
- c)** Se adicionaron 5 µl de Master Mix de Fragmentación a los mismos tubos, obteniendo un volumen final de 55 µl.
- d)** Ambos tubos se mezclaron vigorosamente por 3 segundos y se centrifugó por 30 más previo a la colocación en el termociclador para su fragmentación.
- e)** Las muestras se colocaron en el termociclador utilizando el programa de fragmentación bajo las siguientes condiciones, asegurándose de que esté a temperatura de iniciación previo a la colocación de las muestras:

| Fragment Thermal Cycler Program | |
|--|-------------------------|
| TEMPERATURA | TIEMPO (minutos) |
| 37°C | 35 |
| 95°C | 15 |
| 4°C | ∞ |

f) Se preparó un gel de agarosa al 4%, donde se colocaron 1.5 µl de los productos fragmentados de la muestra y del control, cada una con 4 µl de Buffer de carga, considerándose de buena calidad cuando se obtuvieran fragmentos con longitud menor a 180 pb. En ambos casos se obtuvieron fragmentos con buena calidad con un longitud de 80 y 90 pares de bases respectivamente.

8. Marcaje.

a) Se realizó en el laboratorio principal manteniendo todos los componentes en hielo e inmediatamente después de la fragmentación para evitar la degradación del producto.

b) Se preparó el Master Mix de Marcaje (**Affymetrix®**), en hielo, únicamente para la Muestra 1 utilizando los siguientes componentes:

| Master Mix de Marcaje | Volumen |
|-------------------------------------|----------------|
| TdT Buffer (5x) | 14 µl |
| ADN Labeling Reagent (30 mM) | 2 µl |
| TdT enzyme(30U/µL) | 3.5 µl |
| TOTAL | 19.5 µl |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

c) Se colocaron los 53.5 μ l restantes del producto fragmentado en un tubo de PCR.

d) Se agregaron los 19.5 μ l del master mix de marcaje al tubo, para obtener un volumen final de 75 μ l.

e) La muestra se mezcló vigorosamente por 3 segundos y se centrifugó por 30 más previo a la colocación en el termociclador para su marcaje.

f) Se colocó la muestra en el termociclador con el siguiente programa:

| LABEL PROGRAM | |
|---------------|----------|
| TEMPERATURA | TIEMPO |
| 37°C | 4 horas |
| 95°C | 15 min. |
| 4°C | ∞ |

g) Centrifugar 1 minuto a temperatura ambiente.

h) Se almacenó a -20°C durante la noche.

9. Hibridación.

a) Se realizó en el laboratorio principal.

- b) Se precalentó el horno a 50⁰C y se programó el mismo para oscilar a 60 RPM 1 hora previa a introducir la muestra.
- c) Se colocó el chip a temperatura ambiente y se etiquetó 30 minutos antes de su uso.
- d) Se preparó el Master Mix de Hibridación (**Affymetrix®**), en hielo, utilizando los siguientes componentes:

| Master Mix de Hibridación | Volumen |
|----------------------------------|----------------|
| MES (12X;1.25 M) | 12 µl |
| Denhardt's Solution (50X) | 13 µl |
| EDTA (0.5 M) | 3 µl |
| H Sperm ADN(10 mg/ml) | 3 µl |
| Oligo Control Reagent | 2 µl |
| Human Cot-1 ADN (1 mg/ml) | 3 µl |
| Tween-20 (3%) | 1 µl |
| DMSO (100%) | 13 µl |
| TMACL (5M) | 140 µl |
| TOTAL | 190 µl |

*Nota: En todos los casos se prepara un excedente del Master Mix para evitar errores de pipeteo de acuerdo con el protocolo establecido por el Laboratorio de Genética del Instituto de Oftalmología F.A.P. Conde de Valenciana.

- e) Se transfirió la reacción marcada (75 µl) a un tubo de 1.5 ml.
- f) Se adicionaron los 190 µl de Master Mix de Hibridación a la muestra marcada, para un volumen total de 265 µl.

g) La reacción se mezcló vigorosamente por 3 segundos y se centrifugó por 30 segundos más.

h) Se desnaturalizó la reacción colocándola en el termociclador con las siguientes condiciones:

| TEMPERATURA | TIEMPO |
|-------------|-------------|
| 95°C | 10 min |
| 49°C | 30 segundos |

i) Se homogenizó la muestra mediante pipeteo rápido.

j) Se colocaron 200 µl de la muestra en el chip.

k) Se realizó hibridación a 50°C durante 18 horas a 60 rpm.

l) Después de la hibridación se recuperó la muestra y se colocó a -80°C para trabajar con ella al día siguiente.

10. Lavado de chip.

a) Se realizó en el laboratorio de microarreglos.

b) Se preparó la estación de fluidos con Buffers de lavado Wash Buffer A, B de Quiagen™ y agua, realizando la siguiente mezcla con los reactivos proporcionados por el proveedor para un litro de solución:

| Wash Buffer A | Volumen |
|-----------------------|----------------|
| H ₂ O | 699 ml |
| Tween-20 (10%) | 1 ml |
| 20 X SSP | 300 ml |
| Wash Buffer B | Volumen |
| H ₂ O | 969 ml |
| Tween-20 (10%) | 1 ml |
| 20X SSP | 30 ml |

c) Se colocaron en un tubo Eppendorf de 1.5 ml los siguientes componentes:

| Mezcla Buffer Array Holding | Volumen |
|------------------------------------|----------------|
| SAPE Ambar | 600 µL |
| Anticuerpos anti-Biotina | 600 µL |
| Buffer Array Holding | 1000 µL |

d) Se recuperó la muestra del chip y se almacenó la misma a -80°C.

e) Se llenó el chip con 270 µl de la mezcla Buffer Array Holding (Affymetrix®) (paso 10c).

f) Se inició el lavado del chip con el protocolo Genome WideSNP6_450.

g) Al concluir el lavado se comprobó que no había burbujas dentro del chip.

e) Posterior al lavado se realizó el escaneo del chip en la plataforma designada de Affymetrix® utilizando el programa Mapping250K_Nsp Cell para marcar las regiones homocigotas.

f) Una vez escaneado el microarreglo se verificó el índice “intensity QC”, con un valor de 90.8 (se considera un estudio confiable con un QC de 80 o más).

ANÁLISIS DEL MICROARREGLO

Para el análisis se utilizó el programa Homozygosity Mapper™, mismo que determina el parámetro de pérdida de heterocigosidad (LOH por sus siglas en inglés). Considerando el antecedente de consanguinidad el software permite la identificación de segmentos cromosómicos que potencialmente contienen mutaciones patológicas en estado homoalélico¹⁵ al delimitar regiones de LOH basándose en el número de SNP's homocigotos contiguos, la distancia entre los mismos así como su índice de heterocigosidad¹³, con una probabilidad de error del 2% al contar con solo un paciente afectado proveniente de una familia consanguiniada¹¹.

Cuando la calificación de LOH es alta, se considera fuertemente la posibilidad de que la región es verdaderamente homocigota por descendencia y derivada del mismo ancestro en común. Para primer grado de consanguinidad se considera que el LOH requerido es una secuencia homocigota de un tamaño mayor a 4 MB en 10 microarreglos¹³.

Para la utilización del software se introdujeron los resultados del microarreglo en un formato de texto “.txt” y se establecieron los parámetros del análisis, en base a las características de nuestro caso familiar (Tabla 2).

En el presente estudio se analizaron 250,000 regiones a lo largo de todo el genoma (Mapping 250K), por lo que se requirió configurar el tipo de análisis en el programa para identificar las variantes humanas que coincidieran con mutaciones patológicas homoalélicas (ver Tabla 2).

Si bien la probabilidad de error por heterogeneidad genética en enfermedades poligénicas y multifactoriales es una posibilidad, en este análisis se pudo descartar al encontrar solo una región con LOH alta ¹¹

Debido a que se sospechaba una herencia autosómica recesiva, se realizó la medición de LOH en cada cromosoma (ver Figura 3),

| Tabla 2. Variables de análisis en el programa Homozygosity Mapper™ | |
|---|-------------------------------------|
| Tipo de genoma | Humano |
| Tipo de análisis | Mapping 250KNSP |
| Variables | Motif |
| Caso | Sample 01 |
| Regiones | Analizar homocigotas por cromosoma. |
| Análisis de genes | Todos los fenotipos |
| Genoma de referencia | GRChg38 |

Tabla 2. Variables de análisis de LOH en el programa Homozygosity Mapper™

RESULTADOS Y DISCUSIÓN.

El análisis mediante la plataforma Homozygosity Mapper, mostró una región única que cumplía con la condición de LOH mayor a 4 MB (8 MB en este caso), localizada en 4q13.1 (Figura 2), entre los SNPs rs12644315 y rs10017714 (Figura 3).

Para poder determinar la secuencia contenida en la región fue necesario tomar las coordenadas génicas de los SNPs y comparar su localización con el genoma de referencia más actualizado disponible en las bases de datos: GRCh38.

Las coordenadas genómicas de los SNPs fueron mapeadas en 4:65,555,100 y 4:67,190,300, de acuerdo con el genoma de referencia GRCh38. Utilizando recursos bioinformáticos Homozygosity mapper, Ensembl, GenBank, Genome y UCSC Genome Browser. La región presentó un tamaño de 1,635,200pb.

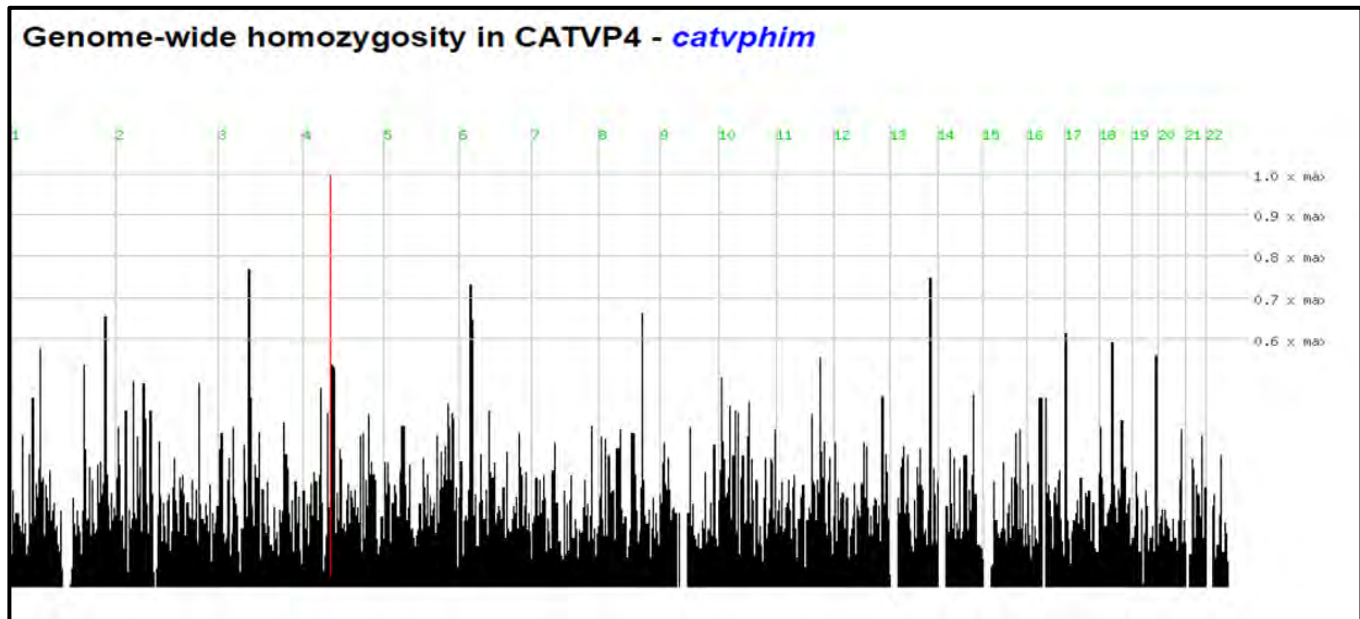


Figura 2. Análisis de pérdida de heterocigosidad por cromosomas. Se aprecia en rojo una región con LOH alto en el cromosoma 4.

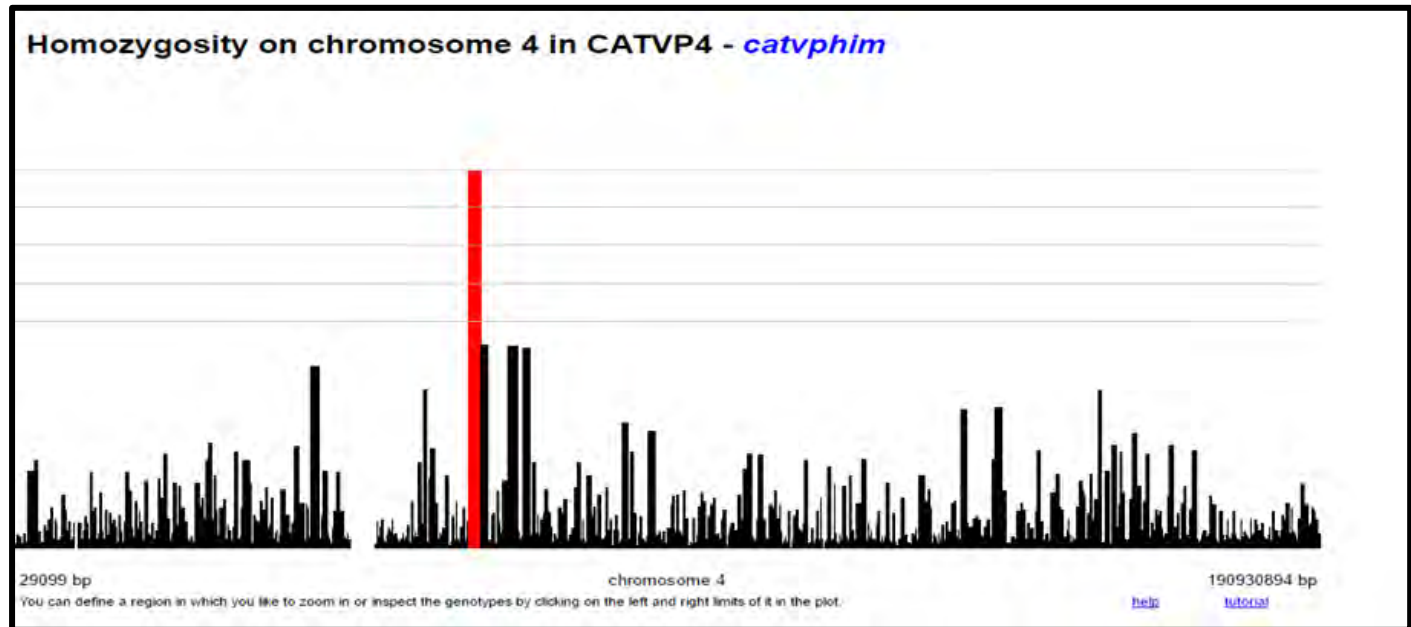


Figura 3. Análisis de LOH en el cromosoma 4. Se aprecia una región homocigota de 1.5 MB en 4q13.1.

La secuencia contenida entre ambos SNPs se caracterizó utilizando las bases de datos Homozygosity Mapper Gene Destiller, UCSC Genome Browser, Genome, Gene, dbSNP y NCBI Map Viewer, lo cual permitió determinar un total de aproximadamente 175 SNPs y 17 secuencias codificantes y reguladoras contenidas en la región (ver Figura 4 y Tabla 3).

Figura 4. Caracterización de la secuencia en la región 4q13.1. Se aprecia el locus y los SNPs contenidos en ella, mostrando todos un LOH alto^{33,34}.

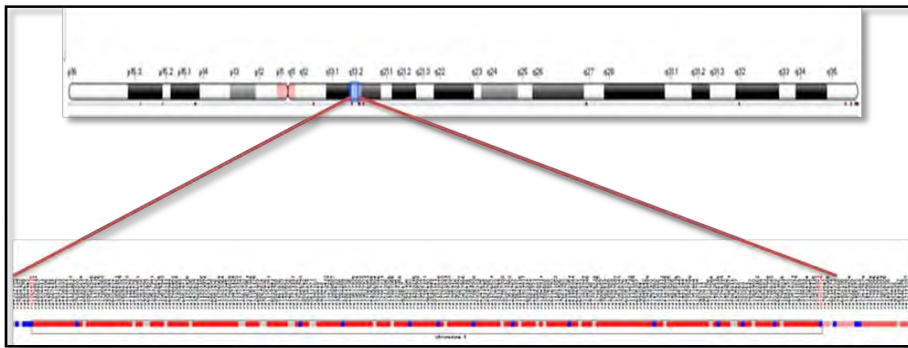


TABLA 3. Características de las secuencias contenidas en la región 4q13.1, respecto al genoma de referencia GRChg38.

| SECUENCIA | BIOTIPO | COORDENADAS GENÓMICAS | LONGITUD (PB) |
|---|---------------------|------------------------------|----------------------|
| <i>EPHA5</i> EPH receptor A5 | Gen | 65,319,563- 65,670,838 | 351,276 |
| <i>RPL6P10</i> Ribosomal proteinL6 pseudogene 10 | Pseudogen | 65,573440- 65574362 | 923 |
| <i>EPHA5-AS1</i> EPHA5 antisense RNA 1 | RNA antisentido | 65,669,961- 65,693,386 | 23,426 |
| RP11-807H7.2 | lincRNA | 65,702,202- 65,705,553 | 3352 |
| LOC105377258 | No caracterizado | 65,780,638- 65,784,672 | 4,035 |
| RNU2-40P RNA, U2 small nuclear 40, pseudogene | snRNA | 65,807,240- 65,807,435 | 196 |
| RP11-51P8.1 | lincRNA | 65,858,761- 65,860,204 | 1,444 |
| LOC105377259 | No caracterizado | 65,868,476- 65,886,184 | 17,709 |
| LOC105377260 | No caracterizado | 65,944,939- 65,948,261 | 3,323 |
| LOC105377261 | No caracterizado | 66,003,201- 66,046,620 | 43,420 |
| RP11-25H12.1 | lincRNA | 65,998,846- 66,150,012 | 151,167 |
| RP11-19C24.1 | LincRNA | 66,003,281- 66,016,792 | 13,512 |

Abreviaturas: PB: pares de bases, ARN: ácido ribonucleico, lincRNA: Long non-coding RNA , snRNA: RNA pequeño nuclear, mRNA: microRNA.

Tabla 3. Continuación. Características de las secuencias contenidas en la región 4q13.1, respecto al genoma de referencia GRCh3.

| SECUENCIA | BIOTIPO | COORDENADAS GÉNICAS | LONGITUD (PB) |
|--|---------------------|----------------------------|----------------------|
| LOC728048 Interferon induced transmembrane protein pseudogene | Pseudogen | 66,094,070- 66,094,686 | 617 |
| MIR1269A | Micro RNA | 66,276,824- 66,276,928 | 105 |
| RPS23P3 | Pseudogen | 66,431,092- 66,431,521 | 430 |
| LOC105377262 | No caracterizado | 66,862,864- 67,077,632 | 214,769 |
| RNU6-699P RNA, U6 small nuclear 699,pseudogene | SnRNA. pseudogen | 66,897,262- 66,897,371 | 110 |

Abreviaturas: PB: pares de bases, ARN: ácido ribonucleico, lincRNA: Long non-coding RNA , snRNA: RNA pequeño nuclear, mRNA: microRNA.

ANÁLISIS DE LA SECUENCIA.

Una vez caracterizada la secuencia flanqueada por los SNPs homocigotos en la región 4q13.1, se procedió a catalogar la información disponible sobre cada uno de los genes y secuencias reguladoras contenidas, para poder establecer su probable relación con la etiología de la CATVP.

Para establecer una probable relación causal con las secuencias codificantes y reguladoras (ver Figura), se realizó una revisión bibliográfica utilizando las bases de datos: Homozygosity Mapper gene destiller, PhenGenI, Gene, Compilation of Aggregate Genomic Data for General Research Use, Pubmed, OMIM y Scielo.

EPHA5 (EPH receptor A5).

Este gen codifica para una proteína de tipo cinasa de tirosina que pertenece a la familia de receptores de Efrina. Esta familia, altamente conservada cuenta con 66 genes ortólogos³⁵, incluidas especies como el ratón (*Hek*)³⁶ y el pollo (*Cek*)³⁷ y adicionalmente se caracteriza por tener una alta homología entre sus integrantes (ver **¡Error! No se encuentra el origen de la referencia.**)³⁸, (ver Figura 4). Aunque se han descrito 16 receptores en vertebrados, divididos en

2 subfamilias: A y B³⁹, solo activan 9 ligandos de efrina, explicando en parte el hecho de que comparten funciones entre ellos⁴⁰.

EPHA5 está conformado por 18 exones, presenta 6 isoformas por corte y empalme alternativo³⁵. En lo que respecta a su estructura, al igual que los otros miembros de la familia, este gen presenta 3 dominios funcionales: dominio catalítico de cinasa de tirosina intracelular, dominio transmembranal y dominio extracelular de unión a ligando³⁶.

Tabla 3. Homología entre las proteínas codificadas por el gen *EPHA5* y sus parálogos³⁴.

| GEN | HOMOLOGÍA CON <i>EPHA5</i> |
|---------------|-----------------------------------|
| <i>EPHA1</i> | 41% |
| <i>EPHA2</i> | 51% |
| <i>EPHA3</i> | 64% |
| <i>EPHA4</i> | 62% |
| <i>EPHA6</i> | 59% |
| <i>EPHA7</i> | 61% |
| <i>EPHA8</i> | 57% |
| <i>EPHA10</i> | 47% |
| <i>EPHB1</i> | 54% |
| <i>EPHB2</i> | 54% |
| <i>EPHB3</i> | 53% |
| <i>EPHB4</i> | 46% |
| <i>EPHB6</i> | 38% |

Esta superfamilia de receptores se ha asociado a diferentes funciones comunes, entre ellas al establecimiento de los segmentos corporales, ya que se sabe que estos receptores se expresan en las somitas y rombómeros del embrión^{37,41}, durante el desarrollo de las extremidades³⁷, desarrollo neurológico y angiogénesis⁴².

De forma específica, *EPHA5* se ha asociado con funciones muy variadas, incluyendo el desarrollo neurológico, expresándose en el cerebro de modelos animales y humanos en todas las etapas de la vida^{37,42}, relacionándose con la plasticidad neuronal y de forma secundaria con diversas alteraciones en el comportamiento y discapacidad intelectual⁴²⁻⁴⁴. De igual manera se ha involucrado en la fisiopatología de la diabetes mellitus tipo 1 y tipo 2, ya que participa activamente en la función ventromedial hipotalámica⁴³ y con la secreción de insulina en los islotes de las células beta pancreáticas⁴⁵. Se ha demostrado su función en relación al metabolismo óseo, al participar como un activador de osteoclastos⁴⁶. También existen reportes que asocian su sobreexpresión con diversos tipos de tumores sólidos malignos, al interferir con la expresión de genes asociados a la reparación celular como ATM (Ataxia Telangiectasia Mutated)⁴⁷. Su regulación a la baja por metilación en su promotor explican este fenómeno⁴⁸⁻⁵⁰. Algunas neoplasias relacionadas incluyen el cáncer pulmonar⁴⁷, prostático, adenocarcinoma ductal pancreático⁵⁰, glioblastoma, cáncer de mama⁴⁹ y colorectal⁵¹.

Su homólogo *EPHB2* se ha asociado con angiogénesis y diferenciación arterio-venosa temprana^{7,52,53}. Por otro lado, en el pasado se ha descrito una afinidad cruzada entre los ligandos de los receptores EPHA4, EPHA5 y EPHB2^{39,52,54}, que se puede explicar por la alta homología entre su dominio de unión a ligando⁴⁰ (ver Figura 4). La alta homología entre los receptores EPHA5 y EPHB2, ha quedado clara en diferentes situaciones, entre ellas la participación similar en el metabolismo óseo, en donde ambos genes se expresan en los osteoclastos, con un efecto inhibitor de la osteogénesis⁴⁶.

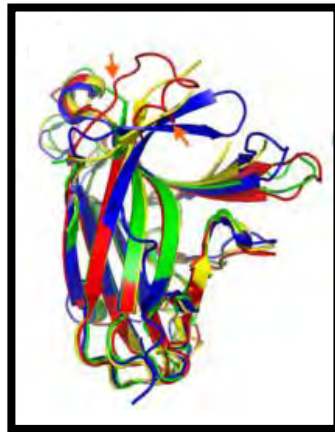


Figura 4. Imagen comparativa de la estructura de los receptores EPH. Se aprecia la superposición del dominio de unión a ligando de EPHA5 (rojo), EPHA2 (verde), EPHA4 (amarillo), EPHB2 (azul). Tomado de Huan X, Shi J, Lim L, *et al.*⁴⁰.

PSEUDOGENES DE PROTEÍNAS RIBOSOMALES

RPL6P10 (Ribosomal proteinL6 pseudogene 10) y
RPS23P3 (Ribosomal protein S23 pseudogene 3).

Pertenecientes a una amplia familia de pseudogenes de proteínas ribosomales (PPR), que está conformada aproximadamente por 2,000 integrantes en el ser humano⁵⁵, se caracterizan por derivar de genes ancestrales funcionales que codifican para diversas proteínas ribosomales, pudiendo ser filogenéticamente rastreables a aproximadamente 60 millones de años⁵⁶. Por ello, que se pueden encontrar ortólogos en distintos organismos como levaduras, bacterias y mamíferos, entre otros^{55,57} y presentan una alta homología con los genes originales (89 a 95%)⁵⁵. Si bien los PPR se expresan en 16 tejidos humanos diferentes, *RPL6P10* solo se ha encontrado sobre expresado en el colon⁵⁵, mientras que *RPS23P3* no se encuentra particularmente expresado en ninguno de los tejidos estudiados hasta el momento⁵⁵.

A pesar de que se sabe que en los seres humanos solo un tercio de los PPR son procesados y por tanto codifican transcritos que regulan la función de otros genes⁵⁵, no se ha demostrado a ciencia cierta si esto ocurre en los pseudogenes de nuestra región en estudio, sin embargo, existe evidencia de que en las levaduras los ortólogos de *Rpl6p10* y *Rps23p3* no solo regulan la expresión génica, si no que presentan funciones similares al servir como promotor en los mismas

proteínas ribosomales (RPS9 y RPL10) asociadas a ventajas evolutivas⁵⁷.

En lo que respecta a la significancia clínica de las variantes en la secuencia de *RPL6P10* y *RPS23P3*, hasta el momento no se han podido asociar a fenotipos específicos, sin embargo han sido analizados en grandes estudios poblacionales donde se ha utilizado secuenciación masiva y se ha comenzado a indagar su posible relación con diferentes fenómenos y patologías⁵⁸.

Algunos de estos estudios de cohorte incluyen al subestudio del Framingham Heart Study^{59,60}: NHLBI Framingham SNP Health Association Resource (SHARe), STAMPEED: Cardiovascular Health Study (CHS) GWAS to identify genetic variants associated with aging and CVD risk factors and events (subestudio del Cardiovascular Health Study (CHS) Cohort)^{60,61}, Compilation of Aggregate Genomic Data for General Research Use⁶², el NHLBI Family Heart Study (FamHS-Visit1 and FamHS-Visit2)⁶³ y el NINDS-Genome-Wide Genotyping in Parkinson's Disease: First Stage Analysis and Public Release of Data.

Los resultados parciales de estos estudios (aun no concluidos), han demostrado la presencia de variantes de secuencia de *RPL6P10* asociadas a individuos con enfermedad coronaria, evento vascular cerebral, disminución de la densidad ósea, ganancia de peso, alteraciones en el fibrinógeno y en la actividad de la proteína quimio

atrayerente de monocitos tipo 1, estos últimos relacionándose a procesos inflamatorios, aunque en la mayoría de los casos se está estudiando el efecto fenotípico de las variantes (ver Tabla 4)⁶².

En estas mismas bases de datos es posible identificar variantes de secuencia de *RPS23P3* asociadas a diversas patologías⁶⁴, por ejemplo enfermedad de Parkinson y aumento en el riesgo para desarrollar esclerosis múltiple⁶⁵ o alteraciones en el examen neurológico⁶⁶ (ver Tabla 4). Sin embargo, algunos de estos polimorfismos se han asociado a las mismas patologías que aquellos descritos en *RPL6P10*.

Vale la pena hacer notar que no se ha descrito una asociación con angiogénesis, cardiogénesis o diferenciación arterio-venosa en ninguno de los dos pseudogenes presentes en la región de estudio.

Tabla 4. Asociaciones fenotípicas de *RPL6P10* y *RPS23P3* en estudios de cohorte.

| FENOTIPO | ESTUDIO | PSEUDOGEN |
|--|---------------|----------------------------------|
| Riesgo de enfermedad coronaria (diámetro de arterias disminuido, cambios en la relación del grosor de la íntima y la media). | NHLBI CAGD | <i>RPL6P10</i> <i>RPS23P3</i> |
| Actividad de la proteína quimioatrayente de monocitos tipo 1 (aumento). | NHLBI CAGD | <i>RPL6P10</i> <i>RPS23P3</i> |
| Variaciones en la medición de fibrinógeno. | NHLBI CAGD | <i>RPL6P10</i> <i>RPS23P3</i> |
| Disminución de la densidad ósea. | SHARe | <i>RPL6P10</i> <i>RPS23P3</i> |
| Ganancia de peso. | NHLBI CAGD | <i>RPL6P10</i> <i>RPS23P3</i> |
| | FHS SHARe | <i>RPS23P3</i> |
| Evento vascular cerebral. | STAMPEED | <i>RPL6P10</i> <i>RPS23P3</i> |
| Conteo plaquetario. | NHLBI CAGD | <i>RPS23P3</i> |
| Variaciones en el colesterol de alta densidad (HDL). | NHLBI CAGD | <i>RPS23P3</i> |

| | | |
|---|----------------|----------------|
| | SHARe FHS | |
| Enfermedad de Parkinson. | NHLBI NINDS | <i>RPS23P3</i> |
| Función respiratoria disminuida (espirometría). | SHARe. FHS | <i>RPS23P3</i> |
| Variaciones en la apolipoproteína B. | SHARe. FHS | <i>RPS23P3</i> |
| Riesgo para Esclerosis Múltiple. | FHS | <i>RPS23P3</i> |
| Alteraciones en el examen neurológico | FHS | <i>RPS23P3</i> |
| Viscosidad de la sangre (aumento). | SHARe FHS | <i>RPS23P3</i> |
| Hipertensión arterial. | SHARe FHS | <i>RPS23P3</i> |
| Albuminuria. | SHARe FHS | <i>RPS23P3</i> |
| Disminución de la Sobrevida en relación a riesgo cardiovascular. | SHARe FHS | <i>RPS23P3</i> |

Abreviaturas: NHLBI: NHLBI Family Heart Study (FamHS-Visit1 and FamHS-Visit2). CAGD: Compilation of Aggregate Genomic Data for General Research Use, SHARe: Framingham Heart Study, NHLBI Framingham SNP Health Association Resource (SHARe), FHAS: Framingham Heart Study .STAMPEED: STAMPEED Cardiovascular Health Study (CHS) GWAS to identify genetic variants associated with aging and CVD risk factors and events. NINDS: NINDS-Genome-Wide Genotyping in Parkinson's Disease: First Stage Analysis and Public Release of Data.

EPHA5-AS1 (EPHA5 antisense RNA 1).

Este es un RNA no codificante compuesto por 4 exones con una gran variabilidad (1536 polimorfismos conocidos), con significancia desconocida⁶⁷. Hasta el momento no se ha asociado con ningún fenotipo⁶⁸.

LOC728048. Interferon induced transmembrane protein pseudogene.

Se describió en el año 2012, durante la realización del estudio VIVA LA FAMILIA, realizado en Texas por la universidad de Baylor para identificar polimorfismos asociados a obesidad infantil en población hispana mediante GWAS⁶⁹, en donde el 1% de los 815 pacientes estudiados expresaban polimorfismos en el pseudogen⁷⁰, se consideró un locus de riesgo para obesidad relacionado con el incremento de la ingesta calórica.

No existen reportes que asocien este pseudogen a otras patologías que expliquen sus interacciones.

MICRO RNA 1269^a (miR1269A).

Los micro RNA (miRNAs), son RNAs no codificantes con una longitud de 19 a 24 nucleótidos, con una función asociada a regulación postraducciona l g é n i c a⁷¹, se han asociado a la patogénesis de múltiples enfermedades como diabetes, hipertensión y particularmente cáncer, al interactuar en los procesos de crecimiento

celular, diferenciación, apoptosis, angiogénesis y metástasis⁷². Estudios actuales buscan utilizar a los miRNAs como marcadores diagnósticos, pronósticos e incluso como blancos de tratamiento⁷¹.

La sobreexpresión de *miR1269A* se ha descrito asociada a la regulación negativa del gen *FOXO1* en el cáncer hepatocelular humano al permitir la proliferación celular, ya que este gen tiene una función importante como activador transcripcional de múltiples genes involucrados en la detención del ciclo celular, apoptosis, reparación del ADN y respuesta a hipoxia⁷¹. Otros estudios han podido relacionar la sobreexpresión de este microRNA con la progresión del cáncer colorectal humano al regular a la alta a TGF-beta cuando esta neoplasia presenta metástasis a otros órganos (estadios III y IV de la clasificación TNM)⁷³.

ÁCIDOS RIBONUCLEICOS LARGOS NO CODIFICANTES (lincRNA: Long non-coding RNA). RP11-807H7.2, RP11-51P8.1, RP11-25H12.1 y RP11-19C24.1.

Los lincRNAs son RNAs no codificantes con una longitud mayor a 200 nucleótidos, que participan en distintos procesos como la compensación de las dosis cromosómicas, regulación epigenética, control del ciclo celular, transcripción, traducción del genoma, y diferenciación celular, entre otras funciones⁷⁴. Fenotípicamente los cambios en la expresión de estos RNAs se han asociado con múltiples

patologías incluyendo cáncer de mama, hepatocelular, gástrico⁷⁴, falla ovárica prematura⁷⁵ y artritis reumatoide⁷⁶, entre otras.

Los pseudogenes de lincRNAs se derivan de parálogos codificantes ancestrales y al igual que el resto de los pseudogenes no son propiamente funcionales⁷⁷, sin embargo, en los últimos años se han comenzado a estudiar sus efectos en la regulación epigenética al generar transcritos capaces de modificar la expresión de los lincRNAs⁷⁸.

A pesar de haber sido descrita la asociación fenotípica (principalmente a cáncer) de algunos pseudogenes de lincRNA, los subtipos encontrados en la región aún no han demostrado significancia clínica⁷⁹.

PSEUDOGENES DE ACIDOS RIBONUCLEICOS PEQUEÑOS NUCLEARES (snRNA)(RNU2-40P Y RNU6-699P)

Los dos pseudogenes *RNU2-40P* (RNA, U2 small nuclear 40, pseudogene) y *RNU6-699P* (RNA, U6 small nuclear 699, pseudogene) encontrados en la región, corresponden a la forma procesada de pseudogenes, es decir, derivan por retrotranscripción de un gen ancestral identificable⁷⁷: La familia de pseudogenes U6 se caracteriza por interactuar con transposones (elementos móviles) tipo LINE 1 (L1), que de manera interesante derivan ancestralmente de los mismos pseudogenes^{80,81} y se identifican en todos los primates.

Esta estrecha relación ha sugerido el uso de estos pseudogenes como un marcadores para estas secuencias L1⁸².

Los pseudogenes de la familia U2 tienden a estar cercanos a los genes U2, participando directamente en su regulación e interaccionando con secuencias repetidas de genoma de tipo *Alu*⁸³. Esta relación con los repetidos *Alu* genera las diversas isoformas de esta familia⁸³.

Fenotípicamente hablando, no se han encontrado asociaciones con patologías.

SECUENCIAS NO CARCTERIZADAS. (LOC105377258, LOC105377259, LOC105377260, LOC105377261 LOC105377262).

Estas secuencias fueron descritas en estudios en donde se realizaron estudios de secuenciación de genoma completo⁸⁴, sin embargo no se ha identificado su biotipo o algún tipo de relación clínica⁵⁸.

DISCUSIÓN Y CONCLUSIONES.

Existen secuencias provenientes de un ancestro en común que segregan con la enfermedad relacionadas con la CATVP en el paciente afectado, con antecedente de consanguinidad por parte de los progenitores al demostrar un LOH alto en el locus 4q13.1, región homocigota de 1.635 Kb, flanqueada por los SNPs rs12644315 y rs10017714.

La técnica utilizada para determinar la región fue el análisis de pérdida de heterocigosidad mediante el uso de un microarreglo de 250k. Tomando en cuenta lo anterior, podemos interpretar los resultados basándonos en que se está identificando un segmento genómico homocigoto en este *locus*.

Se debe de considerar que los genes descritos anteriormente para la CATVP se encontraban dentro de las regiones analizadas a pesar de no haber sido sospechosas en la etiología debido a su forma de herencia y limitaciones étnicas y no presentaron LOH.

Se buscó de manera dirigida las secuencias contenidas dentro de la región, y fueron identificadas 17 secuencias con biotipos diferentes. Al realizar el análisis de la región se logró determinar que dichas secuencias incluyen: un gen codificante (*EPHA5*), tres pseudogenes (*RPL6P10*, *RPS23P3* y *LOC728048*), un RNA antisentido (*EPHA5-AS1*), cuatro RNA largos no codificantes (*RP11-807H7.2*, *RP11-51P8.1*, *RP11-25H12.1*, *RP11-19C24.1*), dos

pseudogenes de RNA pequeños nucleares, un microRNA (*miR1269A*) y cinco secuencias no caracterizadas.

Los resultados del estudio de apoyan fuertemente la noción de una herencia autosómica recesiva para la CATVP presente en esta familia, y sugieren la presencia de una variante causal dentro de este segmento. Así mismo, se realizó un análisis respecto a las funciones, interacciones y relación fenotipo-genotipo de cada una de las secuencias que se encontraban dentro de la región con LOH alto. Este análisis sugiere que la mutación causal no se encuentra dentro de secuencias no codificantes de proteína y reguladoras ya que ninguno de estos elementos se encontró asociado a cardiogénesis, angiogénesis o diferenciación arteriovenosa .

Como resultado de la revisión se determinó que *EPHA5* es la única secuencia que, con la información disponible, puede ser asociada a un fenotipo cardiovascular, al menos de forma indirecta, ya que su homólogo *EPHB2* se encuentra relacionado con angiogénesis y diferenciación arterio-venosa, destacándose una alta homología en el dominio de unión a ligando de ambos genes, permitiendo así que ambos se unan a los mismos ligandos. Adicionalmente se sabe que los miembros de esta familia de receptores presentan pleiotropismo, por lo que no se puede excluir la posibilidad de una nueva función para este gen a pesar de no haber sido reportada en la literatura previamente.

Lo expuesto anteriormente responde a nuestra pregunta de investigación al demostrar que existen secuencias asociadas a un ancestro común que se relacionan con la CATVP y se sugiere como gen candidato al gen *EPHA5*.

PERSPECTIVAS.

En una segunda fase de esta investigación se buscará identificar la variante causal dentro del gen candidato para demostrar su papel en la etiopatología de la CATVP y abre la posibilidad de identificar mutaciones en este mismo gen en otros pacientes con esta patología y explicar la alta prevalencia de la misma en México y latinoamerica. Tomando en cuenta que esta CC, tiene principalmente una presentación aparentemente esporádica en nuestro medio, pero que podrían corresponder a formas de herencia autosómico recesivo no identificados.

Si se encuentran alteraciones similares en la secuenciación de los pacientes esporádicos que en la familia se puede proponer un efecto fundador que explique la alta prevalencia en México e incluso también hacer estudios sobre proporción de portadores, potencialmente útil en consejo genético.

ANEXOS

ANEXO 1. Carta de aceptación por el comité de ética del Hospital Infantil de México “Federico Gómez”.

  
HOSPITAL INFANTIL DE MÉXICO FEDERICO GÓMEZ
INSTITUTO NACIONAL DE SALUD
DIRECCIÓN DE INVESTIGACIÓN

“2018, AÑO DEL TRIENALBIENIO JOSÉ MARÍA MORELOS Y PAVÓN”

México, D.F., a 10 de febrero de 2015 DG/1000/ 103 /2015

Dra. Rocío Sánchez Urbina
Laboratorio de Investigación en Biología del
Desarrollo y Teratogénesis Experimental
Presente

Informo a usted, que los Comités de Investigación, Ética y Bioseguridad, después de haber revisado su protocolo **HIM-2015-036** "Identificación de mutaciones para conexión anómala total de venas pulmonares variedad mendeliana, mediante mapeo de autocigosidad", han emitido el dictamen de:

APROBADO

En los términos y condiciones señalados por dichos Comités. Por lo anterior, se autoriza su desarrollo.

Atentamente


Dr. José Alberto García Aranda
Director General

Con copia:
C.P. Elías Hernández Ramírez, Jefe de la Unidad Auxiliar Administrativa de Proyectos de Investigación.
JAGA/OMH/JGO/bcg

 **INSTITUTO NACIONAL DE SALUD AFILIADO A LA UNAM**
DR. MÁRQUEZ 162, COL. DOCTORES, DEL CUAUHTÉMOC, C.P. 06720 MÉXICO D.F.
CONMUTADOR: 5228-9917 EXT. 4315 Y 4100
www.himfg.edu.mx

BIBLIOGRAFIA:

1. Attie F, Colmenero, Calderón J, Zabal Cerdeira C, Buendía Hernández A. *Cardiología Pediátrica*. 2da e. México: Médica Panamericana; 2013.
2. Básica I, Castellanos LM, Aurelia C, Vargas S, Nivon MK. www.medigraphic.com Estudio morfológico de la conexión anómala total de venas pulmonares. 2007;77(4):265-274.
3. Antonio Manzo-Ríos M, Ernesto Anaya-García J, Ruiz-Gastelum E, et al. Conexión Anómala Total de Venas Pulmonares Infracardiaca, Presentación de un Caso. *Bol Clin Hosp Infant Edo Son*. 2007;24(1):42-47.
4. Stein P. Total anomalous pulmonary venous connection. *AORN J*. 2007;85(3):509-20-4. doi:10.1016/S0001-2092(07)60123-9.
5. Evans WN, Acherman RJ, Ciccolo ML, Castillo WJ, Restrepo H. An increased incidence of total anomalous pulmonary venous connection among Hispanics in southern Nevada. *Congenit Heart Dis*. 2015;10(2):137-141. doi:10.1111/chd.12199.
6. Cervantes-Salazar J, Calderón-Colmenero J, Ramírez-

- Marroquín S, et al. El Registro Mexicano de Cirugía Cardíaca Pediátrica. Primer informe. *es i Rev Investig Clínica*. 2013;65(6):476-482.
7. Rosenthal N, Harvey R. *HEART DEVELOPMENT AND REGENERATION*. First edit. London: Elsevier; 2010.
 8. Carlson MB. *Embiología Humana Y Biología Del Desarrollo*. 2da ed. Madrid: Harcourt; 2000.
 9. Delisle G, Ando M, Calder AL, et al. Total anomalous pulmonary venous connection: Report of 93 autopsied cases with emphasis on diagnostic and surgical considerations. *Am Heart J*. 1976;91(1):99-122.
<http://www.ncbi.nlm.nih.gov/pubmed/1244724>. Accessed September 22, 2016.
 10. Woods CG, Cox J, Springell K, et al. Report Quantification of Homozygosity in Consanguineous Individuals with Autosomal Recessive Disease. *Am J Hum Genet Am J Hum Genet*. 2006;7878:889-896. www.ajhg.org. Accessed October 6, 2016.
 11. Seelow D, Schuelke M, Hildebrandt F, Nürnberg P. HomozygosityMapper - An interactive approach to homozygosity mapping. *Nucleic Acids Res*. 2009;37(SUPPL. 2):593-599. doi:10.1093/nar/gkp369.

12. Keller MC, Simonson MA, Ripke S, et al. Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. *PLoS Genet.* 2012;8(4):e1002656. doi:10.1371/journal.pgen.1002656.
13. Alkuraya FS. The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. *Hum Genet.* 2013;132(11):1197-1211. doi:10.1007/s00439-013-1344-x.
14. Burton BR. generation of CGH microarray. 2013;(July).
15. Alkuraya FS. Homozygosity mapping: One more tool in the clinical geneticist's toolbox. *Genet Med.* 2010;12:236-239. doi:10.1097/GIM.0b013e3181ceb95d.
16. Pickrell WO, Rees MI, Chung S-K. Next generation sequencing methodologies--an overview. *Adv Protein Chem Struct Biol.* 2012;89:1-26. doi:10.1016/B978-0-12-394287-6.00001-X.
17. Bleyl S, Nelson L, Odelberg SJ, et al. A gene for familial total anomalous pulmonary venous return maps to chromosome 4p13-q12. *Am J Hum Genet.* 1995;56(2):408-415. <http://www.ncbi.nlm.nih.gov/pubmed/7847375>. Accessed September 20, 2016.
18. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of

congenital heart disease: the glass half empty. *Circ Res*.
2013;112(4):707-720.
doi:10.1161/CIRCRESAHA.112.300853.

19. Bleyl S, Botto L, Carey J, et al. Analysis of a Scottish founder effect narrows the TAPVR-1 gene interval to chromosome 4q12. *Am J Med Genet Part A*. 2006;140(A):2368–2373.
doi:10.1002/ajmg.a.31471.
20. Cinquetti R, Badi I, Campione M, et al. Transcriptional deregulation and a missense mutation define ANKRD1 as a candidate gene for total anomalous pulmonary venous return. *Hum Mutat*. 2008;29(4):468-474. doi:10.1002/humu.20711.
21. Chung B, Shaffer LG, Keating S, Johnson J, Casey B, Chitayat D. From VACTERL-H to heterotaxy: Variable expressivity of ZIC3-related disorders. *Am J Med Genet Part A*. 2011;155(5):1123-1128. doi:10.1002/ajmg.a.33859.
22. Bleyl SB, Saijoh Y, Bax NAM, et al. Dysregulation of the PDGFRA gene causes inflow tract anomalies including TAPVR: integrating evidence from human genetics and model organisms. *Hum Mol Genet*. 2010;19(7):1286-1301.
doi:10.1093/hmg/ddq005.
23. Raisher B, Dowton S, Grant J. Father and two children with total anomalous pulmonary venous connection. *Am J Med*

- Genet.* 1991;40(1):105-106.
24. Bleyl S, Ruttenberg HD, Carey JC, Ward K. Familial total anomalous pulmonary venous return: a large Utah-Idaho family. *Am J Med Genet.* 1994;52(4):462-466.
doi:10.1002/ajmg.1320520412.
 25. Acquati F, Russo A, Taramelli R, Tibiletti MG, Taborelli M, Camesasca C et al. Nonsyndromic Total Anomalous Venous Return Associated With a De Novo Translocation Involving Chromosomes 10 and 21 t(10;21)(q23.1;q11.2). *Am J Med Genet.* 2000;95(3):285-286.
 26. Harris DL, Siu BL, Hummel M, et al. Mosaic ring 12p and total anomalous pulmonary venous return. *Am J Med Genet A.* 2004;131(1):91-93. doi:10.1002/ajmg.a.30315.
 27. Hasan Mottaghi Moghaddam Shahri a, Reza Jafarzadeh Esfehani a, Nader Yazdan Panah a, Ali, Jafarzadeh Esfehan. Consanguinity and isolated atrial septal defect in North East of Iran.
 28. Solymar L, Sabel KG, Zetterqvist P. Total anomalous pulmonary venous connection in siblings. Report on three families. *Acta Paediatr Scand.* 1987;76(1):124-127.
<http://www.ncbi.nlm.nih.gov/pubmed/3564987>. Accessed September 20, 2016.

29. Wei C, Daniel KB, Robert AS, David HP. Identification and Characterization of a Novel Cytokine-inducible Nuclear Protein from Human Endothelial Cells. *J Biol Chem*. 1995;270(17):10236-10245. doi:doi: 10.1074/jbc.270.17.10236.
30. Lesions MA, Waard V De, Achterberg TAE Van, Beauchamp NJ, Pannekoek H, Vries CJM De. Cardiac Ankyrin Repeat Protein (CARP) Expression in Activin Induces Carp in Smooth Muscle Cells. 2003:64-69. doi:10.1161/01.ATV.0000042218.13101.50.
31. Baxter EJ, Hochhaus A, Bolufer P, et al. leukaemia fuses BCR to PDGFRA. *Hum Mol Genet*. 2002;11(12):1391-1397. doi:10.1093/hmg/11.12.1391.
32. Aruga J, Nagai T, Tokuyama T, et al. The Mouse Zic Gene Family.
33. Gene Data Viewer. https://www.ncbi.nlm.nih.gov/genome/gdv/?context=genome&acc=GCF_000001405.34&chr=4.
34. No Title. Homozygosity Mapper.
35. No Title. Ensembl. http://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000145242;r=4:65319563-65670495. Accessed

June 30, 2016.

36. Fox GM, Holst PL, Chute HT, et al. cDNA cloning and tissue distribution of five human EPH-like receptor protein-tyrosine kinases. *Oncogene*. 1995;10(5):897-905.
<http://www.ncbi.nlm.nih.gov/pubmed/7898931>. Accessed October 9, 2016.
37. Brambilla R, Schnappl A, Casagrande F, et al. Membrane-bound LERK2 ligand can signal through three different Eph-related receptor tyrosine kinases. *EMBO J*. 1995;14(13):3116-3126.
38. No Title. Homozygosity Mapper. homozygositymapper.org. Published 2016. Accessed January 1, 2016.
39. Das G, Yu Q, Hui R, Reuhl K, Gale NW, Zhou R. EphA5 and EphA6: regulation of neuronal and spine morphology. *Cell Biosci*. 2016;6:48. doi:10.1186/s13578-016-0115-5.
40. Huan X, Shi J, Lim L, et al. Unique Structure and Dynamics of the EphA5 Ligand Binding Domain Mediate Its Binding Specificity as Revealed by X-ray Crystallography, NMR and MD Simulations. *Med Res Counc Singapore Natl Institutes Heal*. 2013. doi:10.1371/journal.pone.0074040.
41. Gale NW, Holland SJ, Valenzuela DM, et al. Eph Receptors

and Ligands Comprise Two Major Specificity Subclasses and Are Reciprocally Compartmentalized during Embryogenesis known family of RTKs with at least 13 distinct members family display dynamic and spatially restricted expres. *Neuron*. 1996;17:9-19.

42. Gitanjali Das, Qili Yu, Ryan Hui, Kenneth Reuhl NWG, and Renping Zhou. EphA5 and EphA6: regulation of neuronal and spine morphology. *Cell Biosci*. 2016;6(48). doi:10.1186/s13578-016-0115-5.
43. Szepietowska B, Horvath TL, Sherwin RS. Role of Synaptic Plasticity and EphA5-EphrinA5 Interaction Within the Ventromedial Hypothalamus in Response to Recurrent Hypoglycemia. *Diabetes*. 2014;63:1140-1147. doi:10.2337/db13-1259.
44. Cooper M, Crocket D, Nowakowski R, Gale N, Zhou R. Distribution of EphA5 receptor protein in the developing and adult mouse nervous system. *J Comp Neurol*. 2009;514(4):310-328. doi:10.1002/cne.22030.
45. Konstantinova I, Nikolova G, Ohara-Imaizumi M, et al. EphA-Ephrin-A-Mediated b Cell Communication Regulates Insulin Secretion from Pancreatic Islets. doi:10.1016/j.cell.2007.02.044.

46. Yamada T, Yoshii T, Yasuda H, Okawa A, Sotome S. Dexamethasone Regulates EphA5, a Potential Inhibitory Factor with Osteogenic Capability of Human Bone Marrow Stromal Cells. doi:10.1155/2016/1301608.
47. Staquicini FI, Qian MD, Salameh A, et al. Receptor tyrosine kinase EphA5 is a functional molecular target in human lung cancer. *J Biol Chem*. 2015;290(12):7345-7359. doi:10.1074/jbc.M114.630525.
48. Li S, Zhu Y, Ma C, et al. Downregulation of EphA5 by promoter methylation in human prostate cancer. *BMC Cancer*. 2015;15. doi:10.1186/s12885-015-1025-3.
49. Fu D-Y, Wang Z-M, Wang B-L, et al. Frequent epigenetic inactivation of the receptor tyrosine kinase EphA5 by promoter methylation in human breast cancer. *Hum Pathol*. 2010;41(1):48-58. doi:10.1016/j.humpath.2009.06.007.
50. Giaginis C, Tsourouflis G, Zizi-Serbetzoglou A, et al. Clinical Significance of Ephrin (Eph)-A1, -A2, -A4, -A5 and -A7 Receptors in Pancreatic Ductal Adenocarcinoma. *Pathol Oncol Res*. 2010;16(2):267-276. doi:10.1007/s12253-009-9221-6.
51. Gu S, Feng J, Jin Q, Wang W, Zhang S. Reduced expression of EphA5 is associated with lymph node metastasis, advanced TNM stage, and poor prognosis in colorectal carcinoma. *Histol*

Histopathol. September 2016:11815. doi:10.14670/HH-11-815.

52. Himanen J-P, Chumley MJ, Lackmann M, et al. Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling. *Nat Neurosci.* 2004;7(5):501-509. doi:10.1038/nn1237.
53. Sakamoto A, Kawashiri M, Ishibashi-Ueda H, et al. Clinical Study Expression and Function of Ephrin-B1 and Its Cognate Receptor EphB2 in Human Abdominal Aortic Aneurysm. *Int J Vasc Med.* 2012;7. doi:10.1155/2012/127149.
54. Katoh M, Katoh M. Comparative integromics on Eph family. *Int J Oncol.* May 2006. doi:10.3892/ijo.28.5.1243.
55. Tonner P, Srinivasasainagendra V, Zhang S, Zhi D. Detecting transcription of ribosomal protein pseudogenes in diverse human tissues from RNA-seq data. 2012. doi:10.1186/1471-2164-13-412.
56. Balasubramanian S, Zheng D, Liu Y-J, et al. Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biol.* 2009;10(1). doi:10.1186/gb-2009-10-1-r2.
57. Poidevin L, Andreeva K, Khachatoorian C, Judelson HS. Comparisons of Ribosomal Protein Gene Promoters Indicate

Superiority of Heterologous Regulatory Sequences for
Expressing Transgenes in *Phytophthora infestans*.
doi:10.1371/journal.pone.0145612.

58. phenotype association data PheGenI.
<https://www.ncbi.nlm.nih.gov/gap/phegeni?tab=1&gene=642828#Arteries>. Accessed September 1, 2016.
59. Framingham Heart Study.
https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v28.p10. Accessed September 1, 2016.
60. phenotype association data PheGenI.
61. STAMPEED: Cardiovascular Health Study (CHS) GWAS to identify genetic variants associated with aging and CVD risk factors and events.
https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000226.v5.p1. Accessed September 1, 2016.
62. Compilation of Aggregate Genomic Data for General Research Use. https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000501.v1.p1.
63. NHLBI Family Heart Study (FamHS-Visit1 and FamHS-

Visit2). https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000221.v1.p1.

64. PhenGenI_RPS23P3.
<https://www.ncbi.nlm.nih.gov/gap/phegeni?tab=1&gene=100271091>.
65. Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet.* 2009;18(4):767-778. doi:10.1093/hmg/ddn388.
66. Seshadri S, Destefano AL, Au R, et al. Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham study. *BMC Med Genet* . 2007;8(suppl 1)(15). doi:10.1186/1471-2350-8-S1-S15.
67. GENE. EPHA5-AS1 EPHA5 antisense RNA 1 [Homo sapiens (human)] Gene ID: 100144602.
<https://www.ncbi.nlm.nih.gov/gene/100144602>.
68. GenPhenI. EPHA5-AS1.
<https://www.ncbi.nlm.nih.gov/gap/phegeni?tab=1&gene=100144602>.
69. Comuzzie AG, Cole SA, Laston SL, et al. Novel Genetic Loci

Identified for the Pathophysiology of Childhood Obesity in the Hispanic Population. *PLoS One*. 2012;7(12).
doi:10.1371/journal.pone.0051954.

70. Hindorff LA, MacArthur J (European Bioinformatics Institute), Morales J (European Bioinformatics Institute), Junkins HA, Hall PN, Klemm AK and MT. Catalog of Published Genome-Wide Association Studies LOC728048.
<https://www.genome.gov/gwasstudysnps.cfm?id=7195>.
Published 2016. Accessed October 10, 2016.
71. Shah MY, Ferrajoli A, Sood AK, Lopez-Berestein G, Calin GA. microRNA Therapeutics in Cancer — An Emerging Concept. *EBioMedicine*. 2016.
doi:10.1016/j.ebiom.2016.09.017.
72. Yang X-W, Shen G-Z, Cao L-Q, et al. MicroRNA-1269 promotes proliferation in human hepatocellular carcinoma via downregulation of FOXO1. doi:10.1186/1471-2407-14-909.
73. Bu P, Wang L, Chen K-Y, et al. miR-1269 Promotes Metastasis and Forms a Positive Feedback Loop with TGF- β HHS Public Access. 2015. doi:10.1038/ncomms7879.
74. Jiang C, Li X, Zhao H, Liu H. Long non-coding RNAs: potential new biomarkers for predicting tumor invasion and metastasis. *Mol Cancer*. 2016;15(62). doi:10.1186/s12943-016-

0545-z.

75. Xiong Y, Liu T, Wang S, Chi H, Chen C, Zheng J. Cyclophosphamide promotes the proliferation inhibition of mouse ovarian granulosa cells and premature ovarian failure by activating the lncRNA-Meg3-p53-p66Shc pathway. *Gene*. October 2016. doi:10.1016/j.gene.2016.10.011.
76. Zhang Y, Xu Y-Z, Sun N, et al. Long noncoding RNA expression profile in fibroblast-like synoviocytes from patients with rheumatoid arthritis. *Arthritis Res Ther*. 2016;18(1):227. doi:10.1186/s13075-016-1129-4.
77. Pei B, Sisu C, Frankish A, et al. The GENCODE pseudogene resource. *Genome* . 2012;13(R51). doi:10.1186/gb-2012-13-9-r51.
78. Milligan MJ, Lipovich L, Ørom UA, Alvarez-Dominguez JR. Pseudogene-derived lncRNAs: emerging regulators of gene expression. *Front Genet*. 2015;5(476). doi:10.3389/fgene.2014.00476.
79. GENCODE. www.encode.org. Accessed October 1, 2016.
80. Theissen H, Rinke J, Traver CN, Lihrmann R, Appel B. Novel structure of a human U6 snRNA pseudogene. *Gene*. 1985;36:195-199.

81. Hasnaoui M, Doucet AJ, Meziane O, Gilbert N. Ancient repeat sequence derived from U6 snRNA in primate genomes. *Gene*. 2009;448(2):139-144. doi:10.1016/j.gene.2009.07.015.
82. Elien A, Doucet J, Etan Droc G, Siol O, Er^ Ome Audoux J, Gilbert N. U6 snRNA Pseudogenes: Markers of Retrotransposition Dynamics in Mammals. doi:10.1093/molbev/msv062.
83. Hammarström K, Westin G, Bark C, Zabielski J, Petterson U. Genes and pseudogenes for human U2 RNA. Implications for the mechanism of pseudogene formation. *J Mol Biol*. 1984;179(2):157-169. doi:10.1016/0022-2836(84)90463-7.
84. Gene. <https://www.ncbi.nlm.nih.gov/gene>. Accessed October 1, 2016.