

CARACTERIZACIÓN DE VARIANTES GENÓMICAS ASOCIADAS A CARDIOPATÍA CONGÉNITA EN PACIENTES DE LA CIUDAD DE CALI

Angie Lizeth Grueso Cerón

Universidad del Valle, Apartado Aereo 25360, Cali, Colombia.

correo electrónico: angie.grueso@correounivalle.edu.co

José María Satizábal Soto

Universidad del Valle, Apartado Aereo 25360, Cali, Colombia.

correo electrónico: jose.satizabal@correounivalle.edu.co

Daniela Arturo Terranova

Universidad del Valle, Apartado Aereo 25360, Cali, Colombia.

correo electrónico: daniela.arturo@correounivalle.edu.co

RESUMEN

Las cardiopatías congénitas, corresponden a errores durante la embriogénesis generando malformaciones estructurales y funcionales. Son las malformaciones congénitas más prevalentes y causantes de la mayor morbimortalidad infantil, 8% son debidas a variantes en genes asociados al desarrollo cardíaco. Para establecer la frecuencia poblacional de variantes genómicas causantes de cardiopatías congénitas, se realizó revisión de alcance de genes prevalentes, se utilizaron resultados de secuenciación exómica completa de 320 pacientes sin sospecha de cardiopatías congénitas, se realizó caracterización, clasificación de significancia clínica, y redes de interacción génica. La revisión de alcance permitió determinar que los genes *NKX2-5*, *TBX20*, *GATA4*, *NOTCH1*, *PTPN11* son los más prevalentes, las variantes de mayor frecuencia alélica fueron c.63A>G (0,2844), c.39T>C (0,3406), c.1132A>G (0,0406), c.1669+9T>C (0,3531) y c.854-30T>C (0,0875) respectivamente; se reclasificaron clínicamente las variantes del gen *NKX2-5* c.335-311_335-303del de benigna a probablemente patogénica y del gen *PTPN11*, c.2354-24C>T y c.854-30T>C de benignas a patogénicas. Se identificaron 17% de variantes intrónicas y 4.8% de variantes missense; las redes de interacción permitieron determinar asociaciones entre moléculas pequeñas y genes cercanos con los que se comparten funciones del desarrollo cardíaco. Este trabajo contribuye al conocimiento sobre la frecuencia con la que las variantes genómicas asociadas a cardiopatías congénitas están presentes en la población, generando una herramienta para un diagnóstico temprano, un tratamiento precoz, disminuyendo así la morbimortalidad, con miras a un futuro tamizaje neonatal molecular universal de cardiopatías congénitas.

PALABRAS CLAVE: Bioinformática, Cardiopatía congénita, Genes, Modelos *In silico*, Secuenciación del Exoma

ABSTRACT

Congenital heart disease corresponds to errors during embryogenesis presenting structural and functional malformations. They are the most prevalent congenital malformations and cause the highest infant morbidity and mortality, 8% are due to variants in genes associated with cardiac development. To establish the population frequency of genomic variants that cause congenital heart disease, a scope review of prevalent genes was carried out, complete exome sequencing results of 320 patients without suspicion of congenital heart disease were used; characterization, clinical significance classification, and gene interaction networks were performed. The scope review allowed to determine that the genes *NKX2-5*, *TBX20*, *GATA4*, *NOTCH1*, *PTPN11* are the most prevalent, the ones with the highest allelic frequency were c.63A>G (0.2844), c.39T>C (0.3406), c.1132A>G (0.0406), c.1669+9T>C (0.3531) and c.854-30T>C (0.0875) respectively; The variants of the *NKX2-5* gene c.335-311_335-303del were clinically reclassified from benign to probably pathogenic and of the *PTPN11* gene c.2354-24C>T and c.854-30T>C from benign to pathogenic. 17% of intronic variants and 4.8% of missense variants were identified; the interaction networks made it possible to determine associations between small molecules and nearby genes with which cardiac development functions are shared. This work contributes to knowledge about the frequency with which genomic variants associated with congenital heart disease are present in the population, discovering a tool for early diagnosis, early treatment, thus reducing morbidity and mortality, with a view to future universal molecular neonatal screening. of congenital heart disease.

KEY WORDS: Bioinformatics, Congenital heart disease, Exome sequencing, Genes, In silico models,

INTRODUCCIÓN

El corazón es el primer órgano en formarse en el primer trimestre de la gestación humana; es un proceso complejo que incluye diferenciación, proliferación, migración celular, tabicación, y conducción de impulsos eléctricos. Todos controlados por factores de transcripción, cofactores, moléculas de señalización celular, entre otros. (Kobayashi *et al.*, 2014; Van Weerd y Christoffels, 2016). Los eventos celulares y de señalización son determinantes para el desarrollo normal del corazón, por lo tanto, variantes genómicas en los genes codificantes pueden provocar diferentes cardiopatías congénitas (CC) (Fan *et al.*, 2021).

Las CC son malformaciones estructurales o funcionales del corazón, o grandes vasos, presentes al nacer debidas a errores durante la embriogénesis cardiaca. (Ibáñez *et al.*, 2021; Moreno-Giraldo *et al.*, 2019; Valentín, 2018). Afectan aproximadamente al 1% de los nacimientos anuales (Williams *et al.*, 2019; Meller *et al.*, 2020), siendo las malformaciones congénitas más prevalentes y causantes de la mayor morbilidad infantil (Forero-Manzano *et al.*, 2020). La prevalencia mundial de las CC se estimó del 0.5 al 9 por cada 1.000 nacidos vivos (NV); en Colombia se estimó en 2.42/1000 nacimientos en 2017 (Ibáñez *et al.*, 2021). En el año 2023 el Instituto Nacional de Salud (INS), en el reporte epidemiológico III, registró un total de 2.394 casos de defectos congénitos, en los que las CC se incluyeron dentro de las malformaciones congénitas del sistema circulatorio, siendo las más prevalentes con 39.8/10.000 NV. (INS, 2023). Las principales causas de CC comprenden una etiología poligénica multifactorial que incluyen, trastornos cromosómicos, variantes genómicas puntuales y teratógenos ambientales. (Chung y Rajakumar, 2016).

Los avances en las tecnologías de secuenciación genómica y de la bioinformática han permitido identificar variantes de genes asociadas a CC y realizar un análisis rápido de información genética, que aporta al reconocimiento de secuencias de genes, análisis de expresión génica, organización y relación de información biológica, ayudando a procesar y dilucidar la etiología molecular de estas patologías. (Mullen *et al.*, 2021) (Araújo *et al.*, 2008).

Dentro de la bioinformática, el uso de las herramientas *in-silico* permite la clasificación clínica de las variantes genómicas y su correlación en cuanto a estructura y función proteica, mediante el estudio de cambios en las propiedades físico-químicas, en aminoácidos sustituidos o afectados. (Arturo-Terranova, 2021) Por lo tanto, este trabajo busca establecer la frecuencia alélica poblacional de variantes genómicas que estén asociadas a CC, haciendo uso de técnicas

bioinformáticas y herramientas *in-silico*, para generar una herramienta para un diagnóstico temprano, un tratamiento precoz, disminuyendo así la morbimortalidad, con miras a un futuro tamizaje neonatal molecular universal de cardiopatías congénitas.

METODOLOGIA

TIPO DE ESTUDIO

Estudio transversal, descriptivo, no experimental, donde se realizó revisión de alcance, caracterización genómica de las variantes, cálculo de frecuencia alélica y redes de interacción génica.

REVISIÓN DE ALCANCE

La revisión de alcance se realizó utilizando las bases de datos: Pubmed, Scopus y Lilacs de artículos publicados entre 2010 y 2023, en idioma inglés y español de los genes comúnmente relacionados a CC.

La ecuación de búsqueda que se utilizó fue la siguiente:

((gene) AND (congenital heart disease)) AND (prevalence)

CRITERIOS DE INCLUSIÓN

Paciente con enfermedad compleja, sin sospecha clínica de cardiopatía congénita con resultados de secuenciación de exoma completo obtenido de la base de datos del Instituto de Genética Médica - GENOMICS (Cali).

CRITERIOS DE EXCLUSIÓN

Pacientes con diagnóstico molecular confirmado de cardiopatías congénitas.

DATOS DE ESTUDIO

Se utilizaron los resultados de secuenciación de exoma completo de 320 pacientes sin sospecha de cardiopatías congénitas, y con enfermedad compleja, pertenecientes a la base de datos del Instituto de Genética Médica - GENOMICS (Cali), Utilizando el Editor de VCF 1.0 creado en la tesis doctoral “Representación De La Variabilidad Genómica Del Complejo Mucopolisacaridosis En El Suroccidente Colombiano”, el cual permitió editar el formato VCF de los exomas, y convertirlo en formato Excel, con el objetivo de aplicar los diferentes filtros de búsqueda y obtener un archivo final con la información de interés, visualizando columnas organizadas según cromosoma, posición, aminoácido alterado y registro en Clinvar. (Moreno-Giraldo *et al.*, 2020)

ANÁLISIS BIOINFORMÁTICO

Las variantes encontradas en los genes asociados a CC en la base de datos, se analizaron mediante búsqueda en bases de datos poblacionales, de genomas, de enfermedades y de secuencias con el fin de comprobar si han sido previamente descritas. Sus implicaciones patológicas se determinaron a través de herramientas bioinformáticas, mediante tecnología *in-silico*, que permite predecir las consecuencias de las variantes genéticas en la estructura y función proteicas, además de estudiar características como cambios en propiedades físico-químicas de aminoácidos sustituidos, grado de conservación evolutiva, el entorno de la secuencia de un aminoácido afectado o la alteración en las propiedades estructurales de la proteína.

BASES DE DATOS POBLACIONALES

Proyecto 1000 genomas (<https://www.internationalgenome.org/>)

El proyecto 1000 Genomas (1000GP) describe un catálogo de 84,7 millones de variantes que amplía en un 40 % el número de variantes conocidas del genoma humano. Presenta un acceso gratuito en la web y permite una búsqueda amigable de las variantes reportadas hasta el momento en la población (1000 Genomes Project Consortium, 2015).

Exome Aggregation Consortium (Exac) (<http://exac.broadinstitute.org/>)

Plataforma que busca agregar y organizar los datos de secuenciación de exoma de una variedad de proyectos de secuenciación a gran escala, y poner a disposición de la comunidad científica más amplia datos resumidos. Los datos de este sitio web abarca 60.706 individuos no relacionados secuenciados como parte de varios estudios genéticos específicos de población y enfermedades. (Iek *et al.*, 2016).

Genome Aggregation Database (gnomAD) (<https://gnomad.broadinstitute.org/>)

Es un recurso desarrollado por una coalición internacional de investigadores, con el objetivo de agregar y armonizar los datos de secuenciación del genoma y del exoma de una amplia variedad de proyectos de secuenciación a gran escala, y hacer que los datos resumidos estén disponibles para un público más amplio. El conjunto de datos v2.1.1 (GRCh37/hg19) proporcionado en este sitio web abarca 125 748 secuencias del exoma y 15 708 secuencias del genoma completo de individuos no relacionados secuenciados como parte de varios estudios genéticos de población y específicos de enfermedades. El conjunto de datos v3.1.2 (GRCh38) abarca 76 156 genomas de diversas ascendencias, seleccionados como en v2. (Karczewski *et al.*, 2020)

BASES DE DATOS DE GENOMAS

Ensembl Genome Browser (<http://www.ensembl.org/index.html>).

El proyecto comenzó en 1999 con el fin último de anotar automáticamente el genoma, integrar dicha anotación con otros datos biológicos disponibles, y publicar toda esta información en línea y de manera gratuita. Todas las anotaciones se integran con una enorme cantidad de fuentes externas de referencia, lo que convierte a Ensembl como un recurso integrador único. (Stalker, *et al.*, 2004)

BASES DE DATOS DE ENFERMEDADES

Online Mendelian Inheritance in Man (OMIM) (<https://www.omim.org/>)

Es un compendio completo y autorizado de genes humanos y fenotipos genéticos que está disponible y actualizado diariamente. Las descripciones generales de texto completo referenciadas en OMIM contienen información sobre todos los trastornos mendelianos conocidos y más de 15.000 genes. OMIM se centra en la relación entre fenotipo y genotipo. (Hamosh *et al.*, 2000).

Clinvar (<https://www.ncbi.nlm.nih.gov/clinvar/>)

ClinVar es un archivo público de acceso gratuito de informes de las relaciones entre las variaciones y fenotipos humanos; facilita el acceso y la comunicación sobre las relaciones afirmadas entre la variación humana y el estado de salud observado, procesa las presentaciones que informan variantes encontradas en muestras de pacientes, afirmaciones hechas con respecto a su importancia clínica, información sobre el remitente y otros datos de respaldo. (Landrum *et al.*, 2016).

Varsome (<https://varsome.com/>)

Varsome es un motor de búsqueda de variaciones genómicas humanas que incluye información de 30 bases de datos en una ubicación central. Varsome consta de más de 33 mil millones de puntos de datos que describen más de 500 millones de variantes. (Kopanov *et al.*, 2019).

BASES DE DATOS DE SECUENCIAS

National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/>).

El NCBI es la plataforma pública estadounidense que funciona como referente mundial en información biomédica y genómica. Mediante la búsqueda en dbSNP (repositorio central tanto para sustituciones como pequeñas eliminaciones e inserciones de bases de nucleótidos únicos) se

obtiene información sobre el posible efecto fisiopatológico de la variante y de su frecuencia en población europea (a través de la información de Haplotype Map; HapMap) (Sayers *et al.*, 2020).

Métodos *In Silico*

Por medio de estas herramientas bioinformáticas, es posible predecir las consecuencias de las variantes genéticas en la estructura y función proteicas.

Polymorphism Phenotyping v2 (Polyphen-2) (<http://genetics.bwh.harvard.edu/pph/>)

Polyphen-2 predice el impacto que puede causar una sustitución de un aminoácido en la estructura y/o función de una proteína humana mediante consideraciones comparativas. Dada una sustitución de un aminoácido, PolyPhen-2 extrae varias características relacionadas con la secuencia y la estructura de donde se ha producido la sustitución e introduce estos datos en un clasificador probabilístico para obtener el grado de perjuicio que supone. Clasifica las variantes en benigna, probablemente patogénica y patogénica, basándose en la probabilidad de patogenicidad dada por el clasificador (Adzhubei *et al.*, 2013)

UMD PREDICTOR (<http://umd-predictor.eu/analysis.php>)

UMD predictor es una plataforma independiente que contiene todas las predicciones de todas las sustituciones de cualquier transcripción humana. Esta herramienta proporciona un enfoque combinatorio, para identificar posibles variaciones patógenas, que asocia los siguientes datos: localización dentro de la proteína, conservación, propiedades bioquímicas de los residuos mutantes y de tipo salvaje, y el impacto potencial de la variación en el ARNm. (Salgado *et al.*, 2016)

Algoritmo SIFT (Sorting Intolerant From Tolerant) (<http://sift-dna.org>)

SIFT clasifica las sustituciones de aminoácidos en una proteína dada y predice si estos cambios provocarán un efecto fenotípico en la misma. SIFT se basa en la premisa de que los aminoácidos importantes de una proteína están conservados en la evolución, por lo que cambios en los mismos deben afectar a la funcionalidad de la proteína. Las sustituciones en una posición conservada en el alineamiento serán consideradas como “no tolerables” para la mayoría de los cambios, mientras que las posiciones que no están conservadas en el alineamiento tolerarán mejor los cambios de aminoácido. Las sustituciones con una puntuación menor de 0,05 serán clasificadas como deletéreas, y las de puntuación mayor o igual a 0,05 serán tolerables o neutrales (Kumar *et al.*, 2009)

Human Splicing Finder (HSF) (<http://umd.be/Redirect.html>)

Es una herramienta para predecir los efectos de las mutaciones en las señales de empalme o para identificar motivos de empalme en cualquier secuencia humana. Tiene nuevas matrices de peso de posición para evaluar la resistencia de los sitios de empalme de 5' y 3' y los puntos de ramificación. (Desmet *et al.*, 2009)

CADD.- Combined Annotation Dependent Depletion (<https://cadd.gs.washington.edu/>)

Es una herramienta para calificar el carácter nocivo de las variantes de un solo nucleótido, así como las variantes de inserción/delección en el genoma humano. Los puntajes C se correlacionan fuertemente con la diversidad alélica, la patogenicidad de las variantes codificantes y no codificantes, y los efectos reguladores medidos experimentalmente. Las puntuaciones de CADD-SV en la escala PHRED varían de 0 (potencialmente benigno) a 48 (potencialmente patógeno), lo que indica la posición de la nueva variante dentro de la distribución de puntuación de gnomAD-SV. Los puntajes CADD están disponibles gratuitamente para todas las aplicaciones no comerciales. (Rentzsch *et al.*, 2019).

Clasificación de variantes de acuerdo con American College of Medical Genetics (ACMG)

El colegio americano de genética médica (ACMG por sus siglas en inglés) ha instaurado un sistema de clasificación dependiendo de las investigaciones reportadas en diferentes estudios, análisis familiares, *in-silico* y predicciones sistemáticas que permitan un acuerdo de interpretación en variantes (Richards *et al.*, 2015).

Dicho sistema del ACMG se ha clasificado para obtener información sobre la significancia clínica de la siguiente manera:

- **Patogénica:** variantes que tienen evidencia fuerte de asociación con enfermedad.
- **Probablemente Patogénica:** variantes probablemente implicadas con enfermedad, ya que el cambio nucleotídico presenta una probabilidad de más del 90% de ser patogénica.
- **De significado incierto (VOUS):** variantes con posibles cambios funcionales, pero con evidencia contradictoria o insuficiente. Es decir, no se puede definir si es patogénica o no.
- **Probablemente benigna:** variantes con evidencia que sugiere benignidad, pero con datos débiles en la literatura que no descartan impacto biológico y eventualmente clínico. El cambio nucleotídico presenta una probabilidad de más del 90% de no ser patogénica.
- **Benigna:** variantes genéticas que no alteran funcionalidad.

CALCULO DE FRECUENCIAS ALÉLICAS

La frecuencia alélica (FA) de cada una de las variantes de los genes se calculó por cuenta simple:

$$FA: \frac{\text{Numero de copias del alelo en la población}}{\text{Numero total de copias del gen en la población}}$$

RED DE INTERACCIÓN DE GENES ASOCIADOS A NKX2-5, GATA4, TBX20, NOTCH1, PTPN11

GeneMANIA es un sitio web flexible y fácil de usar para generar hipótesis sobre la función de genes, predicción de su función. Dado un solo gen de consulta, GeneMANIA encuentra genes que probablemente compartan funciones en función de sus interacciones, tiene en cuenta solamente la evidencia encontrada de experimentos, bases de datos proceso biológico, función molecular, ruta o dominio de la proteína que se veía alterada y co-expresión con un nivel de confianza de 0.900 para determinar interacción entre genes cercanos.

RESULTADOS

En la revisión de alcance se encontró que los genes prevalentemente asociados a CC son:

NKX2-5: ubicado en el cromosoma 5q35.1 consta de dos exones, es altamente conservado entre las especies, y uno de los primeros marcadores necesarios para el proceso normal de formación cardíaca. Codifica para una proteína de 324 aminoácidos (NKX2-5), un factor de transcripción homeobox de clase CK, su región codificante presenta un dominio Tinman (TN), un homeodominio (HD) y un dominio específico NK (NK2SD). Juega un papel importante de regulación de proliferación, diferenciación y propiedades electrofisiológicas de células cardíacas (Wang *et al.*, 2019). Las variantes patogénicas en NKX2-5 asociadas a CC incluyen: defectos septales, tetralogía de Fallot, defectos del tracto de salida, hipoplasia del ventrículo izquierdo y defectos del sistema de conducción cardíaca (Monroy *et al.*, 2013).

TBX20: Ubicado en el cromosoma 7p14.2 contiene nueve exones. Codifica el factor de transcripción TBX20 de la familia T-box, subfamilia TBX1, uno de los primeros factores de transcripción que se expresan en el desarrollo cardíaco, en el primer y segundo campo cardiogénico, endocardio, miocardio, almohadillas endocárdicas, válvulas cardíacas y tabique auriculoventricular, actuando principalmente en la proliferación de cardiomiocitos. (Chen *et al.*,

2021). Las variantes patogénicas en *TBX20* asociadas a CC incluyen: defectos en la tabicación, la valvulogénesis y cardiomiopatías (García-flores *et al.*, 2023)

GATA4: ubicado en el cromosoma 8p23.1, contiene siete exones. Codifica la proteína GATA4 miembro de los factores de transcripción dedos de zinc, con dos dominios de transactivación y dos dominios de dedos de zinc clase IV (N-terminal y C-terminal), que se expresan en tejido cardiaco, funcionando como regulador crítico de la diferenciación cardiaca. Actúa en procesos de proliferación, especificación y diferenciación celular, y morfogénesis cardiaca, está presente en la formación del tubo cardiaco, separación y desarrollo del tracto de salida, formación del CAV y desarrollo del sistema de conducción cardiaco (Fan *et al.*, 2021). Las variantes patogénicas en *GATA4* asociadas a CC incluyen: defectos septales, tetralogía de Fallot, fibrilación auricular, estenosis pulmonar (McCulley y Black, 2012)

NOTCH1: ubicado en el cromosoma 9q34.3, contiene 34 exones, codifica para la proteína Notch1 y es miembro de la familia de receptores NOTCH, consta de un dominio extracelular encargado de la interacción con los ligandos y un dominio intracelular responsable de la transducción de la señal; los receptores NOTCH se expresan en estadios tempranos del desarrollo, relacionados con desarrollo y proliferación de las células cardiacas (Ye *et al.*, 2023; Wang *et al.*, 2019). Las variantes patogénicas en *NOTCH1* asociadas a CC incluyen enfermedad de la válvula aórtica calcificada, válvula aórtica bicúspide (BAV), y tetralogía de Fallot (Luxán *et al.*, 2016).

PTPN11: ubicado en el cromosoma 12q24.13, contiene 16 exones y codifica para una proteína tirosina fosfatasa SHP2, miembro de la familia PTP, importante para mantener la homeostasis del ambiente interno, está relacionada con diferenciación de células progenitoras cardiacas, desarrollo de válvulas y almohadillas endocárdicas, maduración y separación de la cavidad cardiaca. *PTPN11* promueve la activación de la vía proteína quinasa activada por RAS mitocondrial (MAPK) y es movilizada directamente por InsR (Receptor de insulina) para mediar su efecto mitógeno, una vez fosforilada inicia cascadas de señalización importante para la diferenciación, proliferación, migración, supervivencia de células cardiacas. (Xu *et al.*, 2022; Saint-Laurent *et al.*, 2022) Este gen está asociado principalmente con síndrome de Noonan, caracterizado por CC que incluyen estenosis de la válvula pulmonar y defectos septales (Xu *et al.*, 2022).

VARIANTES GENÓMICAS ENCONTRADAS:

En los resultados de los 320 exomas completos analizados se encontró:

En el gen **NKX2-5**, 38 variantes, siete reportadas en bases de datos de enfermedades: tres sinónimas, tres intrónicas y una missense (Fig. 1). Mediante herramientas *in silico*, seis variantes fueron clasificadas como benignas y una (la variante c.335-311_335-303del (rs200467566)) fue reclasificada de benigna a probablemente patógena (Fig. 2) ([Apéndice](#)).

Las tres variantes con mayor FA fueron, c.63A>G (rs2277923) (0.2844), c.335-311A>T (rs3095871) (0,2688) y c.335-315A>T (rs55880439) (0,2500), con valores cercanos a las bases de datos poblacionales (tabla1).

La red de interacción de genes para **NKX2-5** presentó interacciones físicas (rosa), y predicciones de asociación (naranja), y asociación principalmente con **GATA4** seguido de **TBX5**, **SMAD4**, **HAND1**, y **MEF2C**. (Fig. 3A).

En el gen **TBX20**, 16 variantes, seis reportadas en bases de datos de enfermedades: dos sinónimas, cuatro intrónicas (Fig. 1). Mediante herramientas *in silico* se clasificaron las seis variantes como benignas (Fig. 2).

Las tres variantes con mayor FA fueron, c.39T>C (rs336283) (0,3406), c.545+13A>G (rs17675148) y c.655-44G>A (rs2072434) (0,2563), y c.381-40A>G (rs148673267) (0,0125), los valores fueron cercanos a los reportados en las bases de datos poblacionales, a excepción de c.381-40A>G que presentó una frecuencia más elevada en este estudio (tabla 1).

La red de interacción de genes para **TBX20** mostró interacciones de dominios compartidos de proteínas (amarillo). Presentó asociación con **TBX5**, **TBX22**, **TBX21**, **EOMES**, **TBX6**, **TBR1** y **TBX3** (Fig. 3B).

En el gen **GATA4**, 27 variantes, seis reportadas en bases de datos de enfermedades: tres sinónimas, dos missense y una intrónica (Fig. 1). Mediante herramientas *in silico* se clasificaron las seis variantes como benignas (Fig. 2).

Las tres variantes con mayor FA fueron c.1132A>G (0.0406) (rs3729856), c.1000+23A>T (rs76808439) (0,0281), y c.1059C>T (rs3729855) (0,02188), con valores cercanos a los reportados en las bases de datos poblacionales (tabla 1).

La red de interacción de genes mostró interacciones físicas (rosa), y predicciones de asociación (naranja); obtuvo asociaciones con **NKX2-5**, **HEY2**, **ZFPM2**, **ZFPM1** y **TBX5** (Fig. 3C).

En el gen **NOTCH1**, 219 variantes, 82 reportadas en bases de datos de enfermedades: 40 intrónicas, 28 sinónimas, 13 missense y 1 UTR (región no traducida) (Fig. 1). Mediante herramientas *in silico* se clasificaron 67 variantes como benignas, 14 se reclasificaron de benignas a VOUS (c.3171+42G>A, c.3325+26G>A, c.5167+189C>A, c.2604C>T, c.663C>T, c.7233A >G, c.3511-10G>A, c.5019-13A>G, c.2587+20G>A, c.2207+10G>A, c.6555C>T, c.3325+21A>G, c.*6G>A, c.5384+34G>A), y una fue reclasificada de benigna a patógena (c.2354-24C>T, rs201516331) (Fig. 2)

La red de interacción génica mostró interacciones físicas (rosa), y asociación con *JAG1*, *RBPJ*, *MAML1* y *NOTCH2* (Fig. 3D).

El gen **PTPN11** 36 variantes, diez reportadas en bases de datos de enfermedades: nueve intrónicas y una sinónima (Fig. 1). Mediante herramientas *in silico* se clasificaron siete variantes benignas, dos se reclasificaron de benigno a VOUS (c.756+1274G>A, c.526-33_526-31del) y una se reclasificó de benigno a patológico (c.854-30T>C, rs144391508) (Fig. 2).

Las tres variantes con mayor FA fueron c.854-30T>C (rs144391508) (0.0875), c.756+1274G>A (rs11066315) (0.0781) y c.854-21C>T (rs41279090) (0.028), cuyos dos últimos valores fueron cercanos a los reportados en las bases de datos poblacionales, mientras que, el primero obtuvo una frecuencia más alta en este estudio (tabla 1).

La red de interacción de genes mostró interacciones físicas (rosa) e interacciones de rutas (azul claro). Según a red este gen solo participa en el desarrollo cardíaco, presentando asociación con el gen *FRS2*, *SPRY1*, *PDGFRB* y *INSR* (Fig. 3E).

La red de interacción entre los cinco genes mostró que los genes *NKX2-5* y *NOTCH1* estaban presentes en todas las funciones cardíacas seleccionadas, seguidos de *GATA4*, *TBX20* y *PTPN11*. (Figura 4)

DISCUSIÓN

Eventos celulares y de señalización genética son determinantes para el desarrollo cardíaco, ya que es un proceso complejo y cualquier variante patológica en estos genes pueden provocar distintas cardiopatías congénitas.

Variantes en el gen *NKX2-5* representan el 4% de las causas de CC (Choquet *et al.*, 2020). La variante c.335-311_335-303del no ha sido reportada previamente en la literatura, fue reclasificada por herramientas *in-silico* como probablemente patológica. La variante c.63A>G, ha sido relacionada con comunicación interauricular (CIA) en pacientes de Indonesia con una

frecuencia del 85.6%, y en China con una frecuencia del 80.18%. (Rozqie *et al.*, 2022; Wang *et al.*, 2019), las variantes c.335-311A>T y c.335-315A>T no han sido reportadas previamente en la literatura.

El gen *NKX2-5* mostró asociación con *GATA4*, *SMAD4*, *TBX5*, *HAND1*, y *SMEF2*. (Riazi *et al.* 2009) sugieren que la asociación con *GATA4* puede deberse a la presencia de sitios de unión a *GATA4* con alta afinidad en regiones potenciadores del gen *NKX2-5*, sugiriendo que está genéticamente posicionado aguas arriba de *NKX2-5*; demostraron también, una regulación positiva de este gen sobre *GATA4* y viceversa. Otros sitios de unión se han encontrado aguas arriba de *NKX2-5* para *SMAD* cercano a los sitios de unión *GATA4*, esta co-asociación física es importante, ya que los factores de transcripción *SMAD1/4* y *GATA4* funcionan como intermediarios de la vía de señalización de BMP que activan *NKX2-5* (Brown *et al.*, 2003; Wang *et al.*, 2011).

La interacción entre genes permite la regulación o expresión de otros genes específicos, por ejemplo, la expresión génica completa de los genes ANF y Cx40 se ve favorecida por la asociación aguas abajo entre *NKX2-5* y *TBX5*, y de su presencia en los sitios de unión específicos en el ADN (Bruneau, 2002); El gen ANF se expresa en los cardiomiocitos auriculares y la capa trabecular de los cardiomiocitos ventriculares (Warren *et al.*, 2011) y Cx40 se expresa en compartimentos auriculares y ventriculares y sistema de conducción cardíaca (Sultana *et al.*, 2014).

Una investigación del control de *NKX2-5* sobre *HAND1* en la asimetría de corazones de murino mostró que *NKX2-5* podría controlar la expresión de *HAND1* directamente al interactuar con su módulo potenciador correspondiente a la señalización izquierda-derecha, o indirectamente ayudado por la regulación de factores intermediarios. *HAND1* se expresa en la región caudal del lado izquierdo del corazón en bucle, específicamente aurícula y ventrículo izquierdo (Biden y Harvey, 1997; Hatemi *et al.*, 2011).

Los genes *NKX2-5* y *MEF2C* participan tempranamente en la cardiogénesis, y pueden homodimerizarse, unirse al ADN y regular otros genes específicos como ANF; actuar cooperativamente e iniciar la transcripción de otros factores como *HAND2* y *GATA4* (Vincentz *et al.*, 2008).

Para las variantes con mayor FA de *TBX20*, (Santillán, 2015) reportó c.39T>C, c.545+13A>G y c.655-44G>A como variantes de genes asociados a enfermedad cardíaca; c.655-44G>A fue

reportada por (Monroy *et al.*, 2015), en pacientes con CIA; c.381-40A>G no ha sido registrada previamente en la literatura.

El gen *TBX20* mostró asociación con *TBX5*, *TBX22*, *TBX21*, *EOMES*, *TBX6*, *TBR1* y *TBX3*. *TBX20* y *TBX5* interactúan físicamente entre sí en los dominios N-terminal y T-box de ambas proteínas, y se requiere que actúen sinérgicamente para el correcto desarrollo del corazón (Brown, 2005) además se expresan al mismo tiempo y en la misma región junto con *GATA4* y *NKX2-5* para promover la especificación de cámaras del miocardio (Paige *et al.*, 2015). *TBX20* y *TBX3* presentan áreas de expresión superpuestas en el tracto de salida, y el canal atri o ventricular (CAV), además la eliminación de *TBX20* conlleva a regulación negativa de *TBX3* (Gravilov *et al.*, 2013). No se encontró reporte de literatura sobre la interacción de *TBX20* con *TBX6*, *TBX22*, *TBR1*, *EOMES* en el desarrollo cardiaco.

Para el gen *GATA4*, (Orjuela *et al.*, 2014) en un estudio de población colombiana de 33 individuos con cardiopatías no sindrómicas, reportó la variante c.1132A>G como “polimorfismo”, en dos individuos, uno con CIA y otro con coartación aortica; (Liu *et al.*, 2017) identificaron esta variante en estado heterocigoto en 2/600 pacientes en una población China Han que presentaban diagnóstico clínico de defectos cono truncales (DCT), uno presentó Tetralogía de Fallot y otro atresia pulmonar con defecto septal ventricular (CIV), sin ser descrita en la población control; indican también que (Al-Azzouny *et al.*, 2016) predijeron que esta variante podría dar lugar a un transcrito no funcional; y (Miuya *et al.*, 2013) la reportan presente en pacientes con CC. La variante c.1000+23A>T fue encontrada en un paciente con insuficiencia tricúspide y reportada en un estudio de pacientes marroquíes con CIA, sin presentar un efecto en el splicing (El Bouchikhi *et al.*, 2020a; Theron *et al.*, 2015). La variante c.1059C>T, fue encontrada en pacientes marroquíes con hipertensión e infarto agudo de miocardio (Miuya *et al.*, 2013).

GATA4 mostró interacciones con los genes, *NKX2-5*, *HEY2*, *ZFPM2*, *ZFPM1* y *TBX5*. *GATA4* funciona como regulador positivo de *NKX2-5* (Riazi *et al.*, 2009). La actividad de *HEY2* se ve favorecida por la asociación con *GATA4* en presencia de *TBX20*, *HEY2* funciona como factor de transcripción indispensable en diferenciación de miocitos ventriculares (Ihara *et al.*, 2020).

GATA4 puede formar un complejo con *ZFPM2/FOG2* importante para el desarrollo del tracto de salida (Zhang *et al.*, 2014), así mismo, (Katz *et al.*, 2003) demostraron la expresión de *ZFPM1/FOG1* en el tracto de salida, co-expresandose con los miembros *GATA4/5/6*. Su interacción con *TBX5* activa la expresión de Cx30.2, que codifica una proteína de unión tipo gap

en el nodo atrio ventricular, necesaria para el retraso normal de este nodo (McCulley y Black, 2012).

En el gen *NOTCH1*, la variante c.2354-24C>T fue reclasificada como patogénica, y teóricamente podría tener un efecto patogénico sobre el empalme según HSF, sin embargo, aún no ha sido reportada en la literatura. Las variantes c.1669+9T>C y c.2970-31A>G fueron identificadas en pacientes diagnosticados clínicamente con válvula aórtica bicúspide (BAV), (Dargis, 2015; Debiec *et al.*, 2022); La variante c.5473-43T>C se encontró en pacientes con BAV/aneurisma aórtico ascendente (Kent *et al.*, 2013), la BAV es una de las anomalías congénitas más frecuentes afectando al 1-2% de la población general; variantes en *NOTCH1* están implicadas en enfermedades de las válvulas aórticas (High *et al.*, 2009).

La asociación de *NOTCH1* con *JAG1* (Jagged1) se debe a que *JAG1* se expresa en la célula vecina produciendo escisión proteolítica del dominio intracelular Notch para ser trasladado al núcleo, formando un complejo transcripcional con Rbp-J y MAML expresando los factores de transcripción *HEY* y *HES* (en ausencia de señal, Rbpj y MAML, no se asocian). MAML1, es una proteína nuclear que actúa como regulador positivo de *NOTCH1* (Kitagawa, 2015, Varadkar *et al.*, 2008). *NOTCH1* y *NOTCH2* se expresan fuertemente en endocardio del CAV y tracto de salida (Niessen y Karsan, 2007).

En el gen *PTPN11*, la variante c.854-30T>C fue reclasificada como patogénica, sin embargo, no ha sido reportada en la literatura en asociación con CC; esta variante presentó la FA más alta; las variantes patogénicas no necesariamente implican enfermedad, pero sí una predisposición a la segregación de la variante a generaciones futuras (Rodriguez *et al.*, 2014). De acuerdo con (El Bouchikhi *et al.* 2020b) la variante c.854-21 C>T fue indicada como recurrente en pacientes con síndrome de Noonan, de los cuales el 84% presentaban CC, principalmente estenosis valvular pulmonar. La variante c.756+1274G>A no ha sido registrada previamente en la literatura.

La asociación de *PTPN11* con *FRS2*, se debe a que *FRS2* contiene sitios de unión a *PTPN11* y *GRB2* formando un complejo en respuesta a la estimulación de receptores de factores de crecimiento fibroblástico (FGFR), además, *FRS2* juega un papel importante en la mediación de señales de la cascada de señalización Ras/MAPK (Ong *et al.*, 2020). En la asociación con *SPRY1*, (Hanafusa *et al.*, 2004) sugieren que *PTPN11* actúa como regulador positivo en vía de señalización de FGF al desfosforilar e inactivar el inhibidor *SPRY1*. *PTPN11* promueve la

activación del receptor proteína tirosina quinasa de la vía de señal del receptor PDGF β en células del músculo liso (Cheng *et al.*, 2018).

En cuanto a la red de interacción de los 5 genes, estos se expresan desde estadios tempranos y se mantienen durante todo el proceso de desarrollo cardíaco. Inician procesos de migración, destino y especificación celular, formando complejos para la expresión de otras moléculas y genes, indispensables para el buen desarrollo del corazón, y especificación de regiones cardíacas. (Paige *et al.*, 2015; Tan y Lewandowski, 2019; Lauriol *et al.*, 2015).

CONCLUSION

En este estudio se identificaron variantes genómicas asociadas a cardiopatías en los diferentes genes implicados en el desarrollo cardíaco en la población del suroccidente colombiano en pacientes con enfermedad compleja y no diagnosticados clínicamente con cardiopatías congénitas. Las variantes intrónicas fueron relevantes en los genes *TBX20*, *NOTCH1* y *PTPN11*, llegando a obtener las mayores frecuencias alélicas, y algunas reclasificadas como patogénicas según las herramientas de predicción *in silico*, lo que sugiere la necesidad de mayores estudios sobre el impacto y frecuencia de estas variantes en población con cardiopatías congénitas a fin de permitir el alcance de una medicina de precisión.

Las redes de interacción génica permitieron determinar las funciones biológicas de los genes estudiados, comprobando su importancia biológica en la participación del desarrollo cardíaco, y la posible activación e inactivación de genes involucrados en dicho proceso. Su correcta interpretación permitirá un acercamiento a la medicina de precisión para un tratamiento específico y una adecuada consejería genética.

Este trabajo contribuye al conocimiento sobre la frecuencia con la que las variantes genómicas asociadas a CC están presentes en la población no diagnosticada clínicamente con esta patología, permitiendo un diagnóstico temprano, para un tratamiento precoz, disminuyendo la morbimortalidad, pensando en un futuro tamizaje neonatal molecular universal de cardiopatías congénitas.

AGRADECIMIENTOS

Agradecer primeramente a Dios por permitir la culminación de este trabajo, al Instituto de Genética Médica-Genomics por contribuir con los datos utilizados de este trabajo, al Dr. José María Satizábal y Daniela Arturo-Terranova por su compartirme su conocimiento y apoyo en el

tema, para un crecimiento personal y académico, a mi esposo Christian Ortiz y mi mamá Olivia Cerón y demás familiares y amigos por su constante acompañamiento y apoyo en este proceso académico, a mi abuelito Rafael Cerón (QEPD) porque siempre tuvo palabras de amor, sabiduría y aliento.

LITERATURA CITADA

- 1000 Genomes Project Consortium (2015). A global reference for human genetic Nature. 526(7571), pp. 68-74. doi:10.1038/nature15393.variation.
- Adzhubei, I., Jordan, D., y Sunyaev, S. (2013). “Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2”. *Current Protocols in Human Genetics*. 76(1), pp.1-52.
- Araújo, N., Farias, R., Pereira, P., Figueirêdo, F., Morais, A., Saldanha, L., y Gabriel, J., (2008). “A era da bioinformática: seu potencial e suas implicações para as ciências da saúde”. *Estudos de Biologia*. 30(70/72), pp. 143-148.
- Arturo-Terranova, D. (2021). Identificación de variantes genómicas del gen GBA en pacientes con patologías diferentes y no diagnosticados clínicamente con enfermedad de Gaucher en el suroccidente colombiano. [online]. Colombia. Universidad del Valle.
- Biben, C. y Harvey, R. P. (1997). “Homeodomain factor Nkx2-5 controls left/right asymmetric expression of bHLH gene eHand during murine heart development”. *Genes & Development*. 11(11), pp.1357–1369.
- Brown, C. O., Chi, X., Garcia-Gras, E., Shirai, M., Feng, X.-H. y Schwartz, R. J. (2003). “The Cardiac Determination Factor, Nkx2-5, Is Activated by Mutual Cofactors GATA-4 and Smad1/4 via a Novel Upstream Enhancer”. *Journal of Biological Chemistry*. 279(11), pp.10659–10669.
- Brown, D. D. (2005). “Tbx5 and Tbx20 act synergistically to control vertebrate heart morphogenesis”. *Development*. 132(3), pp. 553–563.
- Bruneau, B. G. (2002). “Transcriptional Regulation of Vertebrate Cardiac Morphogenesis”. *Circulation Research*. 90(5), pp. 509–519.
- Clinvar Miner (2023). *List of Variants in gene NKX2-5*. <https://clinvarminer.genetics.utah.edu/variants-by-gene/NKX2-5/significance/any> (Accedido: 7 marzo 2023)
- Chen, Y., Xiao, D., Zhang, L., Cai, C.-L., Li, B.-Y. y Liu, Y. (2021). “The Role of Tbx20 in Cardiovascular Development and Function”. *Frontiers in Cell and Developmental Biology*. 9, pp. 1-8.
- Cheng, Y., Yu, M., Xu, J., He, M., Wang, H., Kong, H. y Xie, W. (2018). “Inhibition of Shp2 ameliorates monocrotaline-induced pulmonary arterial hypertension in rats”. *BMC Pulmonary Medicine*. 18(130), pp. 1-10.
- Choquet, C., Kelly, R. G. y Miquerol, L. (2020). “Nkx2-5 defines distinct scaffold and recruitment phases during formation of the murine cardiac Purkinje fiber network”. *Nature Communications*. 11(1), pp 1-12.
- Chung, I., y Rajakumar, G., (2016). “Genetics of Congenital Heart Defects: The NKX2-5 Gene, a Key Player”. *Genes*, 7(2) pp.1-12.
- Dargis, N. (2015). *Identification de variations génétiques associées à la bicuspidie valvulaire aortique chez les hommes et les femmes*. Master's thesis, Université Laval.

- Debiec, R., Hamby, S., Jones, P., Safwan, K., Sosin, M., Hetherington, S., Sprigings, D., Sharman, D., Lee, K., Salahshouri, P., Wheeldon, N., Chukwuemeka, A., Boutziouka, V., Elamin, M., Coolman, S., Asiani, M., Kharodia, S., Skinner, G., Samani, N., Webb, T., Bolger, A. (2022). “Contribution of NOTCH1 genetic variants to bicuspid aortic valve and other congenital lesions”. *Heart*. 2021 pp. 320428.
- Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Bérout, G., Claustres, M. y Bérout, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*. 37(9), número de artículo: e67-e67.
- El Bouchikhi, I., Bouguenouch, L., Moufid, F., Belhassan, K., Samri, I., Chaouti, A., Houssaïni, M., Atmani, S. y Ouldim, K. (2020a). “Absence of GATA4 Mutations in Moroccan Patients with Atrial Septal Defect (ASD) Provides Further Evidence of Limited Involvement of GATA4 in Major Congenital Heart Defects”. *The Eurasian Journal of Medicine*. 52(3), pp. 283–287.
- El Bouchikhi, I., Bouguenouch, L., Moufid, F. Z., Samri, I., Abdouss, F., Melhouf, M. A., Iraqui Houssaini, M., Belhassan, K., Atmani, S. y Ouldim, K., (2020b). “Molecular and environmental characterization of Noonan syndrome in Morocco reveals a significant association with consanguinity and advanced parental age”. *Egyptian Journal of Medical Human Genetics*. 21(1), pp 1-9.
- Fan, D., Pang, S., Chen, J., Shan, J., Cheng, Q. y Yan, B., (2021). “Identification and functional study of GATA4 gene regulatory variants in atrial septal defects”. *BMC Cardiovascular Disorder*, 21(1) PP. 1-8.
- Forero-Manzano, M., Triana-Palencia, E., Figueroa-Rueda, A., Flórez-Rodríguez, X., Castro-Monsalve, M., Quintero-Lesmes, C., y Gamboa-Delgado, M., (2022). “Association of social determinants with the severity of congenital heart disease”. *Pediatric Research*. 93, pp. 1391–1398.
- García-Flores, E., Calderón-Colmenero, J., Borgonio-Cuadra, V. M., Sandoval, J. P., García-Montes, J. A., Cazarín-Santos, B. G., Miranda-Duarte, A., Gamboa-Domínguez, A., Rodríguez-Pérez, J. M., & Pérez-Hernández, N. (2023). “Epigenetic Evaluation of the TBX20 Gene and Environmental Risk Factors in Mexican Paediatric Patients with Congenital Septal Defects”. *Cells*, 12(4), 586 pp. 1-16.
- Gavrilov, S., Harvey, R. P. y Papaioannou, V. E. (2013). “Lack of Genetic Interaction between Tbx20 and Tbx3 in Early Mouse Heart Development”. *PLoS ONE*. 8(7), e70149.
- Hanafusa, H., Torii, S., Yasunaga, T., Matsumoto, K. y Nishida, E. (2004). “Shp2, an SH2-containing Protein-tyrosine Phosphatase, Positively Regulates Receptor Tyrosine Kinase Signaling by Dephosphorylating and Inactivating the Inhibitor Sprouty”. *Journal of Biological Chemistry*. 279(22), pp. 22992–22995.
- Hamosh, A., Scott, A., Amberger, J., Valle, D. y McKusick, V. (2000). “Online Mendelian Inheritance In Man (OMIM)”. *Human Mutation*. 15(1), pp. 57–61.
- Hatemi, A. C., Gulec, C., Cine, N., Vural, B., Hatirnaz, O., Sayitoglu, M., Oztunc, F., Saltik, L., Kansiz, E. y Erginel Unaltuna, N. (2011). “Sequence variations of NKX2-5 and HAND1 genes in patients with atrial isomerism”. *Anadolu Kardiyoloji Dergisi/The Anatolian Journal of Cardiology*. 11(4), pp. 319-328.
- High, F. A., Jain, R., Stoller, J. Z., Antonucci, N. B., Lu, M. M., Loomes, K. M., Kaestner, K. H., Pear, W. S. y Epstein, J. A. (2009). “Murine Jagged1/Notch signaling in the second heart field orchestrates Fgf8 expression and tissue-tissue interactions during outflow tract development”. *Journal of Clinical Investigation*. 119(7), pp. 1986-1996.

- Ibáñez, L., Victoria S., Hurtado, P. (2021). “Prevalencia de cardiopatías congénitas en una cohorte de 54.193 nacimientos entre 2011-2017”. *Revista Colombiana de Cardiología*, 28(1), pp.53-59.
- Ihara, D., Watanabe, Y., Seya, D., Arai, Y., Isomoto, Y., Nakano, A., Kubo, A., Ogura, T., Kawamura, T. y Nakagawa, O. (2020). “Expression of Hey2 transcription factor in the early embryonic ventricles is controlled through a distal enhancer by Tbx20 and Gata transcription factors”. *Developmental Biology*. 461(2), pp.124–131.
- Instituto Nacional de Salud. (2023), Castaño, S. Informe de eventos, defectos congénitos. Periodo epidemiológico III
- Karczewski, K. J. et al., (2020). “The mutational constraint spectrum quantified from variation in 141,456 humans”. *Nature*. 581(7809), pp. 434–443.
- Kitagawa, M., (2015). “Notch signalling in the nucleus: roles of Mastermind-like (MAML) transcriptional coactivators”. *Journal of Biochemistry*. 159(3), pp. 287-294.
- Katz, S. G., Williams, A., Yang, J., Fujiwara, Y., Tsang, A. P., Epstein, J. A. y Orkin, S. H. (2003). “Endothelial lineage-mediated loss of the GATA cofactor Friend of GATA 1 impairs cardiac development”. *Proceedings of the National Academy of Sciences*. 100(24), PP. 14030–14035.
- Kent, K. C., Crenshaw, M. L., Goh, D. L. M. y Dietz, H. C. (2013). “Genotype–phenotype correlation in patients with bicuspid aortic valve and aneurysm”. *The Journal of Thoracic and Cardiovascular Surgery*. 146(1), pp. 158–165.e1.
- Kobayashi, J., Yoshida, M., Tarui, S., Hirata, M., Nagai, Y., Kasahara, S., Naruse, K., Ito, H., Sano, S. y Oh, H., (2014). “Directed Differentiation of Patient-Specific Induced Pluripotent Stem Cells Identifies the Transcriptional Repression and Epigenetic Modification of NKX2-5, HAND1, and NOTCH1 in Hypoplastic Left Heart Syndrome”. *PLoS ONE*. 9(7), pp e102796.
- Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R. y Massouras, A., (2018). “VarSome: the human genomic variant search engine”. *Bioinformatics*. 35(11), pp. 1978–1980.
- Kumar, P., Henikoff, S. y Ng, P. (2009). “Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm”. *Nature Protocols*. 4(7), 1073–1081.
- Landrum, M. J., Lee, J. M, Benson, M, Brown, G, Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-Salomon, R., Rubinstein, W. y Maglott, D. R., (2015). “ClinVar: public archive of interpretations of clinically relevant variants”. *Nucleic Acids Research*. 44(D1), pp. D862—D868.
- Lauriol, J., Jaffré, F. y Kontaridis, M. I. (2015). “The role of the protein tyrosine phosphatase SHP2 in cardiac development and disease”. *Seminars in Cell & Developmental Biology*. 37, pp. 73–81.
- Lek, M. et al., (2016). “Analysis of protein-coding genetic variation in 60,706 humans”. *Nature*. 536(7616), pp.285–291.
- Liu, Y., Li, B., Xu, Y. y Sun, K. (2017). “Mutation Screening of Gata4 Gene in CTD Patients Within Chinese Han Population”. *Pediatric Cardiology*. 38(3), pp. 506–512.
- Luxán, G., D’Amato, G., MacGrogan, D., & de la Pompa, J. L. (2016). “Endocardial Notch Signaling in Cardiac Development and Disease”. *Circulation Research*, 118(1), pp. e1-e8.

- McCulley, D., y Black, B. (2012). “Transcription factor pathways and congenital heart disease”. *Current Topics in Developmental Biology*. 100, pp. 253-277.
- Meller, CH., Grinenco, S., Aiello, H., Córdoba, A., Sáenz-Tejeira, M., Marantz, P., Otaño, L. (2020) “congenital heart disease, prenatal diagnosis and management”. *Archivos Argentinos de Pediatría*. 118(2). Pp. 149-161.
- Monroy, I. Pérez, N. Vargas, G. Ortiz, G. Buendía, A. Calderón, J. Ramirez, S. Cervantes, J. Curi, P. Martinez, N. Rodriguez, J. (2013). “Cambiano el paradigma en las cardiopatías congénitas: de la anatomía a la etiología molecular”. *Gaceta Médica de México*. 149(2), pp. 212-219
- Monroy-Muñoz, I. E., Pérez-Hernández, N., Rodríguez-Pérez, J. M., Muñoz-Medina, J. E., Angeles-Martínez, J., García-Trejo, J. J., Morales-Ríos, E., Massó, F., Sandoval-Jones, J. P., Cervantes-Salazar, J., García-Montes, J. A., Calderón-Colmenero, J. y Vargas-Alarcón, G. (2015). “Novel Mutations in the Transcriptional Activator Domain of the Human TBX20 in Patients with Atrial Septal Defect”. *BioMed Research International*. 2015, pp. 1–7.
- Moreno-Giraldo, L., Herrera-Mafla, C., Satizábal-Soto, J. (2019). “Biología molecular de cardiopatías congénitas”. *Revista de la Asociación Colombiana de Ciencias Biológicas*. 100–105.
- Moreno-Giraldo, L. J., Satizábal-Soto, J. M., & Sánchez-Gomez, A. (2020). “Research Article Genomic variability of the mucopolysaccharidosis complex in southwestern Colombia.” *Genetics and Molecular Research*, 19(2) pp. 1-16.
- Mullen, M., Zhang, A., Lui, G. K., Romfh, A. W., Rhee, J.-W., y Wu, J. C., (2021). “Race and Genetics in Congenital Heart Disease: Application of iPSCs, Omics, and Machine Learning Technologies”. *Frontiers in Cardiovascular Medicine* [en línea]. 8, pp.1-9.
- Muiya, N. P., Wakil, S. M., Tahir, A. I., Hagos, S., Najai, M., Gueco, D., Al-Tassan, N., Andres, E., Mazher, N., Meyer, B. F. y Dzimiri, N. (2013). “A study of the role of GATA4 polymorphism in cardiovascular metabolic disorders”. *Human Genomics*. 7(25), pp.1-11.
- Niessen, K., y Karsan, A. (2007). “Notch signaling in the developing cardiovascular system”. *American Journal of Physiology-Cell Physiology*, 293(1), pp. C1—C11.
- Ong, S. H., Guy, G. R., Hadari, Y. R., Laks, S., Gotoh, N., Schlessinger, J. y Lax, I. (2000). “FRS2 Proteins Recruit Intracellular Signaling Pathways by Binding to Diverse Targets on Fibroblast Growth Factor and Nerve Growth Factor Receptors”. *Molecular and Cellular Biology*. 20(3), pp. 979–989.
- Orjuela, D., Núñez, f., Caicedo, V., Pachón, D., Salazar, M. (2014). “Mutaciones en el gen GATA4 en pacientes con cardiopatía congénita no sindrómica”. *Investigación Clínica*. 55(3), p. 207-216.
- Paige, S. L., Plonowska, K., Xu, A. y Wu, S. M., (2015). Molecular Regulation of Cardiomyocyte Differentiation. *Circulation Research*. 116(2), pp. 341–353.
- Rentzsch, P., Witten, D., Cooper, G., Shendure, J. y Kircher, M. (2018). “CADD: predicting the deleteriousness of variants throughout the human genome”. *Nucleic Acids Research*. 47(D1), pp. D886—D894.
- Riazi, A., Takeuchi, J. K., Hornberger, L. K., Zaidi, S. H., Amini, F., Coles, J., Bruneau, B. G. y Van Arsdell, G. S., (2009). “NKX2-5 Regulates the Expression of β -Catenin and GATA4 in Ventricular Myocytes”. *PLoS ONE*. 4(5), e5698.
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K. y Rehm, H. (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), 405–423.
- Rodríguez, F., Unanue, N., Hernández, I., Heath, K. y Cassorla, F. (2014). "Molecular characterization of Chilean patients with a clinical diagnosis of Noonan síndrome". *Journal of Pediatric Endocrinology and Metabolism*. 27(3-4).
- Rozqie, R., Satwiko, M. G., Anggrahini, D. W., Sadewa, A. H., Gunadi, Hartopo, A. B., Mumpuni, H. y Dinarti, L. K., (2022). "NKX2-5 variants screening in patients with atrial septal defect in Indonesia". *BMC Medical Genomics*. 15(1), pp. 1-6.
- Saint-Laurent, C., Mazeyrie, L., Tajan, M., Paccoud, R., Castan-Laurell, I., Valet, P., Edouard, T., Pradère, J.-P., Dray, C. y Yart, A. (2022). "The Tyrosine Phosphatase SHP2: A New Target for Insulin Resistance?". *Biomedicines*. 10(9), 2139.
- Salgado, D., Desvignes, J., Rai, G., Blanchard, A., Miltgen, M., Pinard, A., Lévy, N., Collod-Bérout, G. y Bérout, C., (2016). "UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution". *Human Mutation*. 37(5), pp. 439–446.
- Santillán, S. (2015). *Aplicación de nuevas tecnologías de secuenciación masiva al diagnóstico de enfermedades genéticas cardíacas heterogéneas*. Tesis doctoral. Universidad de Alicante.
- Sayers, E., Beck, J. Bolton, E., Bourexis, D., Brister, J., Canese, K., Comeau, D., Funk, K. Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T., O’Leary, N., Phan, L., Rangwala, S., Schneider, V., Skripchenko, Y., Wang, J., Ye, J., Trawick, B., Pruitt, K., Sherry, S. (2020). "Database resources of the National Center for Biotechnology Information." *Nucleic Acids Research*. 49, pp10–17.
- Stalker, J. Gibbins, B. Meidl, P. Smith, J. Spooner, W. Hotz, HR. Cox, AV. (2004). "The Ensembl Web Site: Mechanics of a Genome Browser". *Genome Research*. 14(5), pp. 951–955.
- Sultana, N., Nakamura, N., Hirose, S., Kutsuzawa, K., Akaike, T. y Nag, K. (2014). "Congenital Heart Diseases and Biotechnology: Connecting by Connexin". *Advanced Materials Research*. (995), pp. 85–112.
- Tan, C., y Lewandowski, A. (2019). "The Transitional Heart: From Early Embryonic and Fetal Development to Neonatal Life". *Fetal Diagnosis and Therapy*. 47(Suppl. 5), pp. 373–386.
- Theron, A., Pinard, A., Riberi, A. y Zaffran, S. (2015). "An uncommon cause of tricuspid regurgitation: three-dimensional echocardiographic incremental value, surgical and genetic insights". *European Journal of Cardio-Thoracic Surgery*. 50(1), pp.180–182.
- Valentín, A. (2018). "Cardiopatías congénitas en edad pediátrica, aspectos clínicos y epidemiológicos". *Revista Médica Electrónica*, 40(4), pp. 1083-1099.
- Van Weerd, J., y Christoffels, V. (2016). "The formation and function of the cardiac conduction system" *Development*. 143(2), pp.197–210.
- Varadkar, P., Kraman, M., Despres, D., Ma, G., Lozier, J., & McCright, B. (2008). "Notch2 is required for the proliferation of cardiac neural crest-derived smooth muscle cells". *Developmental Dynamics*, 237(4), pp. 1144–1152.
- Vincentz, J. W., Barnes, R. M., Firulli, B. A., Conway, S. J. y Firulli, A. B. (2008). "Cooperative interaction ofNkx2.5andMef2ctranscription factors during heart development". *Developmental Dynamics*. 237(12), pp.3809–3819.

- Wang, H., Liu, Y., Li, Y., Wang, W., Li, L., Meng, M., Xie, Y., Zhang, Y., Yunfeng, Z., Han, S., Zeng, J., Hou, Z. y Jiang, L. (2019). “Analysis of NKX2-5 in 439 Chinese Patients with Sporadic Atrial Septal Defect”. *Medical Science Monitor*. 25, pp. 2756–2763.
- Wang, J., Greene, S. B. y Martin, J. F. (2011). “BMP signaling in congenital heart disease: New developments and future directions.” *Birth Defects Research Part A: Clinical and Molecular Teratology*. 91(6), pp.441–448.
- Wang, Y., Lu, P., Jiang, L., Wu, B., & Zhou, B. (2019). “Control of sinus venous valve and sinoatrial node development by endocardial NOTCH1”. *Cardiovascular Research*, 116(8), pp.1473–1486.
- Warren, S. A., Terada, R., Briggs, L. E., Cole-Jeffrey, C. T., Chien, W. M., Seki, T., Weinberg, E. O., Yang, T. P., Chin, M. T., Bungert, J. y Kasahara, H. (2011). “Differential Role of Nkx2-5 in Activation of the Atrial Natriuretic Factor Gene in the Developing versus Failing Heart”. *Molecular and Cellular Biology*. 31(22), pp. 4633–4645.
- Williams, K., Carson, J., Lo, C., (2019). “Genetics of Congenital Heart Disease”. *Biomolécules*, 9(12), pp. 1-23.
- Xu, Z.-q., Chen, W.-c., Li, Y.-j., Suo, M.-j., Tian, G.-x., Sheng, W. y Huang, G.-y. (2022). “PTPN11 Gene Mutations and Its Association with the Risk of Congenital Heart Disease”. *Disease Markers*. 2022, pp. 1–11.
- Ye, S., Wang, C., Xu, Z., Lin, H., Wan, X., Yu, Y., Adhicary, S., Zhang, J. Z., Zhou, Y., Liu, C., Alonzo, M., Bi, J., Ramirez-Navarro, A., Deschenes, I., Ma, Q., Garg, V., Wu, J. C., & Zhao, M.-T. (2023). “Impaired Human Cardiac Cell Development due to NOTCH1 Deficiency”. *Circulation Research*. 132, pp.187-204.
- Zhang, W., Shen, L., Deng, Z., Ding, Y., Mo, X., Xu, Z., Gao, Q. y Yi, L. (2014). “Novel Missense Variants of ZFPM2/FOG2 Identified in Conotruncal Heart Defect Patients Do Not Impair Interaction with GATA4”. *PLoS ONE*. 9(7), e102379.

Tabla 1. Se reportan las tres variantes con mayor frecuencia alélica por gen, clasificación según el ACMG y comparación con las bases de datos poblacionales.

GEN	Variante	Clasificación	FA en el presente estudio	FA en 1000 genomas	FA en EXAC	FA en gnomAD
NKX 2-5	c.63A>G	Benigna	0,28438	0,53574	0,40522	0,43234
	c.335-311A>T	Benigna	0,26875	0,61100	0,46750	0,39640
	c.335-315A>T	Benigna	0,25000	0,40300	0,44380	0,39330
TBX20	c.39T>C	Benigna	0,34063	0,69389	0,67668	0,67920
	c.545+13A>G	Benigna	0,25625	0,34400	0,38080	0,38900
	c.655-44G>A	Benigna	0,25625	0,34400	0,38110	0,38900
	c.381-40A>G	Benigna	0,01250	0,00300	0,00080	0,00100
GATA 4	c.1132A>G	Benigna	0,04063	0,04300	0,09621	0,09520
	c.1000+23A>T	Benigna	0,02813	0,02300	0,00660	0,02200
	c.1059C>T	Benigna	0,02188	0,03900	0,01123	0,01000
NOTCH1	c.1669+9T>C	Benigna	0,35313	0,88399	0,87042	0,87200
	c.2970-31A>G	Benigna	0,33438	0,76500	0,66270	0,63800
	c.5473-43T>C	Benigna	0,32813	0,74200	0,59200	0,59300
PTPN11	c.854-30T>C	Benigna	0,08750	0,01800	0,01200	0,01800
	c.756+1274G>A	Benigna	0,07813	0,16700	0,21080	0,13100
	c.854-21C>T	Benigna	0,02813	0,03700	0,05919	0,05800

FA: Frecuencia Alélica

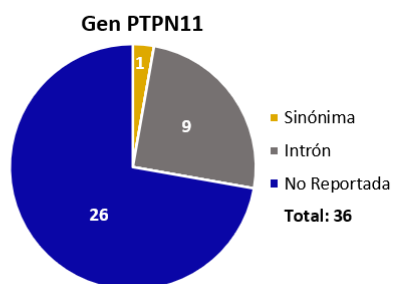
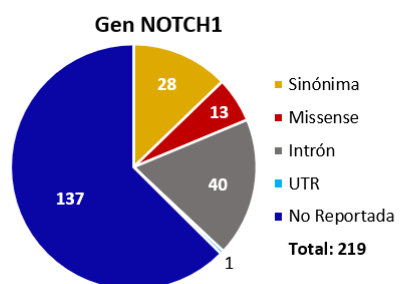
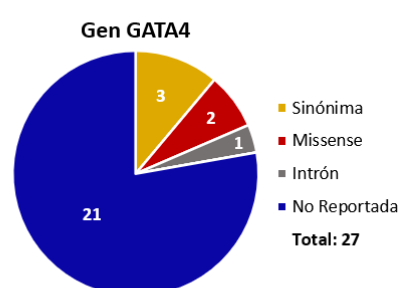
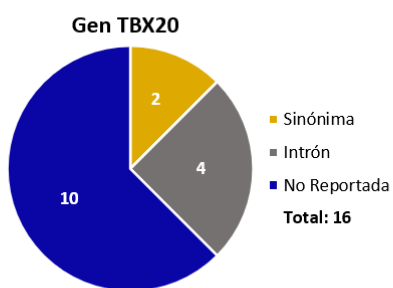
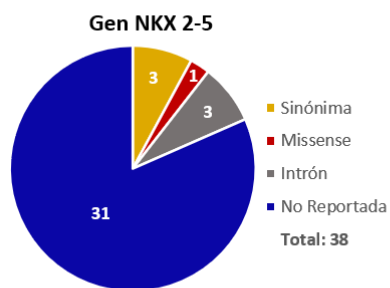
Figura 1. Cantidad de variantes y tipo de cambio para los genes *NKX2-5*, *TBX20*, *GATA4*, *NOTCH1*, y *PTPN11*.

Figura 2. significancia clínica según ACMG, generada por las herramientas *In-silico* para las variantes de los genes *NKX2-5*, *TBX20*, *GATA4*, *NOTCH1*, y *PTPN11*.

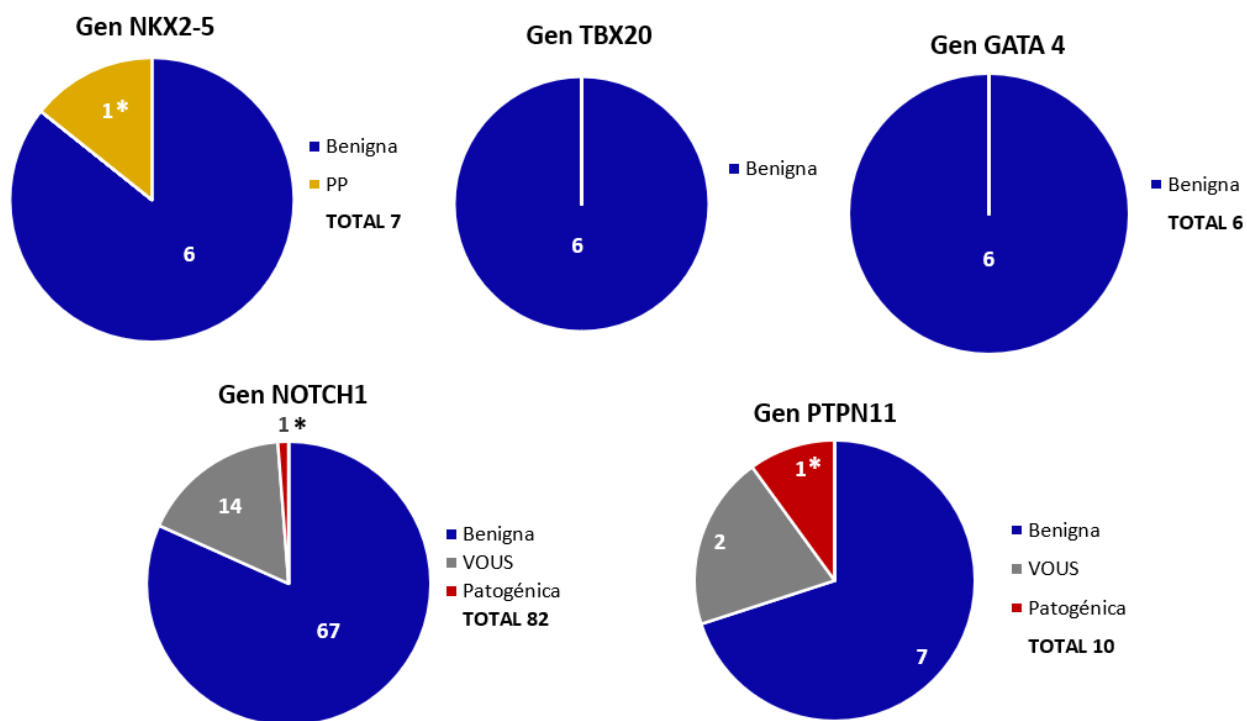
Figura 3. Red de interacción por cada uno de los genes y su interacción con genes o moléculas asociadas al desarrollo cardíaco. Las líneas entre los genes representan las interacciones y los colores dentro del gen representa la función asociada al desarrollo cardíaco. A. Red de interacción del gen *NKX2-5*. B. Red de interacción del gen *TBX20*. C. Red de interacción del gen *GATA4*. D. Red de interacción del gen *NOTCH1*. E. Red de interacción del gen *PTPN11*.

Figura 4. Red de interacción que tiene como nodos los cinco genes en estudio y su relación con otros genes o moléculas. Las líneas de colores representan las interacciones entre ellos y los colores dentro de cada gen representa las funciones asociadas al desarrollo cardíaco.

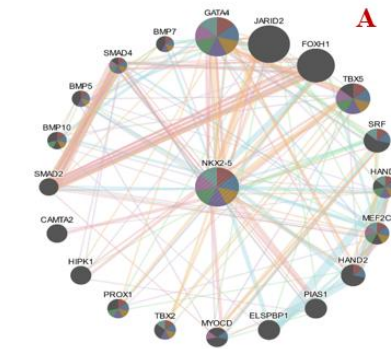
TIPOS DE VARIANTES POR GEN



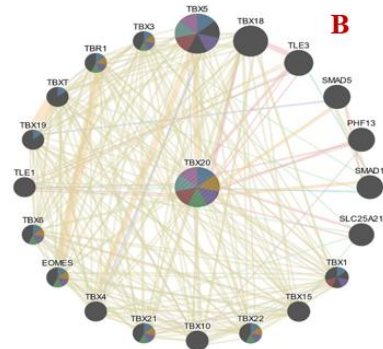
SIGNIFICANCIA CLINICA DE LAS HERRAMIENTAS IN-SILICO



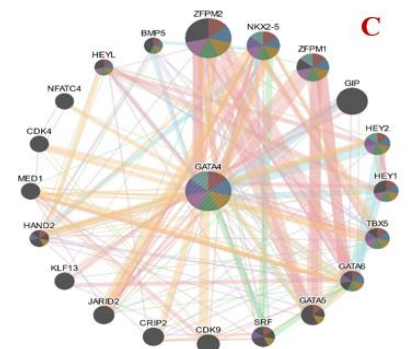
PP: Probablemente patogénica; * Variante Reclasificada.



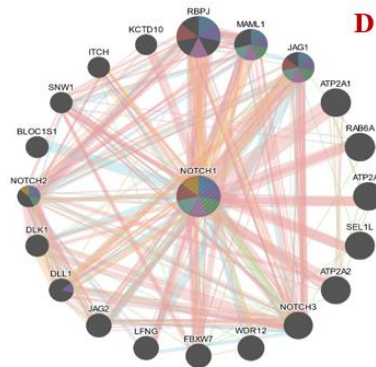
- Morfogénesis del corazón
- Desarrollo del corazón
- Desarrollo de la cámara cardíaca
- Desarrollo del tabique cardíaco
- Desarrollo del ventrículo cardíaco
- Regulación de la diferenciación de cardiocitos
- Morfogénesis del tubo cardíaco embrionario



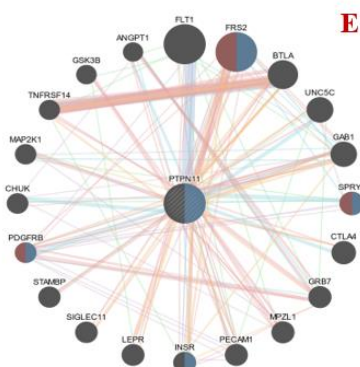
- Desarrollo del corazón
- Morfogénesis del tubo cardíaco embrionario
- Morfogénesis del corazón
- Desarrollo del tubo cardíaco embrionario
- Morfogénesis de la cámara cardíaca
- Regulación de la proliferación de células del músculo cardíaco
- Proliferación de células del músculo cardíaco



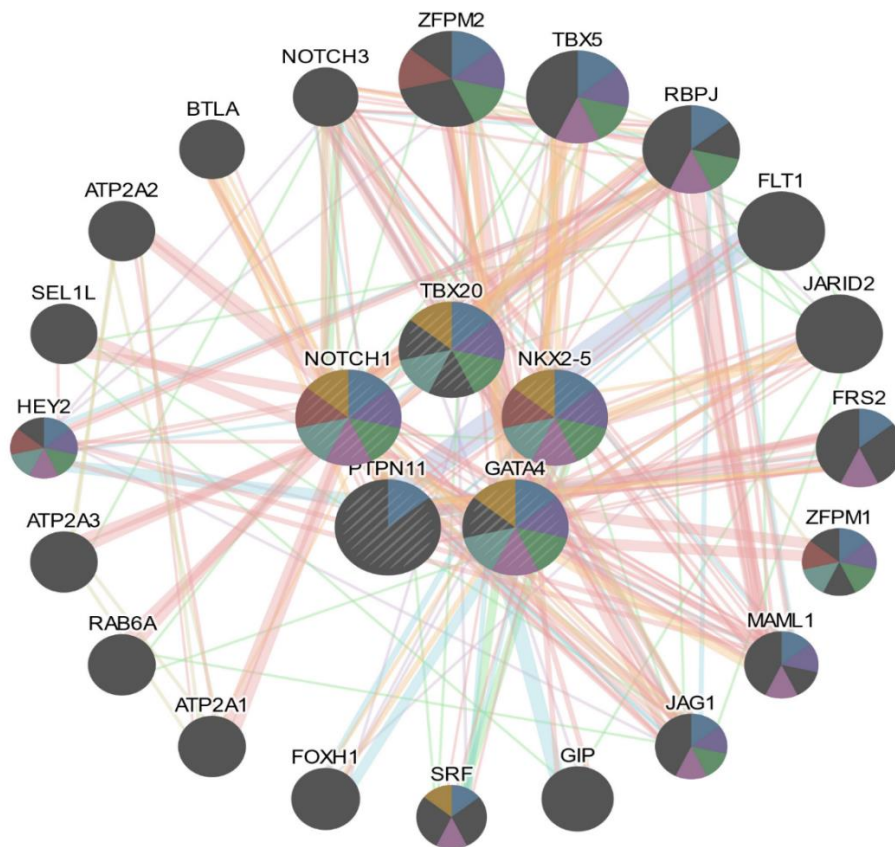
- Morfogénesis del corazón
- Desarrollo del tabique cardíaco
- Desarrollo del corazón
- Desarrollo de la cámara cardíaca
- Desarrollo del ventrículo cardíaco
- Diferenciación de cardiocitos
- Desarrollo del tabique interauricular



- Desarrollo de la cámara cardíaca
- Desarrollo del corazón
- Desarrollo del tabique cardíaco
- Diferenciación de cardiocitos
- Desarrollo de células cardíacas
- Desarrollo del ventrículo cardíaco
- Desarrollo del atrio cardíaco



- Desarrollo del corazón
- Diferenciación de cardiocitos



- Desarrollo del corazón
- Desarrollo del tabique cardíaco
- Desarrollo del ventrículo cardíaco
- Diferenciación de cardiocitos
- Desarrollo del atrio cardíaco
- Desarrollo del tabique ventricular
- Desarrollo del tubo cardíaco embrionario