

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

DEPARTAMENTO DE BIOTECNOLOGÍA



VARIANTES GENÉTICAS IMPLICADAS EN CARDIOMIOPATÍA DILATADA Y SU ASOCIACIÓN CON HIPERTENSIÓN ARTERIAL

TRABAJO FIN DE MÁSTER EN BIOTECNOLOGÍA BIOMÉDICA

ALUMNO/A: María Guerrero Zambrano

TUTOR/A: Felipe Javier Chaves Martínez

Curso Académico: 2020-2021

VALENCIA, 6 de septiembre de 2021

RESUMEN

Las enfermedades cardiovasculares son la principal causa de muerte a nivel mundial, siendo responsables de 17 millones de muertes cada año aproximadamente. Un tipo de enfermedad cardiovascular son las miocardiopatías, que comprenden cambios estructurales y funcionales del músculo cardíaco causando una disfunción mecánica y/o eléctrica del mismo. Hay varios tipos de miocardiopatías, siendo la más frecuente la miocardiopatía dilatada (CMD). La CMD tiene tanto causas genéticas como no genéticas. Aproximadamente el 35% de los casos están causados por genes mutados que codifican proteínas sarcómeras, citoesqueléticas o de envoltura nuclear, además, hay factores ambientales como la presencia de hipertensión arterial (HTA) o diabetes que modifica el riesgo de desarrollar esta patología. El objetivo de este estudio es analizar la importancia de variantes genéticas previamente asociadas con el desarrollo de esta patología en la regulación de los niveles de presión arterial. Para ello se seleccionaron un total de 91 variantes implicadas en el desarrollo de la CMD y que podrían tener cierta implicación con la HTA, y a través de la tecnología de secuenciación de amplicones cortos se llevó a cabo la secuenciación de un total de 1381 individuos pertenecientes al estudio de la población Horteiga. Posteriormente, se realizó un análisis bioinformático que permitió identificar un total de 186 variantes. A pesar de la cantidad de variantes encontradas, tan solo 4 de estas (rs17597065, rs4129218, rs573914858 y rs746745349) salieron significativas en el estudio caso/control con pacientes hipertensos/no hipertensos. Son necesarios otros estudios en poblaciones de mayor tamaño que permitan validar estos resultados y puedan identificar de forma significativa otras variantes asociadas con ambas patologías ya que es importante poder establecer los mecanismos que puedan tener en común el desarrollo de estas dos enfermedades.

PALABRAS CLAVES

Cardiomiopatía dilatada; Hipertensión; Variantes Genéticas; Secuenciación de Segunda Generación; Riesgo Cardiovascular.

ABSTRACT

Cardiovascular diseases are the number 1 cause of death globally, accounting for approximately 17 million deaths each year. One type of cardiovascular diseases are cardiomyopathies which is mainly characterized by structural and functional changes of the heart muscle causing mechanical and/or electrical dysfunction of the heart muscle. There are several types of cardiomyopathies, the most common being dilated cardiomyopathy (DCM). DCM has both genetic and non-genetic causes. Approximately 35% of cases are caused by mutated genes encoding sarcomeric, cytoskeletal or nuclear envelope proteins. Furthermore, there are environmental factors such as the presence of arterial hypertension (AHT) or diabetes that modify the risk of developing this pathology. The aim of this study is to analyze the importance of genetic variants previously associated with the development of this pathology in the regulation of blood pressure levels. A total of 91 variants implicated in the development of DCM and that could have some involvement with hypertension were selected. A total of 1381 individuals belonging to the Horteiga study were sequenced through targeted sequencing. Subsequently, a bioinformatics analysis identified 186 variants, of which only 4 (rs17597065, rs4129218, rs573914858 y rs746745349) were significant in the case/control study with hypertensive/non-hypertensive patients. Further studies in larger populations are needed to validate these results and to significantly identify other variants associated with both pathologies, because it is important to be able to establish the mechanisms that may be common to the development of these two diseases.

KEYWORDS

Dilated Cardiomyopathy; Hypertension; Genetic Variants; Next-Generation Sequencing; Cardiovascular Risk.

AGRADECIMIENTOS

A Javier por haber confiado en mi y darme la oportunidad de formar parte de este proyecto.

A Ana Barbara por haberme ayudado y orientado en todo el proyecto.

A Fran por sus incansables explicaciones y por su disposición en todo momento para ayudarme en todas las dudas que tuviese.

A Azahara por enseñarme un lo que es el mundo de la bioinformática.

A Elena y Evelyn por ser mis compañeras de trabajo, risas y locuras en las tardes más largas de laboratorio.

A Soraya, Iris, Rebeca, en general a todo el laboratorio 23, por haberme hecho mi estancia mucho más amena y por haberme aportado un poco de sus conocimientos para mi aprendizaje.

A Kike, Alicia, Tani por los *runs* y por ayudarnos cuando teníamos alguna duda.

ÍNDICE

1. INTRODUCCIÓN	10
1.1 ENFERMEDADES CARDIOVASCULARES	10
1.2 MIOCARDIOPATÍAS. TIPOS	10
1.3. MIOCARDIOPATÍA DILATADA	11
1.3.1. ETIOLOGÍA Y EPIDEMIOLOGÍA DE LA MIOCARDIOPATÍA DILATADA.....	11
1.3.2 MECANISMO MOLECULAR Y FISIOPATOLOGÍA	12
1.3.3 CAUSAS GENÉTICAS.....	14
1.4. HIPERTENSIÓN ARTERIAL.....	16
1.4.1 TIPOS.....	16
1.4.1.1. HIPERTENSIÓN PRIMARIA O ESENCIAL.....	16
1.4.1.2. HIPERTENSIÓN SECUNDARIA.....	16
1.4.2. ETIOLOGÍA Y EPIDEMIOLOGÍA DE LA HIPERTENSIÓN.....	17
1.4.3. FISIOPATOLOGÍA DE LA HIPERTENSIÓN.....	18
1.4.4. CAUSAS GENÉTICAS	19
1.5. RELACIÓN ENTRE LA MIOCARDIOPATÍA DILATADA Y LA HIPERTENSIÓN.....	20
2. OBJETIVOS Y ANTECEDENTES.....	21
3. MATERIALES Y MÉTODOS.....	21
3.1. POBLACIÓN DE ESTUDIO	21
3.2. DISEÑO EXPERIMENTAL	21
3.3. OBTENCIÓN DE MUESTRAS	22
3.4. PREPARACIÓN DE LIBRERÍAS	23
3.4.1. SELECCIÓN DE GENES Y SNPs	23
3.5. DISEÑO DE CEBADORES.....	23
3.6. OPTIMIZACIÓN PCR	24
3.6.1. PCR1: AMPLIFICACIÓN EN MULTIPLEX.....	24
3.6.2. PCR2: MARCADO CON BARCODES	25
3.7. ELECTROFORESIS CAPILAR.....	26
3.8. PURIFICACIÓN DE LIBRERÍAS	26
3.9. CUANTIFICACIÓN DE LIBRERÍAS	27
3.10. SECUENCIACIÓN	27
3.11. ANÁLISIS BIOINFORMÁTICO	29
3.12. ANÁLISIS ESTADÍSTICO	30
4. RESULTADOS Y DISCUSIÓN.....	31
4.1. CARACTERÍSTICAS DE LA POBLACIÓN HORTEGA.....	31

4.2. DISEÑO DE CEBADORES.....	31
4.3. OPTIMIZACIÓN DE LAS CONDICIONES DE PCR1 Y PCR2.....	31
4.4. VALIDACIÓN DE LAS LIBRERÍAS CON LA POBLACIÓN HORTEGA	34
4.5. CALIDAD DE LA SECUENCIACIÓN.....	34
4.6. ANÁLISIS DE VARIANTES.....	35
4.6.1. rs17597065.....	36
4.6.2. rs4129218.....	37
4.6.3. rs573914858 y rs746745349	38
4.6.4. rs10497520	39
4.7. LIMITACIONES DEL ESTUDIO.....	40
5. CONCLUSIONES.....	42
6. BIBLIOGRAFÍA.....	43

ÍNDICE DE FIGURAS

Figura 1: Diferencias entre las paredes del miocardio de un corazón normal y uno que sufre miocardiopatía dilatada	11
Figura 2: Causas de la CMD	12
Figura 3: Esquema de genes relacionados con la CMD y otras patologías	15
Figura 4: Porcentaje de afectados por la hipertensión arterial (HTA) según su edad	17
Figura 5: Esquema del diseño experimental	22
Figura 6: Condiciones utilizadas en la PCR 1	25
Figura 7: Esquema estructura de los productos obtenidos de ambas PCR con la incorporación de los barcodes	25
Figura 8: Condiciones utilizadas en la PCR 2	26
Figura 9: Ejemplo de una electroforesis capilar y electroferograma a través del sistema Qiaxcel	26
Figura 10: Proceso de purificación de las librerías con bolas magnéticas	27
Figura 11: Ejemplo recta patrón para la cuantificación de librerías	27
Figura 12: Esquema de la amplificación del ADN por PCR-puente para formar los grupos o clusters.	28
Figura 13: Esquema secuenciación por síntesis	28
Figura 14: Resultados de la electroforesis capilar con los cebadores a diferentes concentraciones.	32
Figura 15: Resultados de la electroforesis capilar con diferentes aditivos.....	32
Figura 16: Resultados de electroforesis capilar tras la prueba con diferentes aditivos.	33
Figura 17: Resultados de la electroforesis capilar tras las modificaciones en los aditivos 4 y 5.	33
Figura 18: Análisis de coberturas en los fragmentos amplificados con la condición adecuada para la optimización de la PCR.	34
Figura 19: (A) Validación de muestras de la Población Hortega mediante electroforesis capilar. (B) Comparación entre el pool de la Población Hortega sin purificar y purificado.....	34
Figura 20: Calidad de lecturas antes y después del proceso de trimming.....	35
Figura 21: Porcentaje de variantes encontradas en la población Hortega.	36

ÍNDICE DE TABLAS

Tabla 1: Factores que influyen en padecer Hipertensión primaria.....	16
Tabla 2: Causas relacionadas con el desarrollo de la Hipertensión secundaria.....	17
Tabla 3: Listado de genes, SNPs y CNVs seleccionados.....	23
Tabla 4: Características población Hortega.....	31
Tabla 5: Relación entre la variante rs17597065, los genotipos de la población Hortega y los pacientes con HTA	36
Tabla 6: Relación entre la variante rs4129218, los genotipos de la población Hortega y los pacientes con HTA	37
Tabla 7: Relación entre las variantes, los genotipos de la población Hortega y los pacientes con HTA	38
Tabla 8: Relación entre la variante rs10497520, los genotipos de la población Hortega y los pacientes con HTA	39

ABREVIATURAS

ARVC: Miocardiopatía arritmogénica del ventrículo derecho

HTA: Hipertensión arterial

PA: Presión arterial

ECV: Enfermedades cardiovasculares

PAS: Presión arterial sistólica

PAD: Presión arterial diastólica

VD: Ventrículo derecho

CMD: Miocardiopatía dilatada

VI: Ventrículo izquierdo

MCH: Miocardiopatía hipertrófica

MCR: Miocardiopatía restrictiva

RAAS: Sistema renina-angiotensina-aldosterona

IL-11: interleucina 11

TGF- β : Factor de crecimiento transformante β

NGS: Next Generation Sequencing

ET1: Endotelina 1

NO: Óxido nítrico

SNS: Sistema nervioso simpático

NA: Noradrenalina

LVH: Hipertrofia ventricular izquierda

GWAS: Estudio de Asociación de Genoma Completo

SNP: Polimorfismo de un único nucleótido

WES: Secuenciación masiva de exoma

OXPHOS: Fosforilación oxidativa

ADNmt: ADN mitocondrial

WGS: Secuenciación completa del genoma

PCR: Reacción en cadena de la polimerasa

CNV: Variación en el número de copias

ECG-LV: Masa ventricular izquierda derivada del electrocardiograma

LncRNA: *Long non-coding RNAs*

1. INTRODUCCIÓN

1.1. ENFERMEDADES CARDIOVASCULARES

Las enfermedades cardiovasculares (ECV) se deben a trastornos del corazón y de los vasos sanguíneos. Según la Organización Mundial de la Salud [OMS], las ECV son la principal causa de muerte a nivel mundial, siendo responsables de 17 millones de muertes cada año aproximadamente. Las ECV se producen casi por igual en hombres y mujeres. Además, estas afectan en mayor medida a los países con bajos ingresos, produciéndose en estos más del 80% de las muertes. Se prevé que de aquí a 2030, fallezcan unos 23,6 millones de personas por algún tipo de ECV como pueden ser cardiopatías y los accidentes cerebrovasculares (OMS, 2021).

Esta enfermedad tiene importantes costes económicos que afectan a las personas que la padecen, los sistemas sanitarios y a las sociedades de todo el mundo (Timmis *et al.*, 2020). A pesar de tener una alta mortalidad, en muchos casos se puede prevenir corrigiendo ciertos factores que influyen en el estilo de vida de las personas como pueden ser el tabaquismo, el consumo nocivo de alcohol, la inactividad física, la dieta y la obesidad (OMS, 2012). Además, llevando a cabo un diagnóstico precoz, un buen tratamiento de los episodios agudos y la posterior atención, se puede llegar a mejorar mucho el pronóstico de estas enfermedades (Kivimäki y Steptoe, 2018).

Entre las ECV destacan la insuficiencia cardíaca, enfermedades cerebrovasculares, miocardiopatías, arteriopatías periféricas, cardiopatía coronaria, cardiopatía congénita entre otras. Existen diferentes factores de riesgo para el desarrollo de algunas de ellas como la diabetes, hipertensión o hiperlipidemia. Entre estos, la hipertensión arterial (HTA) es la que presenta una mayor causalidad y tiene una alta prevalencia de exposición (Fuchs y Whelton, 2020).

1.2. MIOCARDIOPATÍAS. TIPOS.

Las miocardiopatías son un tipo de ECV que comprenden cambios estructurales y funcionales del músculo cardíaco causando una disfunción mecánica y/o eléctrica del mismo (Schultheiss *et al.*, 2019). Según las Guías de práctica clínica de la Sociedad Española de Cardiología, las miocardiopatías son un conjunto de procesos muy heterogéneos que tienen en común la afectación del músculo cardíaco. Hay varios tipos de clasificación para las miocardiopatías, en este caso vamos a seguir la clasificación más actual según sus fenotipos morfológicos y funcionales específicos, de forma que se dividen en varios tipos (Perry *et al.*, 2008):

- Miocardiopatía arritmogénica del ventrículo derecho (ARVC): en este caso se produce una disminución del ventrículo derecho (VD) con una infiltración fibrograsa que se puede ver a través de técnicas de imagen (McNally *et al.*, 2013).
- Miocardiopatía dilatada (CMD): se caracteriza por el deterioro de la función sistólica y por la reducción de la función del ventrículo izquierdo (VI) (McNally *et al.*, 2013).
- Miocardiopatía hipertrófica (MCH): la limitación del gasto cardíaco se produce por el engrosamiento del miocardio debido a alteraciones de llenado y salida del flujo sanguíneo. A diferencia de la CMD, en este caso la MCH se caracteriza por el aumento de grosor de la pared del VI que suele afectar al tabique que separa ambos ventrículos (McNally *et al.*, 2013).

- Miocardiopatía restrictiva (MCR): es de las más difíciles de detectar ya que a simple vista el corazón parece normal morfológicamente. Se caracteriza por el deterioro de la función diastólica. En este caso se produce un ligero aumento del grosor de la pared o una pequeña disminución de la fracción de eyección del VI (McNally *et al.*, 2013).
- Miocardiopatías no clasificadas: todas las que no se pueden englobar en los tipos anteriores.

A partir de la clasificación anterior se puede diferenciar las miocardiopatías en dos grupos principales, por un lado, las miocardiopatías primarias que se limitan únicamente al corazón y tienen causas genéticas, no genéticas o mixtas (genéticas y no genéticas), en este grupo se incluye la CMD. Por otro lado, las miocardiopatías secundarias que son las que afectan al corazón como parte de una enfermedad sistémica como por ejemplo puede ser la amiloidosis, hemocromatosis etc. (Reichart *et al.*, 2019).

1.3. MIOCARDIOPATÍA DILATADA

Dentro de los tipos de miocardiopatías, la más frecuente es la CMD (Fundación Española del Corazón [FEC], 2021) que es en la que vamos a centrar este estudio y es uno de los motivos más frecuentes de trasplante cardiaco en niños y adultos (Sanbe 2013).

La CMD es un tipo de enfermedad cardiovascular que se caracteriza por la afectación progresiva del músculo cardíaco produciendo una dilatación del VI y una disfunción sistólica en ausencia de condiciones de carga anormales conocidas o enfermedad de las arterias coronarias significativa (Seferović *et al.*, 2019). De esta forma el músculo cardíaco se debilita y adelgaza siendo incapaz de bombear la sangre de forma normal por todo el cuerpo (Figura 1) (FEC, 2021).

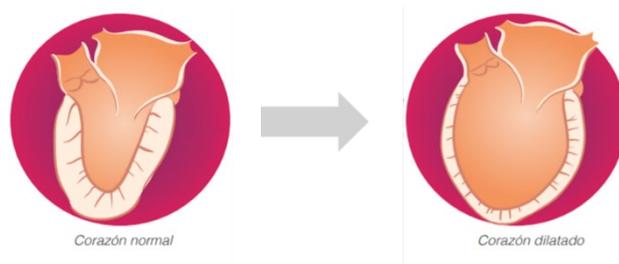


Figura 1: Diferencias entre las paredes del miocardio de un corazón normal y uno que sufre miocardiopatía dilatada (FEC, 2021).

1.3.1. ETIOLOGÍA Y EPIDEMIOLOGÍA DE LA MIOCARDIOPATÍA DILATADA

Esta enfermedad es la tercera causa más común de insuficiencia cardíaca con una prevalencia de 36/100.000 habitantes (Seferović *et al.*, 2019), teniendo una mayor prevalencia en los países subdesarrollados debido a la falta de diagnóstico (Reichart *et al.*, 2019). La mortalidad asociada a esta patología en 2010 fue de 5.9/100.000 habitantes, pero esto ha ido disminuyendo conforme han ido avanzando las formas de diagnóstico y terapias. La EuroHeart Failure Survey II (EHFS II) determinó que la CMD tiene una mayor frecuencia en hombres que en mujeres y que es la miocardiopatía más común entre los niños menores de 18 años (Schultheiss *et al.*, 2019), a pesar de ello esta enfermedad afecta a personas de todas las edades, pero suele ser más frecuente en el rango de 20-50 años (FEC, 2019).

La CMD puede ser para algunos pacientes asintomática, en cambio para otros puede presentar síntomas propios de una insuficiencia cardíaca como la dificultad para respirar, hinchazón en tobillos y abdomen, y cansancio. Los síntomas generalmente aparecen de manera lenta, aunque también pueden manifestarse de forma repentina (FEC, 2020). Aunque la CMD pueden causar los síntomas de insuficiencia cardíaca, cada tipo requiere una estrategia de tratamiento específica, una intervención quirúrgica y unos cambios de estilo de vida diferentes para controlar de forma eficaz la enfermedad (Balije *et al.*, 2016).

Las causas de esta patología son muy heterogéneas y por ello es necesario un estudio diagnóstico detallado para diagnosticar la causa subyacente y excluir otras afecciones que tengan los mismos fenotipos (Reichart *et al.*, 2019). Para el diagnóstico de dicha patología es fundamental llevar a cabo la medición del tamaño del VI y, además, la fracción de eyección tiene que ser <50% (Jordan *et al.*, 2021). Como se ha explicado anteriormente las causas pueden ser tanto genéticas como no genéticas, incluso existiendo superposiciones entre ambos grupos. También puede darse en casos aislados o en pacientes con citopatías mitocondriales y trastornos metabólicos hereditarios (Figura 2) (Perry *et al.*, 2008).

Además, hay causas no genéticas relacionadas con un mayor riesgo de padecer dicha patología como pueden ser infecciones virales, enfermedades autoinmunes, ciertos medicamentos o la exposición a diferentes toxinas, anomalías endocrinas y metabólicas y el embarazo (Figura 2) (Seferović *et al.*, 2019).

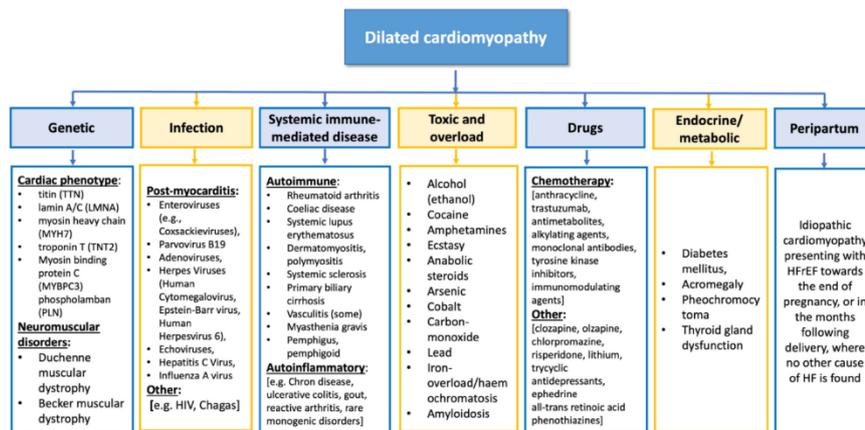


Figura 2: Causas de la CMD (Seferović *et al.*, 2019).

A parte de la etiología para desarrollar esta enfermedad, también existen varios factores de riesgo como son la hipertensión, herencia africana, género masculino, enfermedades infecciosas y factores ambientales (Zeng *et al.*, 2009).

1.3.2. MECANISMO MOLECULAR Y FISIOPATOLOGÍA

La CMD se caracteriza principalmente por la remodelación biomolecular, la producción de matriz extracelular y la hipertrofia de los cardiomiocitos (Barison *et al.*, 2015).

En la CMD existe una lesión miocárdica que produce una disminución del volumen sistólico y del gasto cardíaco, un aumento de la presión diastólica final y un deterioro del llenado ventricular.

Además, se produce una mayor resistencia vascular sistémica, un aumento de la presión venosa y la disminución de la distensibilidad arterial. Tanto la fuerza que ejerce la pared del corazón antes de la contracción (precarga) como la que hace para expulsar la sangre del corazón (poscarga) aumentan, y dicho aumento acaba dando lugar a elevados niveles de tensión en la pared del corazón (Schultheiss *et al.*, 2019).

En las primeras etapas de desarrollo de esta enfermedad, la reducción de la contractilidad se compensa con una remodelación del VI, sin embargo, cuando va pasando el tiempo el VI se dilata y acaba volviéndose disfuncional (Caviedes *et al.*, 2018). En la remodelación del miocardio no solo están implicados los miocitos, sino también las células no miocíticas y la matriz extracelular (Barison *et al.*, 2015).

Para conocer los mecanismos moleculares de esta patología es necesario llevar a cabo experimentos con cultivos celulares (*in vitro*) y con animales (*in vivo*) utilizando técnicas de silenciamiento y sobreexpresión de genes (Dadson *et al.*, 2017). Los experimentos en modelos animales han demostrado que debido a la CMD la dilatación se produce en el VI como consecuencia de la remodelación y fibrosis. En los pacientes se observa que el VI tiene una forma esférica y hay tejido cicatricial ya que al principio de la sobrecarga lo que hace es aumentar la longitud de los sarcómeros para aumentar su contractilidad, pero cuando se sobrepasa el nivel óptimo, se termina promoviendo la hipertrofia (Schultheiss *et al.*, 2019). Con la CMD se produce un fenotipo conocido como el 'VI con parches' en el cual los cardiomiocitos se encuentran intercalados entre calcificaciones intermitentes y parches fibróticos y necróticos (Reichart *et al.*, 2019).

En general no hay distinciones cuando la CMD es adquirida o hereditaria, pero se ha observado que en las formas genéticas se produce un aumento del grosor de la pared del VI, hay unas trabeculaciones prominentes del VI y se produce un agrandamiento de la cámara del VD (Peters *et al.*, 2020).

La regulación de la función cardíaca está controlada principalmente por la regulación neurohormonal que se activa ante la disminución de la fracción de eyección cardíaca (Reichart *et al.*, 2019). El deterioro inicial cardíaco se compensa a través de la estimulación de los sistemas simpático adrenérgico y renina-angiotensina-aldosterona (RAAS) además de la movilización de los péptidos natriuréticos, de esta forma se va aumentando el tono vascular y la retención de agua y sodio (Dadson *et al.*, 2017). En cambio, cuando estos sistemas se desregulan se produce una remodelación hipertrófica y una insuficiencia cardíaca cuyo marcador principal es el péptido natriurético cerebral (Caviedes *et al.*, 2018). Dentro del RAAS se encuentra la angiotensina II que se ha visto que activa dos marcadores implicados en los mecanismos patológicos de remodelación de la CMD como son el factor de crecimiento transformante β (TGF- β) y a la interleucina 11 (IL-11). Ambos se expresan en con bajos niveles cuando se trata de un corazón sano, pero aumentan sus niveles cuando está enfermo. La expresión de ambos está relacionada con proteínas de la matriz extracelular las cuales aumentan en el remodelado del miocardio y la fibrosis del mismo. En algunos tratamientos se usa la inhibición del TGF- β , aunque se ha visto que tiene muchos efectos secundarios. Otro de los enfoques prometedores podría ser la inhibición de la IL-11 (Reichart *et al.*, 2019).

A veces resulta muy complicado intervenir en el proceso de la enfermedad con antelación para prevenir daños irreversibles porque a menudo la progresión es lenta desde el inicio de los daños cardíacos hasta llegar a los síntomas de insuficiencia cardíaca (Schultheiss *et al.*, 2019).

1.3.3. CAUSAS GENÉTICAS

La CMD a diferencia de otras miocardiopatías tiene una gran heterogeneidad de locus con muchos posibles genes implicados en su fenotipo (Jordan *et al.*, 2021). El descubrimiento de los genes responsables de las alteraciones en la CMD hace que haya un diagnóstico más preciso y temprano de la CMD genética. Además, con un diagnóstico precoz se puede mejorar las opciones de tratamientos para estos pacientes (Harakalova *et al.*, 2015). Se han identificado varios genes que cuando mutan se vinculan con el desarrollo de la CMD (Schultheiss *et al.*, 2019). Varios hallazgos indican que las mutaciones genéticas y esporádicas de ciertas proteínas musculares como las estructurales, miofibrilares y los reguladores de Ca²⁺ pueden acabar produciendo CMD (Sanbe 2013).

Aproximadamente el 35% de los casos se producen por causas genéticas destacando las mutaciones en genes que codifican proteínas sarcómeras, citoesqueléticas o de la envoltura nuclear (Reichart *et al.*, 2019). A pesar de esto, distinguir entre variantes que sean patogénicas o no sigue siendo un desafío importante sin resolver (Schultheiss *et al.*, 2019).

La CMD puede ser idiopática cuando aparece de forma espontánea o familiar cuando sucede en dos o más familiares relacionados (Taylor *et al.*, 2006). Normalmente se hereda como un rasgo autosómico dominante, aunque también puede heredarse con un patrón ligado al cromosoma X, autosómico recesivo o mitocondrial (de origen materno) que es poco común (Reichart *et al.*, 2019). Algunos estudios sugieren que entre el 15-30% de los pacientes podrían tener una CMD familiar si se sometiesen a sus familiares a un cribado clínico (Schultheiss *et al.*, 2019).

Los defectos en ciertas proteínas citoesqueléticas pueden dar lugar al desarrollo de la CMD. El citoesqueleto contiene una red de fibras sólidas y tubos que se entrecruzan dando lugar al andamiaje celular que mantiene todo el volumen citoplasmático. Este andamiaje es importante para la protección de los miocitos del estrés mecánico externo y para la transmisión de fuerza desde el sarcómero a la matriz extracelular. Cuando hay defectos en dicho andamiaje se produce una disminución de la transmisión de fuerza y/o la resistencia al estrés mecánico. Además, se sabe que la disfunción contráctil de las miofibrillas tiene un papel importante en la progresión de esta enfermedad (Schultheiss *et al.*, 2019).

Tras los análisis genéticos se han descubierto más de 60 genes relacionados con la progresión de la CMD, aunque solo un total de 12 se clasificaron con una relación clara en el desarrollo de esta enfermedad (*BAG3*, *DES*, *DSP*, *FLNC*, *LMNA*, *MYH7*, *PLN*, *RBM20*, *SCN5A*, *TNNC1*, *TNNT2*, *TTN*) y 7 tenían una implicación moderada en la CMD (*ACTC1*, *ACTN2*, *JPH2*, *NEXN*, *TNNI3*, *TPM1* y *VCL*). De estos genes también hay algunos que se encuentran relacionados con otros tipos de miocardiopatías como son la ARVC o la HCM y en canalopatías incluyendo el síndrome de QT largo y Brugada S (Figura 3) (Jordan *et al.*, 2021). Es difícil distinguir entre las variantes patogénicas y las benignas entre todos los genes asociados con esta enfermedad (Peters *et al.*, 2020).

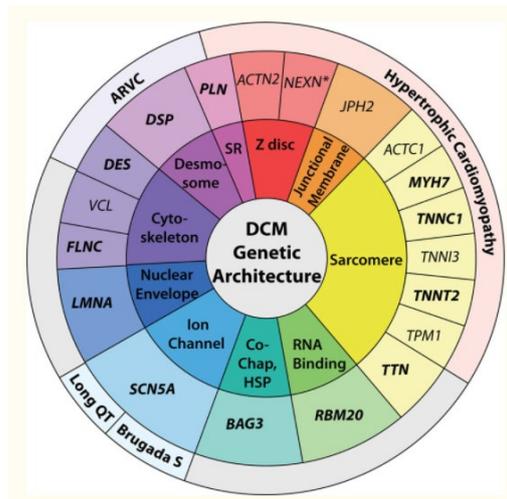


Figura 3: Esquema de genes relacionados con la CMD y otras patologías (Jordan *et al.*, 2021). Los genes resaltados es que tienen mayor implicación en el desarrollo de la CMD mientras que los menos resaltados tienen una menor influencia.

Varias mutaciones en las proteínas del sarcómero se han asociado con la progresión de esta patología, como puede ser en la actina α del músculo cardíaco 1 que es codificada por *ACTC1*; cadenas de miosina codificadas por *MYH6*, *MYH7*, *MYL2* y *MYL3*; la proteína C de unión a miosina codificada por *MYBPC3*. También se asocian con el desarrollo de la CMD las mutaciones en genes que codifican troponinas cardíacas como es el caso de *TNNT2*, *TNNC1* y *TNNI3*. Además, las mutaciones en *MYH7* alteran la función de puente cruzado del citoesqueleto y la unión de actina-miosina ya que codifica a la cadena pesada de la β -miosina, y las que se producen en *TNNT* alteran las propiedades de viscoelasticidad (Schultheiss *et al.*, 2019). Mutaciones en *SCN5A* dan lugar a anomalías de la conducción cardíaca (Japp *et al.*, 2016).

A pesar de que los genes pueden estar implicados en varias vías genéticas al final producen un único fenotipo de CMD. Esto es importante a la hora de la práctica clínica ya que cualquier causa de la CMD subyace en cierta parte de una insuficiencia cardíaca (Hershberger *et al.*, 2013).

Hay proteínas no contráctiles como las chaperonas 70 codificadas por *BAG3* que cuando mutan participan en el desarrollo de la CMD provocando alteraciones en las vías de señalización celular que participan en la respuesta cardíaca. También mutaciones en las proteínas como la desmoplaquina codificada por *DSP* tienen cierta implicación con la CMD aunque se asocian más comúnmente con la ARVC (Schultheiss *et al.*, 2019).

Se necesita una amplia gama de exámenes tanto invasivos como no invasivos para identificar con precisión la etiología de esta enfermedad. El desarrollo de la secuenciación de próxima generación (NGS, *Next Generation Sequencing*) ha permitido el descubrimiento de numerosos genes y variantes genéticas asociadas al desarrollo de esta patología haciendo que el diagnóstico y los recursos médicos para tratar la CMD sean mejores y más personalizados según el paciente (Schultheiss *et al.*, 2019).

1.4. HIPERTENSIÓN ARTERIAL

La hipertensión arterial (HTA) es un trastorno en el cual los vasos sanguíneos están sometidos a una tensión persistente lo que hace que puedan acabar dañándose. La presión que ejerce la sangre sobre las arterias al ser bombeada por el corazón es lo que se conoce como tensión arterial, y cuánto más alta es dicha tensión, más esfuerzo tiene que realizar el corazón para seguir bombeando sangre. Con la hipertensión, se somete al corazón a una resistencia mayor haciendo que este tenga que aumentar su masa muscular, si esto no se controla, a la larga puede producir infartos de miocardio, insuficiencia cardíaca, ensanchamiento del corazón o accidentes cerebrovasculares entre otros (OMS, 2015).

La presión arterial (PA) se calcula en milímetros de mercurio (mm Hg) y se mide en forma de dos números. Por un lado, el primero de ellos corresponde a la presión arterial sistólica (PAS) que es el número más alto y corresponde a la presión máxima que ejerce el corazón cada vez que late sobre las paredes de las arterias. Por otro lado, el segundo número se refiere a la presión arterial diastólica (PAD) que es el número más bajo y es la cantidad de presión que ejerce el corazón sobre las paredes de los vasos mientras descansa entre cada latido. La presión arterial normal en un adulto se define como una PAS de 120 mm Hg y una PAD de 80 mm Hg (OMS, 2015).

Según el Comité Nacional Conjunto de los Estados Unidos de América sobre Prevención, Detección, Evaluación y Tratamiento de la Hipertensión Arterial (JNC), la hipertensión se define como una presión arterial mayor o igual a 140 y una presión diastólica mayor o igual a 90mm Hg en al menos 2 situaciones diferentes separadas por 1 semana al menos (Perumareddi, 2019).

1.4.1. TIPOS

La hipertensión se divide en dos tipos: hipertensión primaria o esencial y la hipertensión secundaria.

1.4.1.1. HIPERTENSIÓN PRIMARIA O ESENCIAL

Se define como una elevación crónica de la PA sin tener alguna causa conocida. En este tipo de hipertensión tienen un papel importante tanto los factores genéticos como el historial familiar, edad, sexo etc y los factores ambientales como puede ser la obesidad, dieta, sedentarismo etc. (Tabla 1) representando aproximadamente el 90-95% de los casos de HTA (Perumareddi, 2019).

Tabla 1: Factores que influyen en padecer Hipertensión primaria.

FACTORES GENÉTICOS	FACTORES AMBIENTALES
Historial familiar	Obesidad
Edad	Dieta
Sexo	Ingesta de sodio
Raza	Sedentarismo

1.4.1.2. HIPERTENSIÓN SECUNDARIA

Es mucho menos frecuente que la hipertensión primaria, tan solo representa entre el 5-10% de los casos. Se define como una elevación persistente de PA superior a 140/90 mm Hg debida a

una causa conocida, es decir, se encuentra asociada a otros estados patológicos o causas exógenas. (Perumareddi, 2019).

Las causas de la hipertensión secundaria suelen ser las siguientes, de mayor a menos prevalencia (Tabla 2): (Perumareddi, 2019).

Tabla 2: Causas relacionadas con el desarrollo de la Hipertensión secundaria.

CAUSAS				
OSA (apnea obstructiva del sueño)	Enfermedad renal	Aldosterismo primario	Alcohol, drogas etc.	Otras causas: Feocromocitoma, hipotiroidismo, acromegalia etc.
25-50%	5-34%	8-20%	2-4%	

En la infancia la forma más prevalente de HTA era la secundaria, pero esto ha ido cambiando y ahora la HTA primaria se está convirtiendo en la dominante, teniendo prácticamente la misma prevalencia que la secundaria en niños mayores de 6 años. Se planteó la hipótesis de que el origen de los adultos que tienen hipertensión está en su infancia, ya que casi la mitad de los adultos con HTA tenían valores elevados de PA en su infancia (Çakici *et al.*, 2020).

1.4.2. ETIOLOGÍA Y EPIDEMIOLOGÍA DE LA HIPERTENSIÓN

En la actualidad, hay un gran porcentaje de personas en riesgo de padecer una enfermedad cardiovascular ya que tienen una PA subóptima o inadecuadamente controlada. Existe una asociación entre el aumento de la presión arterial y el riesgo de padecer una enfermedad cardiovascular y otras enfermedades como pueden ser la insuficiencia cardíaca congestiva o una enfermedad cerebrovascular (Perumareddi, 2019).

La HTA es una de las afecciones más comunes a las que se enfrentan los médicos de atención primaria, siendo una de las principales causas de muerte prematura en todo el mundo (Perumareddi, 2019). Se estima que 1,13 mil millones de personas padecen hipertensión en todo el mundo, y menos de 1 de cada 5 tienen el problema bajo control. La mayoría de los casos se encuentran en países de bajos y medianos ingresos (OMS, 2019). La HTA es un grave problema de salud pública debido a su elevada prevalencia, su impacto clínico y su elevado coste socioeconómico. En Estados Unidos se calcula que 1 de cada 3 adultos son hipertensos. Esta enfermedad tiene una mayor prevalencia según va aumentando la edad, encontrándose el mayor porcentaje en las personas de más de 60 años (Figura 4) (Perumareddi, 2019). Se calcula que complicaciones procedentes de la hipertensión causan 9,4 millones de muertes anualmente en todo el mundo, representado el 45% de muertes por cardiopatías (OMS, 2015). Existe una gran cantidad de pruebas que demuestran que la reducción de la presión arterial reduce sustancialmente la morbimortalidad prematura (Williams *et al.*, 2018).

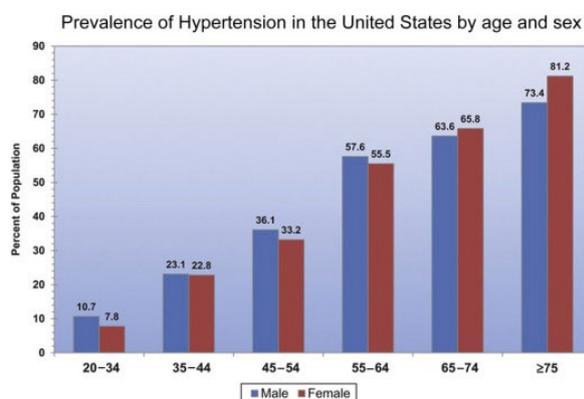


Figura 4: Porcentaje de afectados por la hipertensión arterial (HTA) según su edad (Perumareddi, 2019).

En España, la HTA es una de las enfermedades que más ha crecido desde 2001. Hay aproximadamente 14 millones de personas de las cuales unos 4 millones permanecen sin diagnosticar ya que esta enfermedad no siempre tiene por qué tener claros síntomas. Según la OMS, la HTA es una enfermedad silenciosa, letal y que rara vez produce síntomas en los pacientes, aunque a veces puede provocar dolor de cabeza, vértigos, dificultad respiratoria, palpitaciones del corazón, hemorragias nasales etc. (OMS, 2015). A nivel mundial, una de las prioridades de la salud es diagnosticar, tratar y controlar las enfermedades no transmisibles, entre estas se encuentra la HTA (OMS, 2013).

1.4.3. FISIOPATOLOGÍA DE LA HIPERTENSIÓN

Hay muchos factores ambientales y genéticos que participan en el desarrollo de la HTA. Se sabe que el aumento de la vasoconstricción, los cambios en la pared arterial y a la regulación inadecuada del volumen conducen a un aumento de la PA (Perumareddi, 2019). Hay diferentes vías celulares que contribuyen al desarrollo de la hipertensión (Slivnick y Lampert., 2019).

La HTA se caracteriza principalmente por una disfunción endotelial por el desequilibrio de los factores relajantes de los vasos sanguíneos y los vasoconstrictores. Entre los factores relajante encontramos el factor hiperpolarizante del endotelio y el óxido nítrico (NO) y como vasoconstrictor principalmente endotelinas (Wagner-Grau *et al.*, 2010).

Se produce una disminución de la prostaciclina vasodepresora y un incremento intracelular de los niveles de tromboxano A2 vasoconstrictor (Wagner-Grau *et al.*, 2010).

Las endotelinas son péptidos vasoconstrictores constituidas por tres isoformas, de la que destaca por su capacidad de vasoconstricción la endotelina 1 (ET1). Dicha isoforma tiene ciertas funciones sobre el tono vascular, la producción de matriz celular y la excreción renal de sodio y agua. La ET1, aparte de ser sintetizada en el endotelio puede tener otras fuentes como son las células epiteliales, macrófagos etc., y su concentración extracelular se encuentra regulada principalmente por su internalización, y su aclaramiento por el receptor ETB endotelial. La ET1 puede actuar de forma autocrina o paracrina y gran variedad de factores pueden modular su expresión como son factores de crecimiento, sustancias inflamatorias, agentes vasoactivos etc. Sus principales funciones son la regulación de la proliferación celular y el remodelado vascular, pudiendo producir hiperplasia e hipertrofia del músculo liso vascular (Wagner-Grau *et al.*, 2010).

Una de las cosas que influyen en el mecanismo fisiopatológico para el desarrollo de la hipertensión son los cambios hormonales, sobre todo en mujeres. El estrés neurohormonal también tiene un papel crucial en pacientes hipertensos participando en la remodelación y en la insuficiencia cardíaca (Slivnick y Lampert., 2019). El aumento de PA posmenopausia es probable que se deba a los cambios drásticos hormonales que sufren las mujeres en ese periodo. Los estrógenos inducen vasodilatación, aumentando la disponibilidad NO, modulando el sistema RAAS (aumentando la concentración de angiotensinógeno) y el sistema nervioso simpático (SNS) e inhibiendo procesos de remodelado vascular (Di Giosia *et al.*, 2018). El RAAS puede inducir estrés oxidativo dando lugar a la disfunción endotelial. Dicho sistema comprende 4 angiotensinas (I, II, III y IV) las cuales se forman a partir del angiotensinógeno, destacando la angiotensina II que participa en diversas acciones relacionadas con el desarrollo de la hipertensión como puede ser

la contracción del músculo liso vascular arterial y venosos, secreción de aldosterona, aumento del estrés oxidativo, producción de ET1 y de prostaglandinas vasoconstrictoras como es la tromboxano A2 (Wagner-Grau *et al.*, 2010).

Cuando se produce la estimulación de los β -receptores adrenérgicos se produce un incremento de la frecuencia cardíaca, del gasto cardíaco, de la secreción de renina y de la actividad de RAAS. Sin embargo, se ha demostrado que la estimulación de RAAS aumenta el flujo de salida del SNS promoviendo la liberación de noradrenalina (NA) y amplificando la capacidad de respuesta del receptor adrenérgico (Mann 2018). La liberación de NA desde los terminales nerviosos permiten medir la actividad del nervio simpático, y el aumento de los niveles de NA contribuye al aumento de la PA y a la hipertensión sobre todo en adultos menores de 40 años (Hering *et al.*, 2017).

Hay hormonas gastrointestinales que tienen una importante acción vascular como por ejemplo el péptido vasoactivo, la colecistokinina, la sustancia P que actúan como vasodilatadores y la coherina que actúa como vasoconstrictora. Es probable, que estas hormonas tengan un papel en la regulación de la HTA (Wagner-Grau *et al.*, 2010).

La sobrecarga de presión puede provocar la hipertrofia ventricular izquierda (LVH). Estos cambios dan lugar a la disfunción diastólica que acaban dando lugar a altas presiones de llenado en el VI. Además, los altos niveles de renina y aldosterona juegan un papel en la LVH y remodelación fibrótica (Slivnick y Lampert., 2019).

1.4.4. CAUSAS GENÉTICAS

Se sabe que hay factores genéticos que influyen al desarrollo de la hipertensión, aunque el mecanismo exacto sigue siendo desconocido (Slivnick y Lampert., 2019). La hipertensión se trata de una enfermedad heterogénea en la que además de existir una variación genética hay diversos factores que también afectan como son la edad, sexo, etnia etc. Varias investigaciones en estudios de gemelos y familiares desvelaron que la heredabilidad de esta patología estaba entre el 30 y 50% (Russo *et al.*, 2018).

La mayoría de los estudios se han basado en asociaciones de todo el genoma (GWAS) ya que este tipo de estudio es muy útil para tipificar múltiples polimorfismos de un único nucleótido (SNP) en cohortes muy grandes (Ng *et al.*, 2018). En los últimos años se han identificado más de 100 variantes implicadas en los niveles de PA (Russo *et al.*, 2018). La aclaración de los genes que forman parte de la hipertensión en la herencia mendeliana demuestra que las variantes raras con efectos sustanciales y dichos genes pueden estar asociados con vías que controlan la homeostasis del sodio (Munroe *et al.*, 2013).

En los estudios basados en GWAS se encuentra la hipótesis de “variante común-enfermedad común” pero la considerar todas las variantes con una alta frecuencia, el porcentaje de variabilidad de la PA basada en las variantes genéticas solo representa el 2-3%. Si nos basamos en el escaso control genético sobre las variantes de alta frecuencia había dos hipótesis sobre la asociación de variantes a la PA: una de ellas era que se había estimado una heredabilidad incorrecta y la otra y más probable es que los alelos son muy heterogéneos y poco comunes. Esto hacía que los estudios de GWAS tuviesen grandes limitaciones, por ello con la revolución de las tecnologías NGS se llevó el cambio de enfoque de variantes comunes a variantes raras

(Russo *et al.*, 2018). Los SNPs encontrados a través de GWAS pueden indicar una región cromosómica amplia donde hay más SNPs en desequilibrio. Algunos de los loci que se han relacionado con la PA a través de GWAS son *NPR3*, *SLC4A7* y *SCLC39A8* (Ng *et al.*, 2018).

Otro enfoque de estudio es la secuenciación del exoma completo (WES) en la que solo se secuencian las regiones codificantes del genoma. Se estima que en estas regiones se encuentra el 85% de las mutaciones que producen enfermedades. En este caso, se realizó un estudio con unos 18.000 individuos aproximadamente africanos y europeos y se asociaron 95 variantes de codificación en *CLCN6* con la disminución de PA entre otros (Russo *et al.*, 2018).

La utilización del método de secuenciación completa del genoma (WGS) se considera el estudio más completo de NGS para examinar los efectos de las variantes raras asociadas a la hipertensión. A pesar de ellos, es muy complicado ya que necesita una cantidad de muestra muy grande lo que haría este método muy costoso (Russo *et al.*, 2018).

1.5. RELACIÓN ENTRE LA MIOCARDIOPATÍA DILATADA Y LA HIPERTENSIÓN

La hipertensión puede ser uno de los principales factores de riesgo ya que la CMD puede ir acompañada de esta patología, siendo mayor la prevalencia de hipertensión en pacientes con CMD que en la población general (Zeng *et al.*, 2009). Según la OMS, la CMD pueden provocar morbilidad y mortalidad debido a complicaciones como la insuficiencia cardíaca y arritmias (Reichart *et al.*, 2019). Algunos parámetros invasivos y no invasivos de la función diastólica pueden proporcionar información para el pronóstico de pacientes con CMD y un gran porcentaje de estos pacientes tendrán una mejora con la terapia temprana de la función sistólica del VI (Baliye *et al.*, 2016).

Se sabe que diferentes factores ambientales como la presencia de hipertensión o diabetes modifica el riesgo de desarrollar un gran número de problemas cardíacos, incluida la CMD (Reichart *et al.*, 2019). Se desconoce el pronóstico cuando la CMD está acompañada de fluctuaciones de la presión arterial en ausencia de diabetes y enfermedad coronaria (Baliye *et al.*, 2016). Las causas tóxicas, metabólicas e inmunológicas se han relacionado con la MCD, así como la hipertensión y la valvulopatía. Algunas variantes como el caso de *TTN* en combinación con hipertensión se espera que resulten en una CMD (McNally *et al.*, 2013).

Tras algunos estudios se cree que las mutaciones en el ADN mitocondrial (ADNmt) afectan a la patogenia de la hipertensión y de la CMD, aunque el papel del genoma mitocondrial en ambas enfermedades no está del todo claro. Las mutaciones en el ADNmt que causan un menor funcionamiento de las enzimas de fosforilación oxidativa (OXPHOS) acaban dando como resultado la reducción de la generación de ATP y un aumento de especies reactivas de oxígeno (ROS) que pueden llegar a ser responsables de la aparición de la hipertensión, de la disfunción contráctil cardíaca, insuficiencia cardíaca y finalmente la muerte (Zhu *et al.*, 2016).

2. OBJETIVOS

En este proyecto se realiza un estudio poblacional general de los marcadores descritos previamente asociados con la CMD, utilizando el método de NGS para identificar los polimorfismos de interés. Analizar la importancia de variantes genéticas previamente asociadas con el desarrollo de esta patología en la regulación de los niveles de presión arterial es importante para poder establecer los mecanismos que puedan tener en común el desarrollo de estas dos enfermedades. Así, genes que pueden afectar al desarrollo del corazón o modificarlo en presencia de determinados factores, podrían afectar al desarrollo del sistema circulatorio y provocar el desarrollo de niveles alterados de presión arterial; conociéndose ya que varios de estos genes modulan los niveles de presión arterial.

En este estudio se seleccionaron un total de 91 variantes de 51 genes diferentes que estaban asociados en el desarrollo de la CMD. Estos estudios han permitido seleccionar un panel de genes con el fin de identificar nuevas variantes y validar las descritas anteriormente gracias a la tecnología de secuenciación masiva de amplicones.

3. MATERIALES Y MÉTODOS

3.1. POBLACIÓN DE ESTUDIO

Las muestras seleccionadas provenían del estudio Horteiga, un estudio nacional poblacional impulsado en 1997 a través de una selección aleatoria de personas asignadas al Hospital Universitario Río Horteiga en Valladolid de entre 14 a 85 años. Para la selección de muestras se llevó a cabo unos cuestionarios de los que se obtuvo 11423 respuestas, y de estos, posteriormente se seleccionó aleatoriamente a 1500 pacientes. Las muestras tienen un seguimiento de 14 años en los cuales los pacientes han pasado por diferentes fases para la evaluación de la mortalidad y la incidencia de enfermedades crónicas.

El objetivo principal de este estudio es investigar la asociación transversal y prospectiva de factores genéticos, metabólicos y factores de riesgo ambientales y enfermedades crónicas, con especial interés en las enfermedades cardiovasculares. El estudio fue aprobado por los correspondientes comités de investigación y éticos, y todos los participantes dieron su consentimiento informado.

3.2. DISEÑO EXPERIMENTAL

En este estudio se utiliza la tecnología de secuenciación NGS para identificar los polimorfismos de interés, en concreto la tecnología de secuenciación de amplicones cortos, la cual se trata de una técnica dirigida (*targeted sequencing*) para secuenciar y analizar un número acotado de genes o pequeños fragmentos del ADN donde se localizan los SNPs seleccionados, como es nuestro caso. Este procedimiento permite estudiar variantes exónicas, intrónicas e intergénicas previamente seleccionadas. Esta técnica tiene una buena relación coste-eficacia y al ser dirigida permite obtener una mayor profundidad de secuenciación para reconocer variantes con baja frecuencia. Además, tiene una mayor agilidad en el procesamiento de los datos bioinformáticos

y en la interpretación de los resultados si lo comparamos con otras técnicas como son WGS y WES.

Para empezar con este flujo de trabajo, se diseñaron cebadores específicos para amplificar nuestras regiones de interés y se preparó el material genómico con la concentración deseada. Antes de realizar esta técnica es necesario una optimización de la generación de librerías de amplicones a través de dos amplificaciones por reacción en cadena de la polimerasa (PCR). La primera de ellas se trata de una PCR selectiva la cual permitió seleccionar la región genómica de interés a través de cebadores específicos e incluir una cola de nucleótidos, y la segunda PCR se denomina nominativa o PCR de marcado, cuyo objetivo era identificar cada una de las muestras mediante etiquetas moleculares (*Barcodes*). Para obtener las condiciones más óptimas de la PCR se probaron diferentes concentraciones del pool de cebadores y diferentes aditivos, y se observó la amplificación obtenida a través de una electroforesis capilar. Una vez optimizadas las librerías, se llevo a cabo su purificación y cuantificación para posteriormente secuenciarlas. Tras la secuenciación, se pudo comparar condiciones y se decidió cuál es la que mayor cobertura tenía. Por último, con la condición adecuada se llevó a cabo la amplificación de los SNPs de interés en las muestras poblacionales y su posterior secuenciación. Finalmente, se realizó el análisis bioinformático y la asociación de las variantes obtenidas con diferentes parámetros de riesgo cardiovascular (Figura 5).

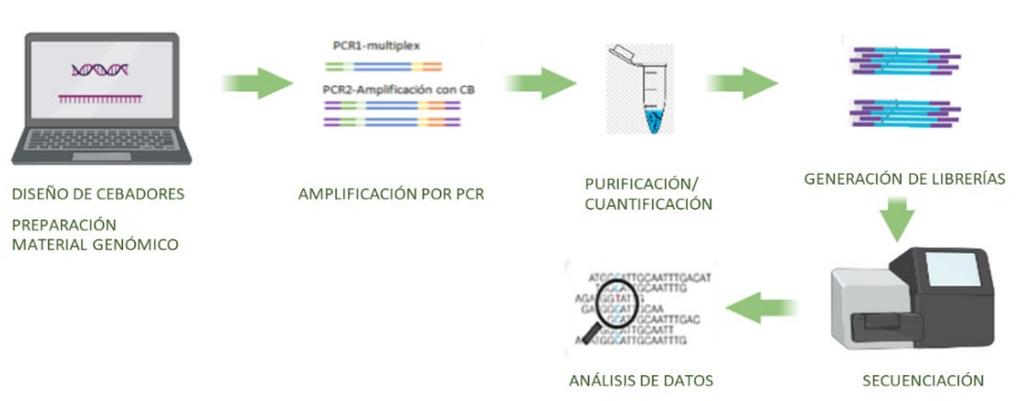


Figura 5: Esquema del diseño experimental. Este diseño consta de varios pasos: diseño de cebadores y preparación del material genómico. A continuación, se lleva a cabo el proceso de generación de librerías, secuenciación y posterior análisis de los datos.

3.3. OBTENCIÓN DE MUESTRAS

El ADN genómico de la población Hortega se extrajo a partir de muestras de sangre a través del kit ADN Maxwell® (Promega®, USA) en el Hospital Clínico de Valencia.

Dichas muestras se encontraban a diferentes concentraciones de partida, por lo que se realizaron diluciones de éstas en placas nuevas con el objetivo de que todas estuviesen a 5 ng/μL, siendo una concentración adecuada para proceder a la preparación de librerías.

3.4. PREPARACIÓN DE LIBRERIAS

3.4.1. SELECCIÓN DE GENES Y SNPs

La selección de genes, de SNPs y variantes de variación en número de copia (CNV) se realizó a través de una búsqueda bibliográfica en estudios previos realizados. Se seleccionaron aquellos que estaban relacionados con la CMD y que podrían estar relacionados con otras ECV como puede ser la HTA (Tabla 3).

SNPs	GEN	SNPs	GEN	SNPs	GEN	SNPs	GEN
rs10497520	TTN	rs1572226	SLC35F1	rs375891215	DSP	rs8031514	
rs1050606	ANXA5	rs17132250	TMEM232	rs3815743	MTRR	rs8031633	
rs1051168	NMB	rs17132261	TMEM232	rs397516037	MYBPC3	rs860867	TMEM117
rs1060501439	MYH7	rs17132292	TMEM232	rs4129218	ACO25419.1	rs878855018	PRKAG2
rs10872167	CEP85L	rs17227124		rs4307206	CEP85L	rs89107	SLC35F1
rs11153730		rs17489969		rs4343	ACE AC113554.1	rs927348	SLC35F1
rs11153768	CEP85L	rs17568359		rs4713854	PPARD	rs9305048	ADGRE2
rs1137101	LEPR	rs17597065		rs4945623	CEP85L	rs937277	
rs11549465-rs1154946	HIF1A	rs1761631		rs4946333	SLC35F1	rs9401060	SLC35F1
rs11603020	SERPING1	rs17825393		rs4966014	IGF1R	rs9481842	CEP85L
rs11724758	FABP2	rs1957757	HIF1A HIF1A-AS3	rs5068	CLCN6 NPPA	rs9489327	SLC35F1
rs11752626	CEP85L	rs199663613	TTN-AS1 TTN	rs5443	GNB3 CDCA3	rs9838915	KLF15
rs11758375	SLC35F1	rs199793620	TTN-AS1 TTN	rs6797133	SCN5A	rs9993110	MYO22
rs11846075		rs201278114	MYBPC3	rs6906287	CEP85L		CNVs
rs11967375	CEP85L BRD7P3	rs201490050	TTN-AS1 TTN	rs6911339			MMGT1-Ex2
rs11968176	CEP85L	rs2059238	WVVOX	rs6912208	SLC35F1		PACRG-Ex2
rs11968293	SLC35F1	rs2074192	ACE2 AC097625.1	rs6927800	SLC35F1		PLAC1-Ex3
rs11970286		rs2106809	ACE2 AC097625.1	rs72650068	TTN-AS1 TTN		QKI-Ex2
rs121912683	SLC25A4	rs2229309	NFATC4	rs731695	SLC35F1		TFDP3-Ex1
rs12619308	ACO92640.1 CCDC14	rs2235487	AL031708.1 JPT2	rs7452622	SLC35F1		
rs12661338	CEP85L	rs2290893	PTGES3	rs7554607	RYR2		
rs1334489		rs2292462	NMB	rs7565161	EPCAM-DT		
rs140037748	MYPN	rs281868	SLC35F1	rs7749263	SLC35F1		
rs149004293	JUP	rs283077	SLC35F1	rs7772845			
rs1558248842	TNNT2 AC119427.2	rs2958155	PTGES3	rs7910620	GRID1		
rs1569180680	JPH2	rs367871780	NEXN	rs8026758			

Tabla 3: Listado de genes, SNPs y CNVs seleccionados.

3.5. DISEÑO DE CEBADORES

Para llevar a cabo nuestro objetivo se diseñaron cebadores que amplifican fragmentos de unos 100-150 pares bases de las regiones de interés del ADN con los polimorfismos seleccionados.

Una vez escogida, mediante revisión bibliográfica, la lista de variantes genéticas con una posible asociación a la condición de interés, deben diseñarse los cebadores que se usarán en la primera PCR (la PCR multiplex) para la generación de la librería de amplicones.

Los cebadores usados en la PCR multiplex deben cumplir una serie de requisitos especiales, puesto que de ellos dependerá en gran parte la cobertura final obtenida para las variantes de interés. De esta forma, estos cebadores deben ser específicos para la región genómica de interés, no estar diseñados encima de ningún otro polimorfismo con una alta frecuencia en la población y ser compatibles para estar en un mismo pool o pocillo de reacción (minimizando la formación de *primer-dimer*), además de los requerimientos típicos de temperatura de melting, contenido GC, etc.

Los pasos para el diseño de los cebadores son los siguientes:

1. Anotación de las posiciones cromosómicas de las variantes mediante la herramienta Biomart (<http://www.ensembl.org/biomart/martview/a66782929e36beaed232aa8f6ddac6fb>)

2. Selección de las secuencias sobre las que diseñar cebadores y generación de un fichero tipo multifasta (estas secuencias están ya enmascaradas para los polimorfismos de alta frecuencia; por ejemplo, los SNPs con frecuencia mayor al 1% están marcados como una "N" dentro de la secuencia del genoma, para poder evitarlos a la hora de diseñar los cebadores).

3. Diseño individual de cada amplicón mediante llamadas automáticas a Primer3, obteniendo múltiples posibles parejas de cebadores (de las que se tendrá que elegir solo una). Primer3 se encarga de controlar los requerimientos individuales de cada pareja (tamaño del producto, contenido GC, temperatura de melting, etc).

4. Filtrado de las parejas de cebadores que no den lugar a productos específicos sobre el genoma humano, o que interaccionen demasiado con las colas de secuenciación de Illumina.

5. Elección de la mejor combinación de parejas de cebadores para minimizar la formación de *primer-dimer* dentro del pocillo de reacción.

En este panel se incluyen, además, 5 variantes de tipo CNV. En este caso para el diseño de los amplicones de los CNVs (forward o reverse) interesan los exones, por ello se intenta coger la máxima parte posible de la parte codificante del exón.

Los pasos 2, 3, 4, y 5 son automáticos y compartidos para ambos tipos de diseño.

Los cebadores venían liofilizados por lo que tuvimos que hidratarlos durante un día entero con la concentración adecuada de TrisHCl. Cada pareja de primer se dispuso en un mismo pocillo a la misma concentración y finalmente se procedió a la unión de todas las parejas de cebadores en un mismo pool ya que no había parejas solapantes entre ellos.

3.6. OPTIMIZACIÓN PCR

3.6.1. PCR1: AMPLIFICACIÓN EN MULTIPLEX

En la técnica de PCR multiplex, para enriquecer las regiones de interés generando amplicones de tamaño corto se utilizan múltiples cebadores específicos en una misma reacción. Con este método se obtienen un número menor de secuencias mapeadas fuera de la región de interés (*off-target*) y tiene una buena relación coste-eficacia.

Una vez obtenido el pool con todos los cebadores se procedió a realizar la primera PCR selectiva en la que se amplificarán las regiones de interés y se introducen las colas universales (CU) para posteriormente poder incluir todas las secuencias necesarias para la reacción de secuenciación. Para poner a punto esta reacción se probaron diferentes condiciones hasta llegar a la óptima utilizando una enzima comercial, con pequeñas variaciones tanto en la concentración de cebadores y aditivos. Las amplificaciones se llevaron a cabo en termocicladores de placas de 384 Veriti® (AppliedBiosystems®, California).

Las condiciones de temperatura, tiempo y ciclos de la PCR elegidas fueron las siguientes: una activación de la ADN polimerasa a 95°C durante 20 minutos. Posteriormente, en primer lugar, se lleva a cabo la desnaturalización del ADN a 98°C para que las dos cadenas se separen. A continuación, se produce la hibridación a 55°C y 62°C para que los cebadores se unan por complementariedad al ADN molde, en este caso es necesario bajar la temperatura para que se

produzca así el alineamiento. Por último, se produce la extensión a 72°C en la que la enzima polimerasa incorpora nucleótidos complementarios a partir del extremo 3' libre de la región en la que han hibridado los cebadores anteriormente. Estos tres pasos se repiten un total de 25 ciclos. La extensión final se realizó a 72°C durante 5 minutos tras el último ciclo de PCR, con esto se asegura la amplificación de cualquier ADN de cadena simple que quede. Finalmente se produce la conservación (*hold*) a 15°C durante un tiempo indefinido para conservar la reacción a corto plazo (Figura 6).

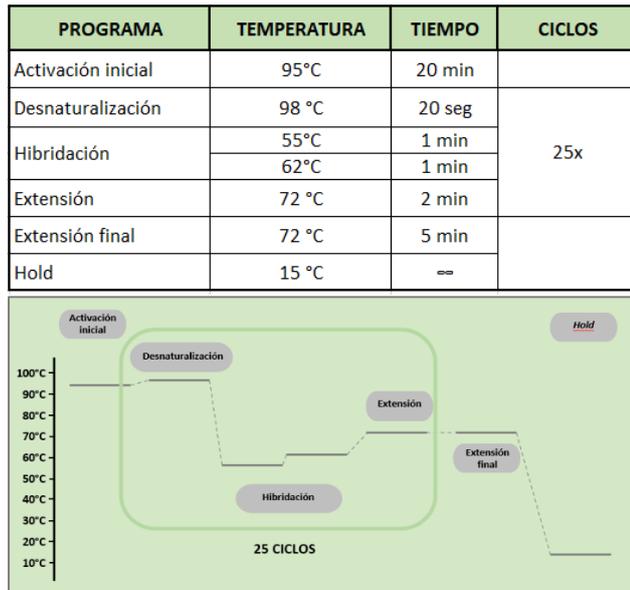


Figura 6: Condiciones utilizadas en la PCR 1

3.6.2. PCR2: MARCADO CON BARCODES

A partir de los productos generados en la PCR1 se lleva a cabo una segunda PCR en la que se incorporan las etiquetas o *barcodes*, que son secuencias conocidas de ADN compatible con la tecnología de Illumina que permite diferenciar cada muestra. De esta forma se puede combinar múltiples muestras en un mismo ensayo sin riesgo a que se produzca un entrecruzamiento entre ellas. Estos *barcodes* contienen la secuencia identificadora (uno para el 3' y otro para el 5'), de tal manera que se pueda identificar cada muestra y pueda ser estudiada de manera individual (Figura 7).

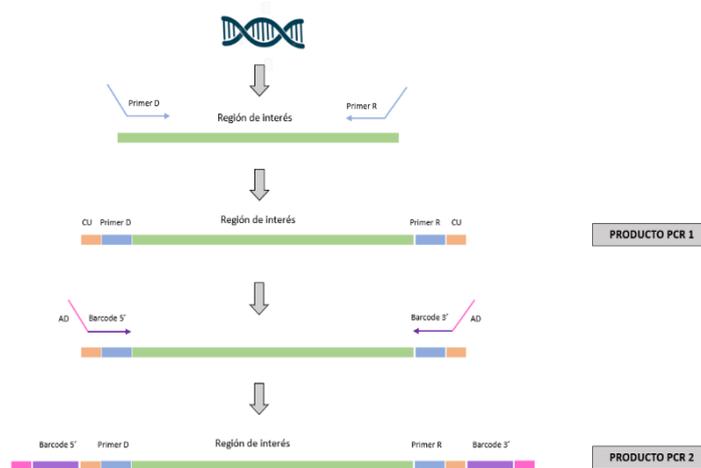


Figura 7: Esquema estructura de los productos obtenidos de ambas PCR con la incorporación de los barcodes.

En la siguiente figura 8 se muestran las diferentes fases y condiciones de temperatura de las muestras:

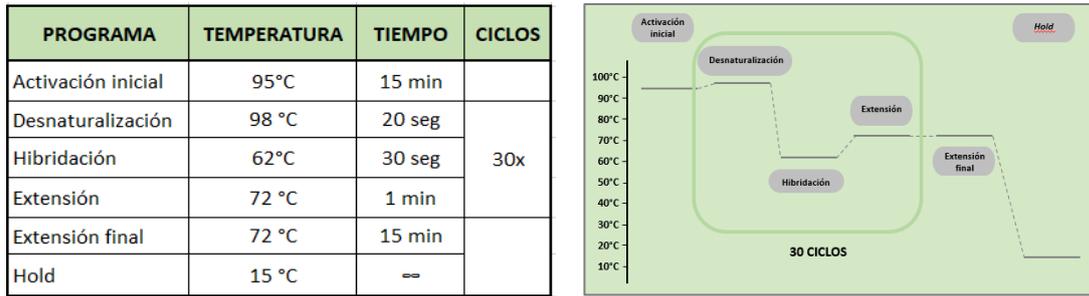


Figura 8: Condiciones utilizadas en la PCR 2.

3.7. ELECTROFORESIS CAPILAR

Para la amplificación de las muestras por electroforesis capilar utilizamos el sistema QIAxcel Advanced (Qiaxcel DNA Screening Kit, Qiagen®, Alemania) que sustituye al tradicional y laborioso análisis en gel de ADN. Este sistema automatiza completamente la electroforesis capilar de alta resolución de hasta 96 muestras por corrida en un periodo moderado de tiempo y con alta sensibilidad. Cuando se ha producido la amplificación de nuestras muestras podemos observar en el programa un gel de electroforesis con sus bandas, pero además se puede ver en el electroferograma de que tamaño es cada fragmento en concreto y así decidir cuál es la condición óptima (Figura 9).

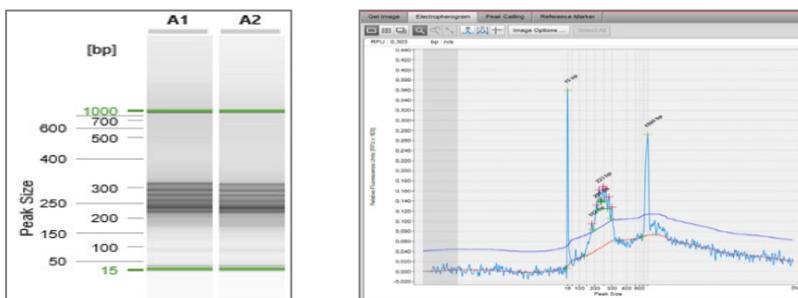


Figura 9: Ejemplo de una electroforesis capilar y electroferograma a través del sistema Qiaxcel. En ambos se observa la librería en torno a 300pb.

3.8. PURIFICACIÓN DE LIBRERÍAS

Tras la amplificación de las muestras, se seleccionaron las que tenían un tamaño adecuado y se procedió a su purificación con el kit comercial Magsi-NGS PREP® (MagnaMedics Diagnostics B.V., Países Bajos). Este kit se basa en perlas magnéticas y sirve tanto para la selección del tamaño adecuado como para eliminar posibles productos inespecíficos y resto de oligonucleótidos que estuvieran en exceso en la reacción.

El kit fue optimizado por el personal de la UGD (Unidad de Genómica y Diabetes) para la eliminación de productos de PCR menores de 200 pb. Para empezar, se añaden las bolas magnéticas a una concentración de 0.6x (siendo x la cantidad de muestra que queremos purificar). Los fragmentos de ADN se unen directamente a la superficie de estas perlas ya que estas se encuentran cargadas positivamente y el ADN negativamente, y de esta forma se quedan en la solución los fragmentos que no interesan. Luego, sobre una placa magnética, se realizan varios lavados con etanol al 70% y finalmente se añade la misma cantidad de agua que de muestra inicial, ya que el agua es un potente nucleófilo que compite con el ADN por unirse a las bolas magnéticas, desplazándolo y permitiendo su recuperación en otro pocillo diferente (Figura 10).

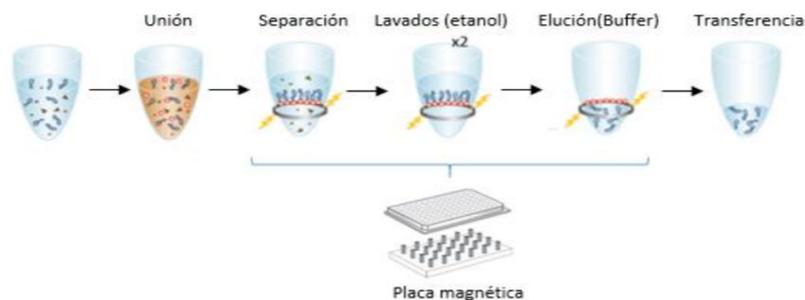


Figura 10: Proceso de purificación de las librerías con bolas magnéticas. Adaptada de (Seda García, 2018)

3.9. CUANTIFICACIÓN DE LIBRERÍAS

Una vez obtenidas las librerías se lleva a cabo su cuantificación con el kit QuantiFluor® dsDNA System (Promega). Este sistema fluorimétrico contiene un colorante fluorescente de unión a ADN que permite medir con precisión y alta sensibilidad pequeñas cantidades de ADN de doble cadena (0,05-200ng/μl de 1μl de muestra original) en una muestra purificada, presentando unos rangos de excitación entre 504nm y 531nm.

Para la cuantificación se realizó una recta patrón a través de diluciones seriadas de una concentración conocida de ADN bicatenario correspondiente al fago Lambda y el buffer Tris-EDTA (TE), el cual también se utiliza como blanco (Figura 11). Una vez añadidas las muestras se le añade el Quantifluor para analizar la fluorescencia a través del espectrofluorímetro GLOMAX Multi-Detection System (Promega®, USA).

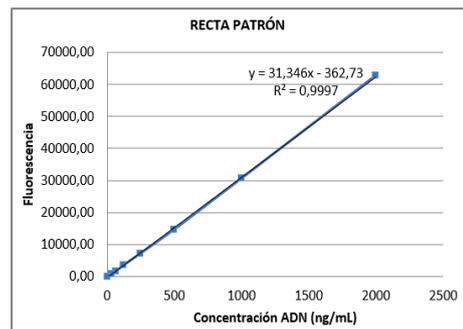


Figura 11: Ejemplo recta patrón para la cuantificación de librerías.

3.10. SECUENCIACIÓN

La secuenciación de nueva generación es un método diseñado para llevar a cabo la secuenciación masiva de una gran cantidad de fragmentos de ADN de forma paralela en el menor tiempo posible y con un bajo coste por base.

En esta parte del trabajo, por un lado, se utilizó el sistema Miseq (Illumina) de INCLIVA para realizar la puesta a punto de las condiciones óptimas de PCR y comprobar las condiciones que proporcionaban una cobertura más homogénea. Y, por otro lado, para analizar las posibles variantes de la población Hortega, al ser muchas muestras se utilizó el sistema HiseqX 150 PE (Illumina) de la empresa Macrogen, Inc (Korea del Sur) para obtener así una mayor cobertura.

Ambas plataformas son de Illumina y utilizan la amplificación clonal. La amplificación de los fragmentos se produce a través de una PCR-puente para generar un grupo o *cluster*. La generación de *cluster* es el proceso por el cual cada fragmento de una biblioteca se clona en miles de copias idénticas. Los fragmentos de ADN se colocan sobre una superficie sólida de vidrio separada por carriles. En cada uno de estos carriles hay oligonucleótidos complementarios a los adaptadores de cada fragmento que se va a secuenciar, de forma que cada fragmento queda anclado a la celda de flujo. A continuación, la polimerasa inicia la copia de la hebra de ADN dando lugar a una hebra reversa complementaria. En este punto, la hebra original es retirada y tan solo se queda la hebra reversa la cual a través de una secuencia terminal se pliega y se ancla a su respectiva secuencia complementaria de oligonucleótidos dando lugar a la forma de puente. Posteriormente, la polimerasa genera una hebra complementaria idéntica a la original que da lugar a dos hebras clonadas del fragmento que teníamos inicialmente. Estos pasos se repiten hasta formar un grupo o *cluster* a partir de la hebra inicial (Figura 12).



Figura 12: Esquema de la amplificación del ADN por PCR-puente para formar los grupos o clusters.

Una vez finalizada la amplificación se retiran las hebras reversas quedando solo las hebras originales y se lleva a cabo la secuenciación por síntesis. En este momento se produce la incorporación de terminadores reversibles modificados con etiquetas fluorescentes específicas que se van uniendo a su base complementaria a la vez que emite una fluorescencia propia que permite su identificación. Estos nucleótidos marcados poseen una modificación química que impide la unión a más de un nucleótido marcado en cada sitio de reacción de forma que se puede ubicar según su posición en la secuencia. Los nucleótidos compiten por elongar la cadena

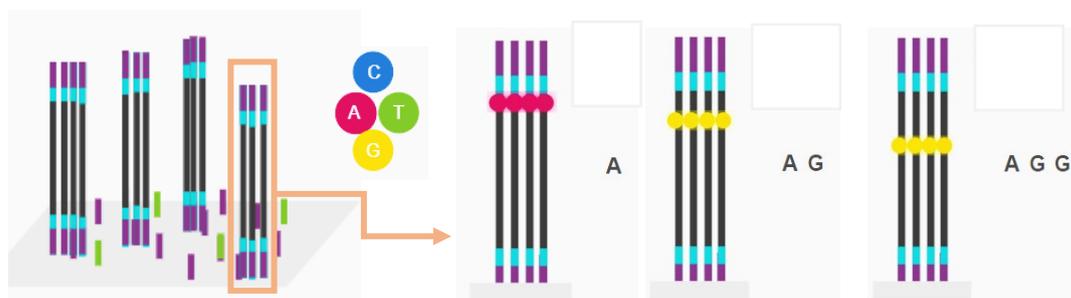


Figura 13: Esquema secuenciación por síntesis. Excitación de los nucleótidos marcados por fluorescencia.

y se unen solo los que son complementarios a la hebra molde, el resto son eliminados. Una vez unido, se excita el fluoróforo del nucleótido añadido y la señal lumínica es registrada por un dispositivo óptico (Figura 13). La etiqueta debe ser eliminada antes de que se produzca la colocación del siguiente nucleótido para evitar que dos bases emitan señal simultáneamente. Al finalizar la primera lectura, el fragmento resultante es retirado. Este paso se repite simultáneamente con todas las hebras del mismo clúster de forma paralela hasta completar la secuenciación.

3.11. ANÁLISIS BIOINFORMÁTICO

En el caso del proceso de optimización de la PCR, una vez terminada la secuenciación se llevó a cabo el demultiplexado. Este proceso consiste en separar cada conjunto de secuencias de cada muestra gracias a sus secuencias *índice*. El MiSeq reporter del secuenciador lleva a cabo este paso y tras éste se generan archivos en formato FASTQ el cual tiene que ser procesado. En estos archivos se recogen las lecturas de secuencia en formato FASTA con un valor de calidad (Q) para cada nucleótido. Por un lado, se eliminaron los nucleótidos de baja calidad en los extremos en un proceso llamado *trimming* o recorte y posteriormente se mapeó contra el genoma de referencia (versión GRCh37/hg19). Estos datos fueron almacenados en archivos en formato BAM (*Binary Alignment Map*), a partir del cual se pudo realizar el cálculo de cobertura por base, que es lo que nos interesaba. En este caso obtenemos dos archivos, por un lado, un archivo con el porcentaje de bases cubiertas que están por encima del límite establecido (en este caso se estableció 100x) y las coberturas máximas y mínimas obtenidas, y por otro lado, un archivo con las coberturas mínima, máximas, media y la longitud de cada amplicón.

El objetivo principal de esto es seleccionar la condición que tenga una mayor cobertura, mayor número de lecturas por amplicón y una distribución más homogénea. Para cada amplicón obtenemos unos resultados diferentes, por ello lo primero que se hace es normalizar los datos, dividiendo las lecturas de cada amplicón individual entre el número total de lecturas de dicha condición. Una vez normalizados, se analizan para observar los fragmentos que no han amplificado, los que tienen una menor cobertura y los que tienen mayor. Con estos datos, se crea una gráfica por condición, observando la distribución de las coberturas por amplicón y de esta forma ver cuál es la más adecuada.

Por otro lado, para el análisis bioinformático de las muestras de la población Hortega, el proceso de demultiplexado se llevó a cabo por el personal de bioinformática de la UGD empleando software libre, y posteriormente los archivos FASTQ que genera pasan por diferentes controles. En el primer control de calidad se utilizó el programa FastQC (v0.11.5, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), que nos analiza la calidad por base de cada FASTQ. Finalmente nos indicará si la secuenciación pasa los filtros de calidad deseados.

Posteriormente, a través del proceso de *trimming*, se llevó a cabo el procesamiento de lecturas crudas. En la tecnología de Illumina, la calidad de secuenciación disminuye conforme avanza el número de ciclos por ello se lleva a cabo el proceso de *trimming* en el cual se eliminan los adaptadores y las bases de baja calidad de los extremos, además de la eliminación de las secuencias adaptadoras. Una vez finalizado este proceso se vuelve a pasar otro control de

calidad con el mismo programa para ver si mejora la calidad. A continuación, se mapeó contra el genoma de referencia (versión GRCh37/hg19) y se almacenaron los datos en formato BAM (*Binary Aligment Map*).

En este caso, para el análisis de variantes, las lecturas alineadas anteriormente se utilizaron para identificar las posiciones que diferían con el genoma de referencia obteniéndose archivos tipo VCF con las variantes encontradas en el ADN de cada paciente, obteniéndose de esta forma el genotipo de cada muestra. Esta información queda recogida en archivos con formato VCF (*Variant Call Format*).

3.12. ANÁLISIS ESTADÍSTICO

Para los estudios de asociación se realizarán estudios estadísticos básicos mediante regresión lineal o logística según los datos sean las variables cuantitativas o dicotómicas, respectivamente. Los datos descriptivos de la población se expresaron como media \pm desviación estándar para las variables cuantitativas y como porcentaje para las variables cualitativas.

Se realizó un estudio de asociación con el programa PLINK para ver si las variantes eran significativas según los casos/control que corresponden a hipertensión/no hipertensión. En este caso se ha utilizado un intervalo de confianza del 95% con un p valor ($p < 0,05$). Además, se analizó si las frecuencias genotípicas cumplían el equilibrio de Hardy-Weinberg, permitiendo descartar aquellas variantes que no estuviesen en equilibrio.

4. RESULTADOS Y DISCUSIÓN

4.1. CARACTERÍSTICAS DE LA POBLACIÓN HORTEGA

En este estudio se incluyó un total de 1381 muestras de la población Hortega del Hospital Universitario Río Hortega en Valladolid. Estos pacientes tienen entre 14-85 años. Entre estos había un total de 345 que padecían hipertensión y ninguno de ellos había recibido tratamiento. La media de presión sistólica entre los hipertensos era de 148,27 mm Hg y la presión diastólica de 86,84 mm Hg (Tabla 4).

POBLACIÓN HORTEGA	
Edad (años)	50.15 ± 18.34
Mujeres (%)	49
Hombres (%)	51
Hipertensión (%)	25
Presión arterial sistólica (mm Hg)	148.27 ± 17.60
Presión arterial diastólica (mm Hg)	86.84 ± 10.13

Tabla 4: Características población Hortega.

4.2. DISEÑO DE CEBADORES

Con el fin de estudiar 95 variantes asociadas a enfermedades cardiovasculares se diseñaron un total de 95 parejas de cebadores, las cuales se unieron en un volumen final de 50 µl a una concentración de 50 µM. Se creó un único pool con estas parejas ya que no eran solapantes, es decir, se mezclaron 5 µl de cada pareja obteniéndose una concentración final de 0,5 µM. El pool final estaba formado por un total de 95 amplicones con una longitud promedio de 279 pares de bases (pb).

4.3. OPTIMIZACIÓN DE LAS CONDICIONES DE PCR1 Y PCR2

Una vez obtenido el pool con todas las parejas de cebadores se llevó a cabo la optimización de las condiciones de la PCR1 y PCR2. Los procesos de ambas reacciones se tomaron de un estudio previo con características similares y se mantuvo la concentración de la enzima comercial.

Para la optimización de la PCR, lo primero que se hizo fue probar los cebadores a diferentes concentraciones utilizando muestras de ADN control con el fin de encontrar la condición que permitiese amplificar el mayor número de amplicones de una forma homogénea. Se probaron 5 concentraciones diferentes de cebadores: 0,5 µM, 0,25 µM, 0,125 µM, 0,0625 µM y 0,05 µM para ver cuál de ellas tenía mejores resultados. En la figura 14 se puede observar los resultados de la electroforesis capilar con las diferentes concentraciones de cebadores, viéndose como la amplificación óptima se lleva a cabo con la concentración inicial de 0,5 µM obteniéndose un tamaño de bases en torno a los 300 pb como se esperaba y una menor suciedad en comparación

con las demás condiciones. La condición con cebadores a $0,25 \mu\text{M}$, aunque es similar, se descartó porque al mirar en el electroferograma tenía mayor suciedad y no era tan claro como la primera opción.

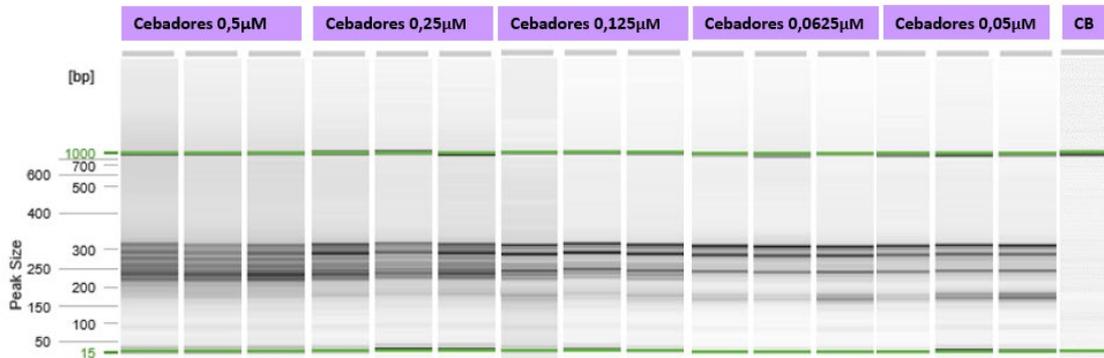


Figura 14: Resultados de la electroforesis capilar con los cebadores a diferentes concentraciones.

Una vez obtenida la concentración adecuada de cebadores para llevar a cabo la amplificación de los fragmentos, se probó diferentes aditivos con distintas concentraciones en la reacción para ver si con ellos se obtiene mejores resultados. Tal y como se observa en la figura 15, los aditivos 4 y 5 son los que muestran la condición más favorable, con una menor suciedad de la reacción y unas bandas más claras.

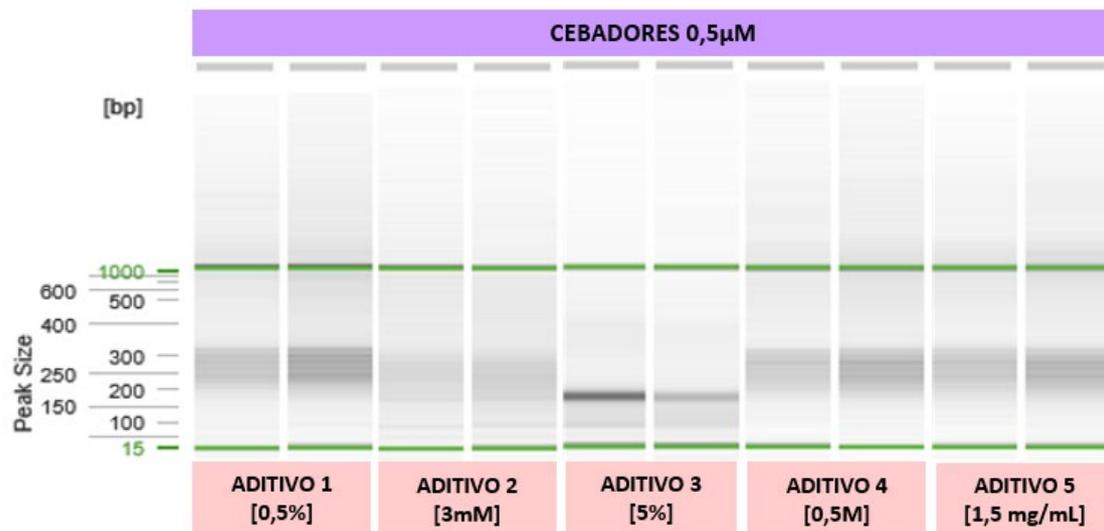


Figura 15: Resultados de la electroforesis capilar con diferentes aditivos.

Una vez decidido los aditivos que mejoraban la reacción, se realizó de nuevo otra prueba en la que se utilizó el doble de concentración de aditivo 4 y 5 y la combinación de ambos en la PCR1. En la figura 16, se puede observar como la reacción salía más sucia, con mayor cantidad de primer-dimer.

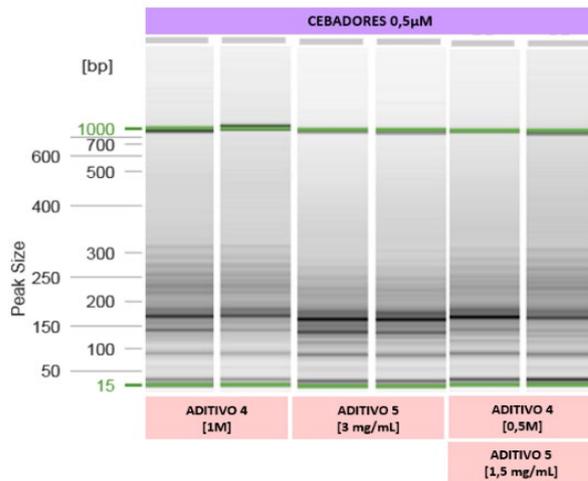


Figura 16: Resultados de electroforesis capilar tras la prueba con diferentes aditivos.

A pesar de ello, se probó estas condiciones añadiendo a la PCR2 también estos aditivos. En este caso, se añadió tanto el aditivo 4 como 5, el doble de concentración de cada uno de ellos y la combinación de ambos. Los resultados obtenidos no fueron los esperados y la amplificación tenía menor calidad además de tener más suciedad en la reacción de PCR (figura 17).

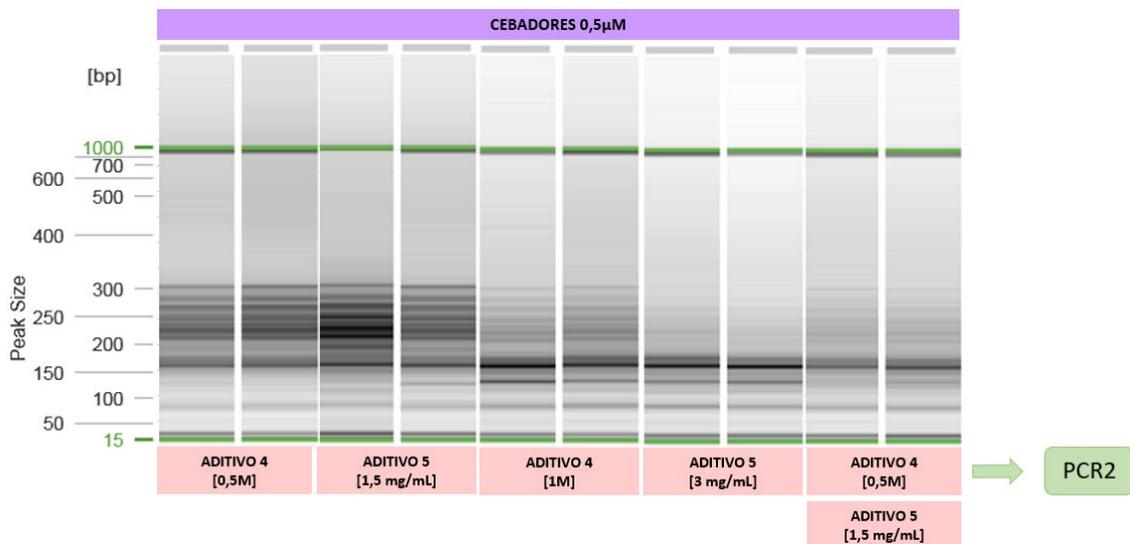


Figura 17: Resultados de la electroforesis capilar tras las modificaciones en los aditivos 4 y 5.

Con el fin de analizar de forma precisa los resultados de estas pruebas para ver que combinación era la más eficiente para la amplificación de forma homogénea de la mayor cantidad de amplicones posibles se secuenciaron varias muestras de estas condiciones para llevar a cabo el análisis de coberturas de estas tres 3 pruebas. Tras la purificación y cuantificación se llevó a cabo la secuenciación en el sistema de secuenciación MiSeq (Illumina) y el posterior análisis de coberturas de todas las condiciones con los datos estandarizados de las coberturas mínimas. Finalmente, se llegó a la conclusión de que la condición más favorable y con la que se obtenía mayor homogeneidad de todos los amplicones (figura 18) era en la que los cebadores estaban a una concentración de 0,5 μ M con el aditivo 5 [1,5mg/mL].

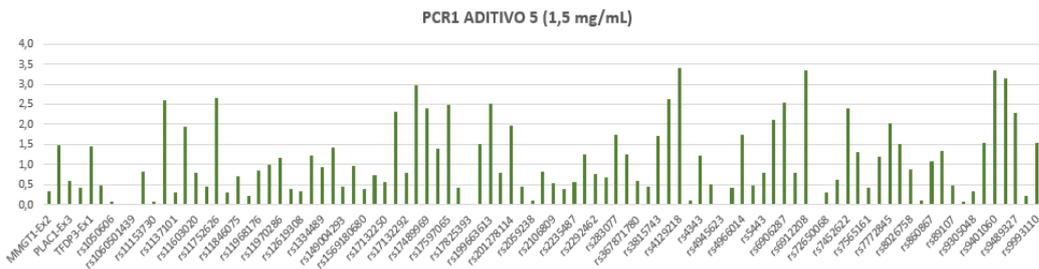


Figura 18: Análisis de coberturas en los fragmentos amplificados con la condición adecuada para la optimización de la PCR.

4.4. VALIDACIÓN DE LAS LIBRERÍAS CON LA POBLACIÓN HORTEGA

Una vez obtenidas las condiciones óptimas de PCR se llevó a cabo la generación de librerías de las muestras de la población Hortega. Antes de enviar todas las muestras a secuenciar se validó a través de su amplificación mediante electroforesis capilar (Figura 19 A).

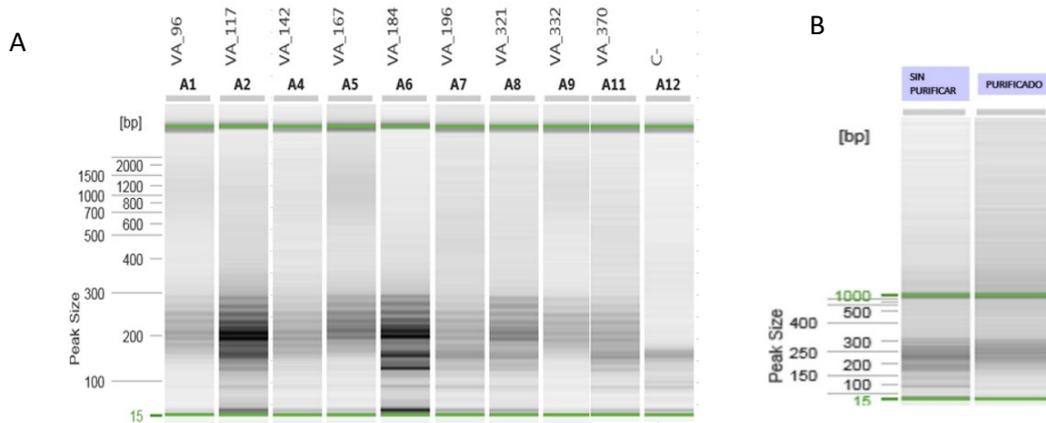


Figura 19: (A) Validación de muestras de la Población Hortega mediante electroforesis capilar. (B) Comparación entre el pool de la Población Hortega sin purificar y purificado.

Al ver que las librerías salían según los esperado, se procedió a realizar el pool con todas mezclando 2 μ l de cada muestra. Posteriormente se purificó para eliminar los restos de reactivos y los posibles dímeros de cebadores que se pudiesen haber formado (Figura 19 B).

4.5. CALIDAD DE LA SECUENCIACIÓN

Una vez realizada la secuenciación de la población Hortega se obtuvo un número medio de lecturas crudas por individuo de 187220.83 ± 79190.03 . Posteriormente se llevó a cabo el proceso de *trimming* de las lecturas crudas para la eliminación de adaptadores y bases de baja calidad. Se aplicaron varios filtros de calidad con el objetivo de eliminar aquellas cuyo valor en escala Phred fuese de $Q < 20$ y/o tuviesen una longitud menor de 80 pb. Tras utilizar estos filtros,

el número medio de lecturas crudas por cada individuo fue de $82281,89862 \pm 67134,51482$. Estos datos pasan por varios controles de calidad para comprobar si la secuenciación pasa los filtros de calidad necesarios. Para ello se utiliza el programa FastQC que analiza varios parámetros. En la figura 20 se puede observar cómo mejora la calidad media de base por muestra tras pasar los diferentes filtros de calidad, en este caso la calidad $Q > 30$ en todas las lecturas tras el *trimming* y bastante homogénea durante todo el proceso, lo que indica que los datos obtenidos tras la secuenciación tienen buena calidad para su posterior análisis.

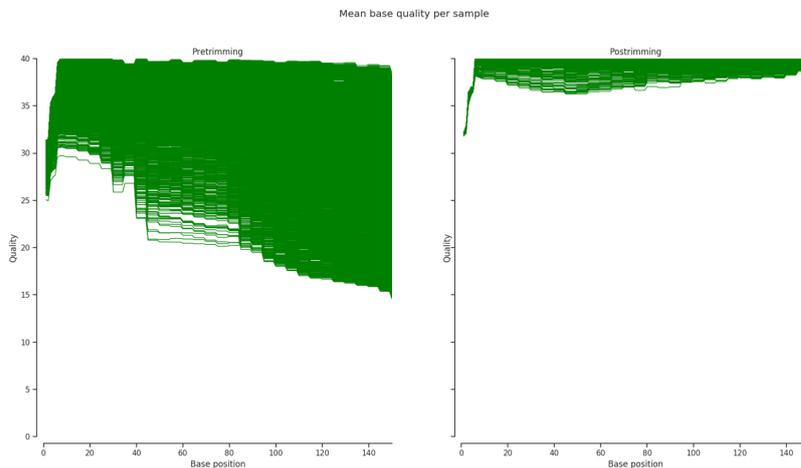


Figura 20: Calidad de lecturas antes y después del proceso de trimming.

El porcentaje de lecturas mapeadas contra el genoma humano es de un 98,90% mientras que el porcentaje de lecturas mapeadas contra las regiones de interés 96,12% lo que quiere decir que no tenemos amplificación inespecífica, es decir, hay pocas regiones *off-target*.

La profundidad de secuenciación fue de 100x ya que hay una gran cantidad de lecturas, permitiendo así identificar variantes tanto de baja como de alta frecuencia. Todas las variantes identificadas, excepto 6 de ellas, cumplen como mínimo la profundidad de 100x en su posición en concreto.

4.6. ANÁLISIS DE VARIANTES

Las variantes se suelen filtrar por calidad y luego son anotadas, es decir, asignadas a un gen utilizando diferentes programas. En este caso se utilizó *VEP*, el cual se base en una serie de criterios según diferentes bases de datos para elegir el transcrito de la anotación. El primer criterio que utiliza es si el transcrito es canónico.

Para llevar a cabo el análisis de variantes y ver cuáles eran significativas se aplicaron diferentes filtros. Por un lado, se priorizaron las variantes que tuviesen una cobertura como mínimo de 100x y una calidad de genotipo de 60x. Por otro lado, se aplicaron filtros entre hipertensos/no hipertensos en la población Hortega para ver cuales eran significativas con $P < 0,05$.

En este análisis incluimos 91 variantes de 51 genes diferentes que estaban asociados en el desarrollado de la CMD. Entre los resultados se encontraron un total de 42 genes asociados y

186 variantes de las cuales la mayoría eran intrónicas (57,5%), seguidas de variantes intergénicas (11,3%), variantes *missense* o con cambio de sentido (8,6%), variantes sinónimas o de bajo efecto (7,5%) y otros tipos (15%) (Figura 21).

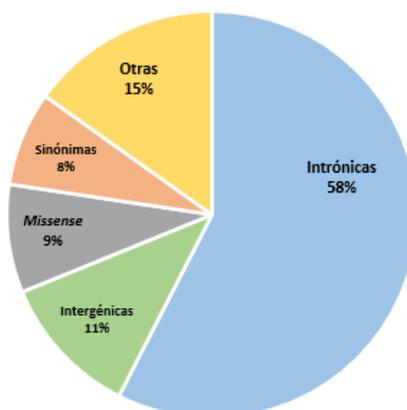


Figura 21: Porcentaje de variantes encontradas en la población Hortega.

El análisis de asociación de variantes encontradas en el grupo de hipertensos/no hipertensos se realizó con PLINK con la ayuda del personal bioinformático de la UGD. Se encontraron cuatro variantes significativas ($P < 0,05$), rs17597065, rs4129218, rs573914858 y rs746745349.

4.6.1. rs17597065

La variante rs17597065 se encuentra en el cromosoma 11. Se trata de una variante genética aguas arriba (*upstream gene variant*), es decir, es una variante de secuencia localizada dentro de una región reguladora, en este caso en la posición 5' de un gen. Según la base de datos Ensembl (v104, mayo 2021) se encuentra entre los genes *TCN1*, *CBLIF* y *MRPL16*. En dicha variante se produce un cambio de A/G con un impacto modificador del riesgo según los criterios de VEP.

Según la base de datos EUR_AF (basada en el estudio de los 1000Genomas), la frecuencia de esta variante en la población es de 4%. Se encuentra presente en el 95% de la población Hortega en homocigosis con el alelo de referencia, es decir, sin presentar ninguna variante. Hay un 4,8% que presentan un genotipo heterocigoto y tan solo el 0,07% tienen un genotipo homocigoto para la variante. En este estudio hay un total de 561 individuos que presentan HTA con el alelo de referencia, 36 tienen HTA con el genotipo heterocigoto de la variante y tan solo 1 presenta el genotipo homocigoto de la variante (Tabla 5).

Tabla 5: Relación entre la variante rs17597065, los genotipos de la población Hortega y los pacientes con HTA.

11-59616200-A-G					
Nº individuos			Nº individuos HTA		
0/0	0/1	1/1	0/0	0/1	1/1
1311	67	1	561	36	1

Varios estudios han asociado esta variante con el peso y la albumina en orina (Kulminski *et al.*, 2019). Además, también hay descrita cierta asociación con los índices de masa ventricular izquierda derivada del electrocardiograma (ECG-LV), en concreto con la suma de voltaje QRS (Arkchand *et al.*, 2013).

A pesar de no haber encontrado relación entre esta variante y la HTA ya descrita, es una de las variantes significativas entre los grupos de hipertensos/no hipertensos de la población Hortega. Además, se ha visto que el gen *MRPL16* está asociado significativamente con la cardiomiopatía séptica. Dicha cardiomiopatía se produce por una disfunción miocárdica sistólica del VI (Huang *et al.*, 2020) la cual puede ser una de las causas de la aparición de la HTA.

El gen *MRPL16* codifica la proteína ribosomal mitocondrial L16 que desempeñan un papel en la OXPHOS. Cuando se produce una alteración en la expresión de las proteínas ribosomales pueden conducir a un trastorno del metabolismo mitocondrial (Huang *et al.*, 2020), como se comentó anteriormente el mal funcionamiento de OXPHOS dan como resultado una disminución de ATP y un aumento de ROS llegando a ser responsables en ciertos casos de la aparición de la hipertensión.

Por un lado, el gen *CBLIF* es un miembro de la familia de proteínas transportadores de cobalamina codificando una glicoproteína secretada por la mucosa gástrica. Esta proteína es necesaria para la correcta absorción de la vitamina B12. Por otro lado, el gen *TCN1* codifica la proteína de unión a la vitamina B12 (Surendran *et al.*, 2018). La vitamina B12 se ha asociado con el desarrollo de la hipertensión, el déficit de esta vitamina se ha relacionado con el desarrollo de HTA (Boachie *et al.*, 2020) por ello la suplementación con B12 podría ser útil para la prevención a la HTA.

La realización de un mayor número de estudios sobre esta variante con respecto a la asociación con la HTA podría esclarecer mejor a través de que vías metabólicas está implicada en el desarrollo de esta patología.

4.6.2. rs4129218

La variante rs4129218 se encuentra en el cromosoma 12. Se trata de una variante intrónica y *non coding transcript* (variante de transcripción no codificante) con un impacto modificador del riesgo según los criterios VEP. En este caso se produce un cambio de G/A. Cada vez tiene mayor aceptación la influencia del transcriptoma no codificante a la fisiología y arquitectura de los mamíferos. Las variantes no codificantes suelen formarse por la omisión del primer o último exón dando como resultado la pérdida del codón de inicio o de stop respectivamente (Dhamija y Mennon, 2018), en este caso se trata de una variante de la transcripción de un gen de ARN no codificante y una variante del transcrito que se produce dentro de un intrón.

Según la base de datos EUR_AF, la frecuencia de esta variante en la población es de 81%. Se encuentra presente en el 3,2% de la población Hortega en homocigosis con el alelo de referencia. Hay un 29,9% que presentan un genotipo heterocigoto mientras que el 62,9% presentan un genotipo homocigoto para la variante, es decir, este SNPs está mucho más presente en la población que el anterior. En este estudio hay un total de 25 individuos que

presentan HTA con el alelo de referencia, 190 tienen HTA con el genotipo heterocigoto de la variante y tan solo 359 presenta el genotipo homocigoto de la variante (Tabla 6).

Tabla 6: Relación entre la variante rs4129218, los genotipos de la población Hortega y los pacientes con

12-65958661-G-A					
Nº individuos			Nº individuos HTA		
0/0	0/1	1/1	0/0	0/1	1/1
45	414	869	25	190	359

Se encuentra localizada en un *long non-coding RNAs* (lncRNA), en concreto en el LOC100507065, entre los genes *MSRB3* y *RPSAP52*. Los lncRNA son transcritos de más de 200 nucleótidos que no codifican proteínas pero que pueden activar o inhibir la expresión de ciertos genes a través de diferentes mecanismos (Park *et al.*, 2015).

Estudios previos muestran una asociación entre esta variante y la LVH en la población caucásica (Arnett *et al.*, 2009). Además, también muestra cierta asociación con el aumento de masa del VI (MVI). Se sabe que el MVI está influida por ciertos factores como puede ser la PA o la obesidad (Park *et al.*, 2015). Una de las causas más común de la LVH es tener una PA alta de forma que si esta variante está asociada con el desarrollo de la LVH podría ser interesante llevar a cabo estudios funcionales para ver de qué forma afecta y como podría repercutir al aumento de PA.

4.6.3. rs573914858 y rs746745349

Por un lado, la variante rs573914858 se encuentra en el cromosoma 6 y se trata de una variante intrónica con un impacto modificador del riesgo según los criterios VEP. En esta variante se produce un cambio de T/G.

Se encuentra presente en el 99,6% de la población Hortega en homocigosis con el alelo de referencia mientras que tan solo el 0,2% que presentan un genotipo heterocigoto con la variante, es decir, es una variante poco común a pesar de está asociada a la HTA. En el caso de los pacientes con HTA, hay 595 que presentan HTA con el genotipo homocigoto, mientras que 3 tienen el genotipo heterocigoto para la variante (Tabla 7 A).

Por otro lado, la variante rs746745349 se encuentra localizada en el cromosoma 6. Al igual que la anterior se trata también de una variante intrónica con un impacto modificador del riesgo según los criterios VEP. En este caso, se produce un cambio C/A. En la población Hortega se encuentra presente en el 99,5% en homocigosis con el alelo de referencia mientras que el 0,3 presentan un genotipo heterocigoto con la variante. Hay un total de 595 que padecen HTA que tienen el genotipo homocigoto sin cambios en la variante (Tabla 7 B).

Tabla 7: Relación entre las variantes, los genotipos de la población Hortega y los pacientes con HTA. La variante rs573914858 se encuentra representada por la tabla A y la variante rs746745349 por la tabla B.

A 6-118566188-T-G						B 6-118578085-C-A					
Nº individuos			Nº individuos HTA			Nº individuos			Nº individuos HTA		
0/0	0/1	1/1	0/0	0/1	1/1	0/0	0/1	1/1	0/0	0/1	1/1
1376	3	-	595	3	0	1374	4	-	595	0	0

Ambas se encuentran dentro del gen *SLC35F1*, que era uno de los genes que se había incluido en el estudio a pesar de que ambas variantes no estuviesen presentes inicialmente. Este gen la proteína de la familia de transportadores de solutos 35 miembros F1 que se trata de una proteína de membrana que pertenece a la familia de transportadores de solutos (Vasan *et al.*, 2009).

A pesar de no haber información previa de ambos SNPs ya descritas, las variantes del gen *SLC35F1* se han asociado en varios estudios con la frecuencia cardiaca en reposo y el tamaño ventricular que podría tener como consecuencia la elevación de la PA (Warren *et al.*, 2017). Una mayor frecuencia cardiaca en reposo se relaciona con un alto riesgo de padecer una enfermedad cardiovascular y una elevada mortalidad (Eijgelsheim *et al.*, 2010). El aumento de la frecuencia cardiaca podría producir un aumento de PA al mismo tiempo, de forma que las variantes presentes en este gen pueden estar implicadas en el desarrollo de la HTA.

No se conoce bien su papel en la fisiología cardiaca, pero se sabe que la proteína se expresa en el tejido cardiaco por lo que su función tiene que estar relacionado con él. También hay estudios que lo asociación con rasgos electrocardiográficos como las dimensiones diastólicas del VI (Vasan *et al.*, 2009).

En base a todo lo anterior descrito, el gen *SLC35F1* y las variantes presentes en él podrían llegar a ser buenos candidatos para futuras investigaciones con el objetivo de evaluar qué papel juega en la genética y fisiopatología de las enfermedades cardiovasculares, en concreto en la HTA y la CMD, para ver que posibles vías metabólicas pueden tener en común ambas patologías.

4.6.4. rs10497520

La variante rs10497520 se encuentra en el cromosoma 2. Se trata de una variante *missense* o de cambio de sentido, en la cual se produce una mutación puntual con un único cambio de nucleótido, dando lugar a un nuevo codón y al mismo tiempo la introducción de un nuevo aminoácido en la proteína. Estas mutaciones pueden no afectar a la proteína o pueden tener varios efectos alterando su estructura o función.

En este caso se produce un cambio T/C con un impacto moderado del riesgo según los criterios VEP. Está presente en el 21% de la población según la base de datos EUR_AF. En la población Horteiga se encuentra en el 1,5% en homocigosis con el alelo de referencia mientras que el 17,2% presentar un genotipo heterocigoto para la variante y el 54,3% un genotipo homocigoto. En cuanto a los individuos que padecen HTA y contienen este SNP, hay 10 de ellos que lo tienen en

Tabla 8: Relación entre la variante rs10497520, los genotipos de la población Horteiga y los pacientes con HTA.

2-179644855-T-C					
Nº individuos			Nº individuos HTA		
0/0	0/1	1/1	0/0	0/1	1/1
21	237	750	10	92	326

homocigosis con el alelo de referencia. En cambio, hay 92 individuos que presentan un genotipo heterocigoto para la variante mientras que 326 tienen el genotipo homocigoto (Tabla 8).

A pesar de no estar entre las variantes significativas se encuentra asociado al desarrollo de la CMD (Priest *et al.*, 2016). Además, esta variante se encuentra en el gen *TTN* que está claramente asociado a esta patología y se encuentra bastante presente en los pacientes hipertensos de la población Hortega.

Este gen codifica la proteína titina que se trata de una proteína del sarcómero que se puede encontrar en los músculos cardíaco y esquelético. Hay varios estudios que respaldan la implicación de este gen en la hipertensión, en concreto de este SNP con la PAD (Hebbar *et al.*, 2017). Además, uno de los factores que contribuyen a la rigidez del VD en la hipertensión arterial pulmonar en la titina que regula la rigidez sarcomérica. (Rain *et al.*, 2013).

La titina es un soporte molecular que abarca medio sarcómero y determina la rigidez muscular presente en diástole. La fosforilación de las dos isoformas de esta proteína determina tanto la elasticidad de la proteína como la rigidez pasiva (diastólica) de los cardiomiocitos. En el caso de la hipertensión arterial pulmonar, se vio que había una menor fosforilación en el tejido del VD en los pacientes que padecían esta patología (Rain *et al.*, 2013).

Varios estudios, asocian ciertas mutaciones que truncan *TTN* con la CMD, ocurriendo en el 25% de los casos familiares y en el 18% de los casos esporádicos aproximadamente (Herman *et al.*, 2012).

Estudios recientes en familias chilenas, también han relacionado esta variante con el riesgo de cáncer de mama. En concreto, se cree que puede tener un efecto protector frente al riesgo de padecer dicha patología (Fernandez-Moya *et al.*, 2020).

La incorporación de otros enfoques para identificar bien las mutaciones de este gen y todas las posibles variantes implicadas en el desarrollo tanto de la CMD como de la hipertensión puede ser clave para futuras terapias, e incluso de otras patologías como el cáncer. La limitación de estos estudios es el gran tamaño de la secuencia codificante de la titina que hace que se estudie de forma incompleta.

4.7. LIMITACIONES DEL ESTUDIO

Una de las limitaciones es que tan solo han salido 4 variantes significativas de todas las candidatas, pero a pesar de ello hemos podido encontrar nuevas asociaciones entre variantes o genes y la hipertensión al realizar el estudio. A pesar de esto, sería conveniente estudiar en profundidad las variantes encontradas en una población más grande de hipertensos/no hipertensos que padezcan CMD para poder encontrar una estrecha asociación entre ambas patologías.

En cuanto al análisis estadístico, por un lado, en el caso de las variantes ha sido limitado dado que el grado de significatividad era muy bajo y no se han podido aplicar correcciones. Por otro lado, ha sido bastante limitado por falta de tiempo. En primer momento estaba pensado hacer el análisis de asociación de estas variantes con otros parámetros relacionados con el riesgo cardiovascular como puede ser la diabetes tipo 2, el BMI etc. utilizando chi-cuadrado en el caso de que fuese cualitativo y si es cuantitativa ANOVA o Kruskal-Wallis. En un futuro llevar a cabo

dichos análisis puede ser beneficioso para ver si estos factores pueden llegar a afectar a desarrollar ambas patologías.

Es evidente la necesidad de seguir trabajando en estos tipos de estudios, profundizando mucho en el análisis de variantes con una cantidad de muestra mayor y con una mayor cantidad de estudios estadísticos.

5. CONCLUSIONES

La secuenciación masiva de amplicones es una buena herramienta para la identificación y validación de variantes en regiones genómicas de interés. No obstante, es necesario implementar un buen protocolo de optimización para conseguir una cobertura media mayor y más homogénea de todos los amplicones para que los resultados no se vean extrapolados.

De las 91 variantes seleccionadas para el estudio de asociación con de la CMD y HTA, todas ellas, excepto 6, se han encontrado tras realizar la secuenciación de 1381 individuos de la población Hortega. A pesar de ello, se han asociado a esta patología un total de 42 genes y 186 variantes, de las cuales la mayoría eran intrónicas. Pero tan solo 4 de ellas se han podido asociar de forma significativa a la HTA en comparaciones caso/control.

Respecto a las variantes estudiadas, los 4 SNPs asociados significativamente con la HTA se ha visto que hay varios estudios que lo relacionan con dicha patología. A pesar de ello, es necesario realizar más estudios para ver que grado de implicación tiene y el papel que puede tener en los mecanismos moleculares o en la genética de estas enfermedades.

A pesar de tener solo 4 variantes asociadas de forma significativa, hay variantes presentes en genes que están claramente implicados en el desarrollo de ambas patologías como es el caso del gen *TTN*.

Todos los estudios realizados de asociación, incluido este, nos muestra que hay numerosas variantes asociadas al desarrollo de la CMD que puede tener un efecto importante en la regulación de los niveles de presión arterial. Para ello se necesita realizar un estudio más exhaustivo con todas las variantes implicadas en otras poblaciones de mayor tamaño que permitan aumentar el poder estadístico. De esta forma se podría llegar a identificar posibles dianas terapéuticas para ambas enfermedades.

6. BIBLIOGRAFÍA

ARNETT D.K et al. 2009. Genome-wide association study identifies single-nucleotide polymorphism in KCNB1 associated with left ventricular mass in humans: the HyperGEN Study. *BMC Med Genet*, 10: 43.

BARISON, A et al. 2015. Myocardial interstitial remodelling in non-ischaemic dilated cardiomyopathy: insights from cardiovascular magnetic resonance. *Heart Fail Rev*, 20: 731–749.

BALIJE, S et al. 2016. Effect of hypertension at presentation on prognosis in patients with dilated cardiomyopathy presenting with normal renal angiogram. *Indian J Med Res*, 144(2): 281- 287.

BOACHIE, J et al. 2020. Low Vitamin B12 and Lipid Metabolism: Evidence from Pre-Clinical and Clinical Studies. *Nutrients*, 12(7): 1925.

ÇAKICI, EK et al. 2020. Clinical assessment of primary and secondary hypertension in children and adolescents. *Archives de Pédiatrie*, 27: 286-291.

CAVIEDES, BP et al. 2018. Dilated cardiomyopathy and severe heart failure. An update for pediatricians. *Arch Argent Pediatr*, 116(3): e421-e428.

DADSON, K et al. 2017. Molecular mechanisms in cardiomyopathy. *Clinical Science*, 131: 1375–1392.

DHAMIJAA, S & MENON, M. B. 2018. Non-coding transcript variants of protein-coding genes - what are they good for?. *RNA biology*, 15(8): 1025–1031.

DI GIOSIA, P et al. 2018. Gender Differences in Epidemiology, Pathophysiology, and Treatment of Hypertension. *Curr Atheroscler, Rep* 20, 13.

EIJGELSHEIM, M et al. 2010. Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum Mol Genet*, 19(19): 3885–3894.

FERNANDEZ-MOYA, A et al. 2020. Germline Variants in Driver Genes of Breast Cancer and Their Association with Familial and Early-Onset Breast Cancer Risk in a Chilean Population. *Cancers (Basel)*, 12(1): 249.

FUCHS, FD & WHELTON PK. 2020. High Blood Pressure and Cardiovascular Disease. *Hypertension*, 75:285–292.

FUNDACIÓN ESPAÑOLA DEL CORAZÓN. (22 de abril de 2021). Miocardiopatía dilatada. <https://fundaciondelcorazon.com/informacion-para-pacientes/enfermedades-cardiovasculares/cardiopatias-familiares-y-genetica/miocardiopatias/miocardiopatia-dilatada.html>

- HARAKALOVA, M et al. 2015. A systematic analysis of genetic dilated cardiomyopathy reveals numerous ubiquitously expressed and muscle-specific genes. *European Journal of Heart Failure*, 17: 484-493.
- HARKCHAND, S.S. 2013. Discovery and application of genetic determinants of cardiovascular disease risk factors. *Tesis doctoral presentada por Sara Harkchand*.
- HEBBAR, P et al. 2017. Genetic risk variants for metabolic traits in Arab populations. *Sci Rep*, 7: 40988.
- HERING, D et al. 2017. Recent advances in the pathophysiology of arterial hypertension: potential implications for clinical practice. *Polish archives of internal medicine*, 127: 195-204.
- HERMAN, D.S et al. 2012. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med*, 366(7): 619-628.
- HERSHBERGER, R et al. 2013. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*, 10: 531-547.
- HUANG, G et al. 2020. Abnormal Expression of Mitochondrial Ribosomal Proteins and Their Encoding Genes with Cell Apoptosis and Diseases. *Int. J. Mol. Sci*, 21: 8879.
- JAPP, AG et al. 2016. The Diagnosis and Evaluation of Dilated Cardiomyopathy. *Journal of the American College of Cardiology*, 26: 2996-3010.
- JORDAN, E et al. 2021. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation*, 144(1): 7–19.
- KIVIMÄKI, M & STEPTOE, A. 2018. Effects of stress on the development and progression of cardiovascular disease. *Nature Cardiology*, 15:215-229.
- KULMINSKI, A.M et al. 2019. Pleiotropic Meta-Analysis of Age-Related Phenotypes Addressing Evolutionary Uncertainty in Their Molecular Mechanisms. *Frontiers in Genetics*, 10: 433.
- MANN, S.J. 2018. Neurogenic hypertension: pathophysiology, diagnosis and management. *Clin Auton Res*, 28: 363–374.
- MCNALLY, EM et al. 2013. Genetic mutations and mechanisms in dilated cardiomyopathy. *J Clin Invest*, 123: 19-26.
- MUNROE, PB et al. 2013. Advances in blood pressure genomics. *Circ Res*, 112(10):1365-79.
- NG, FL et al. 2018. Hypertension genomics and cardiovascular prevention. *Annals of Translational Medicine*, 6: 15.

ORGANIZACIÓN MUNDIAL DE LA SALUD. 2021. Cardiovascular diseases (CVDs). Recuperado de: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).

PARK, JK et al. 2015. Validation study of candidate single nucleotide polymorphisms associated with left ventricular hypertrophy in the Korean population. *BMC Med Genet*, 16: 13.

PERRY, E et al. 2008. Classification of the cardiomyopathies: a position statement from the european society of cardiology working group on myocardial and pericardial diseases. *European Heart Journal*, 29: 270-276.

PERUMAREDDI, P. 2019. Prevention of Hypertension Related to Cardiovascular Disease. *Primary Care: Clinics in Office Practice*, 46: 27–39.

PETERS, S et al. 2020. Familial Dilated Cardiomyopathy. *Heart, Lung and Circulation*, 29: 566-576.

PRIEST, J.R et al. 2016. Early somatic mosaicism is a rare cause of long-QT syndrome. *Proc Natl Acad Sci U S A*, 13(41): 11555-11560.

RAIN, S et al. 2013. Right Ventricular Diastolic Impairment in Patients With Pulmonary Arterial Hypertension. *Circulation*, 128: 2016–2025

REICHART, D et al. 2019. Dilated cardiomyopathy: from epidemiologic to genetic phenotypes. *J Intern Med*, 286: 362– 372.

RUSSO, A et al. 2018. Advances in the Genetics of Hypertension: The Effect of Rare Variants. *Int J Mol Sci*, 19(3): 688.

SANBE, A. 2013. Dilated Cardiomyopathy: A Disease of the Myocardium. *Biol. Pharm. Bull*, 36(1): 18–22.

SCHULTHEISS, HP et al. 2019. Dilated cardiomyopathy. *Nat Rev Dis Primers* 5, 32.

SEDA GARCÍA, E. 2018. Desarrollo de un sistema de detección de mutaciones somáticas de interés clínico mediante secuenciación masiva en muestras oncológicas. *Tesis doctoral presentada por: Enrique Seda García*.

SEFEROVIĆ, PM et al. 2019. Heart failure in cardiomyopathies: a position paper from the Heart Failure Association of the European Society of Cardiology. *European Heart Journal*, 21: 553–576.

SLIVNICK, J & LAMPERT, BC. 2019. Hypertension and Heart Failure. *Heart Failure Clin*, 15: 531-541.

SURENDRAN, S et al. 2018. An update on vitamin B12-related gene polymorphisms and B12 status. *Genes Nutr*, 13: 2.

- TAYLOR, M.R et al. 2006. Cardiomyopathy, familial dilated. *Orphanet journal of rare diseases*, 1:27.
- TIMMIS, A et al. 2020. European Society of Cardiology: Cardiovascular Disease Statistics 2019. *European Heart Journal*, 41:12-85.
- VASAN, R.S et al. 2009. Genetic Variants Associated With Cardiac Structure and Function. *JAMA*, 302(2): 168–178.
- WAGNER-GRAU, P. 2010. Pathophysiology of arterial hypertension. *An Fac med*, 71: 225-229.
- WARREN, H.R et al. 2017. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*, 49(3): 403–415.
- WILLIAMS, B et al. 2018. Practice Guidelines for the management of arterial hypertension of the European Society of Hypertension and the European Society of Cardiology: ESH/ESC Task Force for the Management of Arterial Hypertension. *J Hypertens*, 36: 2.284 - 2.309.
- ZENG, X et al. 2009. Dilated cardiomyopathy with hypertension: prevalence and response to high-dose β 1-adrenoceptor antagonist therapy. *Clinical and Experimental Pharmacology and Physiology*, 36: 945-949.
- ZU, Y et al. 2016. A Mitochondrial DNA A8701G Mutation Associated with Maternally Inherited Hypertension and Dilated Cardiomyopathy in a Chinese Pedigree of a Consanguineous Marriage. *Chin Med J*, 129(3): 259–266.