



DIVISIÓN DE CIENCIAS BIOLÓGICAS Y DE LA SALUD

Investigación científica de medicina genómica:

Estudio de asociación del polimorfismo rs17052958 con huella de selección positiva con factores de riesgo y la enfermedad arterial coronaria en población mexicana

Alumna: Alejandra Orfin García

Matrícula: 218203133

Carrera: Licenciatura en Química Farmacéutica Biológica

Fecha de inicio: 29 de Agosto del 2022

Fecha de término: 01 de Marzo del 2023

Institución: Instituto Nacional de Medicina Genómica (INMEGEN)

---

Asesor externo: Dra. Mayra Domínguez Pérez

---

Asesor interno: Dra. Julia Pérez Ramos

## ÍNDICE

<b>INTRODUCCIÓN</b> .....	1
<b>MARCO TEÓRICO</b> .....	2
<b>Enfermedades cardiovasculares</b> .....	2
<b>Selección natural</b> .....	2
<b>Medicina Evolutiva</b> .....	3
<b>SNP con huella de selección positiva</b> .....	4
<b>Variante rs17052958</b> .....	4
<b>PLANTEAMIENTO DEL PROBLEMA</b> .....	5
<b>JUSTIFICACIÓN</b> .....	5
<b>OBJETIVO GENERAL</b> .....	6
<b>OBJETIVOS PARTICULARES</b> .....	6
<b>METODOLOGÍA</b> .....	6
a) <b>Población de estudio</b> .....	6
b) <b>Extracción de DNA a partir de sangre</b> .....	6
c) <b>Genotipificación</b> .....	6
d) <b>Equilibrio de Hardy Weinberg</b> .....	7
e) <b>Análisis Estadístico</b> .....	7
<b>RESULTADOS</b> .....	7
<b>DISCUSIÓN</b> .....	14
<b>CONCLUSIÓN</b> .....	16
<b>REFERENCIAS</b> .....	17

## INTRODUCCIÓN

Las enfermedades cardiovasculares (ECV) siguen siendo la principal causa de muerte en México y en el mundo (OMS). En México, el Instituto Nacional de Estadística y Geografía (INEGI) reporta que en el año 2021 se presentaron 226 703 defunciones debido a las ECV. La enfermedad arterial coronaria (EAC) representa el tipo más frecuente de las ECV, con un 78.2 % de los casos (INEGI, 2022). Esta enfermedad es un proceso patológico que se caracteriza por la formación de una placa (ateroma) en las arterias coronarias (aterosclerosis), esto conduce a enfermedades cardíacas isquémicas como la angina y el infarto de miocardio (Erdmann et al., 2018). El estudio de la etiología de la EAC es complejo por la participación de factores genéticos y ambientales.

En los últimos años, ha aumentado el uso de escaneos de genoma completo y búsquedas de genes candidatos para identificar asociaciones con la EAC, sin embargo, existe una subrepresentación de poblaciones como la población mexicana. La medicina evolutiva ha surgido para detectar y validar pequeños cambios adaptativos en el genoma humano y entender cómo actuó la selección natural sobre las enfermedades. Así este enfoque se ha convertido en la estrategia que puede explicar el origen de las enfermedades humanas a partir de la incompatibilidad de los genes con el medio ambiente (Naugler, 2008).

Las variantes genéticas o polimorfismos (SNP) son la base de la evolución y combinados pueden ser silenciosos o beneficiosos para un individuo y también pueden contribuir a la enfermedad. La identificación de SNPs que tienen huellas de selección positiva en el genoma humano ha surgido como una estrategia para encontrar variantes potencialmente relevantes en la medicina actual (Biswas & Akey, 2006).

Datos del nuestro grupo identificaron 103 loci con huellas de selección positiva en población indígena de México, como el SNP rs17052958 (Romero-Hidalgo, datos no publicados). Este SNP es una variante intrónica, que se encuentra en el gen delta de sarcoglicano (SGCD, por sus siglas en inglés Sarcoglycan Delta); este gen codifica a la proteína SGCD. La proteína SGCD es un componente del complejo de proteínas transmembrana Distrofina-glicoproteína (DGC) del sarcoglicano. Este complejo se ha relacionado con la distrofia muscular y miocardiopatía dilatada (MCD) (NIH, 2022).

La propuesta de este proyecto es buscar las asociaciones del rs17052958 del gen SGCD con huella de selección positiva con rasgos metabólicos y el riesgo cardiovascular en la cohorte Genética de la Enfermedad Aterosclerosa (GEA).

## MARCO TEÓRICO

### **Enfermedades cardiovasculares.**

Las ECV son aquellos trastornos que afectan al corazón y a los vasos sanguíneos, las enfermedades que la constituyen son: hipertensión arterial, cardiopatía coronaria, enfermedad cerebrovascular, insuficiencia cardíaca, cardiopatía reumática, miocardiopatías, entre otras. Las ECV son la principal causa de mortalidad en el mundo y las que sobresalen son la enfermedad isquémica del corazón o EAC y enfermedad cerebrovascular (Álvarez, et al., 2020).

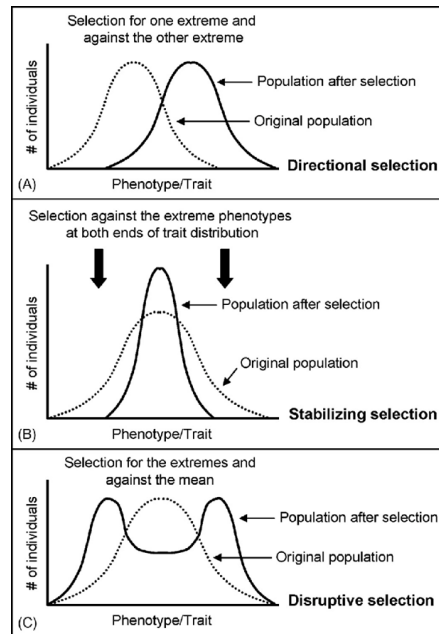
Existen dos tipos de factores de riesgo que pueden promover la enfermedad cardíaca o enfermedad cerebrovascular estos son los no modificables, como el sexo, la edad y los factores genéticos, y los modificables como los hábitos de vida (la mala alimentación, inactividad física, consumo de tabaco y consumo nocivo de alcohol), hipertensión arterial, diabetes, hiperglucemia e hiperlipidemia, sobrepeso, obesidad y sedentarismo (OMS, 2022; Elizondo, et al. 2020). La EAC ha estado presente en las poblaciones humanas durante miles de años, lo que implica serias presiones a las que debió responder la selección natural (Biswas & Akey, 2006)

### **Selección natural**

La selección natural es el proceso que resulta de la adaptación de un organismo y su medio ambiente, en donde los rasgos genéticos favorables se vuelven más frecuentes, mientras que los rasgos genéticos desfavorables se vuelven menos frecuentes dentro de una población, así esta adaptación promueve la reproducción y la supervivencia de la especie (CONABIO, 2023). La selección natural que favorece a las variantes fenotípicas de la población es de tres formas distintas: la selección estabilizadora, la selección direccional y la selección disruptiva.

En la figura 1 se puede observar una curva normal, donde la población con el fenotipo promedio se distribuye bajo la curva y representa la mayor parte de los individuos, y al alejarse del promedio dirigiéndose hacia los fenotipos extremos la cantidad de organismos disminuirá (Iturbe, 2010). La selección estabilizadora preserva a aquellos organismos que tienen un fenotipo intermedio manteniendo la composición genética de la población, la selección direccional se caracteriza por que va a favorecer aquellos

individuos que se encuentran en uno de los extremos fenotípicos respecto a la media de la población y por último la selección disruptiva favorece a aquella distribución fenotípica del organismo bidireccionalmente, correspondiendo aquellos individuos que se encuentran en los extremos de la población, esto ocasiona una separación (Choudhuri, 2014).



**Figura 1. Tipos de selección natural:** Aquellas variantes que favorecen un alelo promoviendo la aparición de nuevos fenotipos, se le conoce como selección positiva o darwiniana y aquellas variantes que perjudican y que son eliminadas son llamadas selección negativa (Choudhuri 2014).

Aquellas variantes que favorecen un alelo promoviendo la aparición de nuevos fenotipos, se le conoce como selección positiva o darwiniana y aquellas variantes que perjudican y que son eliminadas son llamadas selección negativa (Choudhuri, 2014).

## Medicina Evolutiva

La medicina evolutiva o también llamada medicina darwiniana es la disciplina que trata de entender por qué la selección natural ha dejado al organismo vulnerable a las enfermedades y establece que las enfermedades que afectan a las poblaciones son causadas por incompatibilidades entre el medio actual y el medio en el que evolucionaron los humanos. Un ejemplo es el gen ahorrador que se basa en el hecho de que nuestros antepasados se adaptaron a entornos específicos, como la hambruna y debido a que hemos conservado los genotipos que habían sido

ventajosos en ambientes ancestrales, ahora se convierten en una desventaja, en la actualidad cuando los alimentos se han vuelto abundantes y son relativamente constantes resultan en las enfermedades modernas, como la diabetes y las enfermedades cardiovasculares (Moltzau, 2020; Naugler, 2008).

En las últimas décadas, estrategias de la medicina evolutiva, como la búsqueda de huellas de selección positiva, han analizado las enfermedades que actualmente afectan a la humanidad, estas son el resultado de un desajuste entre sus variantes genéticas fundamentales y el entorno en el que evolucionan los humanos modernos (Belbin et al., 2018; Byars et al., 2017).

### **SNP con huella de selección positiva**

Las huellas de selección positiva en poblaciones humanas se identifican mediante dos estrategias generales. El primero usa datos de ramificación para detectar la selección positiva entre especies y el segundo usa datos de polimorfismo de un solo nucleótido (SNP), para detectar la selección positiva dentro de una especie. La identificación de SNP que tienen huellas de selección positiva en el genoma humano ha surgido como una estrategia para encontrar variantes potencialmente relevantes en la medicina actual (Biswas & Akey, 2006; Byars et al., 2017).

Dentro de la estrategia basada en SNP, existen varios métodos estadísticos para realizar dichos análisis, por ejemplo, el índice de fijación. Este método compara frecuencias de alelos dentro de una población de interés con otras poblaciones y se usa como referencia para determinar si alguno de los alelos ha pasado por el proceso de selección (Byars et al., 2017).

### **Variante rs17052958**

El polimorfismo rs17052958 es una variante de intrones, que se encuentra en el gen SGCD (compuesto por 100 kb de DNA genómico), este gen codifica para la glicoproteína delta sarcoglicano de la transmembrana del sarcolema. Delta sarcoglicano pertenece al subcomplejo sarcoglicano que forma parte de un complejo de glicoproteína asociado con la distrofina. Este gen se encuentra activo en músculos estriados y lisos. Los niveles más altos del gen se encuentran en el músculo cardíaco

y esquelético. Sus mutaciones se relacionan con la distrofia muscular y miocardiopatía dilatada (MCD) (Sarohi V. et al., 2022; NIH, 2022).

## **PLANTEAMIENTO DEL PROBLEMA**

El creciente uso de escaneos de genoma completo (GWAS) ha permitido identificar diversas variantes genéticas asociadas a distintas enfermedades, como la EAC. A la fecha se ha estudiado el componente genético de la EAC en las poblaciones caucásicas, con más de 160 loci de susceptibilidad identificados (Erdmann et al., 2018). Sin embargo, en la mayoría de estos estudios existe una sobrerrepresentación de las poblaciones europeas y en las poblaciones latinoamericanas ha sido poco estudiada. Esto ha generado diferencias en la fuerza de la asociación entre variantes y la enfermedad en otras poblaciones, como la población mexicana.

Todas las enfermedades tienen una historia evolutiva, y las señales de esas historias están archivadas en nuestros genomas. En este contexto, una perspectiva evolutiva puede ayudar a detectar y validar la variación adaptativa, ampliando nuestra comprensión de cómo actúa la selección natural en enfermedades y rasgos comunes.

## **JUSTIFICACIÓN**

Las ECV son consideradas como la epidemia del siglo XXI, tan solo en México se reporta como la primera causa de muerte, con un total de 226,703 defunciones en el año 2021. El estudio de ECV como la EAC es un proceso de investigación complejo debido a que involucra factores de riesgo genéticos y ambientales. Sin embargo los factores de riesgo no son suficientes para poder explicar el origen de las ECV, por lo tanto frente al gran problema de salud pública, por ello surge la necesidad de abordar esta enfermedad con diferentes enfoques como la medicina evolutiva.

La identificación de SNPs que tienen huellas de selección positiva en el genoma humano permiten servir como marcadores genéticos de enfermedades específicas como la EAC. Sin embargo, la mayoría de los estudios genómicos se han centrado en personas de ascendencia europea y hasta la fecha es posible que no haya suficientes datos sobre los SNP con huellas de selección positiva de otras poblaciones como la mexicana. En este proyecto se pretende identificar si el rs17052958 con huella de selección positiva en población puede estar asociado con las EAC o factores de riesgo cardiometabólico en la población mestiza mexicana.

## **OBJETIVO GENERAL**

Identificar la asociación de la variante intrónica del gen SGCD (rs17052958) con factores de riesgo cardiometabólicos y EAC en población mexicana.

## **OBJETIVOS PARTICULARES**

1. Identificar las frecuencias alélicas de la variante rs 17052958 del gen SGCD en la cohorte GEA.
2. Calcular el equilibrio de Hardy- Weinberg (HW) de la variante rs17052958 en la cohorte GEA.
3. Buscar las asociaciones del rs17052958 del gen SGCD con rasgos cardiometabólicos y la EAC.

## **METODOLOGÍA**

### **a) Población de estudio**

La población es una cohorte de 2800 personas que pertenecen al proyecto de genética de enfermedad aterosclerosa (GEA). La cohorte consta de dos grupos, controles y pacientes. Los criterios de inclusión y exclusión están reportados en Villarreal-Molina et al., 2012.

### **b) Extracción de DNA a partir de sangre**

Se colocaron 3 mL de la sangre en un tubo de centrifuga de 15 mL, se le adicionó 9 mL de solución de lisis celular y se mezcló 10 veces por inversión. Posteriormente se incubó por 5 min. a temperatura ambiente y se centrifugó por 2 min. a 2000 x g. Se eliminó el sobrenadante, se le añadió 300 µL de solución de lisis celular y se homogeneizó por 10s. Después se agregó 1 mL de la solución de precipitación de proteínas y se homogeneizó, posteriormente se centrifugó por 5 min. a 2000 xg. El sobrenadante se colocó en un tubo de centrifuga de 15 mL con 3 mL de isopropanol frío, y se mezcló por inversión durante 50 veces hasta observar el ADN. A continuación se lavó con 3 mL de etanol al 70 % frío, y se centrifugó durante 1 min a 2000 x g. Se eliminó el sobrenadante y el botón que contiene el ADN se dejó secar de 5 a 10 min. a temperatura ambiente. Finalmente se hidrato con 200 µL de agua grado biología molecular.

### **c) Genotipificación**

Se determinó la frecuencia alélica mediante discriminación del genotipo en las muestras de la cohorte del proyecto GEA utilizando RT-PCR en tiempo real con sonda Taq Man ®. Se analizaron las placas en el equipo de PCR de tiempo real en el termociclador ABI Prism 7900 HT. Las temperaturas que se utilizaron en la PCR fueron: en la etapa de desnaturalización fue de 95 °C durante 10 min, seguido por 45



ciclos de desnaturalización a 95°C por 15 segundos, en alineación y extensión la temperatura fue de 60°C durante 1 min. y finalmente la extensión final fue de 60°C durante 30 s. Al terminar la amplificación se realizó la lectura de la placa en el Software Quant studio, se registraron las mediciones de fluorescencia, se determinaron los alelos por muestra, se observaron en el grafico de discriminación alélica y se determinó su genotipo.

#### d) Equilibrio de Hardy Weinberg

Se realizó el cálculo de las frecuencias genotípicas y alélicas, se calculó el Equilibrio de Hardy Weinberg en la página Gene calc (<https://gene-calc.pl/hardy-weinberg-page>) con una significancia de 0.05.

#### e) Análisis Estadístico

Se utilizó en software SPSS v20 para realizar las pruebas estadísticas. Se hizo la comparación de medias en ANOVA de un factor y prueba T de Student, tomando la significancia de 0.05. El análisis de asociación se realizó bajo tres modelos de herencia: aditivo, dominante y recesivo, además se estratifico por sexo. Las variables con diferencias significativas de acuerdo al genotipo fueron analizadas con modelos lineales ajustados por IMC, edad y sexo. Se excluyeron a los participantes diabéticos y para el análisis de enzimas hepáticas se excluyeron aquellas personas con hígado graso. Finalmente para evaluar el si la variante se asocia a mayor riesgo de EAC, se realizó un análisis de regresión logística multinomial, ajustado por IMC, edad y sexo. Los datos se graficaron utilizando programación R 4.2.2 y el Software GraphPad v9.

## RESULTADOS

Para esta variante se genotificaron 1580 individuos del grupo de control y 957 individuos del grupo de casos. El alelo bajo selección positiva en esta variante fue el alelo “G”, cuya frecuencia fue de 58.20 % en controles y 60.03 % en casos (Tabla 1).

**Tabla 1. Frecuencias alélicas y genotípicas en controles der rs17052958**

Genotipo	Controles			Casos		
	n	Frecuencia genotípica	Frecuencia alélica	n	Frecuencia genotípica	Frecuencia alélica
CC	319	0.202	41.80%	165	0.174	39.97%
GC	683	0.432	-	427	0.451	-
GG	578	0.366	58.20%	355	0.375	60.03%
TOTAL	1580	1	100.00%	947	1	100%

El equilibrio de Hardy Weinberg establece que en una población panmíctica de gran tamaño no existen apareamiento preferencias y en ausencia de una fuerza evolutiva (mutación, migración, deriva génica, selección natural) el proceso de la herencia, por sí mismo, no cambia las frecuencias alélicas ni las frecuencias genotípicas de un determinado locus, por lo tanto permanecen constantes de generación en generación (Universidad de Granada, s.f.; concepto definición, 2021). Por ello se calculó el equilibrio de Hardy Weinberg, en la tabla 2 se observa que en los casos (GEAP) solo hay un equilibrio en la frecuencia de los alelos ( $p= 0.17793$ ). Se realizó la comparación con lo con lo reportado el proyecto de mil genomas en población Amerindia y mexicana nacido en los Ángeles y en la Tabla 2 se observa que existe un equilibrio, mientras que en la cohorte de controles no se encontró este equilibrio ( $p= 0.00005$ ).

**Tabla 2.- Equilibrio de Hardy Weinberg**

	<b>P</b>	<b>X2</b>	<b>n</b>
<b>AMR</b>	0,0749	5,18	347
<b>MXL</b>	0,46637	1,53	64
<b>GEAC</b>	0,00005	19,66	1.580
<b>GEAP</b>	0,17793	3,45	947

Prueba probabilística de chi cuadrada (X2) y la significancia (P) en el proyecto de 1000 genomas y la cohorte GEA. AMR: Amerindio y MXL: mexicano nacido en los Ángeles.

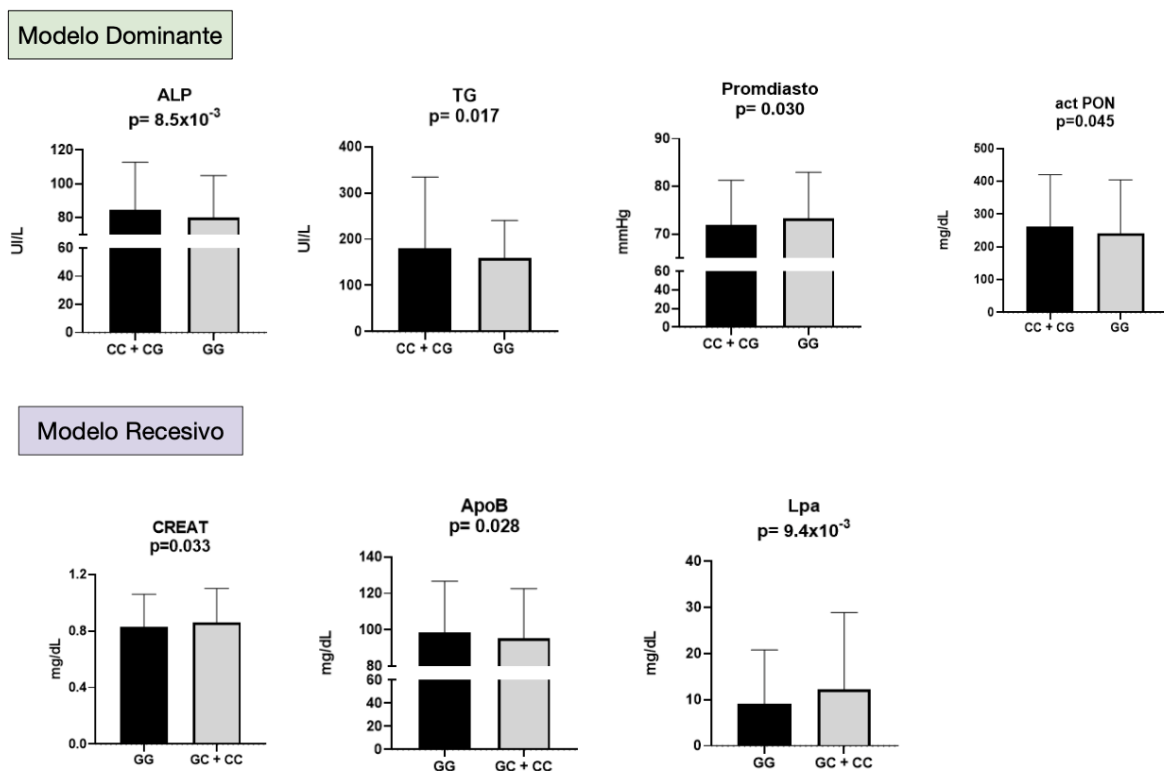
Los resultados de la prueba ANOVA mostraron que individuos con el alelo bajo selección (Tabla 3) presentaron niveles significativamente mayor del valor de la resistencia a la insulina de tejido adiposo (RITA) ( $p = 0.016$ ), así como de niveles séricos de albumina ( $p= 0.048$ ), AST ( $p= 0.035$ ) en población total. Una vez realizada la estratificación por sexo, se observó que en hombres con el alelo “G” presentaron una disminución significativa de los niveles Apolipoproteina A (ApoA) ( $p < 0.05$ ), mientras que las mujeres con este alelo presentaron mayores valores de HOMA, y niveles de insulina, e interleucina 10 (IL-10) ( $p < 0.05$ ). Todas las diferencias reportadas se observaron bajo el modelo aditivo (Tabla 3).

Tabla 3.- Comparación de medias de la variante rs17052958

Variable	Genotipo	Población general 0 =319 1=682 2=578			Hombres 0 =164 1=334 2=279			Mujeres 0 =155 1=349 2=299		
		Media	DE	P	Media	DE	P	Media	DE	P
Insulina	CC	18.98	11.59	0.106	19.98	13.72	0.639	17.09	8.85	<b>0.017</b>
	CG	19.42	10.35		19.23	10.05		19.60	10.64	
	GG	20.66	15.16		20.32	18.21		20.98	11.61	
HOMA	CC	4.57	3.29	0.062	4.94	3.84	0.339	4.17	2.53	<b>0.006</b>
	CG	4.76	3.13		4.77	3.02		4.75	3.24	
	GG	5.48	9.59		5.75	13.26		5.22	3.67	
RITA	CC	10.79	7.73	<b>0.016</b>	10.45	8.64	0.725	11.16	6.57	0.001
	CG	11.31	7.42		10.24	7.38		12.41	7.32	
	GG	12.57	12.35		10.94	14.37		14.14	9.83	
AST	CC	26.94	10.52	<b>0.035</b>	26.99	8.75	0.105	26.90	12.17	0.444
	CG	26.91	10.12		27.70	10.44		26.14	9.75	
	GG	28.52	13.99		29.74	12.83		27.38	14.92	
ApoA	CC	137.61	35.74	0.253	132.55	36.11	<b>0.017</b>	143.01	34.65	0.415
	CG	136.30	34.54		125.00	26.62		147.11	37.69	
	GG	139.68	37.47		131.04	34.98		147.79	37.97	
Albúmina	CC	13.17	28.75	<b>0.048</b>	16.19	35.60	0.244	10.10	19.06	0.039
	CG	17.06	45.87		24.29	63.11		10.15	14.74	
	GG	29.36	165.89		38.31	218.70		20.86	90.38	
IL-10	CC	6.16	68.02	0.305	9.67	95.18	0.460	2.60	11.90	<b>0.045</b>
	CG	3.45	17.71		4.00	21.82		2.92	12.63	
	GG	7.17	45.43		4.51	25.24		9.74	58.59	

Mediante ANOVA de un factor se analizó la comparación de las medias en los diferentes grupos, con o sin la presencia del alelo seleccionado, con parámetros metabólicos en población general. AST: Aspartato aminotransferasa, ALP: fosfatasa alcalina, ApoA: apolipoproteína A, ApoB: Apolipoproteína B, Lpa: Lipoproteína, RITA: Resistencia a la insulina de tejido adiposo, HOMA: índice de resistencia a la insulina e IL\_10: Interleucina 10.

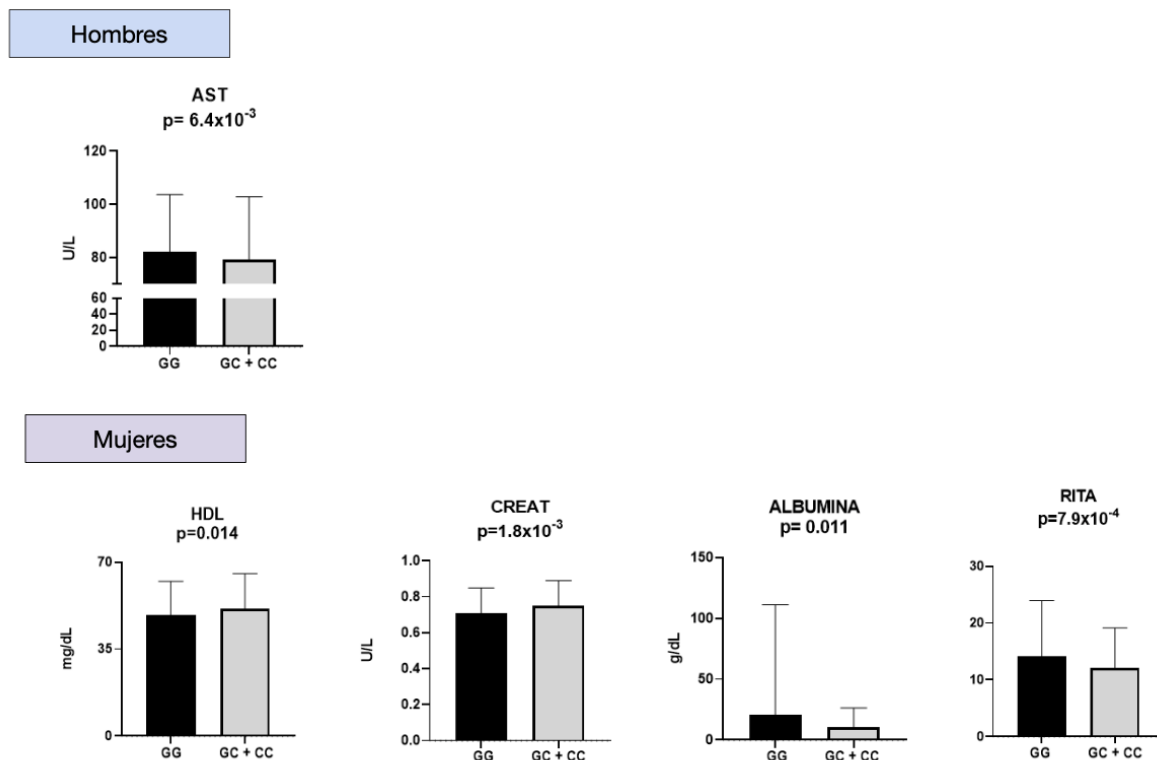
Se realizó el análisis bajo el modelo dominante y el recesivo con ayuda la prueba estadística t de Student en la población total y en la estratificación por sexo. La figura 2 muestra que individuos con el alelo bajo selección “G” en un modelo dominante presentaron niveles significativamente mayor de fosfatasa alcalina (ALP) ( $p= 0.008$ ), triglicéridos (TG) ( $p= 0.017$ ), y actividad enzimática específica de la paraoxonasa (act PON) ( $p= 0.045$ ) y menores valores de presión diastólica ( $p= 0.030$ ) en población general. Mientras que bajo un modelo recesivo, podemos observar que los individuos con el alelo seleccionado presentan una aumento significativo en los niveles séricos de creatinina ( $p= 0.033$ ), Lpa ( $p= 9.4 \times 10^{-3}$ ) y disminución en los niveles de apolipoproteína B ( $p= 0.028$ ).



**Figura 2 .Comparación de medias de la variante rs17052958 en población general bajo modelo dominante y recesivo.** ALP:fosfatasa alcalina, TG :Triglicéridos, Promdiasto: presión diastólica, act PON: Actividad de Paraoxonasa, CREAT: Creatinina, Apo B: apolipoproteína B y Lpa: lipoproteína.

Se estratifico por sexo y solo se encontraron diferencias significativas bajo un modelo recesivo tanto en hombre como en mujeres. La figura 3 muestra que hay una disminución significativa en los niveles de AST ( $p= 6.4 \times 10^{-3}$ ) en los hombres que portan las dos copias del alelo seleccionado “G”, para el caso de las mujeres bajo el modelo recesivo, los datos muestran que los niveles de HDL ( $p= 0.014$ ) y creatinina

( $p= 0.001$ ) disminuyen significativamente, mientras que los niveles de los valores de RITA ( $p= 7.9 \times 10^{-4}$ ) y los niveles séricos de albúmina ( $p= 0.011$ ) se encuentran aumentados (Figura 3).



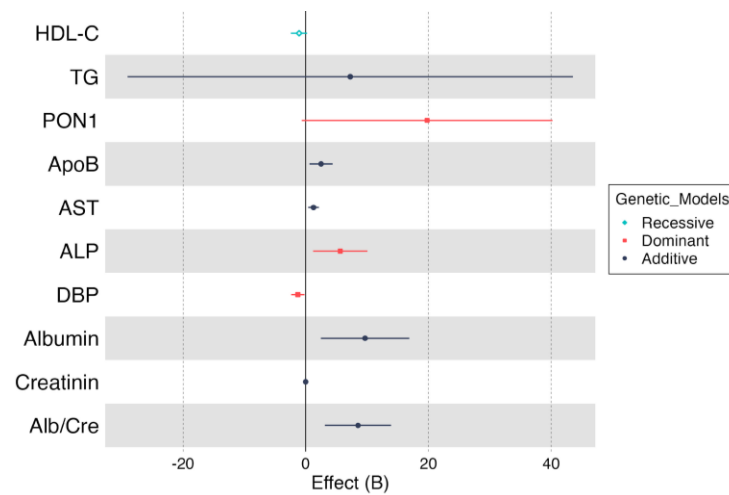
**Figura 3. Comparación de medias de la variante rs17052958 estratificado por sexo bajo un modelo recesivo.**

AST: Aspartato aminotransferasa, HDL: lipoproteínas de alta densidad, TG: Triglicéridos, CREAT: Creatinina, RITA: Resistencia a la inulina de tejido adiposo

Posteriormente, realizamos un análisis de modelos lineales generalizados para encontrar la asociación y calcular el tamaño del efecto del SNP 17052958 sobre las variables analizadas (expresado con la letra B). Solo se reportan las variables con valores de P significativos ( $< 0.05$ ) en población total y estratificadas por sexo.

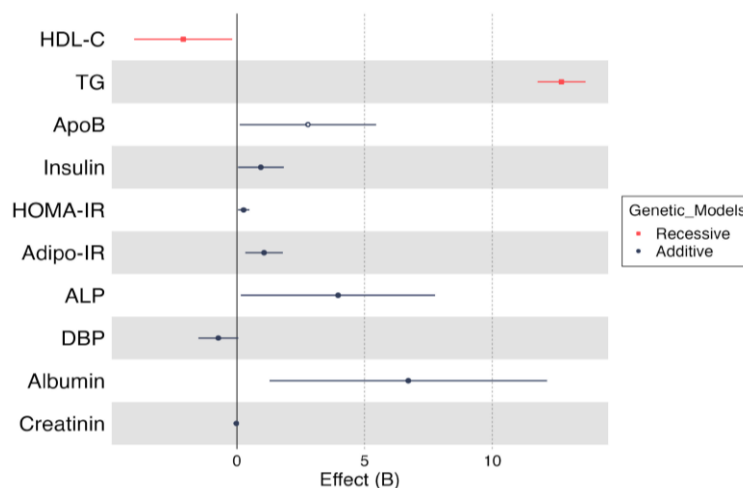
Adicionalmente, todas las variables significativas fueron analizadas bajo sus respectivos modelos de herencia. La figura 4 muestra que la variante rs17052958 en la población total está asociada con altos niveles de TG ( $B= 7.25$  (mg/dL)  $P= 1.5 \times 10^{-3}$ ), Apo B ( $B= 2.50$  (mg/dL)  $P= 0.015$ ), Act PON ( $B= 19.78$   $P= 0.039$ ), AST ( $B= 1.269$  (U/L)  $P= 2.53 \times 10^{-4}$ ), ALP ( $B= 5.625$  (U/L)  $P= 3.2 \times 10^{-3}$ ), albúmina ( $B= 9.673$  (mg/dL)  $P= 0.025$ ) y Alb/Creat ( $B= 8.526$  (mg/dL)  $P= 5.7 \times 10^{-4}$ ) y a menores valores de presión

arterial (B= -1.29 (mmHg) P= 0.0178) y menores niveles de creatinina (B= -0.015 (mg/dL) P= 6.8x10<sup>-3</sup>).



**Figura 4. Asociación del SNP 17052958 con parámetros metabólicos en la cohorte GEA con sus respectivos modelos de herencia.** HDL-C: lipoproteínas de alta densidad, TG: triglicéridos, PON1: Actividad de Paraoxonasa, Apo B: Apolipoproteína B, AST: Aspartato aminotransferasa, ALP: Fosfatasa alcalina, DBP: Presión arterial diastólica branquial.

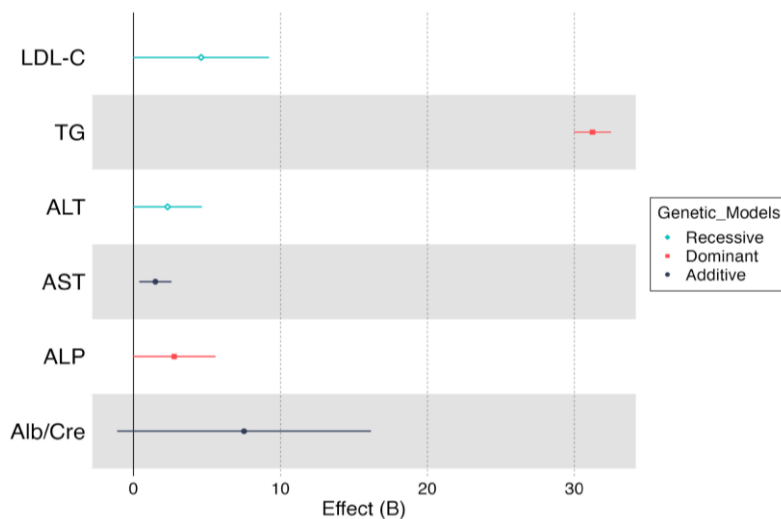
Estratificando por sexo, en mujeres (Figura 5) se asoció significativamente a mayores niveles de TG (B= 12.70 (mg/dL) P= 0.015), HOMA (B= 0.265 (mg/dL) P= 0.024), RITA (B= 1.064 P= 0.011), ALP (B= 3.959 (U/L) P= 0.021), albumina (B= 6.706 (g/dL) P= 0.037) y Alb / Creat (B= 16.42 (mg/dL) P= 0.0044) y a menores niveles de HDL (B= -2.10 (mg/dL) P= 0.026), presión diastólica (B= -0.730 (mmHg) P= 0.041) y creatinina (B= -0.024 (mg/dL) P= 0.0011) .



**Figura 5. Asociación del SNP 17052958 con parámetros metabólicos en la cohorte GEA con sus respectivos modelos de herencia en mujeres.** HDL-C: lipoproteínas de alta densidad, TG: triglicéridos, Apo B: Apolipoproteína B, HOMA-IR: Evaluación de modelo homeostático del índice de resistencia a la insulina,

Adipo-IR: Índice de resistencia a la insulina de tejido adiposo, ALP: fosfatasa alcalina, DBP: presión arterial diastólica branquial.

En hombres (Figura. 6) se asoció significativamente a mayores niveles de TG (B= 31.23 (mg/dL) P= 6.3x10<sup>-3</sup>) AST (B= 1.497 (U/L) P= 0.0205 ), ALP (B= 2.777 (U/L) P= 0.022) y Alb/Cre (B= 7.524 (mg/dL) P= 0.023).



**Figura 6. Asociación del SNP 17052958 con parámetros metabólicos en la cohorte GEA con sus respectivos modelos de herencia en hombres.** HDL-C: lipoproteínas de alta densidad, TG: triglicéridos, ALT: Alanina aminotransferasa, AST: aspartato aminotransferasa, ALP: fosfatasa alcalina.

Para evaluar el riesgo de padecer EAC se determinó mediante un análisis de regresión logística multinomial. El análisis reportó que la variantes rs17052958 aumenta significativamente 1.26 veces el riesgo de EAC (p= 0.35) en la población general bajo un modelo dominante.

**Tabla 4. Riesgo de padecer EAC prematura para la variante rs17052958**

Asociación del gen SGCD (rs17052958) con la EAC prematura.					
	Modelos	Frecuencia Alélica		OR (95% CI)	P
		Controles	EAC		
Cohorte GEA	Aditivo	58.2	60.0	1.26 (0.99-1.61)	0.060
	<b>Dominante</b>			<b>1.26 (1.02-1.57)</b>	<b>0.035</b>
	Recesivo			1.07 (0.90-1.27)	0.469
Hombres	Aditivo	57.5	59.9	0.99 (0.79-1.24)	0.916
	Dominante			1.27 (0.98-1.64)	0.069

	Recesivo			1.06 (0.86-1.31)	0.565
Mujeres	Aditivo	58.8	60.7	1.31 (0.83-2.07)	0.255
	Dominante			1.30 (0.85-1.97)	0.224
	Recesivo			1.09 (0.79-1.51)	0.601

Las asociaciones con la EAC prematura se ajustaron por edad, sexo e IMC. Los valores de p para los modelos aditivo, dominante y recesivo se calcularon mediante regresión logística. OR: odds ratio; IC: intervalo de confianza.

## DISCUSIÓN

La medicina evolutiva es el campo de estudio que busca comprender por qué los organismos desarrollan enfermedades desde la perspectiva de la genética de poblaciones. El objetivo es determinar si las enfermedades que afectan a una población son el resultado de una mala adaptación entre el entorno actual y el entorno en el que evolucionó el ser humano (Montealegre, 2018). Por ello, la estrategia de buscar asociaciones entre variantes genéticas que exhiben señales de selección positivas y fenotipos tienen la ventaja de enfocar el análisis en sitios que probablemente estén asociados con los fenotipos relevantes y, por lo tanto, se corrobora la existencia de variantes funcionales en la enfermedad (Vasseur E. & Quintana L., 2013).

La estrategia de buscar asociaciones entre variantes genéticas que presentan señales de selección positiva y fenotipos tiene la ventaja de enfocar el análisis en sitios que tienen una alta probabilidad de estar relacionados con un fenotipo relevante, lo que respalda la existencia de variación funcional para una enfermedad. Además, este enfoque supera las limitaciones de otros métodos, como el enfoque de genes candidatos (Zhu & Zhao, 2007) y los estudios de GWAS (estudios de asociación del genoma completo, por sus siglas en inglés) (Sabeti et al., 2006), que generan costos elevados debido a la falta de conocimiento previo sobre los genes candidatos en una población extensa (Sabeti et al., 2006; Zhu & Zhao, 2007).

En el análisis de Hardy Weinberg es un principio que establece que la población está en equilibrio cuando sus frecuencias alélicas y genotípicas no cambian en el tiempo, el cálculo realizado en la cohorte de los casos (GEAP) demuestra que no existe un equilibrio de Hardy Weinberg, esto puede deberse al exceso de heterocigotos, como lo menciona Graffelman Jan et.al., 2017 en su artículo.



Se encontraron algunos artículos en donde se reportaron que el gen SGCD se asocia con distintos parámetros que concuerdan con lo obtenido en esta investigación, en el artículo escrito por Kaess B. et al., 2011 reportaron que el gen SGCD se asoció con el tamaño medio de partículas de HDL y el número de partículas de HDL, el rs17052958 si modificó los niveles de HDL en mujeres ya que lo disminuye, asimismo observamos en el análisis un aumento en los niveles de triglicéridos, Apo B y LDL, al tener un aumento en estos parámetros lipídicos no benéficos para la salud, pueden tener un efecto en la EAC siendo más susceptibles a formar la capa de aterosclerosis, asimismo se encontró que Aragam K. et. al., 2022 realizaron un metaanálisis de GWAS en japoneses y de los loci encontrados que tienen significancia para la EAC se encontraba el gen SGCD.

También observamos que tiene un efecto en la presión arterial ya que disminuye la presión diastólica, en el artículo reportado por Surendran et.al., 2020 hicieron un análisis de variantes asociadas a la presión arterial, el gen SGCD se encuentra significativo, este puede ser un factor de riesgo para la enfermedad cardiovascular.

Finalmente Saxena et al., 2103 reportaron que el gen SGCG se asocia a Diabetes Mellitus 2, este gen se encuentra en el mismo complejo de sarcoglicano que el gen SGCD, al encontrarse en este mismo complejo transmembrana puede que compartan la misma asociación para poder ser más susceptibles a tener diabetes mellitus 2 (DM2), en el análisis estadístico aumentan los niveles de insulina, HOMA y RITA en mujeres, estos resultados pueden modificar los niveles de glucosa en sangre e insulinoresistencia y pueden causar la DM2.

## CONCLUSIÓN

El gen SGCD se ha asociado con afecciones cardiacas, como la cardiomiopatía dilatada, sin embargo, este estudio representa la primera evidencia de una asociación entre la variante rs17052958 del gen SGCD con la EAC. La asociación rs17052958 con parámetros cardiometabólicos como lípidos, podrían ser consistentes con la teoría del gen ahorrador. Por otro lado, las asociaciones con el perfil diabético y daño renal podrían tener un impacto en el desarrollo de DM2 y contribuir con la EAC. La estrategia del enfoque evolutivo permitió identificar la asociación rs17052958 con parámetros cardiometabólicos y EAC, por lo que es importante se replique en otros estudios.

## REFERENCIAS

Amir, R. , Ain Q. S., Maqsood W. , Munir F., Nosheen F. A. S., Ahmad J.,(2022) Chapter 6 Pan-genomics of plant pathogens and its applications. Academic Press. <https://doi.org/10.1016/B978-0-12-817076-2.00006-8>

Álvarez, C. L., Frías I., J. Fernández de B. O. J. y Diaz A., M. (2020). Prevalencia de factores de riesgo cardiovascular en trabajadores de un hospital terciario de Madrid. *Revista de la Asociación Española de Especialistas en Medicina del Trabajo*, 29(4), 274-288. [http://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1132-62552020000400274&lng=es&tlng=es](http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1132-62552020000400274&lng=es&tlng=es).

Aragam, K.G., Jiang, T., Goel, A. *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet* 54, 1803–1815 (2022). <https://doi.org/10.1038/s41588-022-01233-6>

Belbin, G. M., Nieves-Colón, M. A., Kenny, E. E., Moreno-Estrada, A., & Gignoux, C. R. (2018). Genetic diversity in populations across Latin America: Implications for population and medical genetic studies. *Current Opinion in Genetics & Development*, 53, 98–104. <https://doi.org/10.1016/j.gde.2018.07.006>

Byars, S. G., Huang, Q. Q., Gray, L. A., Bakshi, A., Ripatti, S., Abraham, G., Stearns, S.C. Inouye, M. (2017). Genetic loci associated with coronary artery disease harbor evidence of selection and antagonistic pleiotropy. *PLoS Genet*, 13(6), e1006328, 1- 27. <https://doi.org/10.1371/journal.pgen.1006328>

Britannica, T. Editors of Encyclopaedia (2022). natural selection. *Encyclopedia Britannica*. <https://www.britannica.com/science/natural-selection>

Biswas, S., & Akey, J. M. (2006). Genomic insights into positive selection. *Trends in Genetics: TIG*, 22(8), 437–446. <https://doi.org/10.1016/j.tig.2006.06.005>

Choudhuri Supratim (2014) *Bioinformatics for Beginners, Genes, Genome, Molecular Evolution, Databases and Analytical Tools*.

CONABIO (2023), Biodiversidad mexicana. <https://www.biodiversidad.gob.mx/especies/queson>

Concepto de definición(2021). Equilibrio de Hardy-Weinberg. <https://conceptodefinicion.de/equilibrio-hardy-weinberg/>

Elselvier.[https://www.researchgate.net/figure/Three-types-of-natural-selection-ADirectional-selection-B-stabilizing-selection\\_fig5\\_340583975](https://www.researchgate.net/figure/Three-types-of-natural-selection-ADirectional-selection-B-stabilizing-selection_fig5_340583975)

Elizondo C. D., Alfaro A.K., Ureña S. F., Guzmán R. A., Lubker C. E. y Alfaro Vellanero M. (2020) Factores de riesgo cardiovascular. *Revista ciencia y salud*.<https://revistacienciaysalud.ac.cr/ojs/index.php/cienciaysalud/article/view/108/188>

Erdmann, J., Kessler, T., Munoz Venegas, L., & Schunkert, H. (2018). A decade of genome-wide association studies for coronary artery disease: The challenges ahead. *Cardiovascular Research*, 114(9), 1241–1257. <https://doi.org/10.1093/cvr/cvy084>

Graffelman J, Jain D, Weir B. A genome-wide study of Hardy-Weinberg equilibrium with next generation sequence data. *Hum Genet*. 2017 Jun;136(6):727-741. doi: 10.1007/s00439-017-1786-7. Epub 2017 Apr 3. PMID: 28374190; PMCID: PMC5429372.

INEGI (2022) Características de defunciones registradas en México durante 2021. [de:https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2022/dr/dr2021\\_07.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2022/dr/dr2021_07.pdf)

Iturbe U. (2010) Adaptaciones y adaptación biológica, 8-9. [https://www.researchgate.net/publication/260363619\\_ADAPTACIONES\\_Y\\_ADAPTACION\\_BIOLOGICA\\_REVISADAS](https://www.researchgate.net/publication/260363619_ADAPTACIONES_Y_ADAPTACION_BIOLOGICA_REVISADAS)

Kaess BM, Tomaszewski M, Braund PS, Stark K, Rafelt S, Fischer M, Hardwick R, Nelson CP, Debiec R, Huber F, Kremer W, Kalbitzer HR, Rose LM, Chasman DI, Hopewell J, Clarke R, Burton PR, Tobin MD, Hengstenberg C, Samani NJ. Large-scale candidate gene analysis of HDL particle features. *PLoS One*. 2011 Jan 21;6(1):e14529. doi: 10.1371/journal.pone.0014529. PMID: 21283740; PMCID: PMC3024972.

Khera, A.,V y Kathiresan, S. (2017) Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*, 18(6):331-344. doi: 10.1038/nrg.2016.160.

Moltzau Anderson J, Horn F.(2020) Defining evolutionary medicine. *Ecol Evol*. 2020 Oct 1;10(20):10930-10936. doi: 10.1002/ece3.6825.

Montealegre, C. de C. (2018). Patología/defensa en medicina evolutiva. *Acta Médica Costarricense*, 60(1), 4–6.

NIH (2022) SGCD Sarcoglycan delta [Homo Sapiens (human)].<https://www.ncbi.nlm.nih.gov/gene/6444> NIH (2022), Single nucleotide

polymorphism (SNPS). <https://www.genome.gov/genetics-glossary/Single-NucleotidePolymorphisms>

Naugler (2008). Evolutionary medicine: update on the relevance to family practice. *Can Fam Physician*. 2008 Sep;54(9):1265-9. PMID: 18791103; PMCID: PMC2553465.

OMS. (2020). Las 10 principales causas de defunción. <https://www.who.int/es/newsroom/fact-sheets/detail/the-top-10-causes-of-death>

OMS(2022), Enfermedades cardiovasculares. [https://www.who.int/es/healthtopics/hypertension/cardiovascular-diseases#tab=tab\\_1](https://www.who.int/es/healthtopics/hypertension/cardiovascular-diseases#tab=tab_1)

Saxena R, Saleheen D, Been LF, Garavito ML, Braun T, Bjornes A, Young R, Ho WK, Rasheed A, Frossard P, Sim X, Hassanali N, Radha V, Chidambaram M, Liju S, Rees SD, Ng DP, Wong TY, Yamauchi T, Hara K, Tanaka Y, Hirose H, McCarthy MI, Morris AP; DIAGRAM; MuTHER; AGEN; Basit A, Barnett AH, Katulanda P, Matthews D, Mohan V, Wander GS, Singh JR, Mehra NK, Ralhan S, Kamboh MI, Mulvihill JJ, Maegawa H, Tobe K, Maeda S, Cho YS, Tai ES, Kelly MA, Chambers JC, Kooner JS, Kadowaki T, Deloukas P, Rader DJ, Danesh J, Sanghera DK. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. *Diabetes*. 2013 May;62(5):1746-55. doi: 10.2337/db12-1077. Epub 2013 Jan 8. PMID: 23300278; PMCID: PMC3636649. <https://pubmed.ncbi.nlm.nih.gov/23300278/>

Shahjehan D. R. & Bhutta S. B. (2022) Coronary Artery Disease. <https://www.ncbi.nlm.nih.gov/books/NBK564304/>

Sarohi V., Srivastava S. & Basak (2022) A Comprehensive Outlook on Dilated Cardiomyopathy (DCM): State-Of-The-Art Developments with Special Emphasis on OMICS-Based Approaches. *J. Cardiovasc. Dev. Dis.* **2022**, *9*, 174. <https://doi.org/10.3390/jcdd9060174>

Surendran, P., Feofanova, E.V., Lahrouchi, N. *et al.* (2020) Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. *Nat Genet* *52*, 1314–1332 (2020). <https://doi.org/10.1038/s41588-020-00713-x>

Talseth B. & Scott R. (2011). Genetic variation and its role in malignancy. *Int J Biomed Sci*. 2011 Sep;7(3):158-71. PMID: 23675233; PMCID: PMC3614837.

Universidad de Granada (s.f.). Genética de poblaciones. Departamento de genética. <http://wpd.ugr.es/~rnavajas/wp-content/uploads/2019/06/poblaciones.pdf>

Vasseur E, Quintana-Murci L. (2013) The impact of natural selection on health and disease: uses of the population genetics approach in humans. *Evol Appl*. 2013

Jun;6(4):596-607. doi: 10.1111/eva.12045. Epub 2013 Jan 21. PMID: 23789027; PMCID: PMC3684741.

Zhu, M., & Zhao, S. (2007). Candidate Gene Identification Approach: Progress and Challenges. *International Journal of Biological Sciences*, 420–427. <https://doi.org/10.7150/ijbs.3.420>