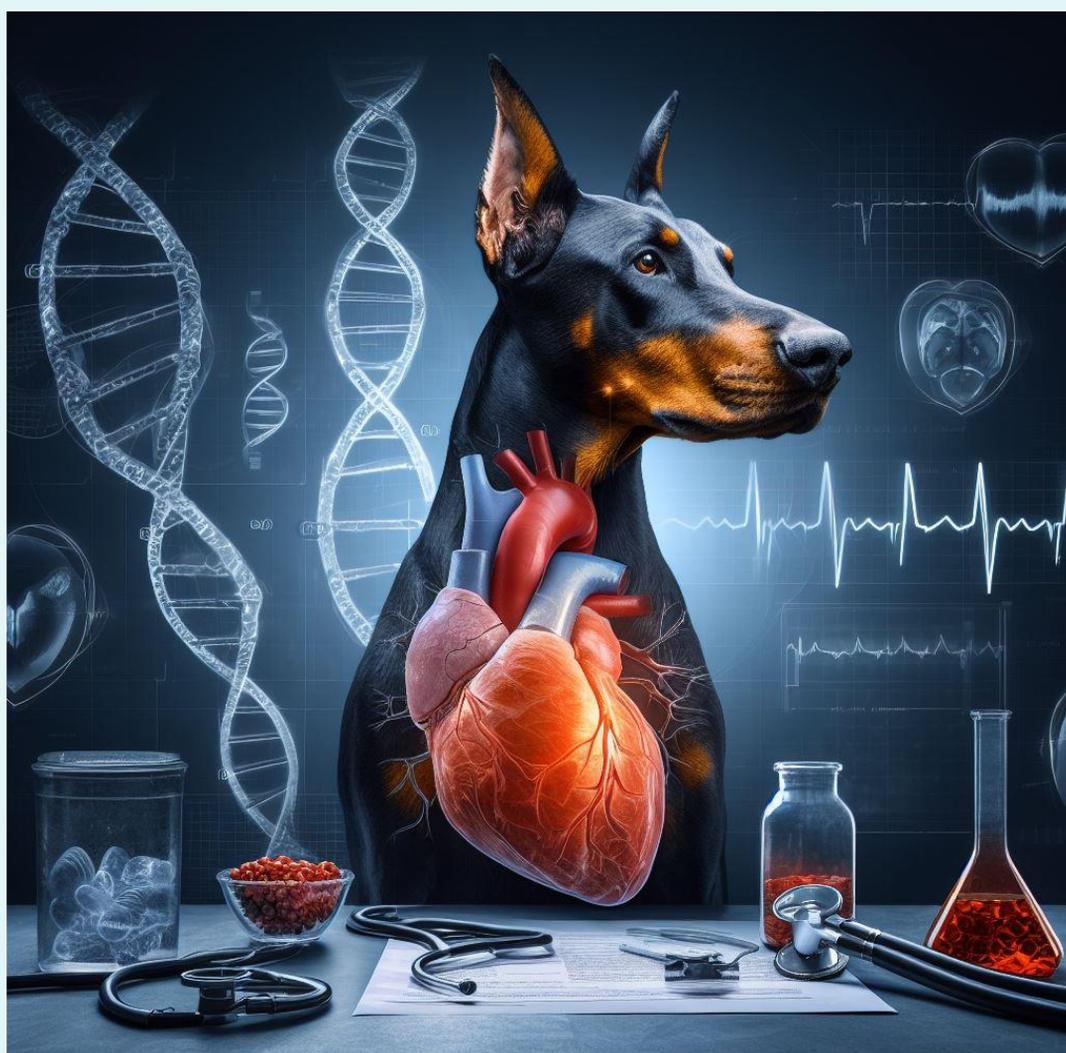


Analía Arizmendi

Tesis Doctoral

MAPEO DE LOS GENES CAUSANTES DE LA CARDIOMIOPATÍA DILATADA EN DOBERMAN PINSCHER



UNIVERSIDAD
NACIONAL
DE LA PLATA



FACULTAD DE CIENCIAS VETERINARIAS
UNIVERSIDAD NACIONAL DE LA PLATA

IGE VET



CONICET
FCV-UNLP



**UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE CIENCIAS VETERINARIAS**

Trabajo de Tesis realizado como requisito para optar al título de
DOCTORA en CIENCIAS VETERINARIAS

**MAPEO DE LOS GENES CAUSANTES DE LA CARDIOMIOPATÍA
DILATADA EN DOBERMAN PINSCHER**

AUTORA: MV ARIZMENDI, Analía

DIRECTOR: Dr. GIOVAMBATTISTA, Guillermo

CO-DIRECTOR: Dr. BATISTA, Pablo Rodrigo

ASESOR CIENTÍFICO: Dr. ARIAS, Daniel Osvaldo

LUGAR DE TRABAJO: Instituto de Genética Veterinaria “Ing. Fernando Noel Dulout (IGEDET), Facultad de Ciencias Veterinarias, Universidad Nacional de La Plata (FCV-UNLP); Servicio de Cardiología, Hospital Escuela, FCV-UNLP.

MIEMBROS DEL JURADO:

Dra. DE PALMA, Viviana

Dr. LIRÓN, Pedro

Dr. MARCOS, Martín

AÑO 2023

A mi familia y amigos...

AGRADECIMIENTOS

Cuando repaso estos años de trabajo y esfuerzo es inevitable pensar en todas las personas que estuvieron conmigo, apoyándome y dándome confianza para concretar este proyecto. Por eso, este logro es tan mío como suyo, GRACIAS.

A mi director, Guillermo Giovambattista, por brindarme la posibilidad de trabajar bajo su dirección, por su ayuda y preocupación constante, y por transmitirme su pasión por la investigación.

A mi codirector y amigo, Pablo Batista, por enseñarme que trabajar en equipo y en lo que uno ama, lo hace todo más fácil. Por su dedicación, su apoyo, su complicidad y sus consejos desde el comienzo de este camino.

A Daniel Arias, por abrirme las puertas de la cátedra, creer en mí e impulsar mi crecimiento. Gracias por estar siempre pendiente, dispuesto a escuchar, pero principalmente, gracias por compartir tus ideas de una educación y un país mejor.

A mi familia, porque son el pilar fundamental de este trabajo, quienes vivieron mis altibajos y nunca me dejaron bajar los brazos, gracias por su amor. Graciela y Carlos, gracias por inculcarme la importancia del estudio y la pasión por el trabajo. A mi mamá por ser mi lugar seguro, a mi papá por guiarme desde arriba. A mis hermanas Ailín y Paula por confiar en mí, apoyarme e impulsarme a mejorar día a día.

A mi amiga Rochi, con quien compartí cada uno de los momentos de este trabajo, los buenos y no tan buenos. Por estar siempre para mí, por crecer juntas, profesional y personalmente, y por enseñarme a luchar por lo que quiero. A mis amigas, Pao y Anita, por ser mis confidentes y compañeras incondicionales, mis sueñitos.

A mis amigas de toda la vida, que me acompañaron y apoyaron a la distancia, entendiendo mis tiempos y siempre atentas a mis avances.

A Luis, por ser mi compañero en el momento más estresante de esta tesis, por su paciencia, pero sobretodo por su amor.

A mis compañeros de MCD por acompañarme, bancar mi intensidad y estar atentos a todo Doberman que apareciera. Especialmente a Mari y a Nico, que me enseñaron y trabajaron a la par, y con quienes compartí muchas anécdotas y odiseas de muestreo.

A mis compañeros del IGEVET por la ayuda para desarrollar este trabajo. Fundamentalmente a Juli, Diego y Egle por su confianza, su apoyo y enseñanza incondicional, y sobretodo por sus horas de terapia.

A la Dra. Silvina Fornasari y su grupo de trabajo, principalmente a Vanesa y Leandro, que tuvieron toda la paciencia del mundo para explicarme y ayudarme desinteresadamente, en el momento cúlmine de la tesis.

Al jurado, por su comprensión y aportes para el desarrollo de esta tesis.

Al Estado Argentino que me permitió acceder a educación pública y gratuita de calidad.

Al Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET) por el financiamiento de la beca doctoral.

A la Facultad de Ciencias Veterinarias UNLP, mi segunda casa, lugar de excelencia donde continúo mi formación y donde conocí grandes personas y amigos.

A los propietarios de las mascotas que integraron este proyecto, por la confianza y la predisposición.

A las mascotas que formaron parte de este estudio y las que me acompañan día a día, brindándome su amor y promoviendo mi deseo de seguir creciendo en esta profesión.

PUBLICACIONES PARCIALES DEL PRESENTE TRABAJO

Publicaciones en revistas científicas internacionales

Arizmendi A, Rudd Garces G, Crespi JA, Olivera LH, Barrientos LS, Peral García P, Giovambattista G. Analysis of Doberman Pinscher and Toy Poodle samples with targeted next-generation sequencing. *Gene*. 2023 15; 853:147069.

Publicaciones en eventos científicos de alcance internacional

Arizmendi A, Batista P, Arias D, Vercellini R, Crespi J, Barrientos L, Giovambattista G. Nuevas perspectivas en el diagnóstico genético de la Cardiomiopatía Dilatada en Doberman Pinscher. 10as Jornadas Internacionales de Veterinaria Práctica. Organizadas por el Colegio de Veterinarios de la provincia de Buenos Aires. 25 y 26 de Agosto de 2017. Mar del Plata, Argentina.

Arizmendi A, Giovambattista G, Tórtora M, Re N, Vercellini R, Arias D, Batista PR. Avaliação eletrocardiográfica em Dobermann Pinscher. III Congresso Brasileiro de Cardiologia Veterinária. Organizado por la Sociedade Brasileira de Cardiologia Veterinária. 07 a 09 de Novembro de 2018. Campos do Jordao- SP Brasil.

Arizmendi A, Barrientos L, Crespi J, Garces G, Giovambattista G. Analysis of clinical samples from Doberman and Toy Poodle dogs with a targeted next-generation genotyping system. 37th International Society of Animal Genetics (ISAG) Conference. Organizado por la International Society of Animal Genetics. 2019. Lleida, España.

Arizmendi A, Batista PR, Crespi J, Tórtora M, Vercellini R, Czernigow M, Arias D, Giovambattista G. Analysis of PDK4 gene deletion in a population of Doberman Pinschers from Argentina. ECVIM-CA Congress. Organizado por el European College of Veterinary Internal Medicine. 2020. Barcelona, España. ***Artículo ganador del premio “Best Poster Presentation” 2020, otorgado por la European Society of Veterinary Cardiology (ESVC).***

Publicaciones en eventos científicos de alcance nacional

Arizmendi A, Garces G, Crespi J, Olivera H, Peral García P, Giovambattista G. Análisis simultáneo de filiación, condiciones genéticas y caracteres cualitativos mediante NGS-Target en perros. XLVII Congreso Argentino de Genética. Organizado por la Sociedad Argentina de Genética. 2020. Ciudad Autónoma de Buenos Aires, Argentina.

Arizmendi A, Batista PR, Re N, Barrena JP, Garces G, Giovambattista G. Evaluación de la prevalencia de la variante del gen PDK4 en una población de Doberman Pinscher de Argentina y su asociación con el desarrollo de Cardiomiopatía Dilatada. XX° Congreso Nacional de AVEACA. Organizado por la Asociación de veterinarios especializados en animales de compañía de Argentina. 2020. Ciudad Autónoma de Buenos Aires, Argentina. ***Artículo ganador del premio “Trabajo científico” 2020, otorgado por la Asociación de veterinarios especializados en animales de compañía de Argentina (AVEACA).***

ÍNDICE DE CONTENIDOS

ABREVIATURAS.....	1
RESUMEN.....	2
SUMMARY	4
INTRODUCCIÓN GENERAL	
1. Enfermedades de origen genético en el perro	6
1.1 Origen de enfermedades genéticas en el perro y variación entre razas	6
1.2 Modo de herencia de enfermedades genéticas en perros.....	8
1.3 Enfermedades cardiovasculares de origen genético en el perro	10
2. Cardiomiopatía dilatada en Doberman Pinscher	11
2. 1 Epidemiología.....	11
2.2 Fisiopatología	12
2.3 Genética y herencia.....	15
3. CMD en otras razas	21
3.1 CMD en el Bóxer.....	23
4. Hipótesis y objetivos	25
Referencias	28
 CAPÍTULO I: Análisis de las variantes genéticas involucradas en el desarrollo de cardiomiopatía dilatada en perros de raza Doberman Pinscher y Bóxer, mediante diferentes estrategias de asociación genómica utilizando marcadores genéticos.	
Introducción.....	39
Materiales y métodos.....	43
Resultados	67

Discusión y conclusiones	89
Referencias	99
CAPÍTULO II: Evaluación <i>in silico</i> del potencial efecto de las variantes causales de CMD sobre la estructura de las proteínas y/o la regulación génica	
Introducción.....	108
Materiales y métodos.....	114
Resultados	126
Discusión y conclusiones	142
Referencias	147
CONCLUSIONES FINALES	153
ANEXOS	156

ABREVIATURAS

ADN: Ácido desoxirribonucleico

CFA: Cromosomas del perro (*Canis Lupus Familiaris*)

CHR5: Cromosoma 5

CMD: Cardiomiopatía dilatada

DP: Doberman Pinscher

EDTA: Etilendiaminotetraacético (anticoagulante)

lpm: Latidos por minuto

mg/ml: Miligramos/mililitros

MHz: Megahertz

ml: Mililitros

mm: Milímetros

mM: milimolar

mm/seg: Milímetros por segundo

ms: Milisegundos

mV: Milivoltios

μl: microlitro

ng: nanogramo

NGS: Next Generation Sequencing

pb: pares de bases (distancia genética)

PCR: Polymerase Chain Reaction (Reacción en cadena de la polimerasa)

PDK4: Piruvato deshidrogenasa quinasa 4

SNP: Single Nucleotide Polymorphism (Polimorfismo de nucleótido simple)

SNV: Single Nucleotide Variant (Variante de nucleótido simple)

TTN: Titina

MAPEO DE LOS GENES CAUSANTES DE LA CARDIOMIOPATÍA DILATADA EN DOBERMAN PINSCHER

Palabras clave: Cardiomiopatía dilatada; Doberman Pinscher; genes candidatos; marcadores genéticos; PDK4; titina

Resumen

El objetivo principal del presente trabajo de tesis fue estudiar polimorfismos en genes candidatos involucrados en el desarrollo de Cardiomiopatía Dilatada (CMD) en una población de Doberman Pinscher (DP) de Argentina, mediante diferentes estrategias de análisis poblacional y de asociación con marcadores genéticos. Se muestreó además una población de Bóxers para determinar la presencia de las variantes en otra raza. Se predijo *in silico* el efecto de estos polimorfismos en la expresión génica de PDK4 y en la función de la proteína titina. Para ello, y en base a los hallazgos del examen físico, ecocardiográfico y electrocardiográfico, se clasificaron los caninos en dos grupos (sanos y enfermos). A partir de la utilización de PCR, técnicas de secuenciación capilar tipo Sanger, pirosecuenciación y targeted NGS, se confirmó la presencia de las variantes candidatas en DP de Argentina (CFA14:g.20,829,667_20,829,682del en el gen PDK4, CFA5:g.53,941,386T>C, CFA36:g.22,321,955C>T en el gen TTN). Se estableció la prevalencia de estos alelos por primera vez en la población local de DP, siendo 0,17 para

CFA14:g.20,829,667_20,829,682del, 0,12 para CFA5:g.53,941,386T>C y 0,58 para CFA36:g.22,321,955C>T. Asimismo, se demostró que la prevalencia de la mutación PDK4 en Argentina es menor que en Estados Unidos (0,37) y similar a Europa (0,16). Sólo la variante CFA36:g.22,321,955C>T mostró asociación significativa con CMD en la población local de DP. En Bóxers se identificó únicamente la presencia del SNP CFA5:g.53,941,386T>C en animales sanos. Los estudios *in silico* predijeron consecuencias perjudiciales de la variante PDK4 en la expresión génica de este gen, posiblemente relacionadas con la menor eficiencia del sitio de empalme, y un impacto negativo en la estabilidad de la proteína titina, en relación a la variante TTN, lo que podría afectar la elasticidad de la proteína y su desenrollamiento. Se concluye que el conocimiento de las variantes genéticas de CMD en la población de DP de Argentina y su potencial efecto biológico, permitió desarrollar métodos moleculares de diagnóstico que permiten identificar animales portadores con el objetivo de realizar el diagnóstico y tratamiento temprano de la enfermedad. De la misma manera, contribuye para la planificación de cruzamientos selectivos con el fin de disminuir la incidencia de CMD en esta raza.

CANDIDATES GENE MAPPING OF DILATED CARDIOMYOPATHY IN DOBERMAN PINSCHER

Keywords: Dilated cardiomyopathy; Doberman Pinscher; candidate genes; genetic markers; PDK4; titin

Summary

The main objective of this thesis was to study polymorphism of candidate genes involved in the development of Dilated Cardiomyopathy (DCM) in an Argentine Doberman Pinscher (DP) population, using genetic population analysis and various association strategies with genetic markers. A population of Bóxers was also sampled to determine the presence of the variants in another breed. The effect of these polymorphisms on PDK4 gene expression and titin protein function was predicted *in silico*. Based on findings from physical, echocardiographic, and electrocardiographic examinations, the dogs were classified into two groups (healthy and affected). Using PCR, Sanger capillary sequencing, pyrosequencing, and targeted NGS, the presence of candidate variants (CFA14:g.20,829,667_20,829,682del in the PDK4 gene, CFA5:g.53,941,386T>C, and CFA36:g.22,321,955C>T in the TTN gen) in Argentine DP population was confirmed. The prevalence of these alleles was established for the first time in the local DP breed population, being 0.17 for CFA14:g.20,829,667_20,829,682del, 0.12 for CFA5:g.53,941,386T>C, and 0.58 for CFA36:g.22,321,955C>T.

Additionally, it was demonstrated that the prevalence of the PDK4 mutation in Argentina is lower than United States (0.37) and similar to Europe (0.16). Only the variant CFA36: g.22,321,955 C>T showed a significant association with DCM in the local DP population. In Bóxers, only the presence of the SNP CFA5:g.53,941,386T>C was identified in healthy animals. *In silico* studies predicted detrimental consequences of the PDK4 variant on the gene expression of this gene, possibly related to lower splice site efficiency, and a negative impact on titin protein stability, relative to the TTN variant, which could affect the elasticity of the protein and its unfolding. In conclusion, knowledge of the genetic variants of DCM in the Argentine DP population and their potential biological effects enabled the development of molecular diagnostic methods to identify carrier animals for early disease diagnosis and treatment. Similarly, this contributes to the planning of selective breeding strategies to reduce the incidence of DCM in this breed.

INTRODUCCIÓN GENERAL

1. Enfermedades de origen genético en el perro

1.1. Origen de enfermedades genéticas en el perro y variación entre razas

El perro doméstico (*Canis Lupus Familiaris*) presenta una alta frecuencia de **enfermedades de origen genético**, es decir, afecciones patológicas causadas por alteraciones en el genoma (Patterson, 1980). Estas condiciones genéticas pueden ser definidas como monogénicas (asociados a la variación de un solo gen), poligénicas (asociados a la variación de múltiples genes y su posible interacción) o multifactoriales (interacción entre factores genéticos y ambientales).

Al menos el 50% de las enfermedades genéticas descritas hasta el momento en esta especie se caracterizan por ocurrir predominante o exclusivamente en una o pocas razas relacionadas (Ostrander y Kruglyak, 2000; Patterson, 2000; Björnerfeldt y col., 2008). Esto puede atribuirse al hecho de que cada raza se encuentra aislada genéticamente por haberse originado a partir de un grupo fundador reducido, sumado a las restricciones generadas por el modo de cría parcialmente consanguíneo (Parker y col., 2004). Los efectos de cuello de botella, reflejados principalmente durante la domesticación del perro y la creación de las diferentes razas, generaron una reducida variabilidad genética intraracial, una alta variabilidad interracial y un tamaño de población efectivo pequeño para cada raza (Mellanby y col., 2013).

El efecto fundador, caracterizado por la selección a favor de ciertas características fenotípicas específicas (estándar racial) y por el uso repetido de los mismos reproductores populares (*“popular sire effect”*) ocasiona que las generaciones subsiguientes sean un reflejo genético de estos perros, fijando tanto características deseables como indeseables. Esto genera un aumento de la frecuencia de alelos recesivos deletéreos y de la probabilidad de aparición de animales homocigotas con alelos indeseables idénticos por descendencia en las próximas generaciones (Calboli y col., 2008; Nicholas, 2010).

La alta predisposición a la aparición de enfermedades específicas en determinadas razas, junto a la baja o nula incidencia en otras, sugiere fuertemente la presencia de alelos de riesgo. Los factores antes mencionados (pequeñas poblaciones fundadoras, aislamiento reproductivo entre razas, barreras y estándares raciales, cuellos de botella, selección extrema y utilización de reproductores populares), sumados al aumento de la homocigosidad a nivel genómico debidos a los altos niveles de endogamia ($> 0,1$; Bannasch y col., 2021), genera que estas variantes alélicas de riesgo presenten una frecuencia significativamente mayor en el grupo de animales afectados en comparación al grupo de animales sanos (Ostrander y Kruglyak 2000; Leroy, 2011). Por estos motivos, cuando se produce una mutación causal espontánea en un animal reproductor (especialmente si es popular), esta tiene una alta probabilidad de transmitirse a las siguientes generaciones y aumentar la frecuencia alélica y la prevalencia de la enfermedad en la población en la que se encuentre.

1.2. Modo de herencia de enfermedades genéticas en perros

En la actualidad, se han reportado un total de 896 alteraciones genéticas asociadas al desarrollo de enfermedades o expresión de caracteres en caninos, de las cuales solo en 513 se conoce la mutación causal. En muchas de estas condiciones genéticas aún no está claro (o es controversial) su modo de herencia (Tabla 1). Es de destacar que por su similitud con alteraciones humanas, 558 de ellas son potenciales modelos para estudios de investigación (*Online Mendelian Inheritance in Animals*, OMIA, 2023, acceso 1 de agosto de 2023). Dentro las enfermedades genéticas caninas es posible encontrar diferentes modos de herencia mendeliana o poligénica, entre las que se puede mencionar:

Autosómicas recesivas: necesita dos copias del alelo mutante para que se exprese la enfermedad. El cruzamiento entre dos portadores (poseen una copia del alelo mutado) genera una proporción de 3:1 de no afectados vs afectados, siempre y cuando tenga penetrancia completa (Parker y col., 2006). Un ejemplo de herencia recesiva es la forma juvenil de cardiomiopatía dilatada en el Perro de Agua Portugués (Dambach y col., 1999).

Autosómicas dominantes: una copia del alelo mutante es suficiente para manifestar la enfermedad (siempre y cuando tenga penetrancia completa), el individuo tiene al menos un progenitor afectado, y teóricamente un 50% de su descendencia puede estar afectada; no hay diferencia en el porcentaje de afectados por sexo (Parker y col., 2006). A modo de ejemplo puede mencionarse la displasia de la válvula tricúspide en Labrador Retriever (Andelfinger y col., 2003).

Ligados al X o ligadas al sexo: la gran mayoría son caracteres recesivos, en los que las hembras portadoras tienen un 50% de probabilidades de transmitir el alelo mutado a su descendencia masculina, mientras que no existe una transmisión macho a macho debido a que las mutaciones causales se encuentran en genes localizados en la región propia del cromosoma X. La gran mayoría de los afectados son machos, aunque raramente pueden encontrarse hembras afectadas (provenientes de un macho afectado y una hembra portadora; Parker y col., 2006). Como ejemplo de este tipo de herencia está la cardiomiopatía dilatada familiar de inicio adulto en el Gran Danés (Meurs y col., 2001a).

Poligénicas y multifactoriales: son enfermedades causadas por interacción de múltiples genes y factores ambientales (Parker y col., 2006). Entre ellas se encuentran el conducto arterioso persistente en Caniches (Patterson y col., 1971) y la enfermedad valvular mixomatosa en Cavalier King Charles Spaniels y Dachshunds (Hägström y col., 1995; Olsen y col., 1999).

Tabla 1: Número de Características/Enfermedades genéticas en caninos basado en las categorías detalladas en la base de datos de OMIA (*Online Mendelian Inheritance in Animals*, agosto, 2023).

Categoría	Cantidad
Total características/enfermedades	896
Condición de herencia mendeliana	410
Condición con mutación causal reconocida	513
Potenciales modelos para humanos	576

1.3. Enfermedades cardiovasculares de origen genético en el perro

Las enfermedades cardiovasculares en caninos corresponden aproximadamente al 11 % de las consultas en los hospitales veterinarios (Detweiler y col., 1965; Buchanan, 1999). Dentro de este grupo existen algunas enfermedades de carácter familiar y/o de alta prevalencia en ciertas razas (Hammer y col., 1996; Meurs, 1998; Sisson y col., 2000; Goodwin y col., 1995). El estudio de estas condiciones genéticas es complejo, debido a que pueden involucrar múltiples genes, penetrancia incompleta o dependiente de la edad, expresividad variable y la influencia de factores ambientales. Por otro lado, la aparición tardía, frecuente en varias de estas enfermedades, es particularmente problemática ya que los individuos a menudo se reproducen antes que se manifieste la enfermedad (Parker y col., 2006). Esto pone en evidencia la importancia que adquieren las investigaciones en el campo de la genética, desarrolladas desde la década del 70 para el estudio del desarrollo de enfermedades cardiovasculares (Patterson y col., 1971; Patterson, 1976).

La Cardiomiopatía Dilatada (CMD) es la segunda enfermedad cardiovascular adquirida más frecuente en perros, luego de la enfermedad degenerativa de la válvula mitral (Detweiler y col., 1965; Tidholm y col., 2001). En la población canina general su prevalencia se estima en 0,5 % (Sisson y col., 2000). Sin embargo, su incidencia aumenta en razas específicas, principalmente en Doberman Pinscher (DP) y Gran Danés, en las que se la considera la enfermedad cardiovascular hereditaria más prevalente (Meurs y col., 2001a; Tidholm y col., 2001; Wess y col., 2010).

2. Cardiomiopatía Dilatada en Doberman Pinscher

2.1. Epidemiología

La cardiomiopatía dilatada (CMD) es una enfermedad que se presenta principalmente en perros de razas grandes y gigantes, siendo el Doberman Pinscher (DP) una de las razas más afectadas (Sisson y col., 2000; Monnet y col., 1995; Tidholm y col., 2001). Según reportes previos, esta cardiopatía ha mostrado una prevalencia en este grupo racial de entre 45 y 63 % en poblaciones de Estados Unidos (Calvert y col., 2000 a lett col. Grad col.) i entras que fue del 58,8 % para ejemplares de Europa (Wess y col., 2010). En Argentina, hasta el momento no hay estudios publicados sobre prevalencia de enfermedades cardiológicas en general, ni de la CMD en particular.

La CMD afecta a animales adultos de entre 5 y 10 años de edad, incrementándose las probabilidades de desarrollo de la enfermedad conforme aumenta la edad (Calvert y col., 1997; Domanjko-Petric y col., 2002; Wess y col., 2010). Del mismo modo, se han reportado casos en los que la CMD afecta a ejemplares de tan sólo 2 o 3 años, con una prevalencia del 10 %, por lo que es de vital importancia la evaluación de los animales desde edades tempranas (Calvert y col., 2000; Wess y col., 2017).

En cuanto a la prevalencia según el sexo, si bien varios estudios han reportado que la mayoría de los pacientes afectados son machos (Calvert y col., 2000), estudios más recientes demostraron que la enfermedad está igualmente distribuida en ambos sexos (Meurs y col., 2007; Wess y col., 2010). La diferencia

en la distribución entre sexos hallada en los estudios anteriores puede deberse a una diferencia en la manifestación clínica de la enfermedad, donde los machos suelen manifestar cambios ecocardiográficos y falla cardíaca congestiva a una edad más temprana que las hembras y por lo tanto estar sobrerrepresentados. En contraposición, las hembras suelen ser más propensas a desarrollar arritmias ventriculares (Wess y col., 2010).

2.2. Fisiopatología

La CMD es un desorden primario del miocardio que se caracteriza por una pérdida progresiva de su capacidad de contracción, lo que genera una disminución del volumen minuto (Sisson y col., 2000). Posteriormente, se produce una dilatación severa de ambas cámaras ventriculares, comenzando por el ventrículo izquierdo (VI), con la consiguiente disfunción sistólica, la reducción de los espesores parietales y el eventual desarrollo de arritmias (supraventriculares y ventriculares). Este cuadro puede derivar en una falla cardíaca congestiva y/o muerte súbita (Calvert y col., 1986; Monnet y col., 1995; Tidholm y col., 1997). La muerte súbita se relaciona con la aparición de taquicardia ventricular que conlleva a la fibrilación ventricular y la subsiguiente disminución severa de gasto cardíaco (Calvert y col., 2000). Mientras que en humanos la CMD se define por la presencia de dilatación del ventrículo izquierdo y disfunción contráctil, en ausencia de condiciones de carga anormales y enfermedad arterial coronaria grave, en ciertas razas de perros puede manifestarse inicialmente solo con

arritmias, como fibrilación atrial (FA) o complejos prematuros ventriculares (CPVs; Wess, 2022).

La progresión de la enfermedad se divide en tres fases según los hallazgos clínicos, morfológicos y eléctricos (Calvert y col., 2000). En la primera fase (etapa A), que puede durar de meses a años, el animal no evidencia signos clínicos de enfermedad cardíaca ni anomalías morfológicas o eléctricas en los registros ecocardiográficos y electrocardiográficos respectivamente (Calvert y col., 2000). Esta etapa, aplica a los perros de razas con predisposición a desarrollar CMD, como Doberman Pinscher, Lebel Irlandés (LI), Gran Danés (GD) y Bóxer (B), y es útil para promover en los propietarios, criadores y veterinarios la realización de exámenes de detección regulares, incluyendo test para detectar mutaciones genéticas conocidas (Wess, 2022).

En la segunda etapa, llamada oculta/silenciosa o preclínica B, se evidencian trastornos morfológicos y/o eléctricos en ausencia de signos clínicos de enfermedad cardíaca (Calvert y col., 2000; Domanjko-Petric y col., 2002; Wess y col., 2010), aunque se ha descrito la posibilidad de manifestar síncope e intolerancia al ejercicio (Wess, 2022). Las alteraciones ecocardiográficas que la caracterizan incluyen agrandamiento del ventrículo izquierdo en primera instancia en sístole y luego en diástole, con posterior adelgazamiento de las paredes del ventrículo. Mediante electrocardiografía puede reconocerse la aparición de arritmias, siendo más frecuentes los complejos prematuros ventriculares (CPVs) y/o fibrilación atrial (FA; Domanjko-Petric y col., 2002). Durante este estadio, que suele durar entre 3 y 4 años, puede producirse muerte súbita en un 25-30 % de los perros afectados (Calvert y col., 2000 Grad col.) . n una

publicación reciente se ha propuesto una subdivisión de esta fase en dos estadios. El estadio B1 incluye perros con anomalías eléctricas sospechadas de ser causadas por CMD. Este estado arritmogénico sin dilatación cardíaca incluye CPVs en ciertas razas como el DP o Bóxer, y FA en Lebrele Irlandés u otras razas gigantes. Los valores ecocardiográficos se encuentran dentro de los rangos de referencia o son inconclusos para el diagnóstico de CMD. El estadio B2 incluye perros con disfunción sistólica del ventrículo izquierdo (dilatación de la cámara del VI en sístole) con o sin dilatación de la cámara del ventrículo izquierdo en diástole concurrente. Las anomalías eléctricas pueden coexistir en este estadio B2. El examen clínico en el estadio B podría revelar amplitud de pulso variable o reducida, ritmos cardíacos irregulares o déficit de pulso causados por FA o CPVs. Otros posibles hallazgos en este estadio pueden incluir soplo leve sobre la válvula mitral o área de la válvula tricúspide, causada por insuficiencia valvular secundaria al agrandamiento diastólico del ventrículo izquierdo, derecho o ambos (Wess, 2022).

La tercera fase (etapa C) comienza cuando se manifiestan los signos de insuficiencia cardíaca izquierda o derecha, consistentes en edema pulmonar o ascitis y derrame pleural respectivamente, culminando con la muerte del animal como consecuencia de sus complicaciones (Calvert y col., 2000; Grad y col., 2004). En esta fase también se describe muerte súbita en un tercio de los perros afectados (Calvert y col., 1997). Por último, Wess (2022) menciona una cuarta fase (etapa D) que incluye a los perros en el estadio terminal de la enfermedad, refractarios al tratamiento médico.

La detección de la enfermedad en estadios tempranos (fase B) permite la implementación de terapéuticas adecuadas para disminuir la progresión de la enfermedad y prevenir los casos de muerte súbita (Summerfield y col., 2012). Asimismo, el reconocimiento temprano de la enfermedad o de la condición de portador de alguna de las mutaciones causales asociadas, permitiría a los criadores llevar adelante programas de crianza que resulten en la eliminación de las mutaciones causales de la población.

2.3. *Genética y herencia*

En humanos, las formas hereditarias de CMD representan hasta un 35 % de los pacientes afectados con esta enfermedad y se transmiten principalmente como rasgos autosómicos dominantes (Mestroni y col., 1999; Grünig y col., 1998). Hasta la fecha, más de 50 *loci* se han asociado a la forma monogénica de CMD, pero estas mutaciones representan solo el 50% de las presentaciones de la enfermedad (Posafalvi y col., 2013). El otro 50% de los casos se cree que está causado por múltiples mutaciones sinérgicas o por interacciones genéticas aún poco conocidas (McNally y col., 2013). De los genes identificados que se relacionan con CMD en humanos, la mayoría codifica para proteínas del citoesqueleto de las células miocárdicas. Las mutaciones en estos genes podrían generar variaciones en dichas proteínas, produciendo un deterioro en la producción o transmisión de la fuerza de contracción del corazón (Hein y col., 2000; Towbin y col., 2000).

En la población canina se ha propuesto que cada raza tiene mutaciones genéticas particulares asociadas al desarrollo de CMD, y que es posible que coexistan más de una mutación causal. La CMD en el DP es una enfermedad hereditaria poligénica de transmisión autosómica dominante con penetrancia incompleta dependiente de la edad (Hammer y col., 1996; Calvert y col., 1992; Meurs y col., 2007). Esta *penetrancia incompleta* implica que no todos los perros que presenten la mutación causal desarrollarán la enfermedad (Meurs y col., 2012). Del mismo modo se ha descrito que esta enfermedad tiene una *expresión variable*, refiriéndose al amplio espectro de signos clínicos y anomalías que pueden ocurrir en perros con esta cardiopatía; una mutación particular puede no afectar de la misma manera con variación de la intensidad (Meurs y col., 2012). Los fenómenos de penetrancia incompleta y expresión variable probablemente resultan de la interacción entre la información genética y la activación diferencial de sus factores epigenéticos reguladores. Diferentes estilos de vida, la dieta y otros factores ambientales pueden alterar la forma en que el código genético determina no solo el estado fisiológico de los órganos de cada perro, sino también las enfermedades que les afectan (Dutton y col., 2018).

Hasta el momento, 14 genes asociados con CMD en humanos (ACTC1, ACTN2, CSRP3, DES, LMNA, MYBPC3, MYH7, PLN, SGCD, TCAP, TNNC1, TNNT2, TPM1, VCL) también han sido estudiados en DP, pero ninguno se ha asociado al desarrollo de la enfermedad en esta raza. La mayoría de estas variantes afectan estructuras esenciales y funciones del corazón, incluidos el sarcómero, disco Z, complejo de distrofina, citoesqueleto, desmosomas, retículo sarcoplásmico, envoltura nuclear, núcleo, canales iónicos, mitocondrias y matriz

extracelular (Meurs y col., 2001b; Stabej y col., 2004; Stabej y col., 2005; Meurs col. 7 Wiersa col. 7 Meurs col. ' Sullivan col. Wess, 2022).

Mausberg y col., (2011) realizaron un estudio de asociación genómica completa (GWAS) en DP de Alemania y Reino Unido, en el que detectaron un locus en el cromosoma 5 que contiene un SNP (alelo "C") asociado al 50 % de los casos de CMD. Los perros con arritmias aisladas (a diferencia de los del grupo de "dilatación") tenían una frecuencia significativamente más alta del alelo C, sin embargo la división en dos sub-fenotipos se ve obstaculizada por el hecho de que la mayoría de los DP afectados por CMD tienen una combinación de ambos síntomas de gravedad variable. No obstante, no fue posible asociar un gen candidato para CMD ligado a esta señal de asociación.

Un estudio similar realizado por Meurs (2012) identificó, en una población de DP de Estados Unidos, una deleción en el sitio de empalme de 16 pares de bases en el gen piruvato deshidrogenasa quinasa 4 (PDK4) ubicado en el cromosoma 14, asociado con el desarrollo de CMD. El modo de herencia de esta mutación es autosómica dominante, con una prevalencia de hasta el 60 % en la población de EEUU y una penetrancia del 68 % (Meurs y col., 2012). Asimismo, la frecuencia alélica del alelo mutado fue de 0,36. Los animales portadores de esta mutación, ya sea en estado heterocigoto u homocigoto, tendrían 10 veces más probabilidades de desarrollar CMD que los perros con genotipo de tipo salvaje (Meurs, 2019). Por otro lado, Owczarek-Lipska y col. (2013) observaron que, en poblaciones europeas de DP, la deleción de ese alelo era levemente más frecuente en los perros afectados. Aunque no se evidenciaba una diferencia estadísticamente

significativa. La prevalencia del alelo con la delección fue menor que la reportada en EEUU (0,16). Estas discrepancias en los resultados pueden deberse a variaciones genéticas entre las poblaciones de América del Norte y Europa. El gen PDK4 codifica para la isoenzima 4 de la piruvato deshidrogenasa quinasa (PDK), una proteína mitocondrial que juega un papel importante en la flexibilidad metabólica, al permitir que las mitocondrias ajusten el uso de la fuente primaria de combustible en función de la disponibilidad de nutrientes. La mutación en el gen PDK4 genera una reducción de la expresión de la isoenzima, lo que conduciría a un control deficiente del complejo piruvato deshidrogenasa, alejando el metabolismo de la oxidación de ácidos grasos y aumentando la glucólisis. La fuente de energía cardíaca principal en el corazón sano son los ácidos grasos (Grynberg y Demaison, 1996), por lo que con el tiempo, la glucólisis no regulada podría conducir potencialmente a un deterioro del transporte de electrones mitocondriales y predisponer a las células al daño mitocondrial (Meurs y col., 2012). Los DP que presentan al menos un alelo mutado mostraron evidencia de estrés mitocondrial y activación de la vía apoptótica intrínseca después de la inanición (Taggart y col., 2017). Sosa y col. (2016) estudiaron la disfunción mitocondrial resultante de la mutación de ese gen en un pequeño número de Dobermans de Estados Unidos encontrando que los Dobermans con CMD y la mutación PDK4 tenían una tasa de consumo de oxígeno mitocondrial basal más baja que los Dobermans sanos. No obstante, y especialmente en vista del pequeño tamaño de la muestra, la importancia de la mutación PDK4 como factor de riesgo de CMD sigue siendo incierta.

En el 2019, Meurs y col. reportaron una segunda mutación en el gen titina (TTN) vinculada al desarrollo de CMD en DP. En este caso se trata de un cambio de un solo nucleótido (C > T) en el gen de la titina, resultando en un cambio en un aminoácido altamente conservado (de glicina a arginina). El modo de la herencia de esta mutación es autosómica dominante, con una penetrancia del 47 %. Este gen codifica para una proteína sarcomérica involucrada en la contracción cardíaca, por lo que se predice que esta mutación generaría un cambio en la estructura de la proteína, asociada a una disminución de la tensión activa en las miofibras. Los autores postularon, según los resultados de diferentes análisis *in silico* que predicen el impacto de la sustitución de un aminoácido o un indel en el funcionamiento biológico de la proteína, que la mutación sería un cambio deletéreo. Al comparar la expresión de la titina en animales de control con animales que eran heterocigotos u homocigotos para la variante, no se encontraron diferencias importantes. Esto demuestra que la fisiopatología de la CMD en relación con la mutación en la titina hasta el momento no se entiende por completo. En referencia al desarrollo de la enfermedad en los perros afectados por la mutación CFA36: g.22,321,955 C>T, no se encontraron diferencias aparentes en la severidad del cuadro, pero la muerte súbita era comúnmente observada. Del total de DP con esta variante de TTN, ya sea en estado de homocigosis o heterocigosis, más de la mitad desarrollaron la enfermedad. En este mismo trabajo se amplió el análisis incluyendo la mutación descrita con anterioridad para PDK4 y reportaron que una o ambas variantes se identificaron en el 99 % de los DP afectados y en el 73 % de los DP sanos de la población evaluada, confirmando la penetrancia incompleta de ambas variantes. Asimismo, del total de DP que

presentaron ambas variantes (homocigota o heterocigota), casi el 60 % desarrolló la enfermedad. Los animales portadores de la variante en el gen TTN, ya sea en estado heterocigoto u homocigoto, tendrían 21 veces más probabilidades de desarrollar CMD que los perros con genotipo de tipo salvaje (Meurs, 2019). A su vez, cuando se combina también con una copia de la variante del gen PDK4 la probabilidad de desarrollar la enfermedad aumenta 30 veces.

En un estudio clínico reciente (Meurs y col., 2020), en el que se evaluó una población de DP con CMD, se identificó que más del 58 % de los perros tenían únicamente la variante TTN, el 20 % de los perros afectados ambas variantes (TTN y PDK4) y sólo un 8 % tenía exclusivamente la variante PDK4. No se encontraron cambios en el desarrollo clínico de la enfermedad en los perros que tenían diferentes mutaciones, sólo se detectó que la edad de presentación en los perros con la variante PDK4 era menor. Este resultado puede estar sesgado por el bajo número de animales con esa mutación, así como el bajo número de perros evaluados y la diferencia entre la cantidad de perros con las diferentes mutaciones. Aún así, queda en evidencia que un 12,5 % de perros con CMD de la población estudiada no presentaban ninguna de estas mutaciones.

Más allá de identificar las mutaciones asociadas al desarrollo de CMD en sí, entender las consecuencias biológicas de estas mutaciones es fundamental para desentrañar los mecanismos subyacentes de las enfermedades genéticas y las variaciones en las características fenotípicas que pueden tener implicaciones significativas en el funcionamiento y la salud de un organismo. Al analizar los efectos de las mutaciones en la función de las proteínas codificadas por los genes o en la expresión génica, se pueden descubrir cómo se alteran las vías bioquímicas

y los procesos celulares vitales. Este conocimiento proporciona información crucial para el diseño de tratamientos médicos precisos, así como para comprender mejor la base genética de los trastornos hereditarios.

A pesar de estos resultados preliminares, aún no se ha podido determinar con certeza los factores genéticos involucrados en la determinación de CMD y las variaciones existentes entre poblaciones de distinto origen genético. El hallazgo de estos factores permitirá identificar los animales portadores con el objetivo de realizar el diagnóstico y tratamiento temprano de la CMD. Asimismo, aportará herramientas para la planificación de cruzamientos selectivos con el fin de disminuir la incidencia de la enfermedad.

3. CMD en otras razas

Como se mencionó previamente esta enfermedad afecta principalmente a razas grandes o gigantes como Bóxer, Gran Danés, Lebel Irlandés, Terranova, Perro de agua Portugués (Sisson y col., 2000). También se ha reportado una alta prevalencia en el Cocker Spaniel, una raza de tamaño mediano (Dutton y col., 2018). En general la edad de presentación es similar en todas las razas, a excepción del Perro de agua Portugués en el que se ha reportado una forma juvenil de CMD a partir de los 2 años de edad (Werner y col., 2008; Dutton y col., 2018). Por otro lado, en los Terranovas la enfermedad se manifiesta a edades más avanzadas y con una progresión más lenta (Dukes-McEwan, 2000), a diferencia del Gran Danés en el que el tiempo de sobrevida una vez manifestada la

enfermedad es muy corto (Martin y col., 2010). En cuanto a la distribución entre sexos, en el Lebel Irlandés se ha demostrado una prevalencia mayor en machos, los cuales a menudo se ven afectados a una edad más joven que las hembras (Vollmar, 2000; Simpson y col., 2016).

Con respecto a los modos de transmisión de la enfermedad, la forma de herencia es particular para cada raza, así como las variantes causales reportadas hasta el momento. Un modo de transmisión autosómico fue sugerido para una línea de Lebreles Irlandeses con CMD (Cobb y col., 1996). Asimismo, dos estudios mostraron que múltiples *loci* tenían asociaciones significativas con el desarrollo de la enfermedad en esta raza (Philipp y col., 2012; Simpson y col., 2016). Dado que la CMD en el Lebel Irlandés tiene una mayor prevalencia en los machos, se estudió el papel de la tafazzin, una proteína producida por un gen ubicado en el cromosoma X (gen TAZ), que se expresa en cantidad en el miocardio (Philipp y col., 2007), aunque no se encontró relación con el desarrollo de la enfermedad. En perros de raza Terranova, el análisis genealógico de un grupo de animales mostró un modo de herencia autosómico dominante con penetrancia incompleta, aunque no se pudo excluir una transmisión autosómica recesiva (Dukes McEwan y Jackson, 2002). Asimismo, en un grupo de perros de esta raza se reportó la baja contractilidad cardíaca como consecuencia de la deficiencia de taurina en la dieta (Backus y col., 2003). En el Gran Danés no se ha identificado ninguna mutación genética asociada al desarrollo de la enfermedad hasta el momento, sin embargo se ha reportado que puede transmitirse como un rasgo recesivo ligado a X (Meurs y col., 2001a). Un análisis genealógico más grande del Reino Unido encontró que el modo de herencia en esta raza fue más

consistente con un rasgo autosómico dominante, aunque la herencia poligénica no pudo ser completamente excluida (Stephenson y col., 2012). En el Cocker Spaniel no se han reportado estudios genéticos hasta el momento, de hecho, no está completamente claro si la CMD del Cocker es una condición hereditaria o si es secundaria a una deficiencia dietética de taurina y carnitina, especialmente en el Cocker Spaniel Americano (Sanderson, 2006). Por otro lado, se sugirió que la forma juvenil de CMD en el Perro de Agua Portugués es transmitida como un rasgo autosómico recesivo (Alroy y col., 2000). Una gran proporción de cachorros tenían bajas concentraciones de taurina en plasma, y al suplementarlos durante dos meses sus valores ecocardiográficos mejoraban, por lo que se planteó una correlación entre un rasgo autosómico recesivo y un metabolismo anormal de la taurina en la infancia de esta raza (Alroy y col., 2005).

3.1. CMD en el Bóxer

Los Bóxers tienen una forma adulta particular de CMD, que no se corresponde con la descripción clásica de la enfermedad. Harpster describió en 1983 tres formas de cardiomiopatía en esta raza, dos de ellas caracterizadas por taquiarritmias ventriculares con o sin manifestaciones clínicas, que actualmente se conoce como cardiomiopatía arritmogénica ventricular derecha (ARVC de la sigla en Inglés *Arrhythmogenic right ventricular cardiomyopathy*; Harpster, 1983). La tercera forma, actualmente considerada una forma de CMD, presenta arritmias ventriculares, supraventriculares y falla cardíaca congestiva, producto de una disfunción sistólica ventricular izquierda y en ocasiones derecha, sumada a la

dilatación de las cámaras cardíacas (Baumwart y col., 2005; Palermo y col., 2011). Wess (2022) describió una subdivisión de la fase 2 de la CMD, en la cual el estadio B1 se caracterizó por anomalías eléctricas sin dilatación cardíaca, lo que incluye la aparición de CPVs en ciertas razas como el DP o Bóxer. La evaluación de una familia de Bóxers con ARVC permitió determinar un modo de herencia autosómico dominante con penetrancia incompleta relacionada a la edad (Meurs y col., 1999).

Entre los estudios genéticos que se realizaron en Bóxers con ARVC, se encontró que la proteína del receptor cardíaco de rianodina (RyR2) y su expresión de ARN mensajero estaban reducidas. Sin embargo, no se demostró ninguna asociación genética con el gen RyR2 (Meurs y col., 2006). Además, se demostró que la expresión del ARN mensajero de calstabin-2 estaba significativamente regulada a la baja en el miocardio, pero tampoco se encontraron mutaciones causales (Oyama y col., 2008).

Meurs y col. (2013), reportaron la asociación entre el desarrollo de la enfermedad en esta raza y la presencia de una delección de ocho pares de bases en el gen de la estriatina (STRN; localizado en el cromosoma 17), lo que genera una reducción en la producción de ARN mensajero para la proteína estriatina. Esta variante está asociada a las otras dos formas de la enfermedad (ARVC) en Bóxers Americanos (Meurs y col., 2010) y en humanos (Beffagna y col., 2005). Esto sugiere que la CMD en muchos de estos animales, como lo descrito en humanos, es otra manifestación de ARVC. Sin embargo, varios de los animales con CMD muestreados en estos trabajos no eran portadores de la variante del gen estriatina, lo que sugiere la posibilidad de que existan otros genes causales en esta

raza. Esto se condice con dos estudios realizados en Reino Unido (Dukes-McEwan y col., 2010; Cattanach y col., 2015), en los que se encontró que la mutación de la estriatina era extremadamente común, pero no explicaba la ocurrencia de la enfermedad en Bóxers afectados.

Hasta la fecha no hay estudios que reporten la presencia de las variantes candidatas descritas en Doberman Pinscher (variante en locus del cromosoma 5, variante en gen PDK4 y variante en gen TTN) en la raza Bóxer. En Argentina, el número de animales registrados anualmente en la Federación Cinológica Argentina es similar al número de DP inscriptos. Por ese motivo se seleccionó esta raza, con el objetivo de conocer si los genes causantes de CMD reportados en el DP participan de la etiología genética en el Bóxer.

4. Hipótesis y objetivos

4.1 Hipótesis

Por todo lo expuesto se hipotetiza que la CMD es una enfermedad genética causada por mutaciones independientes en diferentes regiones cromosómicas. Estas pueden tener además, diferente distribución intra e interracial. Los genes y las mutaciones causales de esta enfermedad podrán ser mapeados mediante la utilización de herramientas genéticas y genómicas.

4.2 Objetivos

Objetivo general:

El objetivo general del presente plan de tesis consiste en mapear el/los gene/s involucrados en el desarrollo de la Cardiomiopatía Dilatada (CMD) en perros de la raza Doberman Pinscher mediante diferentes estrategias de asociación genómica con marcadores genéticos. Para ello se utilizarán técnicas de secuenciación capilar, pirosecuenciación y targeted NGS para genotipar las posibles mutaciones causales en los genes candidatos y evaluar su potencial efecto sobre la estructura de las proteínas y/o la regulación génica.

El hallazgo de estos factores permitirá identificar los animales portadores con el objetivo de realizar el diagnóstico y tratamiento temprano de la CMD. Asimismo, aportará herramientas para la planificación de cruzamientos selectivos, con el fin de disminuir la incidencia de la enfermedad.

Objetivos específicos:

- 1) Genotipificar los diferentes genes y SNPs candidatos reportados previamente asociados al desarrollo de CMD en DP.
 - Desarrollar un método rápido de genotipificación para la variante del gen PDK4 mencionada anteriormente (CFA14: g.20,829,667_20,829,682del) mediante la técnica de secuenciación capilar. Confirmar la presencia de la mutación causal y su asociación

con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.

- Desarrollar un método rápido de genotipificación para la variante del locus en el cromosoma 5 (CFA5:g.53,941,386T>C) mediante la técnica de secuenciación capilar. Confirmar la presencia de la mutación causal y su asociación con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.
 - Desarrollar un método rápido de genotipificación para la variante del gen candidato TTN en el cromosoma 36 (ENSCAFT00000022319; CFA36: g.22,321,955 C>T) mediante la técnica de pirosecuenciación. Confirmar la presencia de la mutación causal y su asociación con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.
 - Genotipificar a partir de target NGS (AgriSeq™ Targeted GBS Thermo Fisher Scientific) la delección de 16pb en el sitio de empalme 5´del intrón 10 del gen PDK4 en el cromosoma 14 (CFA14: g.20,829,667_20,829,682del).
- 2) Validar el método de diagnóstico y determinar la prevalencia de las variantes en poblaciones independientes de Doberman Pinscher.
 - 3) Evaluar la presencia de las variantes anteriormente citadas en una población de Bóxer de Argentina.
 - 4) Evaluar *in silico* el potencial efecto sobre la estructura de la proteína y/o la regulación génica de la variante del gen TTN.
 - 5) Evaluar *in silico* el potencial efecto sobre la estructura de la proteína y/o la regulación génica de la variante del gen PDK4.

Referencias

1. Alroy J, Rush JE, Freeman L, Amarendhra Kumar MS, Karuri A, Chase K, Sarkar S. Inherited infantile dilated cardiomyopathy in dogs: genetic, clinical, biochemical, and morphologic findings. *American Journal of Medical Genetics*. 2000; 95: 57-66.
2. Alroy J, Rush JE, Sarkar S. Infantile dilated cardiomyopathy in Portuguese water dogs: correlation of the autosomal recessive trait with low plasma taurine at infancy. *Amino Acids*. 2005; 28: 51-56.
3. Andelfinger G, Wright KN, Lee HS, Siemens LM, Benson DW. Canine tricuspid valve malformation, a model of human Ebstein anomaly, maps to dog chromosome 9. *J Med Genet*. 2003; 40: 320–4.
4. Backus RC, Cohen G, Pion PD, Good KL, Rogers QR, Fascetti AJ. Taurine deficiency in Newfoundlands fed commercially available complete and balanced diets. *Journal of the American Veterinary Medical Association*. 2003; 223: 1130-1136.
5. Bannasch D, Famula T, Donner J, Anderson H, Honkanen L, Batcher K, Safra N, Thomasy S, Rebhun R. The effect of inbreeding, body size and morphology on health in dog breeds. *Canine Med Genet*. 2021; 8(1):12.
6. Baumwart RD, Meurs KM, Atkins CE, Bonagura JD, DeFrancesco TC, Keene BW, Koplitz S, Luis Fuentes V, Miller MW, Rausch W, Spier AW. Clinical, echocardiographic and electrocardiographic abnormalities in Bóxers with cardiomyopathy and left ventricular systolic dysfunction: 48 cases (1985–2003). *J Am Vet Med Assoc*. 2005; 226: 1102–1104.

7. Beffagna G, Occhi G, Nava A, Vitiello L, Ditadi A, Basso C, Bauce B, Carraro G, Thiene G, Towbin JA, Danieli GA, Rampazzo A. Regulatory mutations in transforming growth factor – B3 gene cause arrhythmogenic right ventricular cardiomyopathy. *Cardiovasc Res.* 2005; 65: 366–373.
8. Björnerfeldt S, Hailer F, Nord M, Vilà C. Assortative mating and fragmentation within dog breeds. *BMC Evol Biol.* 2008; 8: 28.
9. Buchanan JW. Prevalence of cardiovascular disorders. En: Fox P, Sisson D, Moise NS (eds). *Textbook of canine and feline cardiology. Principles and clinical practice* (2nd edition). Philadelphia, Saunders, 1999. p. 457-70.
10. Calboli FCF, Sampson J, Fretwell N, Balding DJ. Population structure and inbreeding from pedigree analysis of purebred dogs. *Genetics.* 2008; 179: 593e601.
11. Calvert CA. Dilated congestive cardiomyopathy in Doberman pinschers. *Compend Contin Educ Prac Vet.* 1986; 8: 417-30.
12. Calvert CA. Update: canine DCM. En: Kirk RW, Bonagura JD (eds). *Current Veterinary Therapy XI.* WB Saunders; 1992. p. 773-92.
13. Calvert CA, Pickus CW, Jacobs GJ, Brown J. Signalment, survival, and prognostic factors in Doberman Pinschers with end-stage cardiomyopathy. *J Vet Intern Med.* 1997; 11: 323-6.
14. Calvert CA, Meurs K. CVT update: Doberman Pinscher occult cardiomyopathy. En: Bonagura JD, Kirk RW, (eds). *Kirk 's Current Veterinary Therapy.* Philadelphia, PA: Saunders; 2000, p. 756-60.

15. Cattanach BM, Dukes-McEwan J, Wotton PR, Stephenson HM, Hamilton RM. A pedigree-based genetic appraisal of Bóxer ARVC and the role of the Striatin mutation. *Vet Rec.* 2015; 176(19): 492.
16. Cobb MA, Brownlie S, Pidduck SE, Batt RM. Evidence for Genetic Involvement in Dilated Cardiomyopathy in the Irish Wolfhound. *Ann Congress BSAVA, Birmingham.* 1996; p 215.
17. Dambach DM, Lannon A, Sleeper MM, Buchanan J. Familial dilated cardiomyopathy of young Portuguese water dogs. *J Vet Intern Med.* 1999; 13: 65-71.
18. Detweiler DK, Patterson DF. The prevalence and types of cardiovascular disease in dogs. *Ann N Y Acad Sci.* 1965; 127: 481-516.
19. Domanjko-Petric A, Stabej P, Zemva A. Dilated cardiomyopathy in Doberman Pinschers: survival, causes of death and a pedigree review in a related line. *J Vet Cardiol.* 2002; 4: 17-24.
20. Dukes-McEwan J. Dilated cardiomyopathy (DCM) in Newfoundland dogs. *ACVIM Forum.* Seattle, WA, USA. 2000; p 118-119.
21. Dukes-McEwan J, Jackson IJ. The promises and problems of linkage analysis by using the current canine genome map. *Mammalian Genome.* 2002; 13: 667-672.
22. Dukes-McEwan J, Stephenson H, Wotton PR, Cattanach B. Cardiomyopathy in Bóxer dogs. *Veterinary Times.* 2010; 40: 6-9.
23. Dutton E, López-Alvarez J. An update on canine cardiomyopathies - is it all in the genes? *J Small Anim Pract.* 2018; p: 455-464.

24. Goodwin JK, Cattiny G, Rouge B. Further characterization of Bóxer cardiomyopathy. Proc. 13th ACVIM FORUM. Lake Buena Vista, FL. 1995.
25. Grynberg A, Demaison L. Fatty acid oxidation in the heart. J Cardiovasc Pharmacol. 1996;28 Suppl 1:S11-7.
26. Grünig E, Tasman JA, Kücherer H, Franz W, Kübler W, Katus HA. Frequency and phenotypes of familial dilated cardiomyopathy. J Am Coll Cardiol. 1998; 31:186-94.
27. Häggström J, Kwart C, Hansson K. Heart sounds and murmurs: changes related to severity of chronic valvular disease in the Cavalier King Charles spaniel. J Vet Intern Med. 1995; 9: 75-85.
28. Hammer TA, Venta PJ, Eyster GE. The genetic basis of dilated cardiomyopathy in Doberman Pinschers. Anim Genet. 1996; 27: 109.
29. Harpster N. Bóxer cardiomyopathy. In: Kirk R (ed.). Current Veterinary Therapy VIII. Philadelphia, PA:WB Saunders; 1983, p. 329–337.
30. Hazlett MJ, Maxie MG, Allen DG, Wilcock BP. A retrospective study of heart disease in Doberman Pinscher dogs. Can Vet J. 1983; 24: 205-210.
31. Hein S, Kostin S, Heling A, Maeno Y, Schaper J. The role of the cytoskeleton in heart failure (Review). Cardiovascular Res. 2000; 45: 273-278.
32. Leroy G. Genetic diversity, inbreeding and breeding practices in dogs: results from pedigree analyses. Vet J. 2011; 189: 177-182.
33. Martin MW, Stafford Johnson MJ, Strehlau G, King JN. Canine dilated cardiomyopathy: a retrospective study of prognostic findings in 367 clinical cases. Journal of Small Animal Practice. 2010; 51: 428-436.

34. Mausberg TB, Wess G, Simak J, Keller L, Drögemüller M, Drögemüller C, Webster MT, Stephenson H, Dukes-McEwan J, Leeb T. A locus on chromosome 5 is associated with dilated cardiomyopathy in Doberman Pinschers. *PloS one*. 2011; 6:e20042.
35. McNally EM, Golbus JR, Puckelwartz MJ. Genetic mutations and mechanisms in dilated cardiomyopathy. *J Clin Invest*. 2013; 123(1): 19-26.
36. Mellanby RJ, Ogden R, Clements DN, French AT, Gow AG, Powell R. Population structure and genetic heterogeneity in popular dog breeds in the UK. *Vet J*. 2013; 196(1): 92–97.
37. Mestroni L, Rocco C, Gregori D, Sinagra G, Di Lenarda A, Miocic S, Vatta M, Pinamonti B, Muntoni F, Caforio AL, McKenna WJ, Falaschi A, Giacca M, Camerini. Familial dilated cardiomyopathy: evidence for genetic and phenotypic heterogeneity. *J Am Coll Cardiol*. 1999; 34: 181-190.
38. Meurs KM. Insights into the heritability of canine cardiomyopathy. *Vet Clin North Am Small Anim Pract*. 1998; 28: 1449-1157.
39. Meurs KM, Spier AW, Miller MW, Lehmkuhl L, Towbin JA. Familial ventricular arrhythmias in Bóxers. *J Vet Intern Med*. 1999; 13(5): 437-439.
40. Meurs KM, Miller MW, Wright NA. Clinical features of dilated cardiomyopathy in Great Danes and results of a pedigree analysis: 17 cases (1990e2000). *J Am Vet Med Assoc*. 2001a; 218: 729-732.
41. Meurs KM, Magnon AL, Spier AW, Miller MW, Lehmkuhl LB, Towbin JA. Evaluation of the cardiac actin gene in Doberman Pinschers with dilated cardiomyopathy. *Am J Vet Res*. 2001b; 62: 33–36.

42. Meurs KM, Lacombe VA, Dryburgh K, Fox PR, Reiser PR, Kittleson MD. Differential expression of the cardiac ryanodine receptor in normal and arrhythmogenic right ventricular cardiomyopathy canine hearts. *Hum Genet.* 2006; 120(1): 111-118.
43. Meurs KM, Fox PR, Horgard M, Spier AW, Lamb A, Koplitz SL, Baumwart RD. A prospective genetic evaluation of familial dilated cardiomyopathy in the Doberman Pinschers. *J Vet Intern Med.* 2007; 21: 1016-1020.
44. Meurs KM, Hendrix KP, Norgard MM. Molecular evaluation of five cardiac genes in Doberman Pinschers with dilated cardiomyopathy. *Am J Vet Res.* 2008; 69: 1050-1053.
45. Meurs KM, Mauceli E, Lahmers S, Acland GM, White SN, Lindblad-Toh K. Genome-wide association identifies a deletion in the 3' untranslated region of striatin in a canine model of arrhythmogenic right ventricular cardiomyopathy. *Hum Genet.* 2010; 128(3): 315-324.
46. Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. *Hum Genet.* 2012; 131: 1319-1325.
47. Meurs KM, Stern JA, Sisson DD, Kittleson MD, Cunningham SM, Ames MK, Atkins CE, DeFrancesco T, Hodge TE, Keene BW, Reina Doreste Y, Leuthy M, Motsinger-Reif AA, Tou SP. Association of dilated cardiomyopathy with the striatin mutation genotype in Bóxer dogs. *J Vet Intern Med.* 2013; 27:1437-1440.
48. Meurs KM, FriedenberG SG, Kolb J, Saripalli C, Tonino P, Woodruff K, Olby NJ, Keene BW, Adin DB, Yost OL, DeFrancesco TC, Lahmers S, Tou S, Shelton

GD, Granzier H. A missense variant in the titin gene in Doberman pinscher dogs with familial dilated cardiomyopathy and sudden cardiac death. *Hum Genet.* 2019; 138(5):515-524.

49. Meurs K. [Internet]. Webinar on dilated cardiomyopathy. Lab at North Carolina State University; 2019. (Consultado en Octubre de 2019). Disponible en: <https://hospital.cvm.ncsu.edu/services/small-animals/genetics/doberman-dcm/>.

50. Meurs KM, Stern JA, Adin D, Keene BW, De Francesco TC, Tou SP. Assessment of PDK4 and TTN gene variants in 48 Doberman Pinschers with dilated cardiomyopathy. *J Am Vet Med Assoc.* 2020; 15;257(10):1041-1044.

51. Monnet E, Orton C, Salman M, Boon J. Idiopathic DCM in dogs: Survival and prognostic indicators. *J Vet Intern Med.* 1995; 9: 12-17.

52. Nicholas FW. Introduction to veterinary genetics. Ames: WileyBlackwell (3th edition). 2010.

53. ' Grad MR o rne R. The prevalence of dilated cardio opath in Doberman Pinschers: A 4.5 year follow-up. *J Vet Intern Med.* 1998; 12:199.

54. ' Grad MR 'Sullivan ML. Dilated cardio opath : An update. *Vet Clin North Am Small Anim Pract.* 2004; 34: 1187-1207.

55. ' Sullivan ML ' Grad MR P le WG Dawson JF “ va luation of genes encoding cardiac proteins in Doberman Pinschers with dilated cardio opath ” *American Journal of Veterinary Research.* 2011; vol. 72, no. 7, p. 932–938.

56. Olsen LH, Fredholm M, Pedersen HD. Epidemiology and inheritance of mitral valve prolapse in Dachshunds. *J Vet Intern Med.* 1999; 13: 448-456.

57. Ostrander EA, Kruglyak L. Unleashing the Canine Genome. *Genome Res.* 2000; 10: 1271–1274.

58. Owczarek-Lipska M, Mausberg TB, Stephenson H, Dukes-McEwan J, Wess G, Leeb T. A 16-bp deletion in the canine PDK4 gene is not associated with dilated cardiomyopathy in a European cohort of Doberman Pinschers. *Anim Genet.* 2013; 44: 239.
59. Oyama MA, Reiken S, Lehnart SE, Chittur SV, Meurs KM, Stern J, Marks AR. Arrhythmogenic right ventricular cardiomyopathy in Bóxer dogs is associated with calstabin2 deficiency. *J Vet Cardiol.* 2008; 10(1): 1-10.
60. Palermo V, Johnson MJS, Sala E, Brambilla PG, Martin MW. Cardiomyopathy in Bóxer dogs: A retrospective study of clinical presentation, diagnostic findings and survival. *J Vet Cardiol.* 2011; 13: 45–55.
61. Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB. Genetic structure of the purebred domestic dog. *Science.* 2004; 304: 5674.
62. Parker HG, Meurs KM, Ostrander EA. Finding cardiovascular disease genes in the dog. *J Vet Cardiol.* 2006; 8(2):115-127.
63. Patterson DF, Pyle RL, Buchanan JW, et al. Hereditary patent ductus arteriosus and its sequelae in the dog. *Circ Res.* 1971; 29: 1-13.
64. Patterson DF. Congenital defects of the cardiovascular system of dogs: studies in comparative cardiology. *Adv Vet Sci Comp Med.* 1976; 20: 1-37.
65. Patterson DF. A catalogue of genetic disorders of the dog. In: Kirk RW (ed.). *Current veterinary therapy VII.* Philadelphia, Saunders; 1980.
66. Patterson D. Companion animal medicine in the age of medical genetics. *J. Vet. Internal. Med.* 2000; 14: 1-9.

67. Philipp U, Broschk C, Vollmar A, Distl O. Evaluation of tafazzin as candidate for dilated cardiomyopathy in Irish wolfhounds. *Journal of Heredity*. 2007; 98: 506-509.
68. Philipp U, Vollmar A, Häggström J, Thomas A, Distl O. Multiple loci are associated with dilated cardiomyopathy in irish wolfhounds. *PLoS One*. 2012; 7: 1-6.
69. Posafalvi A, Herkert JC, Sinke RJ, van den Berg MP, Mogensen J, Jongbloed JD, van Tintelen JP. Clinical utility gene card for: dilated cardiomyopathy (CMD). *Eur J Hum Genet*. 2013; 21(10).
70. Sanderson SL. Taurine and carnitine in canine cardiomyopathy. *Veterinary Clinics of North America: Small Animal Practice*. 2006; 36: 1325-1343.
71. Simpson S, Dunning M, Brownlie S, Patel J, Godden M, Cobb M, Mongan NP, Rutland CS. Multiple genetic associations with Irish wolfhound dilated cardiomyopathy. *BioMed Research International*. 2016; 1-14.
72. Sisson D, Thomas W, Keene B. Primary myocardial disease in the dog. En: *Ettinger SJ, Feldman EC (eds). Textb. Vet. Intern. Med. 5th ed. Philadelphia: Saunders, 2000, p. 874-895.*
73. Sosa I, Estrada AH, Winter BD, Erger KE, Conlon TJ. In vitro evaluation of mitochondrial dysfunction and treatment with adeno-associated virus vector in fibroblasts from Doberman Pinschers with dilated cardiomyopathy and a pyruvate dehydrogenase kinase 4 mutation. *Am J Vet Res*. 2016; 77(2): 156-161.
74. Stabej P, Imholz A, Versteeg SA, Zijlstra C, Stokhof AA, Domanjko-Petric A, Leegwater PA, van Oost BA. Characterization of the canine desmin (DES) gene

and evaluation as a candidate gene for dilated cardiomyopathy in the Doberman. *Gene*. 2004; 340: 241-249.

75. Stabej P, Leegwater PA, Stokhof AA, Domanjko-Petric A, van Oost BA. Evaluation of the phospholamban gene in purebred large-breed dogs with dilated cardiomyopathy. *Am J Vet Res*. 2005; 66: 432-436.

76. Stephenson HM, Fonfara S, López-Alvarez J, Cripps P, Dukes-McEwan J. Screening for dilated cardiomyopathy in Great Danes in the United Kingdom. *Journal of Veterinary Internal Medicine*. 2012; 26: 1140-1147.

77. Su erfield NJ, Boswood A , Grad MR, Gordon SG, Dukes-McEwan J, Oyama MA, Smith S, Patteson M, French AT, Culshaw GJ, Braz-Ruivo L, strada A , Sullivan ML, Loureiro J, Willis R, Watson P. Efficacy of pimobendan in the prevention of congestive heart failure or sudden death in Doberman Pinschers with preclinical dilated cardiomyopathy (the PROTECT Study). *J Vet Intern Med*. 2012; 26: 1337-1349.

78. Taggart K, Estrada A, Thompson P, Lourenco F, Kirmani S, Suzuki-Hatano S, Pacak CA. PDK4 deficiency induces intrinsic apoptosis in response to starvation in fibroblasts from Doberman Pinschers with dilated cardiomyopathy. *Biores Open Access*. 2017; 6: 182-191.

79. Tidholm A, Jönsson L. A retrospective study of canine dilated cardiomyopathy (189 cases). *Journal of the American Animal Hospital Association*. 1997; 33: 544-550.

80. Tidholm A, Haggstrom J, Borgarelli M, Tarducci A. Canine idiopathic dilated cardiomyopathy. Part I: Aetiology, clinical characteristics, epidemiology and pathology. *Vet J*. 2001; 162: 92–107.

81. Towbin JA, Bowles NE. Genetic abnormalities responsible for dilated cardiomyopathy. *Curr Cardiol Rep.* 2000; 2: 475-480.
82. Vollmar AC. The prevalence of cardiomyopathy in the Irish wolfhound: a clinical study of 500 dogs. *Journal of the American Animal Hospital Association.* 2000; 36: 125-132.
83. Werner P, Raducha MG, Prociuk U, Sleeper MM, Van Winkle TJ, Henthorn PS. A novel locus for dilated cardiomyopathy maps to canine chromosome 8. *Genomics.* 2008; 91: 517-521.
84. Wess G, Schulze A, Butz V, Simak J, Killich M, Keller LJM, Maeurer J, Hartmann K. Prevalence of dilated cardiomyopathy in Doberman Pinschers in various age groups. *J Vet Intern Med.* 2010; 24: 533-538.
85. Wess G, Domenech O, Dukes-McEwan J, Häggström J, Gordon S. European Society of Veterinary Cardiology screening guidelines for dilated cardiomyopathy in Doberman pinschers. *Journal of Veterinary Cardiology.* 2017; 19: 405-415.
86. Wess G. Screening for dilated cardiomyopathy in dogs. *J Vet Cardiol.* 2022; 40: 51-68.
87. Wiersma AC, Leegwater PAJ, A Van Oost B, Ollier WE, Dukes-McEwan J. Canine candidate genes for dilated cardiomyopathy: annotation of and polymorphic markers for 14 genes. *BMC veterinary research.* 2007; 3: 28-35.

CAPÍTULO I

ANÁLISIS DE LAS VARIANTES GENÉTICAS INVOLUCRADAS EN EL DESARROLLO DE CARDIOMIOPATÍA DILATADA EN PERROS DE RAZA DOBERMAN PINSCHER Y BÓXER, MEDIANTE DIFERENTES ESTRATEGIAS DE ASOCIACIÓN GENÓMICA UTILIZANDO MARCADORES GENÉTICOS.

Introducción

La cardiomiopatía dilatada (CMD) es una enfermedad hereditaria con alta incidencia en razas como el Doberman Pinscher (DP) y el Bóxer (Tidholm y col., 2001). Esta enfermedad es una afección progresiva que genera una disfunción cardíaca irreversible, derivando en insuficiencia cardíaca y posterior muerte del animal (Calvert y col., 2000). Si bien lo más frecuente es que el diagnóstico se realice durante la fase sintomática de la enfermedad, el empleo de métodos complementarios de diagnóstico permite detectar cambios morfológicos del corazón y/o arritmias en la fase temprana de la enfermedad (también llamada fase oculta; Domanjko-Petric y col., 2002). La CMD se caracteriza por presentarse en animales adultos, es decir en plena etapa reproductiva. Esto se traduce en que los animales pueden cruzarse dejando descendencia antes que la enfermedad pueda ser detectada, contribuyendo a su permanencia en la población (Wess y col., 2010).

Hasta el momento, se han reportado tres variantes candidatas asociadas al desarrollo de esta enfermedad. La primera se trata de un locus (SNP alelo "C") en el cromosoma 5 que se asoció al desarrollo de CMD en el 50% de los DP de Alemania y Reino Unido (Mausberg y col., 2011). Los perros con arritmias aisladas tenían una frecuencia significativamente más alta del alelo C, pero no fue posible asociar un gen candidato para CMD ligado al polimorfismo asociado. Meurs y col. (2012) identificaron, en una población de DP de Estados Unidos, una delección de 16 pares de bases en un sitio de empalme del gen piruvato deshidrogenasa quinasa 4 (PDK4), ubicado en el cromosoma 14, asociado con el desarrollo de CMD. Sin embargo, esta variante candidata fue cuestionada más tarde, ya que no se encontró asociada al desarrollo de CMD en una cohorte de DP de Europa (Owczarek-Lipska y col., 2013). Por otra parte, se reportó un cambio de un solo nucleótido (C > T) en el gen titina (TTN) vinculado al desarrollo de CMD en DP (Meurs y col., 2019). Esta variante no ha sido aún estudiada en otras poblaciones de DP hasta la fecha.

A pesar de estos resultados, no se ha podido determinar con certeza los factores genéticos involucrados en la determinación de CMD y las variaciones existentes entre poblaciones de distinto origen genético. Asimismo, no hay estudios que reporten la presencia de las variantes candidatas descritas en DP en la raza Bóxer, otra de las razas más afectadas por esta enfermedad.

Las variantes descritas son utilizadas como marcadores genéticos, es decir secuencias de ADN que presentan polimorfismos detectables, heredables y que poseen alta variabilidad (Villegas Castagnasso y col., 2021). Una vez que se ha obtenido el ADN, los marcadores genéticos se pueden analizar utilizando

diferentes técnicas de laboratorio, aumentando en primera instancia el número de copias del segmento de ADN que contiene al marcador mediante la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés *Polymerase Chain Reaction*) y posteriormente utilizando una técnica determinada que permita identificar el genotipo buscado. En estudios de asociación caso-control se han utilizado técnicas de genotipificación indirecta, como PCR-análisis de fragmentos en gel de poliacrilamida y PCR alelo-específica (PCR-AS) (Villegas Castagnasso y col., 2015; Villegas Castagnasso y col., 2021), y técnicas directas, como la secuenciación Sanger de ADN de fragmentos de amplificación (PCR-SBT, por *sequencing based typing method*; Sanger y col., 1977), pirosecuenciación (Ronaghi y col., 1998) y secuenciación masiva de próxima generación dirigida (NGS, *New generation sequencing methods*; Mertes y col., 2011). Asimismo, se utilizaron estudios de asociación de genoma completo (GWAS, por sus siglas en inglés *Genome Wide Association Study*).

Si bien estas técnicas contribuyeron a ampliar los conocimientos sobre la etiopatogenia de la CMD, los resultados contradictorios entre las distintas poblaciones plantean la necesidad de profundizar el estudio sobre los genes involucrados en el desarrollo de esta enfermedad. Esto, sumado a la creación de métodos moleculares de diagnóstico utilizando diferentes herramientas de genotipificación permitiría identificar los animales portadores con el objetivo de realizar el diagnóstico y tratamiento temprano de la CMD, aportando herramientas para la planificación de cruzamientos selectivos, con el fin de disminuir la incidencia de la variante.

Por todo lo expuesto en el presente capítulo se desarrollarán los siguientes objetivos:

- 1) Genotipificar los genes y SNPs candidatos reportados previamente asociados al desarrollo de CMD en DP.
 - a) Desarrollar un método rápido de genotipificación para la variante del gen PDK4 (CFA14: g.20,829,667_20,829,682del) mediante la técnica de secuenciación capilar. Confirmar la presencia de la mutación causal y su asociación con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.
 - b) Desarrollar un método rápido de genotipificación para la variante del locus en el cromosoma 5 (CFA5:g.53,941,386T>C) mediante la técnica de secuenciación capilar. Confirmar la presencia de la mutación causal y su asociación con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.
 - c) Desarrollar un método rápido de genotipificación para la variante del gen candidato TTN en el cromosoma 36 (ENSCAFT00000022319; CFA36: g.22,321,955 C>T) mediante la técnica de pirosecuenciación. Confirmar la presencia de la mutación causal y su asociación con el desarrollo de CMD en las poblaciones locales de Doberman Pinscher.
 - d) Genotipificar mediante la técnica de target NGS (AgriSeq™ Targeted GBS, Thermo Fisher Scientific, Waltham, Ma, EE. UU.) la delección de 16pb en el sitio de empalme 5' del intrón 10 del gen PDK4 en el cromosoma 14 (CFA14: g.20,829,667_20,829,682del).

- 2) Validar los métodos de detección y determinar la prevalencia de las variantes en poblaciones independientes de Doberman Pinscher.
- 3) Evaluar la presencia de las variantes anteriormente citadas en una población de Bóxer de Argentina.

Materiales y métodos

El presente estudio fue evaluado y aprobado por el Comité Institucional para el Cuidado y Uso de Animales de Laboratorio (CICUAL) de la Facultad de Ciencias Veterinarias de la Universidad Nacional de La Plata, Argentina (Exp.: 56-1-16T).

1. Animales de experimentación

Población de Dobermans Pinscher

Para el estudio de los polimorfismos de los genes potencialmente involucrados en CMD se utilizaron 72 caninos de raza DP de ambos sexos (43 hembras y 29 machos), de 1 a 14 años de edad ($7,4 \pm 2,9$ años). Estos animales fueron atendidos en el Hospital Escuela (FCV-UNLP), derivados de veterinarias privadas o pertenecientes a distintos criaderos de DP de Argentina. En todos los casos se solicitó por escrito el correspondiente consentimiento informado (Anexo 1) del propietario, detallando el tipo de uso y estudio al que iba a ser sometido el

animal. A todos los animales se les asignó un código identificador, y se obtuvieron datos de reseña y genealógicos, entre ellos: raza, sexo, edad/fecha de nacimiento, peso, pelaje, lugar de dónde provino la muestra y datos de pedigrí (nombre del criadero, padre y madre en caso de conocerlos; Fig. 1).

Figura 1: Planilla de datos de reseña y datos genealógicos.

DATOS DE RESEÑA Y DATOS GENEALÓGICOS		
Código ^a :	Propietario:	Contacto:
Nombre:	Raza:	Sexo:
Edad/Fecha nacimiento:	Peso:	Pelaje:
Origen ^b :	Pedigrí ^c :	

a) Código de laboratorio interno que se le otorgó a la muestra, b) Origen: Lugar de donde provino la muestra (Hospital Escuela de la Facultad, veterinarias particulares, criaderos), c) Pedigrí: nombre del criadero, padre y madre del animal si los conoce.

En base a los hallazgos de la evaluación clínica, ecocardiográfica y electrocardiográfica (ver a continuación) los animales se clasificaron en los siguientes grupos experimentales.

- Grupo caso: caninos Doberman Pinscher \geq año que presentaron los siguientes signos morfológicos: diámetro diastólico del ventrículo izquierdo \geq diámetro sistólico del ventrículo izquierdo \geq (Kittleson, 1998; Domanjko-Petric ' Grad) con o sin alteraciones eléctricas y/o clínicas (n = 27).

- Grupo control: caninos Dober a n Pinscher ≥ 7 años que no presentaron signos clínicos, eléctricos, ni morfológicos (n = 45) (Mausberg, 2011; Steudemann, 2013).

Población de Bóxers

Para corroborar la presencia de las variantes en una población de perros de raza Bóxer se utilizaron 18 caninos de ambos sexos (10 hembras y 8 machos), de 3 a 14 años de edad ($8,5 \pm 3,2$ años). Estos animales fueron atendidos en el Hospital Escuela (FCV-UNLP) o derivados de veterinarias privadas, y se solicitó el correspondiente consentimiento informado (Anexo 1). De todos los animales se obtuvieron los mismos datos que en la población de DP, consignados en la figura 1. En base a los hallazgos de la evaluación clínica, ecocardiográfica y electrocardiográfica (ver a continuación) los animales se clasificaron en los siguientes grupos: grupo sano (sin alteraciones clínicas, eléctricas ni morfológicas), grupo ARVC únicamente con alteraciones eléctricas, sin alteraciones morfológicas (ARVC/estadio B1 de CMD; Harpster, 1983; Wess, 2022) y grupo CMD, con alteraciones morfológicas, con o sin alteraciones clínicas y eléctricas.

2. Evaluación clínica y cardiológica

A cada uno de los perros incorporados en los protocolos experimentales, se le realizó en primera instancia una anamnesis detallada, que indagaba sobre la

presencia de signos de enfermedad cardiovascular como tos, disnea, ascitis, edema periférico, intolerancia al ejercicio, enfermedades concomitantes, estado de ánimo, apetito o algún otro antecedente que el propietario considere relevante. Estos datos incluían evolución, métodos complementarios de diagnósticos realizados con anterioridad y tratamientos a los que fueron sometidos. A continuación se realizó un examen físico general y particular orientado al sistema cardiovascular, considerando la presencia de signos de enfermedad cardíaca (Kittleson, 1998), el que incluía condición corporal, mucosas (coloración, tiempo de llenado capilar), palpación del área cardíaca, pulso femoral (ritmo e intensidad), pulso yugular, frecuencia y tipo respiratorio, frecuencia y ritmo cardíaco, auscultación torácica, presencia de soplo, presencia de ascitis, presencia de edema periférico, temperatura, palpación/percusión abdominal (Fig. 2).

Figura 2: Planilla de evaluación clínica y cardiológica.

EVALUACIÓN CLÍNICA Y CARDIOLÓGICA	
Código del animal ^a :	Condición corporal:
Mucosas ^b :	Palpación área cardíaca:
Pulso femoral ^c :	Pulso yugular:
Frecuencia y tipo respiratorio:	Frecuencia y ritmo cardíaco:
Auscultación torácica:	Soplo:
Ascitis:	Edema periférico:
Temperatura:	Palpación/percusión abdominal:

a) Código de laboratorio interno que se le otorgó al animal, b) Mucosas: coloración y tiempo de llenado capilar, c) Pulso femoral: ritmo e intensidad.

3. Evaluación ecocardiográfica

Las evaluaciones ecocardiográficas bidimensionales y en modo M se realizaron con un ecógrafo SonoScape modelo A5V con transductor microconvex de 4 a 9 MHz (China), posicionando a los pacientes en estación (Chetboul y col., 2005). Para la obtención de los cortes ecocardiográficos, se acopló el transductor a la piel del tórax previamente rasurada. Se utilizaron tanto la ventana paraesternal derecha como la izquierda (Boon, 1998; Thomas y col., 1993). No se requirió sedación previa en ningún caso. A partir de un eje corto del ventrículo izquierdo (VI) a nivel de los músculos papilares, y mediante el modo M, se registraron las medidas del diámetro diastólico (DDVI; mm) y sistólico (DSVI; mm) del VI, espesores diastólico (TIVD; mm) y sistólico (TIVS; mm) del tabique interventricular y espesores diastólico (PLD; mm) y sistólico (PLS; mm) de la pared libre del VI. Para evaluar la función sistólica se calculó la fracción de acortamiento (FA; %) como $(DDVI-DSVI)/DDVI \times 100$. Desde el eje largo de la ventana paraesternal derecha se midió el diámetro atrial izquierdo (AI; mm), desde anterior a posterior, a lo largo de una línea paralela al anillo mitral, que divide al atrio en dos mitades iguales (Boon, 1998; Rishniw y Erb, 2000). Con el mismo corte ecocardiográfico se evaluó con el modo M el punto de separación septal E de la válvula mitral (PSSE). Para minimizar la variabilidad, se realizaron tres mediciones consecutivas de cada parámetro, las cuales fueron realizadas siguiendo las guías de recomendaciones específicas (Thomas y col., 1993; Gottdiener y col., 2004; Nagueh y col., 2009).

4. Evaluación electrocardiográfica

Luego de la evaluación clínica y ecocardiográfica, se realizaron electrocardiogramas (ECG) en 10 derivaciones: 6 derivaciones en plano frontal (DI, DII, DIII, aVR, aVL, aVF) y 4 precordiales (CV6LU, CV6LL, CV5RL, V10). Para el registro de los ECG se posicionó a los pacientes en decúbito lateral derecho sobre una superficie aislante. Los animales fueron sostenidos manualmente con los miembros perpendiculares al cuerpo y ligeramente separados (Ferasin y col., 2010). Los electrodos se acoplaron a la piel con clips metálicos de tipo cocodrilo y se los humedecieron con alcohol para mejorar la transmisión eléctrica. A continuación se procedió a realizar los ECG en un papel calibrado, según reportes previos (Tilley, 1992; Kittleson, 1998), a una velocidad de papel de 50 mm/seg, ajustando la sensibilidad de acuerdo a cada caso. La derivación II se registró nuevamente a una velocidad de 25 mm/seg para el análisis de la frecuencia cardíaca (FC; lpm) y ritmo cardíaco. Se utilizó un electrocardiógrafo Cardio-Técnica digital monocanal modelo RG- 401 (Argentina). Luego de la realización del registro, los ECG fueron analizados. Se midieron manualmente a una velocidad de papel de 50 mm/seg el eje eléctrico medio (EEM; grados), la duración (Pd; ms) y amplitud (Pa; mV) de la onda P, el intervalo P-R (PR; ms), la duración (QRSd; ms) y amplitud (QRSa; mV) del complejo QRS, el intervalo Q-T (QT; ms) y el segmento ST (ST; mV; Tilley, 1992). Además, se midió el intervalo R-R (RR; ms) entre dos complejos consecutivos, con el cual se corrigió el intervalo QT (QTc) utilizando la fórmula de Van de Water ($QTc = QT - 0,087(RR - 1000)$; Tattersall y col., 2006).

5. Extracción de sangre y ADN

Se extrajeron de cada animal 3 ml de sangre con EDTA (ácido etilendiaminotetraacético) como anticoagulante de las venas safena o cefálica. Los tubos con sangre se rotularon con el código de laboratorio asignado a cada perro al momento de registrar sus datos. Las muestras se mantuvieron refrigeradas luego de la extracción y se conservaron en freezers a -20°C hasta su procesamiento.

El ADN genómico se extrajo de las muestras de sangre utilizando la técnica clásica de purificación mediante solventes orgánicos (fenol:cloroformo; Anexo 2.1) o el kit comercial de extracción Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA; Anexo 2.2). La integridad del ADN genómico obtenido se evaluó mediante corridas electroforéticas en geles de agarosa al 1% - TBE 1X durante 30 min. a 100 volts (Anexo 3). Posteriormente, el ADN obtenido se cuantificó utilizando un espectrofotómetro NanoVue™ Plus (GeneAldCare, U.S.A.) o Multiskan™ G (Thermo Scientific, U.S.A.).

6. Métodos de genotipificación de potenciales mutaciones causales de CMD

Con el fin de evaluar la presencia de los alelos causales de las tres mutaciones reportadas hasta el momento, en poblaciones locales de DP, las muestras de ADN se genotipificaron utilizando diferentes técnicas: PCR-análisis de fragmento en geles de poliacrilamida (para la variante CFA14:g.20,829,667_20,829,682del), PCR-alelo específico (PCR-AS; para la variante SNP CFA5:g.53,941,386T>C) y pirosecuenciación (para la variante

CFA36:g.22,321,955C>T). Además, se emplearon las técnicas de secuenciación directa de fragmentos de amplificación (PCR-SBT) y NGS target para validar las metodologías de genotipificación indirecta desarrolladas en el presente plan. La técnica de PCR-SBT se utilizó para validar la variante CFA14:g.20,829,667_20,829,682del del gen PDK4 y el SNP CFA5:g.53,941,386T>C. Asimismo, los resultados de la variante CFA14:g.20,829,667_20,829,682del en un pool de animales fueron validados y comparados con los obtenidos mediante un método de tipificación de enfermedades genéticas basado en NGS Target.

a) PCR-análisis de fragmento en geles de poliacrilamida

La detección de polimorfismos mediante PCR se basa en la amplificación *in vitro* de una secuencia específica de ADN, utilizando una enzima llamada Taq polimerasa, que sintetiza y copia el ADN durante ciclos repetidos de calentamiento y enfriamiento (Villegas Castagnasso y col., 2021). La técnica consta de tres etapas principales: desnaturalización, apareamiento de cebadores y extensión de la cadena. Durante la desnaturalización, las hebras de ADN se separan mediante calor. Luego, en la etapa de apareamiento, los cebadores de ADN se unen a las secuencias objetivo. Finalmente, en la etapa de extensión, la Taq polimerasa sintetiza nuevas hebras de ADN complementarias utilizando nucleótidos libres. La PCR se lleva a cabo en un termociclador. La eficacia de la PCR se evalúa mediante la electroforesis en gel, una técnica que separa las moléculas de ADN según su tamaño y carga eléctrica. Se utilizan marcadores de peso molecular (fragmentos de ADN de tamaño conocido) para estimar el tamaño

de las moléculas amplificadas. Las bandas de ADN amplificado pueden visualizarse utilizando un colorante de ADN (bromuro de etidio) que se intercala entre los pares de bases de ADN y cuando se ilumina con luz ultravioleta emite fluorescencia. La electroforesis en gel de poliacrilamida ofrece una mayor resolución que el gel de agarosa, ya que los poros son más pequeños, y se utiliza para separar moléculas más pequeñas (PCR- análisis de fragmentos en gel de poliacrilamida).

Diseño de un método de PCR-análisis de fragmento en geles de poliacrilamida para la detección de la variante del gen PDK4

En primera instancia se obtuvo la secuencia de ADN depositada en la base de datos GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) correspondiente al gen PDK4 propuesto como causal de la enfermedad CMD. Con esta secuencia de ADN (Gene ID: 482310) se diseñaron un par de primers que flanquean la mutación causal con el software PyroMark Assay Design SX 2.0 (www.qiagen.com) (Tabla 1). En la fig. 3 se muestra un fragmento de la secuencia del gen PDK4 que incluye la mutación reportada asociada al desarrollo de CMD y la localización de los primers Forward y Reverse.

Tabla 1. Secuencia de los *primers* diseñados para la amplificación del fragmento analizado por PCR-análisis de fragmento en geles de poliacrilamida del gen Piruvato Deshidrogenasa Quinasa 4 (PDK4).

Gen	Secuencia 5'-3'	Tamaño	Tm e (°C)	Región
PDK4 (CFA14: g.20,829,667_2 0,829,682del)	Forward 5'-GTTTTGGTTATGGCTTACCAATT- 3'	210pb	52,0	INTRON 10 del gen
	Reverse 5'-ATGGACTCTCTCTCTCAAATA- 3'	194pb	52,2	PDK4

Tm e= temperatura de melting estimada.

Figura 3: Fragmento del gen Piruvato Deshidrogenasa Quinasa 4 (PDK4) del cromosoma 14. Se pueden identificar la variante reportada CFA14: g.20,829,667_20,829,682del (sombreado amarillo) y los primers forward y reverse (sombreado gris).

GCAGCAGGGAGGATAATAACTCCTTCTTACATTTTGTATTTCAATGATC
 ATGTCCACTTCTTTGCCAGTAAGTATCTTTCCCTCTAGGCTGGTTTTG
 GTTATGGCTTACCAATTTCTCGTCTCTATGCCAAGTACTTTCAAGGAGA
 TCTGAATCTCTACTCTATGTCAGGATATGGAACAGATGCCGTCATCTAC
 TTAAAG[GTATCCTTTCAACCCA/-]GTAGAAATAATAATATTCTATTTT
 TAAGATTTAATTAATTAATTAATTTATTTATAAATTTATTTGAGAGAGA
 GAGAGTCCATGCAAGTGGGGGAAGGGGCACAGGGAGAGAGAGAGAA
 AGAGAACTCTTCTTGGCTGAGCATGGAGCCTGATGTAGGGCTTGATCT
 CACAAGCCTGAGATCATGACCTGAGCCAAAATCAAGAGTCAGGCACC
 CAACCAACCAAGCCACCCAAGATGCCCTGAAACAGTAA

Para la genotipificación del polimorfismo se realizaron las reacciones de PCR en un volumen total de 5 µl en una mezcla de reacción que contenía los siguientes componentes:

- Agua destilada µ M

- Buffer Tris- Cl (pH = 8.3) 5 μ M
- dNTP's 5 μ M
- Primer Forward (5 μ M) 5 μ M
- Primer Reverse (5 μ M) 5 μ M
- β -mercaptoethanol 5 μ M
- MgCl₂ 5 μ M
- Taq polimerasa (Invitrogen Carlsbad CA) 1 μ M
- ADN molde (25-50 ng concentración final) 1 μ M

Las condiciones de reacción fueron: 1 ciclo de 5 minutos a 94 °C, 35 ciclos de 30 segundos a 95 °C, 45 segundos a 64 °C (temperatura de annealing), 45 segundos a 72°C y una elongación final de 10 minutos a 72°C. Los productos de amplificación se verificaron en geles de poliacrilamida 6% - 1X TBE, utilizando un marcador de peso molecular (PM) 100 - 1500 pb (INBIO HIGHWAY, Tandil, Argentina) durante 30 min a 150 voltios. Los geles se tiñeron con bromuro de etidio y se visualizaron en un transiluminador UV (Anexo 4).

b) PCR-alelo específico (PCR-AS)

La PCR alelo-específica (PCR-AS) es una variante de la PCR que permite detectar y amplificar secuencias de ADN específicas de alelos de interés. Esta técnica se utiliza para identificar polimorfismos genéticos, mutaciones puntuales o variaciones específicas en el ADN (Villegas Castagnasso y col., 2015). Esta técnica se basa en el uso de cebadores específicos para cada alelo, que difieren en una base para unirse selectivamente a la mutación específica. Se realizan dos

reacciones simultáneas con el mismo ADN molde, utilizando un cebador específico de alelo más un cebador común en un tubo, y el cebador complementario al otro alelo junto con el cebador común en otro tubo. Los productos de PCR se someten a electroforesis y se determina el genotipo del individuo en estudio de acuerdo a la presencia o ausencia de las bandas correspondientes a cada reacción. Si el alelo específico está presente, se observará una banda en el gel, mientras que si está ausente, no se detectará ninguna banda.

Diseño del método PCR-alelo específico para la detección del SNP del locus del CHR5

Teniendo en cuenta la bibliografía consultada se obtuvo la secuencia de ADN del locus del cromosoma 5 asociada al desarrollo de CMD de la base de datos GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>), a partir de la cual se diseñaron los primers que la flanquean (Tabla 2). En la fig. 4 se muestra un fragmento del locus en el CHR5 que incluye el SNP (CFA5:g.53,941,386T>C) reportado asociado al desarrollo de CMD y la localización de los primers Forward y Reverse y Alelos Específicos.

Tabla 2. Secuencia de los *primers* diseñados para la amplificación del fragmento analizado por PCR-alelo específico. El SNP se encuentra resaltado con amarillo.

SNP	Secuencia 5'-3'	Tamaño	Tm e (°C)	Región
CFA5:g.53,941,386 T>C	Forward 5'- TGCTCCATCTGCATAACCAT - 3'	259pb	56,4	Intergénica
	Reverse 5'- AGGGCCTGCTATTTTCAAG - 3'		55,2	
	Alelo 1 5'- ATCAAGGTCCCACAGCTAA - 3'	133pb	54,4	
	Alelo 2 5'- ACCAAGGTCCCACAGCTAA - 3'		54,9	

Tm e= temperatura de melting estimada.

Figura 4: Fragmento del Cromosoma 5 donde se encuentra el locus asociado al desarrollo de CMD (CFA5:g.53,941,386T>C). Se pueden identificar la variante reportada (sombreado amarillo), los primers Forward y Reverse (sombreado gris) y los primers Alelos Específicos (color rojo).

TCTGAGGTAGAACTCAAGCCTCTTGACTCTTGCCTAAGTCTCTTCTCTG
TGTTGGAACCCTGCTCCATCTGCATAACCATAGTCCCATTTAATTGTCA
CAGCCACCATCTGGGGACAGGAGTAAGGAAGAAAGTGGTACCACTGC
CAAATACCACAACGAGGGCTCCGAGGTGAAATGGAGTATAAA[C/T]CA
AGGTCCCACAGCTAAAGTCTGGGCCTTTCTCCACCACAGCCATGCTAT
GAGCTCATTCTTCCCCTGCCTCTCTGCCTATTGCTCCTTGCTTTTATGAG
CCTTAAAAGAGGCTTGAAAATAGCAGGCCCTAATGTTGCAAAGACCTT
CATCTTTATTATTCATCACACTTCTAGAACATTTGGTCAATCTGTCTTCC
CAGTGCAAAAATAAGACCATCCAGGTATGCAAA

Para la genotipificación del polimorfismo se realizaron dos reacciones de PCR en un volu en total de 5 µl cada una. Ade ás de agregar los primers

Forward y Reverse, se incorporó en una de las reacciones el primer alelo específico C (Mix C) y en la otra el primer alelo específico T (Mix T). Las reacciones se prepararon con los siguientes componentes:

- Agua destilada μ M
- Buffer Tris- Cl (p =) 5 μ M
- dNTP's μ M
- Primer Forward (5 μ M) μ M
- Primer Reverse (5 μ M) μ M
- Primer alelo específico C/T (5 μ M) μ M
- *nhancer* 5 μ M
- MgCl μ M
- Taq polimerasa (Invitrogen Carlsbad CA) μ M
- ADN molde (25-50 ng concentración final) μ M

Las condiciones de reacción fueron: 1 ciclo de 5 minutos a 94 °C, 35 ciclos de 30 segundos a 95 °C, 45 segundos a 62 °C (temperatura de annealing), 45 segundos a 72°C y una elongación final de 10 minutos a 72°C.

Los productos de amplificación se verificaron en geles de poliacrilamida 6% - 1X TBE, utilizando un marcador de peso molecular (PM) 100 - 1.500 pb (INBIO HIGHWAY) durante 30 min a 150 voltios. Los geles se tiñeron con bromuro de etidio y se visualizaron en un transiluminador UV (Anexo 4).

c) Genotipificación mediante secuenciación directa de fragmentos de amplificación (PCR-SBT)

La secuenciación directa de fragmentos de amplificación (PCR-SBT) es otra herramienta de genotipificación que utiliza como punto de partida los amplicones obtenidos de las PCR mencionadas previamente, y que se utilizó en el presente estudio con el fin de validar los resultados de estas metodologías de genotipificación indirecta. Los fragmentos resultantes se purifican para separarlos de los restos de los componentes remanentes de la PCR (nucleótidos, primers). A continuación se desnaturaliza el ADN para separar las cadenas y se realiza la secuenciación directa mediante la técnica de Sanger (*dideoxy chain termination method*; Sanger y col., 1977). Esta secuenciación se basa en la síntesis controlada de ADN utilizando terminadores dideoxi marcados. La secuencia de ADN de interés se usa como molde para PCR de terminación de cadena, que se diferencia de la estándar por la adición de nucleótidos (dNTP) modificados llamados dideoxirribonucleótidos (ddNTP). Los ddNTP, que carecen del grupo 3'-OH necesario para la formación de enlaces fosfodiéster, son incorporados por la ADN polimerasa aleatoriamente y detienen la extensión. En la secuenciación automática en secuenciadores capilares, los cuatro tipos de ddNTP, cada uno con una marca fluorescente diferente, se mezclan en una reacción y se detecta qué nucleótido termina la polimerización. Los fragmentos se separan por tamaño mediante electroforesis en gel capilar, y se utiliza un láser que excita las marcas fluorescentes en cada fragmento que pasa por la ventana del capilar y una computadora registra la luz emitida resultante. La secuencia original del ADN se reconstruye a partir de los colores de las marcas registrados en el detector, y los datos se presentan en forma de cromatograma, que muestra los picos de intensidad de fluorescencia de cada nucleótido a lo largo de la secuencia.

Diseño del método PCR-SBT para la validación de las variantes CFA14:g.20,829,667_20,829,682del y CFA5:g.53,941,386T>C

En el presente trabajo de tesis, la variante CFA14:g.20,829,667_20,829,682del del gen PDK4 y el SNP CFA5:g.53,941,386T>C se validaron mediante la técnica de PCR-SBT. Los productos de PCR se purificaron mediante la técnica de precipitación diferencial con polietilenglicol 8.000 (Anexo 5) para eliminar los remanentes de primers y nucleótidos no incorporados durante la reacción. Posteriormente, los productos purificados se corrieron con un secuenciador capilar Genetic Analyzer ABI 3500 (Applied Biosystem, USA). Este secuenciador utiliza la tecnología de electroforesis capilar que proporcionan una alternativa clara al sistema basado en geles de acrilamida/bisacrilamida donde este soporte ha sido sustituido por un polímero que se inyecta de forma automática en un capilar antes de cargar el producto purificado de la reacción de secuenciación; las muestras se van analizando una a una. Una vez secuenciadas las muestras de ADN, se determinó su identidad mediante el alineamiento con secuencias de referencia.

Las secuencias de ADN obtenidas con el secuenciador automático Genetic Analyzer ABI 3500 (Applied Biosystem) se editaron para corregir los errores de lectura automática, y se alinearon utilizando el programa de alineación múltiple CLUSTAL OMEGA (Sievers y Higgins, 2014), y el programa BLAST de búsqueda de similitudes (Altschul y col., 1990). Para el análisis individual de las secuencias en los sitios específicos de los polimorfismos causales de las

enfermedades se utilizó el software Chromas Lite v2.1.1 (Technelysium, http://technelysium.com.au/?page_id=13).

d) Pirosecuenciación

La pirosecuenciación (Ronaghi y col., 1998) es una técnica de secuenciación muy útil para genotipificación, que permite determinar la secuencia de ADN de un fragmento amplificado que incluye el polimorfismo de interés mediante luminiscencia. Esta técnica se basa en la detección del pirofosfato (PPi), liberado durante la incorporación de un nucleótido al ADN molde, que forma parte de una cascada de reacciones enzimáticas cuyo producto final es la generación de luz visible. Mediante la actividad de una polimerasa se van incorporando nucleótidos de a uno, a partir de un cebador unido a un ADN molde. Cada vez que se incorpora un nucleótido complementario a la cadena molde se libera un PPi que se transforma en ATP por acción de una ATP sulfurilasa. El ATP se acoplará a la luciferina, formando un complejo el cual será utilizado por la enzima luciferasa para producir una señal luminosa de intensidad proporcional al número de nucleótidos incorporados, que es captada por una lente ubicada en el secuenciador, reproduciéndose como un pico en un pirograma. De esta forma se puede deducir la secuencia del ADN molde partiendo del tamaño de los picos obtenidos (Peral García y col., 2015). La gran ventaja de este método es que se puede detectar cualquier nuevo polimorfismo y analizar un gran número de individuos simultáneamente. La desventaja es que se requiere de un equipamiento automatizado específico.

Diseño del método de pirosecuenciación para la detección de la variante CFA36: g.22,321,955 C>T del gen TTN

Esta técnica se utilizó para genotipificar la variante del gen TTN (ENSCAFT00000022319; CFA36: g.22,321,955 C>T; Meurs y col., 2019). Para ello se empleó la secuencia de ADN depositada en la base de datos GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) correspondiente al gen TTN, teniendo en cuenta la bibliografía consultada. Con esta secuencia de ADN (Gene ID: 478820) se diseñaron los primers que flanqueaban la mutación causal utilizando el programa PyroMark Assay Design SX 2.0 (www.qiagen.com; Tabla 3). El algoritmo del software seleccionó como mejor opción los primers que se unen a la cadena complementaria. Uno de los primers (reverse) fue modificado agregándole en el extremo 5' la secuencia complementaria para el primer universal de M13. A su vez se agregó a la reacción un primer M13 marcado con la biotina para ser utilizado para la posterior purificación de una de las cadenas de ADN del producto de amplificación. Por último se diseñó un tercer primer (interno) a partir del cual se inicia el procedimiento de pirosecuenciación. En la fig. 5 se muestra un fragmento de la secuencia del gen estudiado (TTN), que incluye la mutación reportada asociada al desarrollo de CMD y la localización de los primers.

Tabla 3. Secuencia de los *primers* diseñados para la amplificación del fragmento del Titina (TTN) que incluye la mutación candidata (ENSCAFT00000022319; CFA36: g.22,321,955 C>T) analizada por pirosecuenciación.

Gen	Secuencia 5'-3'	Tamaño	Tm e(°C)	Región
TTN CFA36: g.22,321,955 C>T	Forward 5'-GTGGTTTAAGGATGGAAAAGAGC-3'	212pb	54,1	EXON 90 del gen TTN
	Reverse 3'-M13- AGATGGAACTTCCCATACACGT-5'	212pb	84,72	
	Interno 5'-AACGTGGCACCCAGTGACAG-3'	130pb	72,35	
M13	AGCGCATAACAATTCACACAGG			

Tm e = temperatura de melting estimada.

Figura 5. Fragmento del gen Titina (TTN) ubicado en el cromosoma 36 (cadena complementaria). Se pueden identificar la variante reportada CFA36: g.22,321,955 C>T (sombreado amarillo), los primers forward y reverse (sombreado gris) y el primer interno para el Pirosecuenciador (color rojo).

GAATGTAATTTATTATATCCCGTTCCAGAACCTGCGGCAATTGTGGAA
AAGCCAGAATCCATAAGAGTTACCACAGGAGATACCTGTACCTTGGAA
TGCACAGTCACTGGCACACCTGAACTCAGTACTAAGTGGTTTAAGGAT
GGAAAAGAGCTGACGAGTGACAGCAAATACAAAATAAGTTTCTTCAA
CAAAGTATCTGGCCTTAAGATTATC**AACGTGGCACCCAGTGACAGTG/**
AGGGTATACAGTTTCGAGGTGCAAACCCTGTTGGCAAAGACAGCTGT
ACAGCCTCAGTACAAGTTTCAGGTTGGTTGGTTGGACTTTTATCTTAAT
GTTATTGTGATGTAGTTCTGATTTACATATACGTGTATGGGAAGTTCC
ATCTTCTACAGTGTGACATGAGTCTGACACCCTCTGCATCAGCAAAT
CAGTGGTAACGTACCCATAGGGTTGGAAGAAAATGAAGTCCAAAAGT
TATATCCTGGGGTGCCTGGTGGCTCAGTTTGTGAGCATCTACCTTTGG
CTCAGGTCATGATCCTGGGGTCCACGGATCGAGCCCCACGTTGGGTTT

Para la genotipificación de esta variante se realizó la reacción de PCR en un volumen total de 5 μ l con los siguientes componentes:

- Agua destilada μ M
- Buffer Tris- Cl (pH =) 5 μ M
- dNTP's μ M
- Primer Forward (5 μ M) 5 μ M
- Primer Reverse (5 μ M) 5 μ M
- γ -nha ncer 5 μ M
- MgCl₂ 5 μ M
- Taq polimerasa (Invitrogen Carlsbad CA) μ M
- ADN molde (25-50 ng concentración final) μ M

En cuanto a las condiciones de ciclado (Anexo 6), la elongación se dividió en dos etapas: la primera de 15 ciclos con una hibridación específica para cada fragmento, utilizando los primers forward y reverse-M13 (30 seg a 94°C, 45 seg a 58°C (temperatura de annealing) y 45 seg a 72°C), y en la segunda etapa se incorporó el primer biotinilado M13-Biot con 30 ciclos adicionales y una temperatura de hibridación de 52°C (30 seg a 94°C, 45 seg a 52°C y 45 seg a 72°C). Los protocolos para purificar los amplificadores biotinilados fueron los recomendados por el proveedor (Anexo 6). Las muestras se tipificaron en un pirosecuenciador PSQTM96 System Instrument (QIAGEN) y los resultados se analizaron utilizando el software Pyromark ID 1.0 (QIAGEN), que utiliza el propio Pirosecuenciador.

e) Técnicas de secuenciación masiva o de nueva generación (NGS), variante target NGS

La secuenciación masiva de próxima generación (NGS, del inglés *Next Generation Sequencing*) es una técnica que permite determinar el orden de los nucleótidos en genomas completos, la transcriptómica de una muestra o de regiones específicas de ADN. Se basa en la ejecución de una gran cantidad de diminutas reacciones de secuenciación de Sanger al mismo tiempo (en paralelo). Gracias a esta paralelización y a la pequeña escala, se pueden secuenciar grandes cantidades de ADN de forma mucho más rápida y económica. La secuenciación masiva de nueva generación dirigida (targeted NGS) tiene como objetivo lograr un "enriquecimiento dirigido" de las regiones de interés del genoma para reducir significativamente la secuenciación de *loci* genómicos (cobertura de secuenciación), aumentar el número de lecturas por posición genómica (profundidad de secuenciación) y reducir costos y esfuerzos (Mertes y col., 2011). Este método secuencia simultáneamente millones de fragmentos de ADN, de manera que permite analizar desde el genoma completo hasta varios genes o regiones de genes con una sola prueba (Yohe y Thyagarajan, 2017). De esta manera se genera un conjunto de datos más pequeño y manejable, reduciendo los costos de secuenciación y el tiempo de respuesta. A su vez es un método especialmente sensible y poderoso para identificar variantes y mutaciones, incluidas variantes raras, siendo menos susceptibles a las pérdidas de alelos y a la falta de datos.

En el presente trabajo de tesis doctoral se utilizó la tecnología NGS AgriSeq™ Targeted GBS (Thermo Fisher Scientific), en la que el enriquecimiento se basó en la utilización de un pool de primers diseñados para amplificar específicamente las regiones de interés, los que son posteriormente secuenciados. Las muestras se pueden multiplexar o agrupar, lo que implica agregar un código de barras (índice) a las muestras durante la construcción de las librerías de ADN para que puedan identificarse después de la secuenciación. Las muestras utilizadas para la secuenciación del amplicón deben transformarse en bibliotecas y enriquecerse individualmente mediante amplificación por PCR antes de que puedan multiplexarse para la secuenciación. Después de la amplificación por PCR de las regiones genómicas objetivo, los cebadores restantes se digieren y los adaptadores con código de barras se ligan a cada muestra, codificando cada muestra con una identificación única, para que puedan secuenciarse al mismo tiempo. Luego, la biblioteca se purifica utilizando perlas magnéticas y está lista para la secuenciación simultánea de cientos de muestras en una única corrida de un equipo de NGS de Ion Torrent o similar (Thermo Fisher Scientific).

Esta técnica se utilizó para la genotipificación de la delección de 16 pb en el sitio de empalme 5' del intrón 10 del gen PDK4 ubicado en el cromosoma 14 (CFA14: g.20,829,667_20,829,682del) en 45 Dobermans.

Análisis de las secuencias de ADN obtenidas por NGS target

Se genotipificaron 45 muestras de ADN de DP utilizando la versión beta del panel Canine Targeted GBS (Thermo Fisher Scientific). El panel incluía 383

marcadores: 228 SNV de parentesco, 137 marcadores de condiciones genéticas y 18 marcadores de rasgos de pelaje. Dentro de los marcadores de condiciones genéticas se encontraba la variante CFA14: g.20,829,667_20,829,682 del del gen PDK4 asociada al desarrollo de CMD (Meurs y col., 2012). Los ADN obtenidos se utilizaron para construir las bibliotecas NGS, que se secuenciaron utilizando la plataforma Ion Gene Studio™ S5 (Thermo Fisher Scientific) en el laboratorio de desarrollo de Thermo Fisher (Austin, TX, EE. UU.). El análisis de los archivos FASTq se realizó siguiendo el flujo de trabajo recomendado por Genome Analysis Toolkit (GATK, <https://gatk.broadinstitute.org/>), siendo analizados por calidad utilizando el programa Galaxy (Li y col., 2009). Posteriormente, se alinearon con la secuencia del genoma canino de referencia CanFam 3.1. y se crearon los archivos BAM / SAM. Se calculó el *call rate* de cada muestra, definido como el porcentaje de marcadores que se pueden genotipificar para una muestra específica, y el *call rate* de cada marcador, que indica el porcentaje de muestras que se pueden genotipificar para un marcador específico. Las variantes genéticas se visualizaron a través del software IGV (Robinson y col., 2011).

7. Análisis poblacional

Para determinar la frecuencia genotípica y alélica para cada una de las variantes asociadas al desarrollo de CMD en la población de Argentina se agregaron a los DP citados previamente 87 muestras de ADN de animales provenientes del biobanco del IGEVET. De estos animales no se obtuvieron datos fenotípicos en cuanto al desarrollo o no de la enfermedad en cuestión. Las

frecuencias genotípicas y alélicas para cada mutación, dentro de las poblaciones analizadas, se calcularon con el software GENEPOP 4.7 (Rousset y col., 1997; Rousset, 2007). La heterocigosidad observada (h_o) y heterocigosidad esperada (h_E) para cada locus se estimaron de acuerdo a Nei (1978) usando el software de análisis genéticos poblacionales ARLEQUIN 3.5 (Schneider y col., 2000). Las posibles desviaciones del equilibrio de Hardy-Weinberg (HWE) se estimaron para cada locus y subpoblación mediante el test exacto de Fisher incluido en el programa GENEPOP 4.7.

8. Análisis estadístico

Se verificó la normalidad de los datos en cada grupo mediante el test de Kolmogorov-Smirnov. Se compararon los valores de FC, EEM, Pd, Pa, PR, QRSd, QRSa, QT, ST, QTc, DDVI, DSVI, TIVD, TIVS, PLD, PLS, FA, AI y PSSE entre ambos grupos mediante un test de Student. La presencia de signos clínicos, el ritmo cardíaco y la prevalencia de la enfermedad según el sexo se analizaron mediante un test de Chi cuadrado. Todos los datos fueron presentados como $\text{media} \pm \text{EEM}$ (error estándar de la media). Todas estas estimaciones se realizaron utilizando el programa SPSS 23.0 (SPSS, Chicago, IL, USA).

Para llevar a cabo el análisis genético de asociación basado en los datos de genotipificación de caso-control, el test de Chi-cuadrado, Test de Fisher y Odds Ratio (riesgo relativo estimado) se calcularon utilizando el paquete de R (<http://www.r-project.org/>), el paquete Plink (Purcell, 2009) y el software SPSS 23.0. Asimismo, considerando únicamente los DP enfermos, se analizó la

asociación entre las variantes reportadas con el desarrollo de arritmias, mediante un test de Chi-cuadrado. Por último se evaluó si el ser portador de más de una mutación aumenta la posibilidad de desarrollar CMD mediante regresión logística binaria, utilizando el software SPSS 23.0. Se consideró como significativo un valor de $p < 0,05$.

Resultados

Resultados en la población de doberman pinscher

En base a los hallazgos de la evaluación clínica, electrocardiográfica y ecocardiográfica, los animales se distribuyeron de la siguiente manera en cada uno de los grupos: En el grupo control (sanos) se incluyeron 45 DP, 13 machos y 32 hembras, de 7 a 14 años de edad ($8,5 \text{ años} \pm 1,8 \text{ años}$). Ninguno de estos animales presentó alteraciones clínicas, electrocardiográficas ni ecocardiográficas. En el grupo caso (enfermos) se incluyeron 27 DP, 16 machos y 11 hembras, de 1 a 13 años de edad ($6,3 \pm 3,8 \text{ años}$). En este grupo, se encontró una relación estadísticamente significativa entre el sexo y el desarrollo de la enfermedad ($p < 0,05$) donde los machos fueron más propensos a padecer la enfermedad. Con respecto a los hallazgos clínicos, de estos animales, el 55% presentó uno o más signos de alteración cardiovascular al momento de la evaluación (tabla 4). Asimismo, el 74% de los DP enfermos evidenció algún tipo de alteración electrocardiográfica (tabla 5). Sin embargo, sólo el 41% de ellos presentaron

arritmias patológicas de relevancia clínica: complejos prematuros ventriculares (CPVs), fibrilación atrial, o una combinación entre ambas.

Todos los valores electrocardiográficos (FC, EEM, Pd, Pa, PR, QRSd, QRSA, QT, ST, QTc) y ecocardiográficos (DDVI, DSVI, TIVD, TIVS, PLD, PLS, FA, AI, PSSE y Ao) mostraron distribución normal. Se observaron valores mayores de FC ($p < 0,01$), DDVI ($p < 0,01$), DSVI ($p < 0,01$), AI ($p < 0,01$) y PSSE ($p < 0,01$) en el grupo caso. En cambio, se observaron valores menores de FA ($p < 0,01$), TIVD ($p < 0,05$), TIVS ($p < 0,01$), PLD ($p < 0,01$) y PLS ($p < 0,01$) en el grupo caso versus el grupo control. No se observaron diferencias para las variables Ao, ritmo cardíaco, duración onda P, amplitud onda P, intervalo PR, duración QRS, amplitud QRS, duración QT, duración QTc, segmento ST y EEM ($p > 0,05$). Esto demostró una apropiada diferenciación entre ambos grupos, lo que permitió realizar el consiguiente análisis de asociación genética.

El análisis clínico y cardiovascular desarrollados en el presente estudio, evidenció una prevalencia estimada de la enfermedad en la población estudiada de DP de 37,5%.

Tabla 4: Frecuencia de signos clínicos cardiológicos en los DP del grupo caso.

Signos clínicos cardiológicos	Nº de casos
Mucosas color rosa pálido	7
Intolerancia al ejercicio	7
Pulso femoral débil	6
Ascitis	5
Presencia de soplo	5
Disnea	2
Taquipnea	2
Edema periférico	2
Pulso asincrónico	2
Anorexia	2
Decaimiento	2
Tos	1

Tabla 5: Frecuencia de alteraciones electrocardiográficas en los DP del grupo caso.

Alteraciones electrocardiográficas	Nº de casos
Complejos prematuros ventriculares	11
Trastornos en la repolarización ventricular	4
Fibrilación atrial	3
Taquicardia sinusal	2
Sobrecarga atrial izquierda	2
Bloqueo de rama derecha	1
Microvoltaje	1

Métodos de genotipificación de potenciales mutaciones causales de CMD

Técnicas de genotipificación indirecta

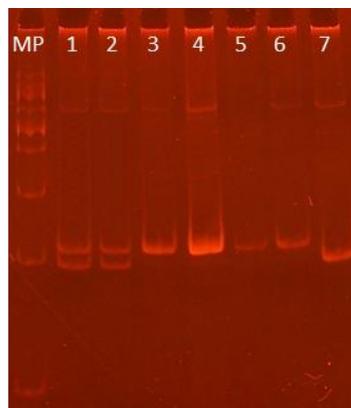
El primer paso para el estudio de las variantes consistió en el diseño de cebadores para la amplificación de las regiones que incluyen las mutaciones causales con el software PyroMark Assay Design SX 2.0 (www.qiagen.com). A continuación se realizaron las PCR mencionadas previamente para cada variante. Esto permitió obtener fragmentos amplificados del tamaño esperado tanto para PDK4, para el locus del CHR5, y para TTN (210 pb y 194 pb; 259 pb y 133 pb; y 212 pb y 130 pb, respectivamente).

Con la técnica de PCR- análisis de fragmentos se confirmó la presencia de la variante **CFA14: g.20,829,667_20,829,682del** del gen PDK4 en la población de DP local. En el gel de poliacrilamida se pueden observar bandas de diferente peso molecular, pudiendo identificar animales con los tres posibles genotipos para este polimorfismo (Fig. 6):

- En animales homocigotas para el alelo salvaje se puede observar una única banda de 210 pb de longitud correspondiente al fragmento completo.
- En animales homocigotas mutados, es decir con dos copias del alelo mutado, se puede observar también una única banda, pero en este caso de 194 pb de longitud, debido a presencia de una delección de 16 pb en este alelo

- En animales heterocigotas se pueden observar 2 bandas, una de 210 pb de longitud correspondiente al alelo salvaje y una de 194 pb correspondiente al alelo mutado.

Figura 6: Gel de poliacrilamida 6% - 1 X TBE. Fragmento del gen PDK4 amplificado por PCR.



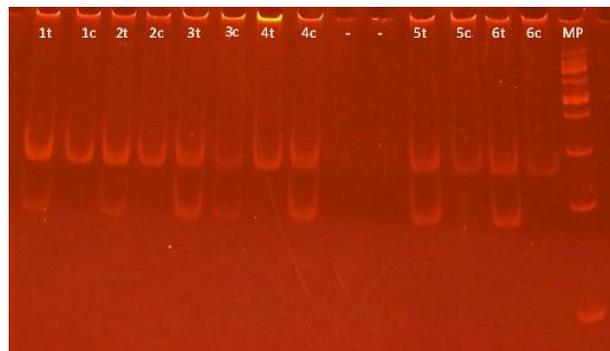
MP: Marcador de peso molecular (Marcador de ADN 100 - 1500 pb – INBIO HIGHWAY), 1 y 2: animales heterocigotas (W/Del), 3 a 6: animales homocigotas para el alelo salvaje (W/W), 7: animal homocigota para el alelo mutado (Del/Del).

Por otro lado, con la técnica de PCR-AS se confirmó la presencia del SNP **CFA5: g.53,941,386T>C** en la población de DP local, identificando animales con los tres posibles genotipos para la variante (Fig. 7). Como se mencionó previamente, en este tipo de PCR se realizaron 2 reacciones de amplificación, una utilizando el cebador Reverse T para el alelo salvaje (T) y otra utilizando el cebador Reverse C para el alelo mutado (C). Asimismo, ambas reacciones incluyeron los cebadores Forward y Reverse del fragmento completo de 259 pb de longitud. Los productos de cada PCR se sembraron en calles paralelas, en una

calle el resultado para la reacción T y en la otra el resultado para la reacción C (dos calles por animal), observando los siguientes resultados:

- En animales homocigotas para el alelo salvaje (dos copias del alelo T) se puede observar en la calle correspondiente a la reacción T la presencia de 2 bandas, una de 259 pb de longitud (fragmento completo) y otra de 133 pb de longitud (fragmento alelo específico T). Sin embargo, en la calle de la reacción C sólo se verá la banda correspondiente al fragmento completo.
- En los animales homocigotas para el alelo mutado (dos copias del alelo C) se puede observar en la calle correspondiente a la reacción C la presencia de 2 bandas, una de 259 pb de longitud (fragmento completo) y otra de 133 pb de longitud (fragmento alelo específico C). Sin embargo, en la calle de la reacción T sólo se verá la banda correspondiente al fragmento completo.
- En animales heterocigotas (una copia del alelo T y una copia del alelo C) se pueden observar en las dos calles las 2 bandas, una de 259 pb de longitud correspondiente al fragmento completo y una de 133 pb correspondiente al fragmento alelo específico T y C, respectivamente.

Figura 7: Gel de poliacrilamida 6% - 1 X TBE. Fragmento del locus del CHR5 amplificado por PCR.



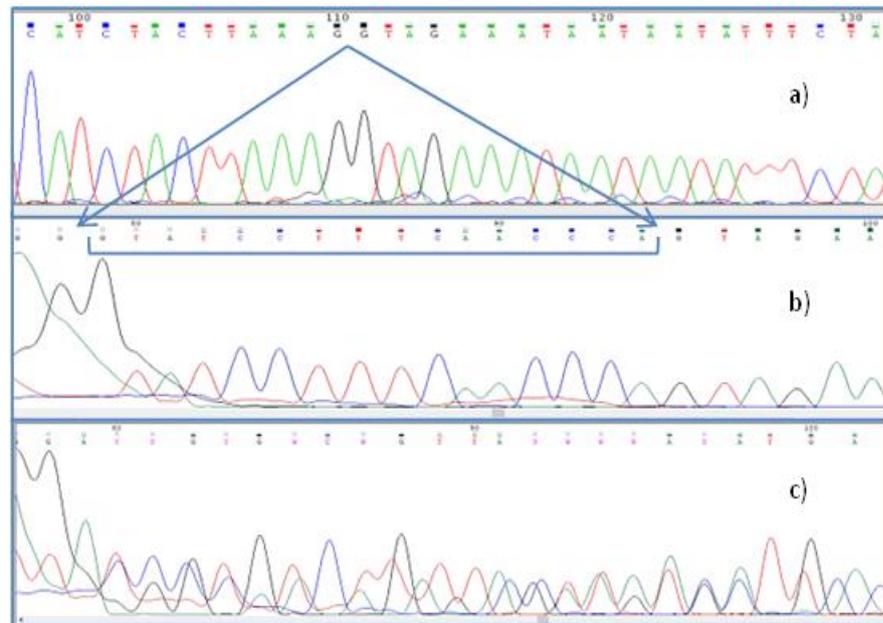
1, 2, 5 y 6: animales homocigotas para el alelo salvaje (T/T), 3: animal heterocigota (T/C), 4: animal homocigota para el alelo mutado (C/C). t: reacción de PCR con alelo específico T; c: reacción de PCR con alelo específico C; -: negativo; MP: Marcador de peso molecular (Marcador de ADN 100 - 1500 pb – INBIO HIGHWAY).

Técnicas de genotipificación directa

Posteriormente, se procedió a realizar los protocolos de purificación de los amplificados para la secuenciación directa tipo Sanger y de los amplificados biotinilados para la pirosecuenciación.

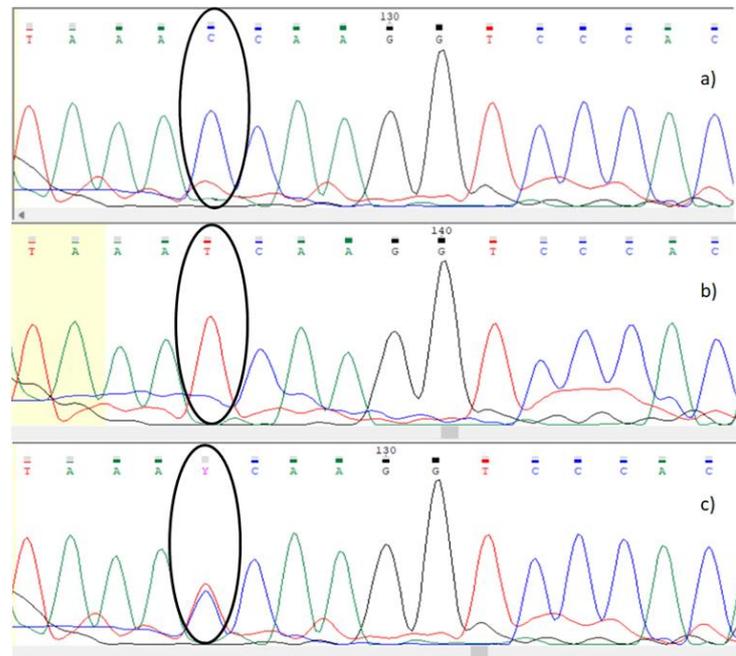
Los productos de amplificación del gen PDK4 y del locus del CHR5 de un subset de animales se secuenciaron en un secuenciador de ADN de tecnología capilar Genetic Analyzer ABI 3500 (Applied Biosystem) obteniéndose los electroferogramas correspondientes (Figs. 8 y 9). De esta manera se confirmaron los resultados obtenidos por las técnicas de genotipificación indirectas en las que se observó la presencia de ambas variantes, con sus tres genotipos posibles (Tabla 6).

Figura 8: Electroferogramas obtenidos por secuenciación del fragmento del gen PDK4 que incluye la delección de 16 pb.



a) genotipo homocigota para el alelo mutado Del/Del; b) genotipo homocigota para el alelo normal (W/W); c) genotipo heterocigota (W/Del).

Figura 9: Electroferogramas obtenidos por secuenciación del fragmento de ADN que incluye el polimorfismo del locus en el CHR5.



a) genotipo homocigota para el alelo mutado C/C; b) genotipo homocigota para el alelo normal (T/T); c) genotipo heterocigota (T/C).

Tabla 6: Genotipos hallados en la población de DP para la deleción en el gen PDK4 y el SNP del CHR5 mediante secuenciación directa de los productos de amplificación.

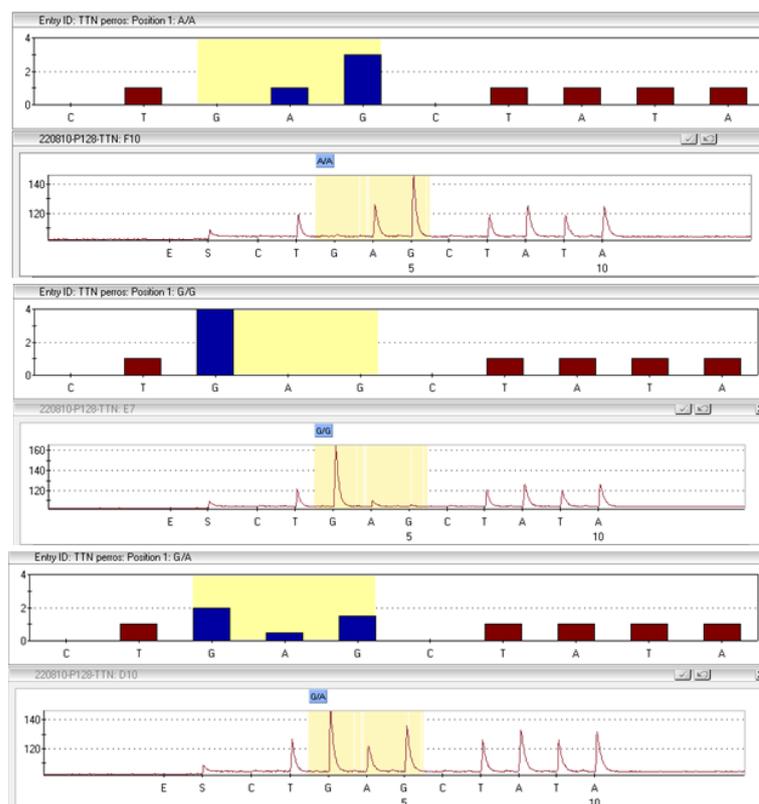
Genotipo	CFA14:g.20,829,667_20,829,682del	CFA5:g.53,941,386T>C
Homocigota para el alelo normal	GTATCCTTTCAACCCA/GTATCCTTTCAA CCCA	T/T
Heterocigota	GTATCCTTTCAACCCA/-	C/T
Homocigota para el alelo mutado	-/-	C/C

Por otro lado, se secuenció mediante la técnica de Pirosecuenciación, la variante **CFA36:g.22,321,955 C>T** del gen TTN. Esto permitió confirmar la presencia de la mutación en la población local, obteniendo los pirogramas correspondientes a los tres genotipos esperados (Tabla 7, Fig. 10).

Tabla 7: Genotipos hallados en la población de DP para la variante CFA36: g.22,321,955 C>T del gen TTN mediante pirosecuenciación.

Genotipo	TTN CFA36:g.22,321,955 C>T
Homocigota para el alelo normal	C/C
Heterocigota	T/C
Homocigota para el alelo mutado	T/T

Figura 10: Pirogramas obtenidos por pirosecuenciación del fragmento del gen TTN que incluye el polimorfismo CFA36:g.22,321,955 C>T (cadena complementaria).



a) genotipo homocigota para el alelo mutado T (A); b) genotipo homocigota para el alelo salvaje C (G); c) genotipo heterocigota C/T (G/A).

Finalmente, se utilizó la técnica de NGS target para validar la mutación CFA14: g.20,829,667_20,829,682del del gen PDK4 en un subset de 45 animales que ya habían sido tipificados por la técnica de PCR-análisis de fragmento en geles de poliacrilamida. Este proceso sirvió como un nuevo método de comprobación de la eficacia de la técnica desarrollada. Se utilizó la versión beta del panel Canine Targeted GBS (Thermo Fisher Scientific) que incluía, dentro de los marcadores de condiciones genéticas, la variante del gen PDK4 asociada al

desarrollo de CMD (Meurs y col., 2012). Este marcador tuvo un desempeño bajo (call rate < 56,8%), mostrando resultados en solo 28 de los 45 DP analizados. Los resultados obtenidos por esta técnica mostraron una total coincidencia entre ambas técnicas (Tabla 8).

Tabla 8: Genotipos hallados en el subset de muestras de DP para la variante CFA14: g.20,829,667_20,829,682del en el gen PDK4 realizada mediante NGS target en Thermo Fisher Scientific.

Genotipo PDK4			Genotipo PDK4		
Muestra	PCR*	NGS target	Muestra	PCR*	NGS target
CMD 3	Ins/Ins	Ins/Ins	CMD 26	Ins/Del	Ins/Del
CMD 5	Ins/Ins	Ins/Ins	CMD 27	Ins/Ins	Ins/Ins
CMD 6	Ins/Ins	Ins/Ins	CMD 29	Ins/Del	Ins/Del
CMD 8	Ins/Del	Ins/Del	CMD 32	Ins/Del	Ins/Del
CMD 11	Ins/Del	Ins/Del	CMD 33	Ins/Del	Ins/Del
CMD 12	Ins/Del	Ins/Del	CMD 34	Ins/Del	Ins/Del
CMD 15	Ins/Ins	Ins/Ins	CMD 35	Ins/Ins	Ins/Ins
CMD 16	Ins/Del	Ins/Del	CMD 36	Ins/Del	Ins/Del
CMD 17	Ins/Ins	Ins/Ins	CMD 38	Ins/Del	Ins/Del
CMD 18	Ins/Del	Ins/Del	CMD 40	Ins/Del	Ins/Del
CMD 21	Ins/Del	Ins/Del	CMD 41	Ins/Ins	Ins/Ins
CMD 22	Ins/Del	Ins/Del	CMD 42	Ins/Ins	Ins/Ins
CMD 24	Ins/Ins	Ins/Ins	CMD 44	Ins/Del	Ins/Del
CMD 25	Ins/Ins	Ins/Ins	CMD 46	Ins/Del	Ins/Del

*PCR-análisis de fragmento

Análisis poblacional

Para la variante CFA14: g.20,829,667_20,829,682 del del gen PDK4 se analizaron en total 158 DP, encontrando una frecuencia alélica del alelo mutado de 0,17. Tres animales resultaron homocigotas para el alelo mutado (Del/Del), mientras que 106 animales presentaron el genotipo homocigota salvaje (Ins/Ins) y 49 fueron heterocigotas (Ins/Del). La estimación del EHW en la población total no evidenció un desvío significativo ($p = 0,42$) con respecto a las proporciones teóricas esperadas para el equilibrio. El valor negativo del coeficiente de consanguinidad (F_{IS}) (-0,08) sugiere un exceso de heterocigotos en la población en comparación con lo esperado bajo equilibrio genético. Sin embargo, como el valor de p es mayor a 0,05, no hay suficiente evidencia estadística para considerar que esta desviación del equilibrio genético es significativa (Tabla 9).

Con respecto al SNP localizado en el CHR5 en una muestra de 134 DP, se encontró una frecuencia alélica del alelo mutado de 0,12. Ciento siete DP resultaron homocigotas para el alelo salvaje, 12 fueron heterocigotas y los 6 restantes presentaron un genotipo homocigota para el alelo mutado. La estimación del EHW en la población total evidenció un desvío significativo ($p = 0,05$) con respecto a las proporciones teóricas esperadas para el equilibrio. El valor de F_{IS} positivo (0,28) sugiere un déficit de heterocigotos en la población en comparación con lo esperado bajo el EHW (tabla 9).

Por último, para la variante CFA36: g.22,321,955 C>T del gen TTN se analizaron en total 120 DP, encontrando una frecuencia alélica del alelo mutado de 0,58. Catorce perros fueron homocigotas para el alelo salvaje, 72 resultaron

heterocigotas y 34 homocigotas para el alelo mutado 34. La estimación del EHW en la población total, evidenció un desvío significativo ($p = 0,01$) con respecto a las proporciones teóricas esperadas para el equilibrio. El valor negativo de $F_{IS} - 0,23$ para este polimorfismo, con un valor de p menor a $0,05$, indica un aumento significativo de animales heterocigotas en la muestra analizada de DP (Tabla 9).

Tabla 9: Frecuencias genotípicas y alélicas de las variantes asociadas al desarrollo de CMD, obtenidas en la población de DP analizada.

Variante	n°	Frecuencias genotípicas			Frecuencias alélicas		HWE	Valor de F_{IS}	h_O	h_E
		W/W	W/m	m/m	p	Q				
CFA14^a	158	106	49	3	0,83	0,17	0,42	-0,08	0,31	0,29
CFA5^b	134	107	21	6	0,88	0,12	0,05	0,28	0,16	0,22
CFA36^c	120	14	72	34	0,42	0,58	0,01	-0,23	0,6	0,49

“W” alelo salvaje “ ” alelo mutado deno inados p q respectiva ente. n° = número de animales analizados; HWE = Equilibrio de Hardy Weinberg; FIS = coeficiente de endogamia o consanguinidad; h_O : heterocigosidad observada; h_E : heterocigosidad esperada; a: CFA14:g.20,829,667_20,829,682del; b: CFA5:g.53,941,386T>C; c: CFA36:g.22,321,955 C>T

Análisis estadístico

Una vez identificada la presencia de los tres polimorfismos en la población de DP locales se prosiguió a determinar si existe asociación entre la presencia de cada una de las variantes y el desarrollo de CMD. Para este análisis se utilizaron

los grupos experimentales caso ($n = 27$) y control ($n = 45$). Además del test de Chi-cuadrado, Test de Fisher y Odds Ratio (riesgo relativo estimado) calculados con el paquete de R y el software SPSS 23.0 (SPSS Inc. Chicago, IL, USA), con el paquete de PLINK se realizaron 3 modelos de asociación basados en la prueba exacta de Fisher, que incluyen el efecto genético dominante (dominant gene action), efecto genético recesivo (recessive gene action) y método de Cochran-Armitage (Cochran-Armitage trend; Tablas 10 y 11).

Para la delección de 16 pb en el gen PDK4 se genotipificaron en total 72 animales, 27 correspondientes al grupo caso y 45 al grupo control, encontrando los dos alelos (salvaje/mutado) en ambos grupos experimentales. La frecuencia del alelo mutado fue mayor en el grupo caso (0,26) que en el control (0,19), sin embargo, no se encontró una asociación estadísticamente significativa entre la variante y el desarrollo de la enfermedad ($p > 0,05$). Asimismo la presencia de la variante no se considera un factor de riesgo para desarrollar la enfermedad (OR = 1,50). Considerando que se reportó que CMD tiene un modo de transmisión autosómica dominante se agruparon los animales en mutados ($n = 29$), que incluyen homocigotas para el alelo salvaje y heterocigotas, y no mutados, homocigotas para el alelo salvaje ($n = 43$) para evaluar el efecto genético dominante, obteniendo resultados similares, que descartan una asociación entre la variante y la enfermedad ($p > 0,05$; OR = 1,68). Por último se evaluó el efecto genético recesivo, sin encontrar tampoco asociación ($p > 0,05$; OR = 1,69).

Para el SNP C/T localizado en el locus del CHR5 se genotipificaron en total 72 animales, 27 correspondientes al grupo caso y 45 al grupo control, y se detectaron los dos alelos (salvaje/mutado) en ambos grupos experimentales. Como

en el caso anterior, la frecuencia del alelo mutado fue mayor en el grupo caso (0,24) que en el control (0,14), sin embargo, no se encontró una asociación estadísticamente significativa entre la variante y el desarrollo de la enfermedad ($p > 0,05$). Asimismo, la presencia de la variante no se considera un factor de riesgo (OR = 1,87). Al evaluar la acción genética dominante (mutados $n = 20$; no mutados $n = 52$), se obtuvieron resultados similares, que descartan una asociación entre la variante y CMD, pero se evidencia un efecto protector al alelo salvaje (T) ($p > 0,05$; OR = 2,06). Por último, se evaluó la acción genética recesiva, sin encontrar asociación ($p > 0,05$; OR = 1,75).

Con respecto a la variante C > T en el gen TTN se genotipificaron en total 71 animales, 26 correspondientes al grupo caso y 45 al grupo control, y se detectaron los dos alelos (salvaje/mutado) en ambos grupos experimentales. La frecuencia del alelo mutado fue mayor en el grupo caso (0,71) que en el control (0,51), y se evidenció una asociación estadísticamente significativa entre la variante y el desarrollo de la enfermedad ($p < 0,05$). Asimismo, la presencia de la variante mutada se considera un factor de riesgo (OR = 2,36). Sin embargo, al considerar el modo de transmisión como autosómico dominante (mutados $n = 64$; no mutados $n = 7$), se obtuvieron resultados contradictorios, ya que estadísticamente se descarta una asociación entre la variante y la enfermedad ($p > 0,05$; OR = 1,50). Por último, se evaluó la acción genética recesiva (mutados $n = 19$; no mutados $n = 52$), encontrándose en este caso asociación significativa, confirmando que la presencia del alelo mutado (A) es un factor de riesgo ($p < 0,01$; OR = 6,50).

Por otro lado, al considerar únicamente los perros con CMD se identificó que el 30% tenía únicamente la variante TTN, un 22% tenía las 3 variantes, otro 22% registró las variantes TTN y PDK4, un 15% las variantes TTN y el SNP del CHR5, un 7% no poseía ninguna de las 3 variantes y un 4% tenía exclusivamente la variante de PDK4. No se encontraron cambios en el desarrollo clínico, eléctrico ni morfológico de la enfermedad en los perros que tenían diferentes mutaciones. Partiendo de un total de 27 animales enfermos, 12 con arritmias patológicas de relevancia clínica y 15 sólo con cambios morfológicos y/o alteraciones electrocardiográficas sin relevancia clínica, no se encontró relación entre la presencia de las variantes CFA14: g.20,829,667_20,829,682del, CFA5: g.53,941,386T>C y CFA36: g.22,321,955 C>T y el desarrollo de arritmias en la patología ($p > 0,1$). Por último, no se encontró relación entre la presencia de más de un gen y el desarrollo de la enfermedad ($p > 0,1$).

Tabla 10: Análisis estadístico de asociación entre el desarrollo de CMD y las variantes reoportadas, calculado con el paquete de R y el software SPSS 23.0.

Variantes	Alelos			Genotipos
	χ^2 (p)	F (p)	OR	
PDK4	0,32	0,40	1,50	0,57
SNP CHR 5	0,15	0,18	1,88	0,40
TTN	0,02	0,02	2,36	0,004

PDK4: CFA14:g.20,829,667_20,829,682del; SNP CHR 5: CFA5:g.53,941,386T>C; TTN: CFA36:g.22,321,955 C>T; χ^2 (p): Test de Chi cuadrado χ^2 (valor de p); F (p): Test Exacto de Fisher (valor de p); OR: Odds Ratio.

Tabla 11: Modelos de asociación basados en la prueba exacta de Fisher realizados con el paquete de PLINK.

Variantes	Efecto genético dominante			Efecto genético recesivo			Cochran-Armitage	
	χ^2 (p)	F (p)	OR	χ^2 (p)	F (p)	OR	F (p)	OR
PDK4	0,29	0,33	1,68	-	1,00	1,69	0,33	1,68
SNP CHR 5	0,17	0,19	2,06	-	0,66	1,75	0,18	1,88
TTN	-	1,00	1,50	0,001	0,002	6,50	0,007	6,79

PDK4: CFA14:g.20,829,667_20,829,682del; SNP CHR 5: CFA5:g.53,941,386T>C; TTN: CFA36:g.22,321,955 C>T; χ^2 (p): Test de Chi cuadrado χ^2 (valor de p); F (p): Test Exacto de Fisher (valor de p); OR: Odds Ratio.

Resultados en la población de Bóxers

De los 18 Bóxers incluidos en este estudio, tres presentaron CMD, dos machos de 9 años cada uno y una hembra de 14 años de edad. El 100% de los animales afectados presentaron uno o más signos clínicos cardiológicos al momento de la evaluación, como la presencia de pulso asincrónico, intolerancia al ejercicio, mucosas color rosa pálido, pulso femoral débil y síncope. Asimismo, los tres Bóxers enfermos presentaron arritmias patológicas de relevancia clínica: CPVs (n = 1), taquicardia ventricular paroxística (n = 1) y fibrilación atrial (n = 1). Por otro lado, cuatro de los animales muestreados (3 hembras de 6, 10, 14 años y 1 macho de 10 años), presentaron alteraciones clínicas y electrocardiográficas, compatible con cardiomiopatía arritmogénica ventricular derecha (ARVC)

(CPVs; n = 4 y bloqueo atrioventricular de 2do grado; n = 1), pero sin manifestar alteraciones ecocardiográficas.

Análisis estadístico de la evaluación clínica, electrocardiográfica y ecocardiográfica.

Para el análisis estadístico de la evaluación clínica, electrocardiográfica y ecocardiográfica de los Bóxers se dividió a los animales en tres grupos (sanos, n = 11; ARVC, n = 4; CMD, n = 3) sólo con fines descriptivos. Los valores ecocardiográficos no arrojaron diferencias significativas entre los grupos sanos y ARVC. Se observaron valores mayores de DSVI ($p < 0,01$), AI ($p < 0,01$), PSSE ($p < 0,01$) y FC ($p < 0,05$) en el grupo CMD. En cambio, se observaron valores menores de FA ($p < 0,01$), TIVS ($p < 0,01$), PLD ($p < 0,01$) y PLS ($p < 0,01$) en el grupo CMD versus los grupos ARVC y sanos. No se observaron diferencias para las variables FC ($p > 0,05$), ritmo cardíaco ($p > 0,1$), duración onda P ($p > 0,1$), amplitud onda P ($p > 0,1$), intervalo PR ($p > 0,1$), duración QRS ($p > 0,1$), amplitud QRS ($p > 0,1$), duración QT ($p > 0,1$), duración QTc ($p > 0,1$), segmento ST ($p > 0,1$), EEM ($p > 0,1$), DDVI ($p > 0,1$), TIVD ($p > 0,1$) y Ao ($p > 0,1$).

Métodos de genotipificación de potenciales mutaciones causales de CMD

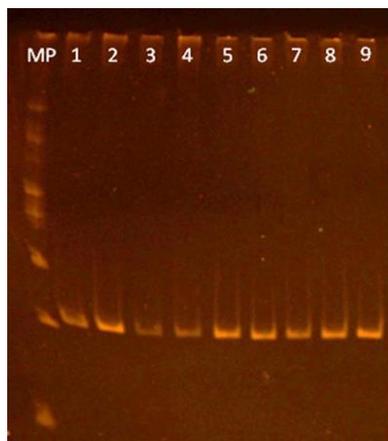
Con el fin de evaluar la presencia de las tres variantes asociadas a CMD reportadas en DP en una población local de Bóxers, se genotipificaron los animales siguiendo las técnicas antes descritas: PCR-análisis de fragmento en

geles de poliacrilamida (para la variante CFA14:g.20,829,667_20,829,682del), PCR-alelo específico (PCR-AS) (para la variante SNP CFA5:g.53,941,386T>C) y pirosecuenciación (para la variante CFA36:g.22,321,955C>T).

A partir de los cebadores diseñados previamente, se realizaron las PCR antes mencionadas para cada variante, obteniendo los fragmentos amplificados del tamaño esperado tanto para PDK4, para el locus en el CHR5, y para TTN (210pb y 194pb; 259 pb y 133 pb; y 212pb y 130 pb, respectivamente).

Con la técnica de PCR- análisis de fragmentos se descartó la presencia de la variante **CFA14: g.20,829,667_20,829,682del** del gen PDK4 en la población de Bóxers local, obteniendo animales homocigotas para el alelo salvaje (en el gel de poliacrilamida 1 única banda de 210 pb de longitud correspondiente al fragmento completo, Fig. 11).

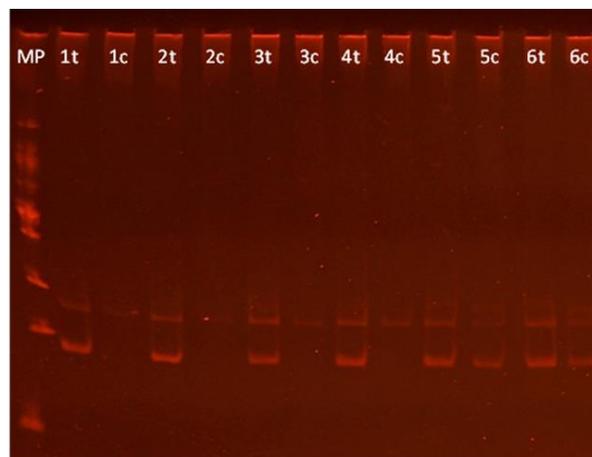
Figura 11: Gel de poliacrilamida 6% - 1 X TBE. Fragmento amplificado por PCR que incluye el polimorfismo CFA14:g.20,829,667_20,829,682del del gen PDK4 donde se observan todos animales homocigotas para el alelo salvaje (W/W; 1 a 9).



MP: Marcador de peso molecular 100 - 1500 pb (INBIO HIGHWAY).

Por otro lado, con la técnica de PCR-AS se confirmó la presencia del SNP **CFA5:g.53,941,386T>C** en la población de Bóxers local, identificando animales con dos de los posibles genotipos para la variante: homocigota para el alelo salvaje (T/T) y heterocigota (C/T) (Fig. 12).

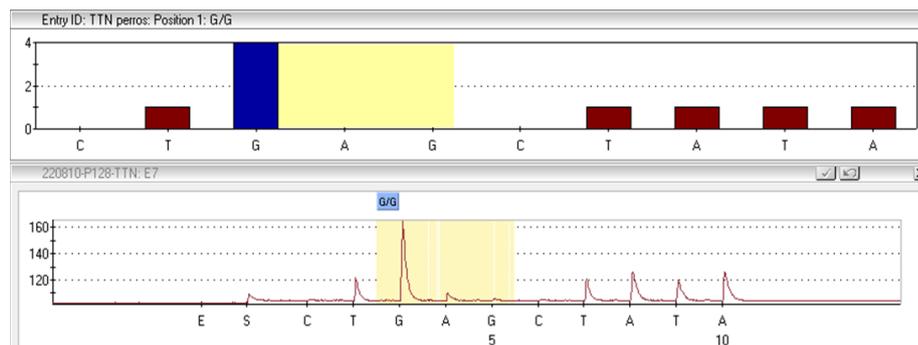
Figura 12: Gel de poliacrilamida 6% - 1 X TBE. Fragmento amplificado por PCR que incluye el SNP **CFA5:g.53,941,386T>C** ubicado en el CHR5.



Muestras 1, 2, 3 y 4: animales homocigotas para el alelo salvaje (T/T), 5 y 6: animales heterocigotas (T/C), t: reacción de PCR con alelo específico T; c: reacción de PCR con alelo específico C; MP: Marcador de peso molecular (Marcador de ADN – INBIO HIGHWAY (100 pb)).

Por último, se secuenció mediante la técnica de Pirosecuenciación la variante **CFA36:g.22,321,955 C>T** del gen TTN. Esto permitió descartar la presencia de la mutación en la población local, obteniendo los pirogramas correspondientes a animales homocigotas salvajes (C/C) (Fig. 13).

Figura 13: Pirograma obtenido por pirosecuenciación del fragmento del que incluye el SNP CFA36:g.22,321,955 C>T del gen TTN, donde se observa el genotipo homocigota para el alelo salvaje C (visualizado como G por utilizar la cadena complementaria).



Análisis poblacional y estadístico de la variante CFA5:g.53,941,386T>C

Se analizó únicamente la variante **CFA5:g.53,941,386T>C** del CHR5 en los Bóxers, ya que los otros dos polimorfismos resultaron monomórficos en la muestra analizada. La frecuencia alélica del alelo mutado fue de 0,11. Catorce animales presentaron un genotipo homocigota para el alelo salvaje, mientras que cuatro fueron heterocigotas. La estimación del EHW en la población total no evidenció un desvío significativo con respecto a las proporciones teóricas esperadas para el equilibrio (Tabla 12).

Tabla 12: Frecuencias genotípicas y alélicas de la variante CFA5:g.53,941,386T>C obtenidas en la población de Bóxers analizadas.

Variante	n°	Frecuencias genotípicas		Frecuencias alélicas		HWE	h _O	h _E
		W/W	W/m	P	Q			
CFA5^a	18	14	4	0,88	0,11	1	0,22	0,20

a: CFA5:g.53,941,386T>C; n° = número de animales; “W” alelo salvaje “ ” alelo mutado, denominados p y q, respectivamente. HWE = Equilibrio de Hardy Weinberg; h_O: heterocigosidad observada; h_E: heterocigosidad esperada.

Para el análisis estadístico de asociación de la variante del SNP del CHR5 al desarrollo de CMD en Bóxers se los dividió en 3 grupos (sanos n = 11; ARVC n = 4; CMD n = 3). De los 18 Bóxers genotificados los 4 animales heterocigotas (C/T), portadores del alelo mutado, pertenecían al grupo sano. Esto no permite realizar un análisis de asociación confiable. El tamaño muestral de cada grupo es reducido y se requiere aumentar el número de muestras para realizar un análisis estadístico confiable.

Discusión y conclusiones

El análisis clínico y cardiovascular desarrollado en el presente estudio arrojó una prevalencia estimada de la enfermedad en la población estudiada de DP

de 75 . Este valor es menor al reportado en otros países (5 para poblaciones de Estados Unidos (Calvert y col., 1997a; Lett y col.,

Grady y col., 1998), y 58,2% en ejemplares de Europa (Wess y col., 2010). La menor prevalencia hallada en la población de DP de Argentina puede estar relacionada a varios factores. En primer lugar, se puede atribuir al efecto fundador que ocurrió durante la creación inicial de la población de DP en este país, donde se introdujo un número limitado de individuos reproductores procedentes del extranjero. Con el desarrollo de la cría en Argentina, los niveles de consanguinidad aumentaron. Además, debido al reducido flujo genético con el exterior, la población local no se homogeneizó con la población de origen, lo que resultó en poblaciones diferenciadas y estructuradas a nivel mundial. Finalmente, se ha observado una disminución en la cría de esta raza durante los últimos años. Según la Federación Cinológica Argentina, la cantidad de caninos de la raza DP registrados en Argentina ha descendido de 580 nuevos ingresos en 2015 a 250 en 2020 (<http://www.fca.org.ar>).

La edad promedio de detección de la enfermedad fue de $6,3 \pm 3,8$ años, con un rango de 1 hasta 13 años de edad. Esto está en línea con lo reportado previamente donde se informó una edad de presentación de entre 5 y 10 años de edad, incrementándose las probabilidades de desarrollo de CMD conforme esta aumenta (Calvert y col., 1997a; Domanjko-Petric y col., 2002; Wess y col., 2010). A su vez, el hallazgo de cuatro DP afectados con una edad de entre 1 a 3 años enfatiza la importancia de la evaluación de los animales desde edades tempranas (Calvert y col., 1997b; Wess y col., 2017). Por otro lado, la asociación entre el sexo y el desarrollo de la enfermedad donde los machos fueron más propensos a

padecer la enfermedad, concuerda con varios estudios previos (Calvert y col., 1997b). El desarrollo de las manifestaciones clínicas en los animales que desarrollaron la enfermedad, sumado a los hallazgos electrocardiográficos y ecocardiográficos están en línea con lo previamente reportado para esta raza (Domanjko-Petric y col., 2002; Wess, 2022).

En el presente estudio se utilizaron diferentes técnicas de genotipificación indirectas, como PCR-análisis de fragmento en geles de poliacrilamida y PCR-alelo específico, y directas como pirosecuenciación y secuenciación directa de fragmentos de amplificación. Esta última se realizó para validar las metodologías de genotipificación indirecta antes mencionadas. Así, se confirmó la presencia de las variantes CFA14: g.20,829,667_20,829,682del, CFA5: g.53,941,386T>C y CFA36: g.22,321,955 C>T en la población de DP de Argentina, encontrando los tres genotipos posibles para cada una de los polimorfismos en cuestión. El desarrollo de estas técnicas para cada mutación permitió obtener métodos de diagnóstico rápidos, confiables, repetibles y a un costo relativamente bajo, de acuerdo a las circunstancias y realidad en la que se encuentra el sector productivo relacionado a la cría de perros en Argentina. Asimismo, en este estudio se utilizó la tecnología de NGS target, técnica que permite la secuenciación de millones de pequeños fragmentos de ADN en paralelo. Esta tecnología permite analizar simultáneamente un gran número de polimorfismos y muestras, aunque el costo de la corrida es una limitante si no se tiene una escala de trabajo suficiente. Sin embargo, si consideramos el costo polimorfismo / animal se transforma en una técnica muy eficiente (Hall, 2007). En nuestro país, la demanda de este tipo de análisis (enfermedades genéticas) aún no llega a tener un volumen que justifique

el uso de estas tecnologías, aunque es de esperar que el uso de esta tecnología vaya aumentando significativamente durante los próximos años.

PDK4 es una de las cuatro isoenzimas de la piruvato deshidrogenasa quinasa (PDK), y se expresa predominantemente en el corazón y el músculo esquelético (Matkovich y col., 2010). La proteína PDK4 tiene un papel regulador en el metabolismo de energía cardíaca mediante la fosforilación del piruvato, y por lo tanto, disminuyendo la oxidación de la glucosa durante los períodos en que la oxidación de ácidos grasos puede generar energía de manera más eficiente. Lopaschuk y col. (2010) propusieron que la presencia de la variante CFA14:g.20,829,667_20,829,682del produciría una proteína con reducción de su actividad, una menor capacidad para fosforilar y por ende, regular negativamente la oxidación de la glucosa. Sin embargo, no se ha podido evaluar adecuadamente la actividad de PDK4 (Meurs y col., 2012). La asociación entre la delección de 16 pb en el gen PDK4 y el desarrollo de CMD reportada en una población de DP estadounidense (Meurs y col., 2012) no fue confirmada en la cohorte de DP de Argentina. De la misma manera la frecuencia alélica total fue menor en la población local (0,17) que en la población estadounidense (0,37). Los resultados de este estudio coinciden con reportes previos realizados en una población de DP europea, en donde no se encontró asociación entre CFA14:g.20,829,667_20,829,682del y el desarrollo de CMD, y donde la frecuencia alélica fue similar (0,16) a la del presente trabajo (Owczarek-Lipska y col., 2013). Sin embargo, en este estudio y en el resto de los reportes mencionados, se evidenció una frecuencia alélica mayor del alelo mutado en la población de perros enfermos que en los perros sanos, lo que plantea la necesidad

de desarrollar nuevos estudios. Por otro lado, se analizó el valor del F_{IS} de esta variante, el cual arrojó un valor negativo aunque no significativo, lo que indicaría un leve exceso de heterocigotas en la población (Royo, 2020). Por ese motivo, los resultados sugieren que la variante se encuentra en equilibrio de Hardy Weinberg, aunque puede existir cierto grado de desviación que no ha sido detectado con suficiente poder estadístico por el número de animales analizados (Royo, 2020). La evaluación de la actividad de la PDK4 y el análisis de asociación en poblaciones de DP de diferente origen son aspectos importantes de futuros estudios, y más aun considerando que el gen PDK4 aún no se ha informado como un gen causante de la cardiomiopatía dilatada en seres humanos.

1 SNP CFA5:g.5 T> C (alelo "C") reportado en el cromosoma 5 fue asociado al desarrollo de CMD en DP y se lo consideró responsable de casi el 50% de los casos afectados (Mausberg y col., 2011). Sin embargo, la señal de asociación no fue ligada a ningún gen candidato para CMD. Al analizar esta variante en la población de DP local, no se encontró evidencia de una participación en la etiología de la CMD, en contraposición con lo reportado en los DP de Alemania y Reino Unido (Mausberg y col., 2011). A diferencia de lo reportado en esos países, los DP locales con arritmias no presentaron una frecuencia significativamente más alta del alelo C. Por más que la frecuencia del alelo mutado fue mayor en el grupo caso que en el control, la frecuencia total del alelo mutado fue menor que la reportada por Mausberg y col. (2011). La variante en la población de DP de Argentina sugiere una desviación del equilibrio genético de Hardy Weinberg. Sumando a esto, el valor positivo de F_{IS} puede indicar que el número de heterocigotas en la población sería menor de lo esperado. Esto puede

ser causado por factores como la endogamia, la consanguinidad, o la presencia de selección en contra de ciertos genotipos (Falconer y Mackay, 1996) El cuello de botella genético en perros de la raza DP es responsable de un desequilibrio de ligamiento de relativamente largo alcance (LD). El largo LD facilita la detección de mutaciones causales con una cantidad limitada de marcadores, aunque simultáneamente limita la resolución del mapeo (Mausberg y col., 2011). Por lo tanto, la verdadera mutación causal subyacente a la señal de asociación detectada podría estar ubicada en cualquier lugar en una región de 1 Mb. Investigaciones adicionales son necesarias para poder identificar un nuevo gen relacionado a este locus del cromosoma 5, que podría ser de relevancia para una fracción de los casos de CMD en humanos que aún no tienen explicación genética.

TTN es el gen más comúnmente asociado con CMD en seres humanos (de Gonzalo-Calvo y col., 2017; Franaszczyk y col., 2017; Herman y col., 2012; Kimura, 2016; Le Winter y Granzier, 2014). Este gen codifica para la transcripción de titina, la proteína más grande conocida, que se expresa tanto en músculo cardíaco y esquelético. Esta proteína actúa como un resorte molecular que contribuye a la rigidez pasiva del músculo y contracción activa, así como detección biomecánica y señalización (de Gonzalo-Calvo y col., 2017; Herman y col., 2012; Kimura, 2016; Linke, 2018; Granzier y col., 2009). La variante CFA36: g.22,321,955 C>T, analizada en este estudio, fue identificada en humanos dentro de un dominio, similar a la inmunoglobulina, denominado I71 (Bang y col., 2001). Este dominio compromete uno de los tres elementos de resorte dentro de la titina cardíaca y está incluido en la isoforma N2BA de titina cardíaca (Herman y col., 2012; Schafer y col., 2017; Bang y col., 2001; Labeit y col., 2006). Aunque

la variante humana que causa la CMD no ha sido identificada en esta ubicación exacta, se ha observado una variante missense dentro el exón ortólogo en humanos asociada a cardiomiopatía arritmogénica del ventrículo derecho (Taylor y col., 2011). En este estudio, se demostró una asociación estadísticamente significativa entre la variante CFA36: g.22,321,955 C>T en el gen TTN y el desarrollo de CMD en la población de DP de Argentina, en coincidencia con lo reportado por Meurs y col. (2019), evidenciando una frecuencia del alelo mutado mayor en el grupo caso que en el control. La frecuencia alélica total del alelo mutado fue ligeramente superior a la encontrada en otros reportes (0,58 vs. 0,53, respectivamente; Meurs y col., 2019). Meurs y col., (2007) reconocieron en la población de DP estudiada un modo de herencia autosómico dominante. Por más que en este estudio se demostró que la presencia de la variante actúa como un factor de riesgo para el desarrollo de la enfermedad, no fue posible confirmar el efecto dominante de la mutación. Esto podría deberse a diferentes causas. En primera instancia el diseño experimental es distinto, ya que en los estudios previos utilizaron pedigrís para el análisis de herencia, y en este trabajo se analizaron perros provenientes de poblaciones independientes. Asimismo, la determinación del modo de transmisión de la variante es difícil a nivel poblacional, debido a sus características de penetrancia incompleta y expresividad variable, por lo que sería necesario evaluar un número de animales aún mayor al estudiado en este estudio. Por otro lado, la variante en esta población mostró una desviación del equilibrio genético de Hardy Weinberg, con un valor de F_{IS} negativo, indicando que puede haber factores que están afectando la distribución esperada de genotipos en la población, presentando un mayor número de heterocigotas (Hardy, 1908;

Weinberg, 1908). Esto puede deberse a factores como la presencia de migración genética, selección natural o estructura genética entre subpoblaciones (Royo, 2020). Es importante destacar que el desequilibrio de Hardy-Weinberg no siempre indica una situación problemática o anormal en la población. En algunos casos, el desequilibrio puede ser indicativo de factores como la presencia de enfermedades genéticas, selección artificial o procesos de fragmentación de la población, lo que puede tener implicaciones importantes para la salud y la viabilidad de la población (Royo, 2020). Esta estructuración de la población podría explicar el aumento de heterocigotas en la población de Argentina. Por otro lado, el desequilibrio de Hardy-Weinberg puede ser más notable en muestras pequeñas debido a los efectos de la deriva genética (cambios aleatorios en las frecuencias genéticas que pueden ocurrir en poblaciones pequeñas debido a la variación aleatoria en la transmisión de los alelos de una generación a otra; Gillespie, 1998).

La penetrancia y la expresión de la cardiomiopatía dilatada varían según la mutación causal y otros factores como edad, sexo, ambiente, estilo de vida y factores genéticos modificadores, tanto en perros como en humanos, lo que puede afectar el desarrollo de la enfermedad (Hershberger y col., 2013; Meurs y col., 2007; Burkett y Hershberger 2005). Los fenómenos de penetrancia incompleta y expresión variable probablemente resultan de la interacción entre la información genética y la activación diferencial de sus factores epigenéticos reguladores. Diferentes estilos de vida, la dieta y otros factores ambientales pueden alterar la forma en que el código genético determina no solo el estado fisiológico de los órganos de cada perro, sino también las enfermedades que les afectan (Dutton y col., 2018). La CMD tiene una penetrancia incompleta dependiente de la edad,

con expresión variable, por lo que las mutaciones particulares pueden tener una amplia variabilidad en el efecto y la gravedad, incluso dentro de la misma familia. Aunque la mutación en TTN demostró una asociación significativa y un riesgo relativo al desarrollo de la enfermedad, algunos de los perros afectados (12% de la población) no presentaron la mutación. Este hallazgo es consistente con la heterogeneidad genética observada en CMD en seres humanos, en donde más de 30 genes han sido asociados con el desarrollo de cardiomiopatía dilatada familiar y, sin embargo, sólo se ha identificado alrededor del 30-35% de la etiología genética (Hershberger y col. 2010). Asimismo, en la población de DP evaluada, en la que hay perros de diferentes familias, muchos de ellos tenían la variante, pero no desarrollaron la enfermedad. Esto podría indicar que los perros no comparten el mismo *background* genético o factores ambientales que influyen en la aparición de la enfermedad. La penetrancia incompleta y la variabilidad de expresión en seres humanos y perros sugieren un posible papel de factores ambientales, modificadores genéticos o ambos, que influyen en el curso natural de la enfermedad. Por este motivo, se requieren grandes cohortes para detectar asociaciones con alelos de riesgo de baja frecuencia o con efecto menor (Stark y col., 2010).

Por último, al ser esta una enfermedad poligénica, se evaluó la interacción entre las variantes reportadas, pero no encontrando relación estadística que implique una relación entre las mutaciones (Meurs y col., 2007). Este resultado puede estar limitado por la diferencia entre la cantidad de perros con las diferentes mutaciones (TTN = 8; 3 variantes = 6; TTN y PDK4 = 6; TTN y SNP CHR5 = 4;

sin mutaciones = 2; PDK4 = 1) y la magnitud del efecto de cada polimorfismo y de sus interacciones en el desarrollo de CMD.

En la población de Bóxers analizada en este estudio se pudieron identificar tanto caninos con CMD, caninos con ARVC y caninos sanos. No se detectaron las variantes de los genes PDK4 y TTN en los tres grupos examinados. Estos resultados concuerdan con trabajos en los que no encontraron la variante PDK4 en cuatro Bóxers americanos (Meurs y col., 2012) y un Bóxer europeo (Owczarek-Lipska y col., 2013), ni la variante TTN en 25 Bóxers americanos (Meurs y col., 2019). Por otro lado, con la técnica de PCR-AS se confirmó la presencia del SNP CFA5:g.53,941,386T>C en la población de Bóxers local, identificando animales con dos de los posibles genotipos, homocigota para el alelo salvaje (T/T) y heterocigota (C/T). El alelo mutado se encontró únicamente en el grupo sano, probablemente porque el número de animales es mayor que en el resto de los grupos, lo que impide la realización de estudios de asociación.

Los resultados de este trabajo enfatizan que las técnicas de detección genética desempeñan un papel fundamental en el estudio de enfermedades cardiovasculares, ofreciendo una comprensión más profunda de las variantes genéticas asociadas y proporcionando información valiosa para la prevención, el pronóstico, el tratamiento temprano y la planificación de cruzamientos selectivos, con el fin de disminuir la incidencia de la enfermedad. Sin embargo, las pruebas genéticas deben verse como un método para evaluar el *riesgo* de enfermedad en un animal reproductor, considerando que la penetrancia incompleta y la naturaleza poligénica y multifactorial de las cardiomiopatías complican la interpretación de los resultados de las pruebas.

Por todo lo expuesto se concluye que en la población de DP de Argentina están presentes las variantes CFA14: g.20,829,667_20,829,682del, CFA5:g.53,941,386T>C, CFA36: g.22,321,955 C>T, las que fueron detectadas mediante el desarrollo de métodos rápidos de genotipificación. Asimismo, se determinó la prevalencia de los polimorfismos antes mencionados, encontrando una frecuencia alélica de 0,17, 0,12 y 0,58, respectivamente, valores que varían con respecto a los reportados para DP en otras regiones. Sin embargo, sólo la variante en el gen TTN evidenció una asociación significativa con el desarrollo de la enfermedad, siendo considerada la presencia del alelo mutado un factor de riesgo. Por último, en la población de Bóxers sólo se pudo confirmar la presencia del SNP CFA5:g.53,941,386T>C, sin poder relacionarlo con la etiología genética de la enfermedad.

Referencias

1. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990; 215(3): 403-410.
2. Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circulation Research.* 2001; 89: 1065–1072.

3. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21: 263-265.
4. Boon J. The echocardiographic examination. En: Boon J (ed). *Manual of veterinary echocardiography*. Media, Williams and Wilkins, 1998, p. 35-128.
5. Burkett EL, Hershberger RE. Clinical and genetic issues in familial dilated cardiomyopathy. *J Am Coll Cardiol*. 2005; 45: 969–981.
6. Calvert CA, Pickus CW, Jacobs GJ, Brown J. Signalment, survival, and prognostic factors in Doberman Pinschers with end-stage cardiomyopathy. *J Vet Intern Med*. 1997; 11: 323-6.
7. Calvert CA, Meurs K. CVT update: Doberman Pinscher occult cardiomyopathy. En: Bonagura JD, Kirk RW, (eds). *Kirk 's Current Veterinary Therapy*. Philadelphia, PA: Saunders; 2000, p. 756-60.
8. Chetboul V, Tidholm A, Nicolle A, Sampedrano CC, Gouni V, Pouchelon JL, Lefebvre HP, Concordet D. Effects of animal position and number of repeated measurements on selected two-dimensional and M-mode echocardiographic variables in healthy dogs. *J Am Vet Med Assoc*. 2005; 227: 743–747.
9. de Gonzalo-Calvo D, Quezada M, Campuzano O, Perez-Serra A, Broncano J, Ayala R y col. Familial dilated cardiomyopathy: a multidisciplinary entity, from basic screening to novel circulating biomarkers. *Int J Cardiol*. 2017; 228:870–880.
10. Domanjko-Petric A, Stabej P, Zemva A. Dilated cardiomyopathy in Doberman Pinschers: survival, causes of death and a pedigree review in a related line. *J Vet Cardiol*. 2002; 4: 17-24.
11. Dutton E, López-Alvarez J. An update on canine cardiomyopathies - is it all in the genes? *J Small Anim Pract*. 2018 Apr 17.

12. Falconer DS, Mackay TFC. Introduction to quantitative genetics. Oxford: Blackwell's Oxford .
13. Ferasin L, Ferasin H, Little CJ. Lack of correlation between canine heart rate and body size in veterinary clinical practice. *J Small Anim Pract.* 2010; 51: 412-418.
14. Franaszczyk M, Chmielewski P, Truszkowska G, Stawinski P, Michalak E, Rydzanicz M y col. Titin truncating variants in dilated cardiomyopathy—prevalence and genotype-phenotype correlations. *PLoS One.* 2017; 12: e0169007.
15. Gillespie JH. Population Genetics: A concise guide. The Johns Hopkins University Press, Baltimore and London, 1998.
16. Gottdiener JS, Bednarz J, Devereux R, Gardin J, Klein A, Manning WJ, y col. American Society of Echocardiography recommendations for use of echocardiography in clinical trials. *J Am Soc Echocardiogr.* 2004; 17: 1086–1119.
17. Granzier HL, Radke MH, Peng J, Westermann D, Nelson OL, Rost K, King NMP col. Truncation of titin's elastic P VK region leads to cardio path with diastolic dysfunction. *Circ Res.* 2009; 105: 557–564.
18. Hall N. Advanced sequencing technologies and their wider impact in microbiology. *The Journal of Experimental Biology.* 2007; 209: 1518-1525.
19. Hardy GH. Mendelian proportions in a mixed population. *Science.* 1908; 28(706): 49-50.
20. Harpster N. Bóxer cardiomyopathy. In: Kirk R (ed.). *Current Veterinary Therapy VIII.* Philadelphia, PA:WB Saunders; 1983, p. 329–337.
21. Hazlett MJ, Maxie MG, Allen DG, Wilcock BP. A retrospective study of heart disease in Doberman Pinscher dogs. *Can Vet J.* 1983; 24: 205-210.

22. Herman DS, Lam L, Taylor MRG, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *New England Journal of Medicine*. 2012; 366: 619–628.
23. Hershberger RE, Hedges DJ, Morales A. Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nat Rev Cardiol*. 2013; 10: 531–547.
24. Kimura A. Molecular genetics and pathogenesis of cardiomyopathy. *J Hum Genet*. 2016; 61: 41–50.
25. Kittleson MD. Primary Myocardial Disease - Primary Myocardial Disease Leading to Chronic Myocardial Failure. En: Kienle RD, Kittleson MD, (eds). *Small Animal Cardiovascular Medicine Textbook*. USA, Mosby, 1998, p. 319-46.
26. Labeit S, Lahmers S, Burkart C, Fong C, McNabb M, Witt S, Witt C, Labeit D, Granzier H. Expression of distinct classes of titin isoforms in striated and smooth muscles by alternative splicing, and their conserved interaction with filamins. *J Mol Biol*. 2006; 362: 664–681.
27. LeWinter MM, Granzier HL. Cardiac titin and heart disease. *J Cardiovasc Pharmacol*. 2014; 63: 207–212.
28. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25(16); 2078-2079.
29. Linke WA. Titin gene and protein functions in passive and active muscle. *Annu Rev Physiol*. 2018; 80: 389–411.
30. Lopaschuk GD, Ussher JR, Folmes CDL, Jaswal JS, Stanley WC. Myocardial fatty acid metabolism in health and disease. *Physiol Rev*. 2010; 90: 207–258.

31. Matkovich SJ, Van Booven DJ, Hindes A, Kang MY, Druley TE, Vallania FL, Mitra RD, Reilly MP, Cappola TP, Dorn GW. Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy revealing HSPB7 polymorphisms associated with disease. *J Clin Invest.* 2012; 120: 280–289.
32. Mausberg TB, Wess G, Simak J, Keller L, Drögemüller M, Drögemüller C, Webster MT, Stephenson H, Dukes-McEwan J, Leeb T. A locus on chromosome 5 is associated with dilated cardiomyopathy in Doberman Pinschers. *PloS one.* 2011; 6:e20042.
33. Meurs KM, Fox PR, Horgard M, Spier AW, Lamb A, Koplitz SL, Baumwart RD. A prospective genetic evaluation of familial dilated cardiomyopathy in the Doberman Pinschers. *J Vet Intern Med.* 2007; 21: 1016-1020.
34. Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. *Hum Genet.* 2012; 131: 1319-1325.
35. Meurs KM, FriedenberG SG, Kolb J, Saripalli C, Tonino P, Woodruff K, Olby NJ, Keene BW, Adin DB, Yost OL, DeFrancesco TC, Lahmers S, Tou S, Shelton GD, Granzier H. A missense variant in the titin gene in Doberman pinscher dogs with familial dilated cardiomyopathy and sudden cardiac death. *Hum Genet.* 2019; 138(5):515-524.
36. Mertes F, Elsharawy A, Sauer S, van Helvoort JM, van der Zaag PJ, Franke A, Nilsson M, Lehrach H, Brookes AJ. Targeted enrichment of genomic DNA regions for next-generation sequencing. *Brief Funct Genomics.* 2011; 10(6): 374-386.

37. Nagueh SF, Appleton CP, Gillebert TC, Marino PN, Oh JK, Smiseth OA, Waggoner AD, Flachskampf FA, Pellikka PA, Evangelista A. Recommendations for the evaluation of left ventricular diastolic function by echocardiography. *J Am Soc Echocardiogr.* 2009; 22: 107–133.
38. Nei, M. *Molecular Evolutionary Genetics*. Columbia University Press. New York, USA. 1989, p12.
39. 'Grad MR, 'orne R. The prevalence of dilated cardiomyopathy in Doberman Pinschers: A 4.5 year follow-up. *J Vet Intern Med.* 1998; 12:199.
40. 'Grad MR, 'Sullivan ML. Dilated cardiomyopathy: An update. *Vet Clin North Am Small Anim Pract.* 2004; 34: 1187-1207.
41. Owczarek-Lipska M, Mausberg TB, Stephenson H, Dukes-McEwan J, Wess G, Leeb T. A 16-bp deletion in the canine PDK4 gene is not associated with dilated cardiomyopathy in a European cohort of Doberman Pinschers. *Anim Genet.* 2013; 44: 239.
42. Peral García P, Giovambattista G, Ripoli MV. *Genética forense no- humana*. EDULP (Ed. Universidad de La Plata), Argentina, La Plata, 2015; 3:90-92.
43. Purcell S. PLINK. Center for Human Genetic Research (CHGR), Massachusetts General Hospital. Broad Institute of Harvard & MIT. 2009; 1: 7.
44. Rishniw M, Erb HN. Evaluation of four 2-dimensional echocardiographic methods of assessing left atrial size in dogs. *J Vet Intern Med.* 2000; 14: 429-435.
45. Robinson J, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol.* 2011; 29: 24–26.
46. Ronaghi, M., Uhlen, M., Nyren, P. A sequencing method based on real-time pyrophosphate. *Science.* 1998; 281: 363-365.

47. Rousset F, Raymond M. Statistical analysis of population genetic data: old tools, new concepts. *Trends Ecol. Evol.* 1997; 12: 313-317.
48. Rousset F. Inferences from spatial population genetics. In: Balding DJ, Bishop M, Cannings C (eds). *Handbook of statistical genetics*. 3th edition. Wiley, Chichester, U.K., 2007, p. 945-79.
49. Royo JL. Hardy Weinberg Equilibrium Disturbances in Case-Control Studies Lead to Non-Conclusive Results. *Cell J.* 2021; 22(4): 572-574.
50. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977; 74(12): 5463-5467.
51. Schafer S, de Marvao A, Adami E, Fiedler LR, Ng B, Khin E y col. Titin-truncating variants affect heart function in disease cohorts and the general population. *Nature Genetics.* 2017; 49: 46–53.
52. Schneider S, Roessli D, Excoffier L. Arlequin Version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
53. Sievers F, Higgins DG. Clustal omega. *Curr Protoc Bioinformatics.* 2014; 48:3.13.1-16.
54. Stark K, Esslinger UB, Reinhard W, Petrov G, Winkler T. Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. *PLoS Genet.* 2010; 6: e1001167.
55. Steudemann C, Bauersachs S, Weber K, Wess, G. Detection and comparison of microRNA expression in the serum of Doberman Pinschers with dilated cardiomyopathy and healthy controls. *BMC veterinary research.* 2013; 9: 12-25.

56. Tattersall ML, Dymond M, Hammond T, Valentin JP. Correction of QT values to allow for increases in heart rate in conscious Beagle dogs in toxicology assessment. *J Pharmacol Toxicol Methods*. 2006; 53: 11-19.
57. Taylor M, Graw S, Sinagra G, Barnes C, Slavov D, Brun F, Pinamonti B, Salcedo EE, Sauer W, Pyxaras S, Anderson B, Simon B, Bogomolovas J, Labeit S, Granzier H, Mestroni L. Genetic variation in titin in ARVC-overlap syndromes. *Circulation*. 2011; 124: 876–885.
58. Thomas WP, Gaber CE, Jacobs GJ, Kaplan PM, Lombard CW, Moise NS, Moses BL. Recommendations for standards in transthoracic two-dimensional echocardiography in the dog and cat: Echocardiography Committee of the Specialty of cardiology, American College of Veterinary Internal Medicine. *J Vet Intern Med*. 1993; 7: 247–252.
59. Tidholm A, Haggstrom J, Borgarelli M, Tarducci A. Canine idiopathic dilated cardiomyopathy. Part I: Aetiology, clinical characteristics, epidemiology and pathology. *Vet J*. 2001; 162: 92–107.
60. Tilley LP. The approach to the electrocardiogram. En: Tilley LP (ed.). *Essentials of canine and feline electrocardiography*. Pennsylvania, Estados Unidos, Lippincott Williams & Wilkins, Media, 1992, p. 40-50.
61. Villegas Castagnasso EE, Francisco EI, Crespi JA, Diaz S, Corbi Botto CM, Ripoli MV. Métodos de genotipificación. En: Peral García P, Giovambattista G, Ripoli MV (eds). *Genética forense no-Humana*. Editorial de la Universidad Nacional de La Plata (UNLP), La Plata, Argentina, 2015.
62. Villegas Castagnasso EE, Posik DM, Crespi JA. Marcadores genéticos: Introducción al análisis y su aplicación en diversas áreas biológicas. En: Catanesi

CI, Villegas Castagnasso EE (eds). Elementos de Genética para estudiantes de Ciencias Biológicas. Editorial de la Universidad Nacional de La Plata (UNLP), La Plata, Argentina, 2021.

63. Weinberg W. About proof of heredity in humans. Germany: The association for national natural history in württemberg. 1908; 64: 368-382.

64. Wess G, Schulze A, Butz V, Simak J, Killich M, Keller LJM, Maeurer J, Hartmann K. Prevalence of dilated cardiomyopathy in Doberman Pinschers in various age groups. J Vet Intern Med. 2010; 24: 533-538.

65. Wess G. Screening for dilated cardiomyopathy in dogs. J Vet Cardiol. 2022; 40: 51-68.

66. Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. Arch Pathol Lab Med. 2017; 141(11): 1544-1557.

CAPÍTULO II

EVALUACIÓN IN SILICO DEL POTENCIAL EFECTO DE LAS VARIANTES CAUSALES DE CMD SOBRE LA ESTRUCTURA DE LAS PROTEÍNAS Y/O LA REGULACIÓN GÉNICA

Introducción

Los avances en biología molecular y el creciente acceso a las tecnologías de secuenciación de material genético, contribuyen a la obtención de abundante cantidad de datos de enfermedades genéticas. De cada análisis surgen varios polimorfismos, por lo que es fundamental procesar, identificar y clasificar aquellas variantes genéticas que son relevantes para el desarrollo de las enfermedades, lo que requiere una gran experiencia y poder computacional. Estas variantes pueden ocurrir en las regiones codificantes o no codificantes de proteínas, las cuales pueden influir en las porciones reguladoras que controlan la expresión génica y regiones no traducidas que afectan la traducción de ARNm.

Existen diferentes tipos de polimorfismos genéticos que pueden afectar el funcionamiento de los genes y, en consecuencia, la funcionalidad y/o estructura de las proteínas. Las variantes de nucleótido simple (SNV) representan aproximadamente el 90% del total de variantes, y pueden tener diferentes efectos en las proteínas. Algunas pueden truncar o elongar la proteína (*Stop Gain/loss* y *Start Loss/gain*), pueden ocurrir con corrimiento del marco de lectura, mientras

que otras no causan cambios en su secuencia (sinónimas) o alteran un aminoácido específico (no sinónimas, nsSNV) (Katsonis y col., 2022). Estas últimas pueden afectar diversas características de las proteínas, como su estabilidad de plegamiento, afinidad de unión, interacciones con otras moléculas, actividad catalítica, modificaciones postraduccionales y su solubilidad, entre otras (Katsonis y col., 2022). Otro tipo de variante genética afecta los sitios de empalme, que son secuencias específicas y altamente conservadas presentes en los intrones de los genes eucariotas. Estas secuencias son reconocidas por complejos de ribonucleoproteínas llamados espliceosomas, que permiten la eliminación de los intrones y la unión de los exones para generar un ARN mensajero maduro que se traduce en proteínas. Las mutaciones en los sitios de empalme pueden afectar la capacidad de los espliceosomas para reconocer y adherirse a estos sitios, lo que lleva a errores en la unión del ARN. Esto puede resultar en la inclusión incorrecta o la omisión de exones, lo que altera la secuencia del ARNm y, por lo tanto, la secuencia de aminoácidos de la proteína final. Estos cambios pueden afectar la estructura tridimensional de la proteína y, a su vez, alterar su función biológica (Tang y col., 2016).

El análisis *in silico* de las variantes genéticas implica el uso de herramientas computacionales para comprender su impacto funcional en un entorno virtual (García y col., 2022). Estas técnicas permiten predecir el efecto de las variantes genéticas, como mutaciones y polimorfismos, en la estructura y el funcionamiento de las proteínas y la regulación génica (García y col., 2022). Son herramientas rápidas, económicas y confiables, lo que nos permite analizar grandes cantidades de datos genómicos y/o proteómicos y filtrar las variantes

relevantes para un fenotipo o enfermedad específica (Singh y Mistry, 2016). Sin embargo, es importante tener en cuenta que los resultados son predicciones y deben interpretarse con cuidado. Es necesario realizar estudios adicionales de validación y confirmación *in vitro* o *in vivo* (Singh y Mistry, 2016).

Existen diferentes herramientas *in silico* para predecir el impacto de las sustituciones no sinónimas, como también los causados por cambios en los sitios de empalme, entre otras. El efecto de las variantes no sinónimas depende de varios factores, como la conservación evolutiva de un aminoácido o nucleótido, la ubicación y el contexto dentro de la secuencia proteica y la consecuencia bioquímica de la sustitución del aminoácido (Richards y col., 2015). Los algoritmos *in silico* utilizan características estructurales para clasificar y evaluar el impacto de las sustituciones en la estructura y función de las proteínas, como también pueden hacer uso de información evolutiva. Las herramientas que predicen el impacto de las mutaciones en los sitios de empalme tienen una buena sensibilidad (~90-100%) en la detección de anomalías en lo que se refiere a la creación o pérdida de sitios de empalme (Richards y col., 2015).

Por otro lado, cada proteína se pliega en un ensamble con un número variable de conformaciones que le permiten llevar a cabo su función biológica. Por ese motivo, el conocimiento de la estructura tridimensional de las proteínas y sus distintas conformaciones es fundamental ya que la presencia de una variante, puede generar cambios significativos (Priya Doss y col., 2014). Para analizar los cambios causados por una variante en su entorno cercano, es necesario contar con el conjunto de conformaciones tanto de la proteína nativa como de la portadora de la variante. El modelado de estructuras proteicas nos permite obtener esquemas

del plegamiento tridimensional de proteínas aún no cristalizadas, mapear los residuos aminoacídicos de interés y evaluar las sustituciones en el contexto estructural (Priya Doss y col., 2014). Sin embargo, estas metodologías aportan un modelo de una única configuración estructural relativamente rígida en un estado específico, sin considerar los movimientos internos y los cambios conformacionales que juegan un papel esencial en su función (Karplus y McCammon, 2002). Esta descripción incompleta podría subsanarse contando con una serie de estructuras experimentales obtenidas en diferentes situaciones (presencia/ausencia de ligandos, cambios de concentración, o de condiciones del entorno). Asimismo, resulta muy importante sumar al análisis estructural las simulaciones de dinámica molecular, que brindan una visión del comportamiento dinámico de la proteína. Proporcionan información acerca del movimiento de los átomos que la conforman y permiten evaluar las interacciones que se establecen entre la proteína y los elementos de su entorno, tales como solvente, iones, ligandos, otras proteínas, etc. De esta manera, las simulaciones sirven como un microscopio computacional, que permite determinar mecanismos moleculares a escalas espaciales y temporales que resultarían difícilmente observables experimentalmente (Dror y col., 2012).

Hasta la fecha se han reportado variantes en dos genes asociados al desarrollo de CMD en el DP. Una de ellas se trata de una delección del sitio de empalme de 16 pares de bases en el gen piruvato deshidrogenasa quinasa 4 (PDK4) ubicado en el cromosoma 14 (Meurs y col., 2012). El gen PDK4 codifica para la isoenzima 4 de la piruvato deshidrogenasa quinasa (PDK), una proteína mitocondrial que juega un papel importante en la flexibilidad metabólica, al

permitir que las mitocondrias ajusten el uso de la fuente primaria de combustible en función de la disponibilidad de nutrientes. La mutación en el gen PDK4 genera una reducción de la expresión de la isoenzima, lo que conduciría a un control deficiente del complejo piruvato deshidrogenasa, alejando el metabolismo de la oxidación de ácidos grasos y aumentando la glucólisis. La glucólisis no regulada podría conducir potencialmente a un deterioro del transporte de electrones mitocondriales y predisponer a las células al daño mitocondrial (Meurs y col., 2012). No obstante, y especialmente en vista del pequeño tamaño de la muestra, no se ha podido evaluar adecuadamente la actividad de PDK4, por lo que la importancia de la mutación como factor de riesgo de CMD sigue siendo incierta (Meurs y col., 2012). Asimismo, cabe resaltar los resultados contradictorios de esta variante candidata, ya que no se encontró asociada al desarrollo de CMD en una cohorte de DP de Europa (Owczarek-Lipska y col., 2013), ni en la población de DP estudiada en el marco de esta tesis doctoral (ver capítulo 1).

Por otro lado, se identificó un cambio de un solo nucleótido (C > T) en el gen de la titina (TTN), resultando en un cambio en un aminoácido altamente conservado (de glicina a arginina; Meurs y col., 2019). Dicha sustitución fue hallada en la población de Dobermans estudiada en el marco de esta tesis doctoral y se asoció al desarrollo de CMD en esta raza (ver capítulo 1). El gen TTN codifica la proteína más grande encontrada en los mamíferos, que se expresa tanto en el músculo cardíaco como en el esquelético, y que presenta 2 isoformas en el corazón, N2B y N2BA (LeWinter y col., 2007). Las regiones funcionales de la titina se denominan según sus ubicaciones dentro del sarcómero, incluyendo las regiones del disco Z, banda I, banda A y banda M (Granzier y Labeit 2002; Gerull

y col., 2006). La mutación C>T se identificó en la banda I, dentro de un dominio similar a la inmunoglobulina, en humanos denominado I71 (Bang y col., 2001), que está incluido tanto en la titina del músculo esquelético como en la isoforma N2BA de la titina cardíaca (Herman y col., 2012; Schafer y col., 2017; Labeit y col., 2006). La banda I de la titina ha sido caracterizada experimentalmente como una región excepcionalmente flexible de la titina, con propiedades elásticas derivadas de su arquitectura multidominio compuesta por dominios similares a Ig, PEVK, N2B y otros dominios desordenados (Tskhovrebova y col., 2003). La elasticidad de la banda I de la titina presenta tres regímenes complementarios de extensión: 1) un régimen de elasticidad entrópica suave debido a las regiones desordenadas N2B y PEVK; 2) un régimen de elasticidad suave, llamada elasticidad de estructura terciaria, debido a la flexión y torsión de dominio a dominio, así como a estiramientos menores de dominios; y 3) una elasticidad rígida, llamada elasticidad de estructura secundaria, debido a la ruptura de la estructura de la hélice β de los dominios Ig individuales (Lee y col., 2007). Al ser la titina una proteína sarcomérica involucrada en la contracción cardíaca, se propone que la mutación generaría un cambio en la estructura de la proteína, asociada a una disminución de la tensión activa en las miofibras. Los autores postularon, según los resultados de diferentes softwares que predicen el impacto de la sustitución de un aminoácido en el funcionamiento biológico de la proteína, que la sustitución sería un cambio deletéreo (Meurs y col., 2019). Al comparar la expresión de la titina en animales de control con animales que eran heterocigotos u homocigotos para la variante, no se encontraron diferencias importantes en la expresión de ambas isoformas N2AB/N2B. Sin embargo, la proporción de T2

(producto de degradación de la proteína) fue mayor en los animales portadores de la variante, lo que sugiere que podría haber una mayor degradación de proteínas en los animales portadores (Meurs y col., 2019). Esto demuestra que la fisiopatología de la CMD en relación con la mutación en la titina hasta el momento no se entiende por completo.

Aunque en el presente trabajo sólo se observó asociación entre la presencia de la variante en el gen TTN y el desarrollo de CMD en DP, es fundamental conocer cómo las variantes reportadas podrían generar un cambio en la estructura de las proteínas o en la regulación génica. Por este motivo en este capítulo se desarrollarán los siguientes objetivos:

- Evaluar *in silico* el potencial efecto sobre la regulación génica de la variante del gen PDK4.
- Evaluar *in silico* el potencial efecto sobre la estructura de la proteína de la variante del gen TTN.

Materiales y métodos

Predicción del impacto funcional de la delección en el sitio de empalme del gen PDK4

La interpretación de una predicción de sitio de empalme se basa en la comparación de los puntajes de los sitios de empalme generados utilizando las secuencias salvajes y las secuencias variantes. Para analizar el potencial efecto de

la delección de 16 pb en el sitio de empalme donante 5' del intrón 10 del gen PDK4, se utilizaron dos de las herramientas computacionales de libre acceso con mejor desempeño en términos de sensibilidad y especificidad (Tang y col., 2016). Se considera como valor de corte para la interpretación de una puntuación del sitio de empalme una reducción del 10% en la puntuación de predicción causada por la variante (Tang y col., 2016). Los programas utilizados se describen a continuación:

- Max Entropy Scan (MES) (http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html; Yeo y Burge 2004): tiene en cuenta las dependencias entre las posiciones dentro de las secuencias que se analizan. Utiliza un modelo de Entropía Máxima para predecir sitios de empalme basados en la información de las secuencias.
- Alternative Splice Site Predictor (ASSP) (<http://wangcomputing.com/assp/>; Wang, 2006): utiliza una combinación de modelos de preprocesamiento y redes de retropropagación, utilizando una matriz de peso de posición para el sitio aceptor y un modelo de descomposición de máxima dependencia para el sitio donante.

Alineamiento de la secuencia de titina

En primer lugar se identificó la variante CFA36: g.22,321,955 C>T, que resulta en un cambio del aminoácido glicina (G) a arginina (R) en la secuencia de la proteína titina, siguiendo la ubicación descrita por Meurs (2019) que coincide

con el dominio I71 de la titina humana (p.8898G > R; ENST00000589042.5). Para analizar la secuencia de aminoácidos en el perro y compararla con otras especies se utilizó la herramienta BLASTp (Basic Local Alignment Search Tool; <https://blast.ncbi.nlm.nih.gov/>; Altschul y col., 1990). Con el fin de evaluar el grado de conservación del aminoácido de la proteína nativa se consideraron especies domésticas como gato (*Felis catus*), cerdo (*Sus scrofa*), caballo (*Equus caballus*), vaca (*Bos taurus*), cabra (*Capra hircus*) y oveja (*Ovis aries*), animales frecuentemente utilizados en experimentación como conejo (*Oryctolagus cuniculus*), ratón (*Mus musculus*), rata (*Rattus norvegicus*) y chimpancé (*Pan troglodytes*), y la especie humana (*Homo sapiens*), además del perro (*Canis lupus familiaris*). El programa de alineamiento BLAST realiza una búsqueda basada en las secuencias de aminoácidos de interés, la compara con estructuras presentes en diferentes bases de datos (NCBI nr (non-redundant database), UniProtKB (Universal Protein Resource Knowledgebase), PDB (Protein Data Bank), RefSeq, entre otras) y calcula la importancia estadística de las coincidencias. La base de datos contra la que se desea correr la secuencia de interés debe ser especificada de antemano, en este caso se utilizó NCBI nr.

Por otro lado, debido a que el rendimiento y la precisión de las herramientas de predicción pueden verse afectados por la longitud de la secuencia y la disponibilidad de secuencias homólogas para realizar comparaciones, se trabajó con un solo dominio de la proteína. Al ser la titina la proteína más grande conocida en mamíferos (34350 aminoácidos en humanos; Tskhovrebova y col., 2003), se utilizó la secuencia del dominio donde se ubicaba la variante CFA36: g.22,321,955 C>T. Como referencia se utilizó una porción de la secuencia de la

proteína de perros (Q7YRF5_CANLF) disponible en UniProtKB (<https://www.uniprot.org/>; Fig. 1), identificando el dominio tipo inmunoglobulina que abarca del aminoácido 1643 al 1734 (dominio I1643-1734).

Figura 1: Secuencia del dominio I1643-1734, identificado en una porción de la proteína cardíaca titina canina (Q7YRF5_CANLF) disponible en la base de datos UniProtKB.



Sequence: PAAIVEKPEsirVTTGDTCTLECTVTGTPELSTKwFKDgKELTSDSKYKISFFNKVSGLKIINVAPSDSGVYSFEVQNPVgKDSCTASVQVS

A partir del dominio I1643-1734, una vez identificada la glicina sobre la que se produce la mutación (G1712R), y para analizar el impacto de la diferencia, se procedió a comparar la secuencia con la de la proteína mutada: PAAIVEKPEsirVTTGDTCTLECTVTGTPELSTKwFKDgKELTSDSKYKISFFNKVSGLKIINVAPSDSRVYSFEVQNPVgKDSCTASVQVS

Predicción del impacto funcional de la sustitución aminoacídica no sinónima en la proteína titina

Para analizar el potencial efecto de la nsSNV sobre la estructura, función y estabilidad de la proteína codificada por TTN, se utilizaron múltiples herramientas de libre acceso que realizan enfoques basados en la secuencia y la estructura de la proteína nativa. Cada uno de los programas tiene sus propias fortalezas y debilidades dependiendo del algoritmo, y en muchos casos, la calidad del resultado puede variar según la secuencia del gen y la proteína, motivo por el cual es ideal utilizar más de uno. A continuación se describen las herramientas utilizadas:

- SIFT: Sorting Intolerant From Tolerant (<https://sift.bii.a-star.edu.sg/>; Ng y Henikoff, 2003): Una herramienta que predice el potencial deletéreo de las variantes en proteínas basándose en la conservación de aminoácidos en la secuencia utilizando un enfoque evolutivo. A partir de la secuencia problema, SIFT realiza una búsqueda de secuencias homólogas para realizar un alineamiento múltiple. A continuación calcula la probabilidad normalizada para todas las posibles sustituciones de cada posición de dicho alineamiento. El programa asigna un valor (score) para clasificar los cambios observados en cada columna del alineamiento: dañinos (score 0 - 0,05), potencialmente dañinos (score 0,05 - 0,10) y neutrales (score > 0,10).
- PolyPhen-2: Polymorphism Phenotyping v2 (<http://genetics.bwh.harvard.edu/pph2/index.shtml>; Adzhubei y col.,

2010): Se trata de algoritmo que predice el impacto funcional de las variantes en proteínas utilizando características estructurales, evolutivas y bioquímicas. Además de realizar un alineamiento de secuencias homólogas como SIFT, PolyPhen 2 incorpora una serie de características y anotaciones tomadas de distintas bases de datos (como Protein Data Base (PDB)) para analizar con mayor profundidad el potencial efecto de las sustituciones. Finalmente, toda esta información es evaluada por un clasificador entrenado por métodos de aprendizaje computacional (machine learning). Las variantes se clasificaron como probablemente dañinas ($\geq 0,5$) posiblemente dañinas ($0,5 - 0,85$) y benignas ($0 - 0,15$).

- PANTHER-PSEP - Position-Specific Evolutionary Preservation - cSNP Scoring (<http://www.pantherdb.org/tools/csnpscoreForm.jsp>; Thomas y col., 2022): se utilizó para profundizar el análisis del efecto de mutaciones en la función de una proteína desde un punto de vista evolutivo. Este método se basa en la generación de árboles filogenéticos, alineamientos y la reconstrucción de secuencias ancestrales, para evaluar la preservación de un residuo específico a lo largo del tiempo. Las mutaciones se clasificaron como probablemente dañinas cuando el tiempo de preservación de un residuo era mayor a los 450 millones de años, posiblemente dañino cuando estaba entre los 200 y 450 millones de años, y probablemente benigno si era menor a los 200 millones de años.
- MutPred 2 (<http://mutpred.mutdb.org/>; Pejaver y col., 2020): se utilizó para calcular la probabilidad que tiene una sustitución aminoacídica de ser patogénica (scores mayores a 0,8) o no patogénica (scores menores a 0,8),

utilizando una metodología de machine learning que integra información genética y molecular, en función a un entrenamiento previo con sets de datos conocidos.

Modelado estructural del dominio I1643-1734 de titina

El Protein Data Bank (PDB) es una base de datos en la cual se reporta información estructural de biomoléculas obtenida mediante distintas técnicas experimentales, tales como cristalografía de rayos X, resonancia magnética nuclear y electro-crio-microscopía. Dado que en ella no se encuentra publicada la estructura de la proteína titina canina ni la del dominio de interés para este trabajo I1643-1734, resulta necesario generar un modelo estructural para su estudio. Por esta razón, la variante nativa del dominio I1634-1734 fue modelada utilizando AlphaFold2 (Jumper y col., 2021), un predictor basado en aprendizaje automático que modela la estructura tridimensional de la proteína a partir de su secuencia. Para lograr esto hace uso de información proveniente tanto de alineamientos múltiples de secuencias como de la información estructural depositada en el PDB. Así, mediante procedimientos de entrenamiento basados en restricciones evolutivas, físicas y geométricas de las estructuras de las proteínas, proporciona modelos cuya resolución tiene una precisión en la escala de Angstroms y son comparables con los resultados que se obtienen mediante experimentos.

Para generar los modelos se utilizó el servidor AlphafoldCollab2 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). Este servidor utiliza como input la secuencia de la proteína que se

desea modelar, y provee cinco modelos como resultados. Adicionalmente, provee los valores de dos métricas utilizadas para evaluar la calidad de los mismos, que son el Per-residue Local Distance Difference Test (pLDDT por sus siglas en inglés) y el Predicted Aligned Error (PAE por sus siglas en inglés). El valor de pLDDT es una medida que proporciona información sobre la confiabilidad de la predicción de la posición de cada residuo en la estructura tridimensional de la proteína. Un valor alto de pLDDT indica una predicción confiable y cercana a la estructura real, mientras que un valor bajo sugiere una mayor incertidumbre en la predicción. Los valores de rango típicos para pLDDT en AlphaFold2 son:

- 90-100: Región de alta confianza, indicando una predicción estructural muy precisa y cercana a la estructura real.
- 70-89: Rango intermedio, las predicciones siguen siendo confiables pero pueden tener algunas desviaciones menores.
- 50-69: Valores bajos, indicando predicciones menos confiables y con desviaciones significativas en la estructura.
- Menos de 50: Región de baja confianza, donde las predicciones son menos precisas y pueden estar considerablemente desviadas de la estructura real.

Por otra parte, el parámetro PAE está relacionado con la capacidad de AlphaFold para predecir correctamente la posición de los dominios de una proteína. Un bajo PAE refleja una predicción más precisa y confiable de la disposición de los dominios, mientras que un alto PAE podría sugerir cierta inexactitud en estas predicciones.

Estudios basados en dinámica molecular

Las simulaciones de Dinámica Molecular clásica se basan en la resolución de las ecuaciones de movimiento de Newton con el objetivo de obtener la trayectoria de un conjunto de partículas que interactúan entre sí. En simulaciones de biomoléculas, estas partículas pueden ser los átomos o grupos de átomos que las constituyen, como así también iones y moléculas de su entorno. Así, las partículas que conforman el sistema son sometidas a la acción de un campo de fuerzas. Un campo de fuerzas está determinado por las ecuaciones que se utilizan para modelar las diferentes contribuciones a la energía potencial de un sistema y todos los parámetros asociados para realizar este cálculo. Este potencial alberga toda la información esencial para describir la energía potencial asociada al movimientos de los átomos (estiramientos de enlace, flexiones de ángulos, rotaciones) como así también a las interacciones no ligadas (coulómbicas y de Van der Wals) y su elección depende del tipo de sistema a simular, ya que existen campos de fuerza específicos para distinto tipo de biomoléculas además de proteínas, tales como ADN, ARN y lípidos. En todos los casos, los parámetros que emplean para la representación de interacciones se obtienen a partir de datos experimentales y/o de cálculos mecano-cuánticos (Nerenberg y Head-Gordon, 2018).

En la simulación, a partir de la derivada del campo de fuerzas en función de las coordenadas del sistema, se calcula la fuerza a la que es sometida cada partícula del sistema partiendo de su posición inicial y se actualiza la posición, velocidad y aceleración de las mismas. Esta integración de las ecuaciones de

movimiento es repetida iterativamente. A medida que el sistema evoluciona, se registran las posiciones y velocidades de cada paso, que normalmente tiene una duración del orden de los femtosegundos. La extensión de la simulación dependerá del tipo de proceso o movimiento en estudio, pudiendo alcanzar el orden de los milisegundos para eventos como plegamiento de proteínas o cambios conformacionales a nivel global (Hospital y col., 2015).

Desde el punto de vista práctico, las simulaciones de proteínas constan de una serie de etapas que son minimización, calentamiento, equilibración y producción. La minimización tiene por objetivo hacer que el sistema inicial evolucione a un mínimo de energía local, ya que puede haber distancias de enlace alejadas de los valores de equilibrio o solapamiento de átomos que surgen al generar un entorno de solvente. A continuación, se realizan el calentamiento y la equilibración. En estas etapas la presión, temperatura y densidad del sistema adquieren los valores adecuados para la producción. Finalmente, durante la producción se recolectan los valores de posición y velocidad cada cierto número de pasos de simulación. De esta manera, se obtiene una trayectoria que posteriormente podrá ser visualizada y procesada para el estudio del comportamiento dinámico del sistema.

En este trabajo se realizaron simulaciones exploratorias para conocer el comportamiento del dominio I1643-1734, tanto de la titina nativa como de su variante G1712R. Las mismas fueron realizadas con el paquete de software de dinámica molecular AMBER (*Assisted Model Building with Energy Refinement*; Case y col., 2021). Los sistemas que constituyen el punto de partida de las simulaciones fueron generados empleando el plugin Leap de AMBER. En el caso

de la variante nativa, se utilizó uno de los modelos obtenidos a partir de AlphaFold, mientras que la estructura de la variante no sinónima fue generada mediante edición manual del residuo G1712R y posterior procesamiento con LeaP. Cada estructura fue neutralizada y solvatada en una caja de moléculas de agua modelo TIP3P. A continuación, cada sistema fue minimizado, calentado a 311 K y equilibrado adaptando el protocolo descrito por Racigh y col. (2021). El campo de fuerzas empleado fue amber ff14SB. A partir de cada estructura equilibrada, se realizaron tres producciones en paralelo de 100 ns cada una muestreando las posiciones cada 2 fs y se obtuvieron en total 3000 confórmeros para cada una.

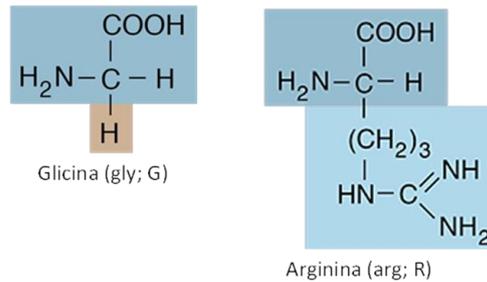
Las trayectorias fueron visualizadas empleando el software VMD (<http://www.ks.uiuc.edu/Research/vmd/>; Humphrey y col., 1996). Se empleó el software cpptraj de AMBERTools21 (Case y col., 2021) para calcular el RMSD (desviación cuadrática media; *RMSD* por las siglas en inglés *Root Mean Square Deviation*) de los carbonos del esqueleto (*backbone*) de la proteína (C, O, CA y N) durante la trayectoria, a fin de corroborar la estabilidad de la estructura secundaria del dominio en sus formas nativa y mutada. Utilizando VMD, se examinaron las regiones específicas de la proteína donde ocurrió la sustitución para identificar cualquier cambio en la conformación local como diferencias en la orientación de residuos individuales, cambios en la estructura secundaria (hélices alfa, láminas beta, etc.) o alteraciones en bucles y enlaces disulfuro. De la misma manera, se analizaron las interacciones que surgen entre el residuo en la posición 1712 en cada caso y los aminoácidos de su entorno, para evaluar nuevas interacciones que puedan surgir como producto del reemplazo de glicina por

arginina en dicha posición. Además, se calculó el área superficial y el volumen de cada estructura a fin de compararlas. Los gráficos correspondientes a los distintos parámetros determinados se realizaron con el programa Gnuplot 5.4 (<http://www.gnuplot.info/>).

A continuación con el software PropKa 2.0 (<https://www.ddl.unimi.it/vegaol/propka.htm>; Hui y col., 2005) se realizó el cálculo del pKa (potencial de disociación ácida) que proporciona información sobre la acidez o basicidad de los grupos funcionales presentes en la proteína y cómo estos grupos pueden estar protonados o desprotonados a diferentes valores de pH. Esto permite predecir cómo cambios en la secuencia proteica pueden afectar la estabilidad y las interacciones dentro de una proteína. En este caso este aspecto resulta relevante ya que la variante G1712R implica la sustitución de glicina por arginina, y la cadena lateral de la arginina puede participar en procesos ácido-base debido a su naturaleza química (Fig. 2). A su vez, el programa realiza el cálculo de pkmodel, un valor de ajuste que tiene en cuenta el entorno local del residuo en la proteína.

Por último, empleando el programa FoldX (<https://foldxsuite.crg.eu/>; Schymkowitz y col., 2005) y la totalidad de los conformeros obtenidos para cada estructura se calculó la diferencia de los valores de δG asociados al desplegamiento de cada estructura ($d\delta G$). Al haber un único aminoácido diferente entre los dos modelos, la magnitud y el signo de este valor reflejan la contribución a la energía de la estructura producto de las nuevas interacciones que pueden tener lugar a causa de la mutación G1712R. El resultado de este análisis será contextualizado en base a las nuevas interacciones observadas.

Figura 2: Estructuras esquemáticas del aminoácido original (glicina) a la izquierda, y el resultante de la sustitución (arginina) a la derecha. El esqueleto, que es el mismo para cada aminoácido, se encuentra dentro del recuadro azul. La cadena lateral, única para cada aminoácido, se encuentra dentro del recuadro rosa (glicina) y celeste (arginina).



Resultados

Predicción del impacto funcional de la delección en el sitio de empalme del gen PDK4

La delección de 16 pares de bases eliminó el sitio donante GT 5', además de 14 bases adicionales hasta el siguiente GT en el intrón, creando un posible sitio de empalme críptico. El sitio GT críptico obtuvo un puntaje Max Ent más bajo (6,33) que el sitio de empalme constitutivo (8,33). De la misma manera, al analizar los puntajes de los modelos de procesamiento con ASSP, la fuerza del sitio de empalme del GT críptico fue menor (9,89) que la del sitio constitutivo (12,29), con una confianza de 0,9 y 0,92, respectivamente. La confianza varía

entre 0 (indeciso) y 1 (clasificación perfecta). Ambos programas arrojaron una diferencia mayor al 10 % en la puntuación de predicción.

Alineamiento de la secuencia de titina

En el presente trabajo se utilizó el dominio I1643-1734 de TTN de perro, que coincide con el dominio I71 humano (Meurs, 2019), donde se localiza la variante en estudio. El alineamiento, realizado con la herramienta BLASTp, permitió confirmar que dentro del alto grado de identidad global entre las secuencias comparadas, también la glicina está conservada (Tabla 1; Fig. 3).

Tabla 1: Resultados de BLASTp el valor de los parámetros obtenidos por la búsqueda de similitud entre el dominio I1643-1734 del perro y otras regiones de titinas de diferentes especies.

Descripción	Nombre científico	Puntaje máximo ^a	Cobertura ^b	E value ^c	% Identidad ^d
cardiac titin	Canis lupus familiaris	187	100%	3e-55	100.00%
titin isoforma X29	Felis catus	187	100%	7e-55	98.91%
titin isoform X46	Oryctolagus cuniculus	187	100%	7e-55	98.91%
LOW QUALITY PROTEIN: titin	Sus scrofa	187	100%	7e-55	98.91%
titin isoform X6	Equus caballus	182	100%	2e-53	96.74%
titin isoform X1	Bos taurus	181	100%	7e-53	95.65%
titin isoform N2BA	Homo sapiens	180	100%	1e-52	94.57%
titin isoform X45	Pan troglodytes	180	100%	1e-53	94.57%
PREDICTED: titin isoform X1	Capra hircus	176	100%	3e-51	92.39%
titin isoform X6	Mus musculus	176	100%	4e-51	91.30%
titin isoform X1	Ovis aries	176	100%	5e-51	91.30%
titin isoform X47	Rattus norvegicus	173	100%	3e-50	90.22%

a) puntuación de similitud más alta entre la porción de secuencia ingresada y las de la base de datos; b) porcentaje de la secuencia en cuestión que se alinea con las secuencias similares de la base de datos; c) medida estadística que estima la probabilidad de obtener un alineamiento con un puntaje igual o mejor simplemente por azar (cuanto más chico el valor mayor significancia de similitud); d) porcentaje de similitud entre la secuencia en cuestión con la totalidad de las secuencias de la base de datos.

Cabe aclarar que esta comparación de secuencias realizada entre especies de mamíferos va a mostrar una alta similitud de secuencia meramente por la distancia evolutiva corta entre las especies. Por un lado, esta alta similitud permitió resaltar el efecto del cambio fisicoquímico notorio al reemplazar glicina por arginina, pero al mismo tiempo implicó tener que completar la evaluación del impacto de esta variante en la estructura mediante dinámica molecular, como se detallará más adelante.

Figura 3: Comparación entre la secuencia de aminoácidos del dominio I1643-1734 de TTN nativa y con la variante (var.) estudiada en el presente estudio.

Perro (var.)	PAAIVEKPEsirVTTGDTCTLECTVTGTPELSTKWFKDGKELTSDSKYKISFFNKVSGGLKIINVAPSDS	RVYSFEVQNPVGKDSCTASVQVS
Perro	PAAIVEKPEsirVTTGDTCTLECTVTGTPELSTKWFKDGKELTSDSKYKISFFNKVSGGLKIINVAPSDS	GVYSFEVQNPVGKDSCTASVQVS
Gato	Q.....G.....
Conejo	T.....G.....
Cerdo	T.....G.....
Caballo	Q.....G.....G.....R.....
Vaca	T.....S.....T.....N.....G.....
Humano	..T.....K.....A.....N.....	G.....L.....
Chimpancé	..T.....K.....A.....N.....	G.....L.....
CabraG.....K.....L.....T.....R.....N.....	A.....G.....
RatónK.....S.....N.....	V.....G.....G.....KV.....I.....
OvejaG.....K.....A.....S.....T.....R.....N.....	A.....G.....
Rata	..T.....K.....M.....S.....G.....	S.....G.....G.....V.....I.....

*Los aminoácidos en rojo representan cambios en los que a pesar de ser diferentes, tienen alguna o varias propiedades fisicoquímicas similares (carga, polaridad, hidrofobicidad, volumen de cadenas laterales) y pueden tener una función o estructura conservada. Los identificados en verde indican que el cambio de aminoácidos genera una diferencia significativa entre los residuos en esas posiciones.

Predicción del impacto funcional de la sustitución aminoacídica no sinónima en la proteína titina

Todos los programas de predicción utilizados en este estudio obtuvieron resultados similares. SIFT le asignó a la variante un valor de “dañina” PolyPhen-2 y PANTHER la clasificaron como “probablemente dañina” MutPred le otorgó un valor de “patogénica” (Tabla 2).

Tabla 2: Resultados del análisis individual de la nsSNV CFA36: g.22,321,955 C>T en el gen TTN con las herramientas *in silico*: SIFT, PolyPhen-2, PANTHER y MutPred 2.

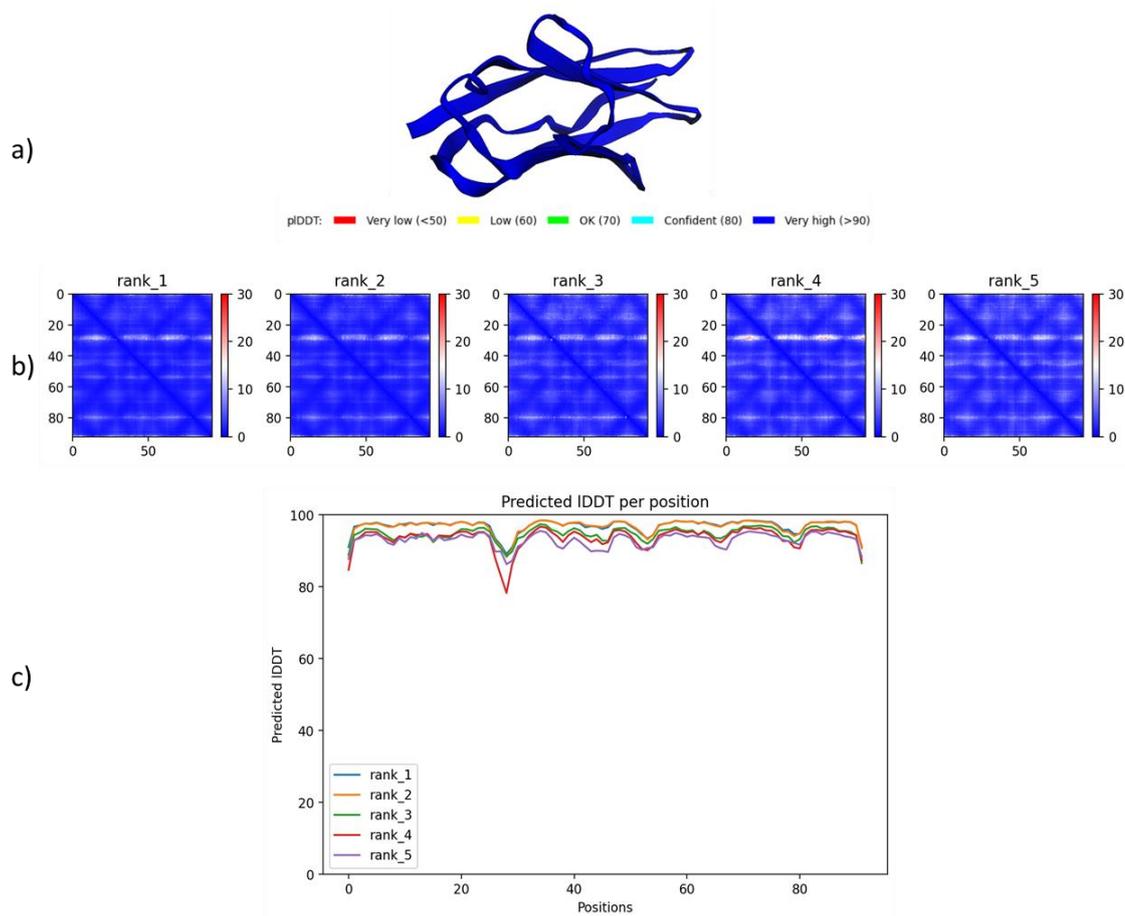
Posición		g.22,321,955
Residuo en la referencia		C
Residuo alternativo		T
Herramientas	SIFT	0,00*
	PolyPhen-2	1*
	PANTHER (millones de años)	797*
	MutPred 2	0.854*

*Indica los resultados significativos de cada análisis. a) p-value < 0.05: dañino, 0.05-0.10: potencialmente dañino, > 0.10: neutral. b) probabilidad > 0.85: probablemente dañino, 0.15-0.85: posiblemente dañino, < 0.15: benigno. c) tiempo de preservación > 450: probablemente dañino, 200-450: posiblemente dañino, < 200: probablemente benigno. d) probabilidad > 0.8: patogénico, < 0.8: no patogénico.

Modelado estructural del dominio I1643-1734 de titina

Utilizando el predictor de estructuras Alphafold, a partir de la secuencia del dominio I1643-1734 se obtuvieron cinco modelos estructurales. La superposición de estas estructuras muestra que son prácticamente idénticas. Considerando el ranking provisto por Alphafold que asigna un valor de 1 a 5 a los modelos de acuerdo a los valores de PAE (Predicted Aligned Error) y pLDDT (Predicted Local Difference Test) descritos anteriormente, se optó por utilizar la estructura que ocupa la primera posición de este ranking (Fig. 4). Los valores bajos de PAE obtenidos para cada modelo generado indican que las predicciones son altamente precisas y confiables. Asimismo, al considerar los valores de pLDDT se observa que Alphafold2 asigna valores de alta confianza a todas las posiciones de la secuencia de los 5 modelos generados.

Figura 4: Modelos estructurales predichos por Alphafold. a) Estructura terciaria del dominio I1643-1734; b) Gráfico de PAE; c) Valores de pLDDT de la secuencia para cada modelo generado.



Estudios basados en dinámica molecular

La comparación visual de las trayectorias de la variante salvaje y de G1712R obtenidas a partir de las simulaciones permitió descartar la ocurrencia de cambios en la estructura secundaria de láminas beta del dominio I1643-1734 a causa de la sustitución (Fig. 5). Asimismo, al observar los alineamientos

tridimensionales se pudo apreciar una alta coincidencia estructural, es decir, se presenta superposición estructural entre ambas proteínas (Fig. 6). Para corroborar cuantitativamente esta observación, en cada caso se analizó la movilidad de la estructura durante las simulaciones a través de cálculo del RMSD. La desviación cuadrática media se calcula para medir la magnitud de las diferencias entre las posiciones de los átomos estructuralmente equivalentes en dos estructuras. Un RMSD bajo indica una superposición cercana y similitud estructural entre ambas formas, mientras que un RMSD más alto sugiere mayores diferencias conformacionales. En este caso, se calculó para los átomos del *backbone* de la forma salvaje y de la variante, utilizando como referencia las posiciones atómicas de la estructura inicial de las dinámicas. Los valores de RMSD obtenidos para la forma salvaje oscilan entre 0,6 y 1,6 Å con un valor promedio de 0,8 Å, mientras que los de la variante G1712R comprenden un rango entre 0,6 y 1,7 Å siendo el promedio 0,79 Å. En la Fig. 7 se muestra a modo de ejemplo el valor de RMSD calculado para una de las trayectorias generada a partir de cada estructura. La amplitud de los rangos de RMSD indica que en ambos casos las conformaciones exploradas por la estructura durante las simulaciones no se desvían significativamente de la conformación inicial de cada modelo, reflejando la estabilidad de la estructura secundaria del dominio I1643-1734.

Figura 5: Estructura secundaria del dominio I1643-1734 de la proteína nativa (a) y de la proteína con la mutación no-sinónima G1712R (b). En violeta podemos identificar el residuo 1712 (glicina vs arginina).

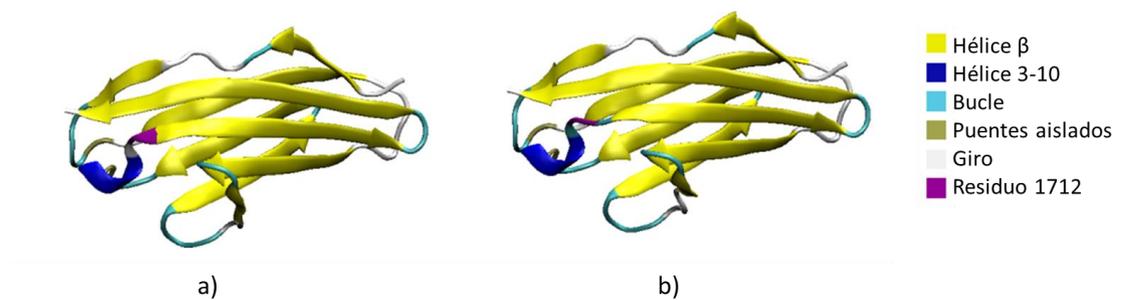


Figura 6: Alineamiento del dominio I1643-1734 de la proteína nativa (azul) y de la proteína con la mutación no-sinónima G1712R (verde).

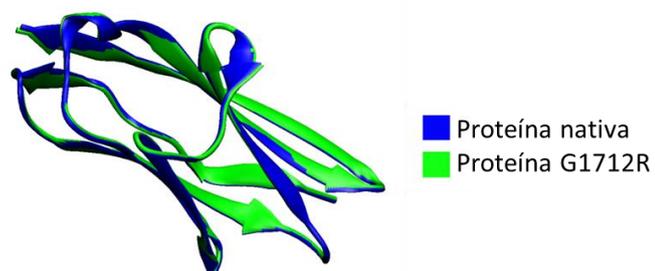
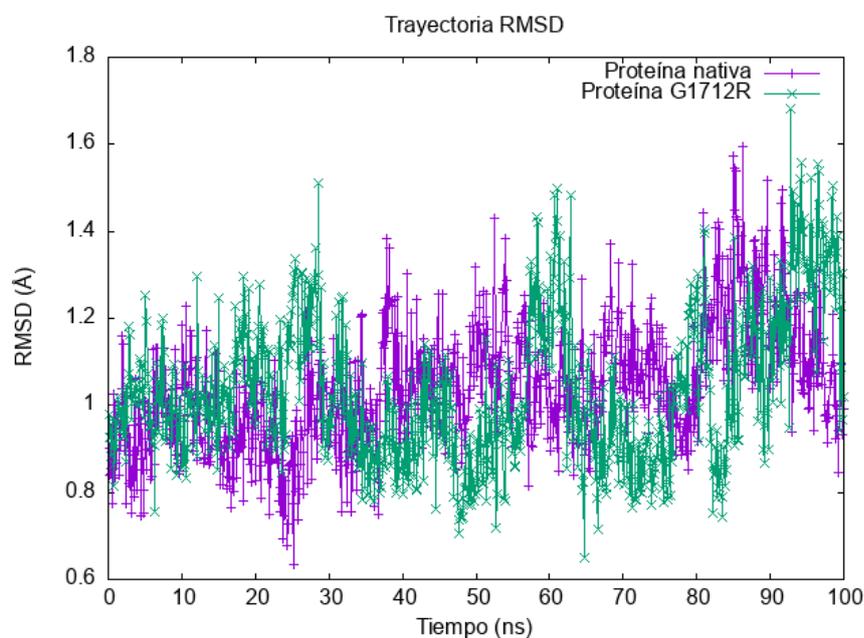


Figura 7: Gráfico de la evolución de la desviación cuadrática media (RMSD) de las proteínas salvaje y G1712R.



A partir de los conformeros obtenidos en las simulaciones de las formas salvaje y con la sustitución, se identificaron los aminoácidos que se encuentran en un radio menor a 3,5 Å de la posición 1.712 y que por lo tanto pueden interactuar no covalentemente con el residuo en esta posición (Fig. 8). En la Tabla 3 se listan los residuos en contacto identificados en cada caso. El análisis reveló que tanto para la forma salvaje como para la variante, prevalecen los puentes de hidrógeno que forman el backbone de la proteína durante las simulaciones y además son los mismos en ambos casos (Fig. 9; a, b y c). Por un lado, esto resulta esperable ya que la sustitución de glicina por arginina sólo modifica la composición de la cadena lateral de la posición 1.712, y por el otro, este resultado concuerda con la evaluación de RMSD que indica que la estructura

secundaria de lámina beta es estable y se conserva durante las simulaciones. Las estructuras secundarias, como se sabe, están estabilizadas por puentes de hidrógeno del *backbone*. Además, en la tabla se indica que la arginina interacciona con cuatro aminoácidos adicionales: lisina 1.679, prolina 1.708, serina 1.709 y tirosina 1.714. Con respecto a la lisina 1.679 y la tirosina 1.714, se observó la ocurrencia de interacciones de tipo hidrofóbico entre las cadenas laterales de estos residuos y la cadena carbonada de la arginina, la que no está presente en la glicina y es por esto que en la forma salvaje no se observan dichas interacciones. Por otra parte, en el caso de la prolina 1.708 y la serina 1.709, las interacciones que tienen lugar son de tipo puente de hidrógeno y, a diferencia de la glicina, son posibles debido a la longitud de la cadena lateral de este residuo que además contiene átomos de hidrógeno unidos a nitrógeno en su grupo guanidinio. El procesamiento de las simulaciones indica que en ambas formas (salvaje y variante) el aspartato 1.680 se encuentra dentro del radio de corte establecido para el análisis de interacciones. Sin embargo, en la forma salvaje se establecen interacciones dipolo-dipolo con G1712, mientras que en la estructura de la variante se establecen interacciones puente de hidrógeno con los hidrógenos del grupo guanidinio de la arginina 1.712 y de tipo carga-carga ya que el aspartato posee dos cargas negativas y el guanidinio de la arginina una carga positiva. Estas nuevas interacciones que emergen entre los residuos producto de la sustitución resultan particularmente relevantes ya que involucran aminoácidos que no se encuentran cercanos en la secuencia, sino que se encuentran próximos desde el punto de vista estructural, a diferencia de la prolina 1.708 y la serina 1.709 que son prácticamente adyacentes (Fig. 10, a y b). Debido a que la ocurrencia de

interacciones puente de hidrógeno depende fuertemente de la orientación de las cadenas laterales de los residuos, se analizó esta interacción en los 3.000 conformeros de la variante y se observó que es dinámica, ya que entre los átomos de oxígeno de la cadena lateral del aspartato y los hidrógenos del guanidinio (Fig. 10c) el número de puentes de hidrógeno puede oscilar entre 0 y 3, con un valor promedio de 1.

Figura 8: Imagen obtenida por VMD en la que se representan los aminoácidos que se encuentran en un radio menor a 3,5 Å de la posición 1.712 del a) residuo salvaje (glicina) y b) del residuo variante (arginina), ambos graficados en amarillo.

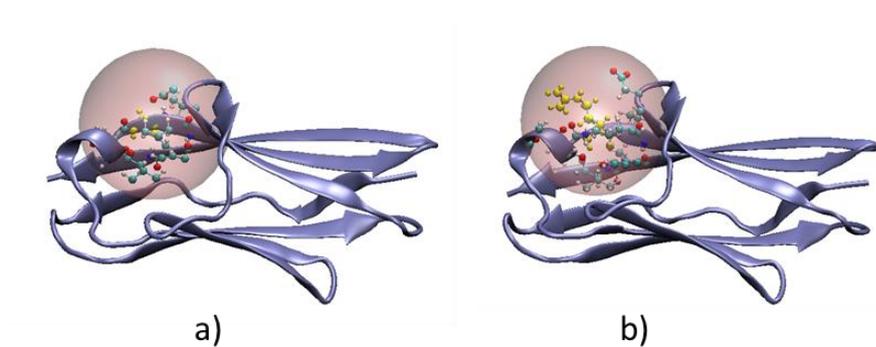
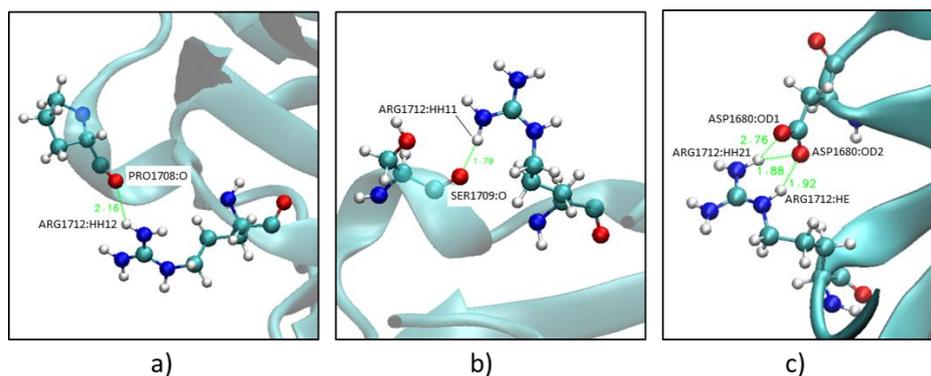


Tabla 3: Residuos que se encuentran en un radio menor a 3,5 Å de la posición 1.712 y tipo de interacción con este aminoácido en caso de que existiera.

Residuo	Contactos Glicina 1712		Contactos Arginina 1712	
	<3,5 Å	Interacción	<3,5 Å	Interacción
1679: Lisina	X	-	X	Hidrofóbico
1680: Aspartato	X	-	X	Puente de hidrógeno
1708: Prolina	-	-	X	Puente de hidrógeno
1709: Serina	-	-	X	Puente de hidrógeno
1710: Aspartato	X	Dipolo-dipolo	X	Puente de hidrógeno/Carga-carga
1711: Serina	X	-	X	-
1713: Valina	X	-	X	-
1714: Tirosina	X	-	X	Dipolo-dipolo/Hidrofóbico
1730: Serina	X	Puente de hidrógeno	X	Puente de hidrógeno
1731: Valina	X	Puente de hidrógeno	X	Puente de hidrógeno

X = indica presencia; - = indica ausencia.

Figura 10: Imágenes obtenidas por VMD que representan las interacciones propias de la proteína variante G1712R. a) interacción con la prolina 1.708; b) interacción con la serina 1.709; c) interacción con el aspartato 1.680. En verde se identifican las distancias (Å) de las interacciones.

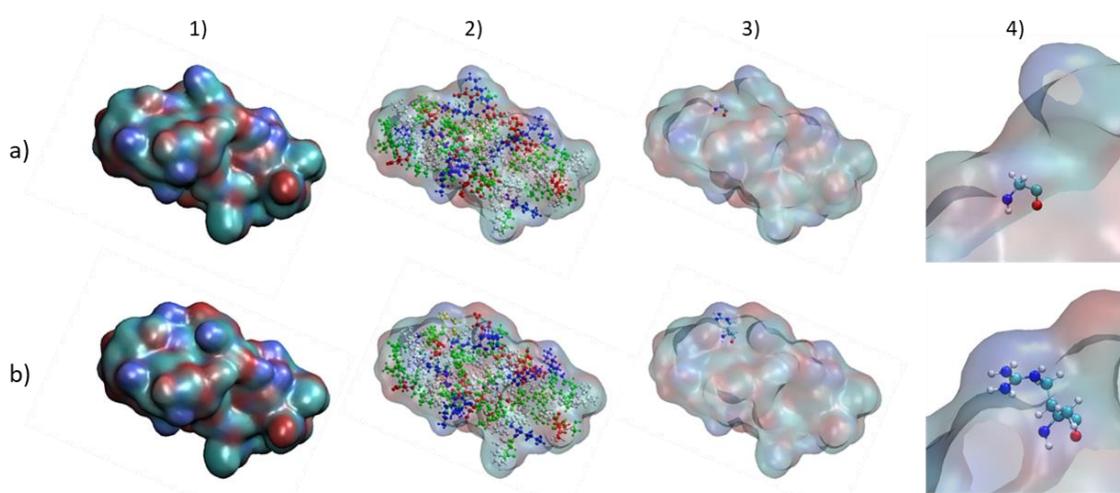


El análisis del efecto de la mutación sobre la estructura de la proteína, a partir de la determinación de los parámetros del área superficial y del volumen no indica un cambio significativo (Tabla 4). Tanto el área superficial como el volumen de los modelos equilibrados aumentaron en el caso de la variante G1712R, pero en un valor que representa tan solo el 1% del valor total (Fig. 11).

Tabla 4: Características estructurales (área superficial y volumen) del dominio I1643-1734 de la proteína nativa y para la variante G1712R.

Dominio I1643-1734	Área superficial	Volumen (Å ³)
Proteína nativa	6.682,99	11.485,66
Proteína variante G1712R	6.755,64	11.602,56

Figura 11: Área superficial del dominio I1643-1734 de la proteína nativa (a) y del correspondiente a la proteína variante G1712R (b). 1) representa el área superficial del dominio, 2) incluye la distribución de los residuos en su interior, 3) se identifica el residuo 1712 en la proteína, 4) zoom al residuo 1712 glicina (4a) y arginina (4b).



En el análisis de pKa de la arginina 1.712 de la variante G1712R se obtuvo un valor de 12,22 para este residuo que coincide con el valor reportado en la literatura (Nelson y Cox, 2018). Además, el valor de pkmodel (12,50) fue ligeramente mayor que el pKa, lo que sugiere que los aminoácidos del entorno local de la arginina en la estructura favorecen a la forma protonada de este residuo, en comparación a su forma en solución que es a la que refiere el valor de literatura. A su vez, el programa indica que hay dos interacciones que podrían ser las responsables de esta estabilización adicional, que son las correspondientes a los residuos aspartato 1.680 y tirosina 1.714 descritas en el apartado anterior. Los

valores de energía asociados a estas interacciones son -0,08 Kcal/mol y -0,92 Kcal/mol, respectivamente.

Por último, el cálculo mediante FoldX de la energía de estabilidad total entre ambas estructuras no mostró diferencias significativas ($p > 0,05$). Sin embargo, al analizar los diferentes tipos de energía que componen la energía de estabilidad total, se identificaron cambios significativos en la energía de enlace de los puentes de hidrógeno de las cadenas laterales ($p < 0,05$), y en la energía de interacción electrostática ($p < 0,01$) a favor de la proteína variante. La primera es una medida asociada a la formación y estabilidad de los puentes de hidrógeno entre las cadenas laterales de los aminoácidos. Los puentes de hidrógeno son interacciones débiles pero cruciales en la estabilización de la estructura tridimensional de las proteínas. Los resultados obtenidos para las variantes nativa y G1712R (valor1 Kcal/mol versus valor2 Kcal/mol) indican un aumento en la estabilidad de la estructura del dominio a raíz de los puentes de hidrógeno que surgen al sustituir el átomo de hidrógeno de la glicina por la cadena lateral de la arginina que incluye un grupo guanidinio. Esto puede interpretarse en el contexto del análisis de las interacciones realizado, donde se encontró que la arginina 1.712 forma puentes de hidrógeno adicionales con la prolina 1.708, la serina 1.709, y con el residuo aspartato 1.680 con el cual puede formar hasta tres puentes de hidrógeno. Por otro lado, al sustituir glicina por arginina se introdujo una carga positiva en el dominio. Este cambio resultó en un valor de energía de interacción electrostática favorable para la estabilidad estructural de la variante G1712R con respecto a la nativa (valor1 Kcal/mol versus valor2 Kcal/mol).

El programa FoldX calcula esta contribución a la energía total en base a las interacciones entre cadenas laterales cargadas eléctricamente y a las interacciones entre dipolos presentes en la proteína. Entre las interacciones identificadas entre la arginina 1.712 y los residuos de su entorno se encuentran las de tipo dipolo-dipolo con el residuo tirosina 1.714 y las interacciones entre cargas con el aspartato 1.680. En cambio, en la variante nativa la glicina sólo establece una interacción dipolo-dipolo con el aspartato 1.680, que es más débil que la interacción carga-carga que establece con la arginina cuando se encuentra en la misma posición. Considerando que las interacciones electrostáticas entre cargas opuestas son atractivas y estabilizadoras, esto explica la diferencia entre valores observada entre ambas variantes.

Discusión y conclusiones

En este capítulo se describieron los métodos empleados con el objetivo de comprender el impacto de las variantes reportadas asociadas al desarrollo de CMD en la regulación génica o en la estructura de las proteínas, en caninos. En primer lugar se evaluó la delección en el sitio de empalme donante del gen PDK4, que parece resultar en el uso de un sitio gt crítico. Las fortalezas tanto del sitio de empalme 5' original como del sitio crítico se compararon con MES y ASSP. Ambos programas arrojaron como resultado que el sitio de empalme original tenía una puntuación más alta que la asignada al sitio crítico, con una diferencia mayor al 10 % en ambas predicciones. Este resultado concuerda con los hallazgos de

Meurs y col. (2012) en los que realizaron predicción con MES. La mayor fuerza del sitio de empalme constitutivo podría sugerir una preferencia por el sitio original, por lo que aunque es probable que el sitio crítico pueda ser utilizado, se usaría con una preferencia más baja y posiblemente con una eficiencia disminuida (Eng y col., 2004). La presencia de la variante del gen PDK4 produciría una proteína con reducción de su actividad, una menor capacidad para fosforilar y por ende, regular negativamente la oxidación de la glucosa (Lopaschuk y col., 2010).

En segundo lugar se analizó el efecto de la variante G1712R en la estructura y función del dominio I1643-1734 de la proteína titina de *canis lupus familiaris*, y los resultados a los que se arribó con este estudio. Si bien existen numerosos reportes acerca de las alteraciones producidas por diferentes mutaciones en la titina de distintas especies, hasta donde sabemos, este es el primer estudio acerca de los cambios estructurales y funcionales del dominio I1643-1734 de la titina canina, estudiada mediante herramientas de análisis estructurales y simulaciones de dinámica molecular.

El gen TTN codifica para la proteína más grande encontrada en los mamíferos, la titina, que se expresa entre otros lugares, en el músculo cardíaco. La mutación (C > T) descrita por Meurs y col. (2019) en DP, y que fue asociado al desarrollo de CMD, produce la sustitución del aminoácido glicina por arginina, lo que podría tener efectos importantes (estructurales o funcionales) debido a las diferencias en las propiedades químicas de estos dos aminoácidos (Meurs y col., 2019). Sumado a que cada aminoácido tiene su propio tamaño, carga y valor de hidrofobicidad específico, la glicina en esta posición de la titina se encuentra altamente conservada entre diferentes especies.

Los algoritmos de predicción del impacto funcional de la mutación G1712R arrojaron como resultado que la variante sería un cambio perjudicial, lo que está en línea con lo reportado por Meurs y col. (2019). Estos programas *in silicos*, generalmente evalúan el impacto de mutaciones en la función y la estructura, pero no proporcionan detalles específicos sobre la causa del daño predicho (Richards y col., 2015). Para esto es necesario realizar un análisis más profundo y considerar otros factores, como la posición de la mutación en la secuencia de aminoácidos, la estructura tridimensional de la proteína y las simulaciones de dinámica molecular.

La sustitución de una glicina por una arginina en la posición 1.712 implica el reemplazo de un átomo de hidrógeno por una cadena lateral larga y voluminosa que contiene un grupo guanidinio con una carga positiva a pH fisiológico. Para comprender el efecto de esta sustitución en el dominio I1634-1734, se realizó una caracterización estructural y dinámica de las variantes nativa y G1712R. A partir del modelado de estructuras y trayectorias obtenidas mediante simulaciones de dinámica molecular, fue posible calcular el valor de distintos parámetros asociados a la forma salvaje y a la variante y observar sus diferencias de comportamiento dinámico. Por un lado, el cálculo del volumen y área superficial de las estructuras muestra que la sustitución no produce diferencias significativas en relación a estos parámetros, ya que la diferencia es menor al 1 %. Esto contrasta con la propuesta de Meurs y col. (2019), quienes propusieron que la introducción de esta mutación podría ocasionar cambios en el volumen de la estructura nativa, vinculando esta hipótesis a la alteración de la función de la proteína. Otra propuesta realizada en ese mismo estudio, respecto al efecto de la

sustitución en la estructura, hace referencia al desplegado de este dominio producto de la sustitución. Sin embargo, el análisis de las trayectorias de dinámica molecular realizada en el marco de esta tesis doctoral muestra que la estructura secundaria de láminas beta del dominio es estable para ambas formas, de acuerdo a los valores de RMSD calculados.

Por otra parte, se analizaron las interacciones entre el residuo 1.712 y los aminoácidos de su entorno tanto en la forma salvaje como en la variante. Debido al tamaño y composición química de su cadena lateral, se observó que la arginina establece interacciones adicionales con respecto a la glicina en esa misma posición. Además, el cálculo del valor de pKa de la arginina en el entorno proteico reveló que la forma protonada de este residuo se encuentra estabilizada por interacciones con el entorno. Se evaluaron las interacciones emergentes producto de la sustitución y se encontró que la arginina participa en interacciones hidrofóbicas, carga-carga y de tipo puente de hidrógeno, siendo de especial relevancia las interacciones entre cargas y de tipo puente de hidrógeno que establece con el aspartato 1.680, el que se localiza en una lámina beta adyacente en la estructura. El procesamiento de las trayectorias indica que la interacción por puentes de hidrógeno es dinámica (el número de puentes de hidrógeno entre estos residuos varía entre 0 y 3) y los cálculos de energía muestran que junto a las interacciones entre cargas de las cadenas laterales estabilizan a la estructura de la variante G1712R.

La importancia de los resultados descritos en el marco de este capítulo puede relacionarse con el mecanismo de desenrollamiento de la proteína titina que confiere elasticidad pasiva al músculo cardíaco reportados por Lee y col. (2010).

En este trabajo, utilizaron el segmento seis de la titina Ig6 (PDB: 3B43) del conejo, que incluye los dominios I65 al I70, cuyas estructuras cristalizadas fueron determinadas por difracción de rayos X. La respuesta elástica de la titina al estiramiento forzado proviene de las propiedades mecánicas de sus muchos dominios constituyentes. La flexibilidad de estos dominios contribuye a dos regímenes distintos de elasticidad: 1) un régimen suave caracterizado por la reorganización de la estructura terciaria de la proteína y el desenrollamiento de segmentos desordenados durante la elongación de la proteína, y 2) a fuerzas fisiológicamente extremas, un régimen caracterizado por la ruptura de pliegues de estructura secundaria de dominios individuales (Lee y col., 2007). Las simulaciones de dinámica molecular dirigida de dominios individuales de titina han contribuido a esta comprensión al revelar que la red de enlaces de hidrógeno, que abarca las hebras b terminales de los dominios Ig individuales, es la que rige la elasticidad de la estructura secundaria de la titina. La región de la proteína titina utilizada por Lee y col. (2010) es justamente la secuencia de aminoácidos anterior a la del dominio estudiado en este trabajo experimental. Al analizar la similitud entre la secuencia 3B43 sumada a la del dominio I1643-1734 de ambas especies (conejo y perro) mediante la herramienta de alineamiento BLASTp, se confirmó el alto grado de similitud entre las secuencias salvajes (100 % cobertura, 98,91 % de identidad). Estos hallazgos permiten extrapolar los resultados y estimar que la mutación C>T que produce el cambio de aminoácido de glicina a arginina podría interferir en la elasticidad de la proteína, dificultando el desenrollamiento del dominio debido al aumento de interacciones tipo puente de hidrógeno detectado en la proteína mutada (G1712R) y a la presencia de interacciones entre cargas que

estabilizan su estructura secundaria. Estos resultados se contradicen a lo reportado por Meurs y col. (2019), quienes postularon que el dominio afectado podría estar permanentemente desplegado o se podría desplegar fácilmente cuando se ejerce fuerza durante la diástole, momento en el que los sarcómeros se estiran.

Estos hallazgos subrayan la importancia de evaluar no solo la estructura estática, sino también las simulaciones de dinámica molecular y las interacciones, para comprender completamente cómo una sustitución puede afectar la función de una proteína. Esta información es esencial para ampliar nuestra comprensión de las bases moleculares de diversas enfermedades, ya que los cambios aquí descritos podrían afectar la conformación total de la proteína titina y sus propiedades funcionales. Por este motivo será necesario realizar futuros estudios con diferentes estrategias de evaluación, por ejemplo, realizar simulaciones de dinámica molecular de otras variantes previamente reportadas u obtenidas por análisis de las sustituciones ocurridas al comparar las secuencias de diferentes especies, así como el desarrollo de estudios experimentales de validación y confirmación *in vitro*.

Por todo lo expuesto se concluye que ambas variantes causales de CMD en los genes PDK4 y TTN podrían tener un efecto negativo, tanto en la expresión del gen PDK4 (por una menor eficiencia del sitio de empalme crítico) como en la estabilidad de la proteína titina. Con respecto a la variante G1712R podemos concluir que a pesar que el volumen y la estructura secundaria no varían al producirse la sustitución, el número de residuos con los que interactúa la proteína variante es mayor (principalmente puentes de hidrógeno), los que estabilizan aún

más a la estructura de la variante G1712R y podrían interferir en la elasticidad de la proteína, dificultando su desenrollamiento.

Referencias

1. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nature Methods*. 2010; 7(4): 248-249.
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology*. 1990; 215: 403-410.
3. Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, Witt CC, Labeit D, Gregorio CC, Granzier H, Labeit S. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circulation Research*. 2001; 89: 1065–1072.
4. Case DA y col. *Amber 2021 Reference Manual*. University of California, San Francisco, 2021.
5. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics*. 2012; 41: 429–452.
6. Eng L, Coutinho G, Nahas S, Yeo G, Tanouye R, Babaei M, Dork T, Burge C, Gatti RA. Nonclassical splicing mutations in the coding and noncoding regions of

the ATM gene: Maximum entropy estimates of splice junction strengths. *Hum Mut.* 2004; 23: 67–76.

7. Garcia FAO, de Andrade ES, Palmero EI. Insights on variant analysis *in silico* tools for pathogenicity prediction. *Frontiers in Genetics.* 2022; 13: 1010327.

8. Gerull B, Atherton J, Geupel A, Sasse-Klaasen S, Heuser A, Frenneaux M, McNabb M, Granzier H, Labeit S, Thierfelder L. Identification of a novel frameshift mutation in the giant muscle filament titin in a large Australian family with dilated cardiomyopathy. *J Mol Med.* 2006; 84: 478–483.

9. Granzier H, Labeit S. Cardiac titin: an adjustable multi-functional spring. *Journal of Physiology.* 2002; 541: 335–342.

10. Herman DS, Lam L, Taylor MRG, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *New England Journal of Medicine.* 2012; 366: 619–628.

11. Hospital A, Goñi JR, Orozco M, Gelpi JL. Molecular dynamics simulations: advances and applications. *Adv. Appl. Bioinforma. Chem.* 2015; 8: 37.

12. Hui Li, Robertson AD, Jensen JH. Very Fast Empirical Prediction and Interpretation of Protein pKa Values. *Proteins.* 2005; 61: 704-721.

13. Humphrey W, Dalke A, Schulten K. "VMD - Visual Molecular Dynamics". *J. Molec. Graphics,* 1996, 14: 33-38.

14. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021; 596: 583–589.

15. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nature Structural Biology.* 2002; 9: 646–652.

16. Katsonis P, Wilhelm K, Williams A, Lichtarge O. Genome interpretation using *in silico* predictors of variant impact. *Human Genetics*. 2022; 141: 1549–1577.
17. Labeit S, Lahmers S, Burkart C, Fong C, McNabb M, Witt S y col. Expression of distinct classes of titin isoforms in striated and smooth muscles by alternative splicing, and their conserved interaction with filamins. *J Mol Biol*. 2006; 362: 664–681.
18. Lee EH, Hsin J, Schulten K. Secondary and tertiary structure elasticity of titin Z1Z2 and a titin chain model. *Biophysical Journal*. 2007; 93: 1719–1735.
19. Lee EH, Hsin J, von Castelmur E, Mayans O, Schulten K. Tertiary and secondary structure elasticity of a six-Ig titin chain. *Biophysical Journal*. 2010; 98(6): 1085-1095.
20. LeWinter MM, Wu Y, Labeit S, Granzier H. Cardiac titin: structure, functions and role in disease. *Clin Chim Acta*. 2007; 375 (1-2): 1-9.
21. Lopaschuk GD, Ussher JR, Folmes CDL, Jaswal JS, Stanley WC. Myocardial fatty acid metabolism in health and disease. *Physiol Rev*. 2010; 90: 207–258.
22. Meurs KM, Lahmers S, Keene BW, White SN, Oyama MA, Mauceli E, Lindblad-Toh K. A splice site mutation in a gene encoding for PDK4, a mitochondrial protein, is associated with the development of dilated cardiomyopathy in the Doberman pinscher. *Human Genetics*. 2012; 131: 1319-1325.
23. Meurs KM, FriedenberG SG, Kolb J, Saripalli C, Tonino P, Woodruff K, Olby NJ, Keene BW, Adin DB, Yost OL, DeFrancesco TC, Lahmers S, Tou S, Shelton GD, Granzier H. A missense variant in the titin gene in Doberman pinscher dogs

with familial dilated cardiomyopathy and sudden cardiac death. *Hum Genet.* 2019; 138(5): 515-524.

24. Nelson DL, Cox M. Lehninger. *Principios de Bioquímica.* 7^a ed. Editorial Omega, 2018.

25. Nerenberg PS, Head-Gordon T. New developments in force fields for biomolecular simulations. *Current Opinion in Structural Biology.* 2018; 49: 129–138.

26. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research.* 2003; 31(13): 3812-3814.

27. Owczarek-Lipska M, Mausberg TB, Stephenson H, Dukes-McEwan J, Wess G, Leeb T. A 16-bp deletion in the canine PDK4 gene is not associated with dilated cardiomyopathy in a European cohort of Doberman Pinschers. *Animal Genetics.* 2013; 44: 239.

28. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nature Communications.* 2020; 11: 5918.

29. Priya Doss CG, Chakraborty C, Chen L, Zhu H. Integrating *In Silico* Prediction Methods, Molecular Docking, and Molecular Dynamics Simulation to Predict the Impact of ALK Missense Mutations in Structural Perspective. *Biomed Res. Int.* 2014; 1–14.

30. Racigh V, Pierdominici-Sottile G, Palma J. Ion Selectivity in P2X Receptors: A Comparison between hP2X3 and zfP2X4. *The Journal of Physical Chemistry B.* 2021; 125(49): 13385–13393.

31. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*. 2015; 17(5): 405-424.
33. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic Acids Research*. 2005; 33.
34. Singh PK, Mistry KN. A computational approach to determine susceptibility to cancer by evaluating the deleterious effect of nsSNP in XRCC1 gene on binding interaction of XRCC1 protein with ligase III. *Gene*. 2016; 576: 141–149.
35. Tang R, Prosser DO, Love DR. Evaluation of Bioinformatic Programmes for the Analysis of Variants within Splice Site Consensus Regions. *Advances in Bioinformatics*. 2016; 5614058.
36. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou L-P, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science*. 2022; 31(1): 8–22.
37. Tskhovrebova L, Trinick J. Titin: properties and family relationships. *Nature Reviews Molecular Cell Biology*. 2003; 4: 679–689.
38. Wang M, Marin A. Characterization and Prediction of Alternative Splice Sites. *Gene*. 2006; 366: 219-227.
39. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology*. 2004; 11(2-3): 377-394.

CONCLUSIONES FINALES

Actualmente existe escasa información acerca de la prevalencia de las enfermedades cardiológicas en caninos de Argentina, principalmente aquellas con origen genético dentro de las que se incluye a la cardiomiopatía dilatada. Asimismo no se conocen las causas genéticas de estas enfermedades, ni la frecuencia de las mutaciones asociadas en la población de perros locales. Tal vez motiva esto, la falta de test diagnósticos que posibiliten un rápido resultado y con un costo accesible para los criadores. La posibilidad de conocer el genotipo para muchas de las enfermedades presentes en perros que forman parte del circuito de reproducción y cría en el país, constituye un gran aporte al objetivo de ir reduciendo las prevalencias de este tipo de enfermedades.

Por otro lado, la utilización del perro como modelo animal para el estudio de enfermedades genéticas humanas se ha visto incrementado debido a las más de 360 enfermedades genéticas compartidas entre ambas especies, siendo aproximadamente el 46% específicas de ciertas razas caninas. Estos trastornos en caninos a menudo reflejan aspectos fisiológicos, manifestaciones clínicas y respuestas a tratamientos similares a las enfermedades humanas, lo que ha llevado a su empleo como modelos experimentales desde el siglo XVII. La similitud en factores ambientales y la alta homología entre los genomas canino y humano han fortalecido esta elección. La estructura de la población canina, con razas derivadas de un pequeño grupo de fundadores y crías endogámicas, simplifica la cartografía genética de enfermedades y permite identificar mutaciones patológicas con relativa facilidad. Asimismo, la expectativa de vida del perro es lo suficientemente

corta, lo que acelera la observación de fenotipos. Entre estas razas, el Doberman Pinscher se destaca como un modelo ideal para la investigación de la cardiomiopatía dilatada. A pesar de la menor heterogeneidad genética en comparación con los humanos, su susceptibilidad a la enfermedad no se limita a una sola mutación, lo que lo convierte en un modelo valioso que comparte aspectos diagnósticos, fisiopatológicos, de modo de herencia y de respuesta a terapias con la enfermedad humana.

Este Trabajo de Tesis aporta información acerca de las bases genéticas de la cardiomiopatía dilatada en la población de Doberman Pinscher local, colaborando con el desarrollo de métodos moleculares de diagnóstico accesibles para el mercado local. La aplicación de técnicas moleculares como PCR, secuenciación Sanger y pirosecuenciación confirmó la presencia de las variantes candidatas en Argentina (CFA14:g.20,829,667_20,829,682del, CFA5:g.53,941,386T>C, CFA36:g.22,321,955C>T). La prevalencia de estos alelos se estableció por primera vez en las poblaciones locales de esta raza, registrando valores de 0,17, 0,12 y 0,58, respectivamente. Esto permitió demostrar que la prevalencia de la mutación PDK4 en Argentina fue menor (0,17) que en la población de DP de Estados Unidos (0,37) y similar a la de Europa (0,16). No se encontró asociación entre las variantes CFA14:g.20,829,667_20,829,682del y CFA5:g.53,941,386T>C y el desarrollo de CMD en esta población, a diferencia de la variante en el gen TTN que se asoció significativamente con la enfermedad. Por más que sólo se haya encontrado relación estadísticamente significativa entre una de las variantes reportadas (CFA36: g.22,321,955 C>T) y el desarrollo de CMD en la población local de DP, es necesario conocer el genotipo de los

animales lo más precisamente posible para realizar cruzamientos reproductivos selectivos internacionales. Contar con métodos de detección basados en el análisis del ADN, permite realizar las determinaciones de animales portadores de las mutaciones en el país, sin necesidad de realizar estos estudios en el exterior. Por otro lado, permite el monitoreo cardiológico de aquellos animales portadores para detectar de forma temprana si presentan o no signos de cardiopatía e instaurar un tratamiento médico preventivo. Asimismo, se estudiaron las consecuencias biológicas de estas variantes para tratar de comprender los mecanismos subyacentes por los cuáles estas mutaciones podrían predisponer al desarrollo de la enfermedad. Mediante la utilización de programas de predicción *in silico*, se observaron efectos perjudiciales de la variante en el gen PDK4 que podría relacionarse con una menor eficiencia del sitio de empalme críptico. Asimismo, se encontró un efecto negativo de la variante en el gen TTN que afectaría la estabilidad de la proteína titina. A pesar que el volumen y la estructura secundaria no varían al producirse la sustitución de aminoácidos producto de la variante no sinónima, el número de residuos con los que interactúa la proteína variante es mayor (principalmente puentes de hidrógeno), los que estabilizan aún más a la estructura de la variante G1712R y podrían interferir en la elasticidad de la proteína, dificultando su desenrollamiento.

ANEXOS

Anexo 1. Consentimiento informado

n e l día de la fecha:...../...../.....

Yocon DNI Nro.....

Con do i cilio en: Calle:..... Nro:.....

Localidad:..... Pcia:.....

CP..... Teléfono.....

Tras haber recibido información verbal clara y sencilla sobre el procedimiento.

Autorizo voluntariamente a la realización de estudios cardiológicos (examen físico, electrocardiograma y ecocardiograma) y la extracción de muestra de sangre entera con EDTA del paciente:

No br e..... spe cie.....

Sexo..... Ra a.....

da d..... Pelaje.....

Nro. de ist oria Clínica.....

Para su utilización en la determinación de las mutaciones causales de la Cardiomiopatía Dilatada

A realizarse en el IGEVET (Instituto de Genética Veterinaria), Facultad de Ciencias Veterinarias (FCV), Universidad Nacional de La Plata (UNLP) y en el Servicio de Cardiología del Hospital Escuela FCV-UNLP.

Firma:

Firma:

Aclaración:

Aclaración:

VETERINARIO

PROPIETARIO

Anexo 2. Protocolos de extracción de ADN

2.1- Protocolo de extracción orgánica de ADN a partir de sangre periférica:

- Alicuotar 5 μ l de la muestra de sangre en tubos de 5 ml.
- Lavar con 1 ml de agua pura centrifugar a 1000 g por 1 minuto.
- Descartar el sobrenadante volcando suavemente el tubo y secar sobre papel la última gota. Repetir el lavado una vez más o hasta que el pellet quede blanco.
- Resuspender el pellet en 50 μ l de Buffer de extracción S (5 mM de Cl-Tris, 25 mM de DTT, 2 % de N-Lauroylsarcosine)
- Agregar 10 μ l de proteinasa K (10 mg / l) e incubar durante 2 horas a 55 °C o toda la noche a 37 °C.
- Llevar la muestra a temperatura ambiente agregar 100 μ l de cloroformo. Mezclar por inversión. Centrifugar a 10000 g por 1 minuto.
- Transferir la fase acuosa, que contiene el ADN (superior), a un nuevo tubo y descartar la fase orgánica.
- Repetir la extracción orgánica una vez más.
- Transferir la fase acuosa a otro tubo agregar 100 μ l de Acetato de Amonio 1 M. Centrifugar a 10000 g por 1 minuto. Transferir la fase acuosa a otro tubo y agregar 1 ml de etanol (preenfriado). Mezclar suavemente, precipitar durante 1 hora en hielo o en el freezer (se puede dejar toda la noche).
- Centrifugar a 14000 g por 15 minutos. Luego descartar el sobrenadante volcando suavemente y secar la boca del tubo sobre papel.
- Lavar el tubo con 50 μ l de etanol 70 % frío, centrifugar a 14000 g por 15 minutos y volver a descartar el sobrenadante.

- Secar el pellet a T° ambiente.
- Agregar 100 – μ l de agua pura e incubar toda la noche en un baño a 7 °C para hidratar el ADN.

2.2 Protocolo de extracción de ADN a partir de sangre utilizando Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA)

- Cargar μ l de Solución de lisis celular en tubos estériles de 5 l .
- Agitar el tubo con la muestra de sangre varias veces para homogeneizar la muestra y luego transferir μ l al tubo con la solución de lisis. Meclar invirtiendo 5-6 veces.
- Incubar 10 minutos a T° ambiente (invirtiendo 2-3 veces en ese tiempo) para lisar los glóbulos rojos.
- Centrifugar 20 segundos a 14000 g a T° ambiente.
- Descartar el sobrenadante y agitar el pellet en el vortex 15 segundos para resuspender los glóbulos blancos. La eficacia de esta técnica depende de la correcta resuspensión de la muestra.
- Agregar a cada tubo μ l de la Solución de lisis nuclear. Cargar y descargar la pipeta en la solución 5-6 veces para homogeneizar la muestra y lisar así los glóbulos blancos. La solución se pondrá viscosa.
- Agregar a cada tubo μ l de Solución para precipitar proteínas. Agitar en vortex durante 10-20 segundos.
- Centrifugar a 14000 g por 3 minutos a T° ambiente. Se formará un precipitado marrón oscuro de proteínas y un sobrenadante con el ADN.

- Precipitación del ADN: Colocar en tubos limpios μl de isopropanol a T° ambiente. Volcar en ellos el sobrenadante y descartar el tubo con el precipitado marrón.
- Mezclar la solución invirtiendo el tubo suavemente, varias veces hasta observar la formación de pequeños hilos de ADN.
- Centrifugar a 15000 g por 1 minuto a T° ambiente.
- Lavado del ADN: Descartar el sobrenadante volcando el tubo agregar μl de etanol 70 %. Hacer girar el tubo horizontalmente para lavar bien las paredes.
- Centrifugar a 15000 g por 1 minuto a T° ambiente.
- Aspirar el etanol cuidadosamente con una pipeta, evitando aspirar el pellet.
- Dejar el tubo destapado e invertido sobre un papel absorbente limpio durante 15 minutos.
- Agregar a cada tubo μl de Solución de Rehidratación de ADN e incubar a 65° por una hora o a 37° por 24 horas.

Anexo 3. Análisis en gel de agarosa

Preparación de geles de Agarosa 1 % (volumen final 100 ml con Agua Destilada)

La preparación se agita y se coloca en el microondas o en autoclave hasta su disolución completa. Esta solución se coloca en moldes, se coloca un peine para que forme las fosas de siembra y se deja solidificar durante aproximadamente 1 hora.

Preparación de las muestras para su siembra

Se toman 3-5 μ l de la extracción de ADN se le agregan μ l de DYE se echa y se coloca una por una en cada fosa en el gel. La corrida electroforética se realiza a 100 voltios, constantes durante 1 hora o hasta que el DYE haya migrado 3 o 4 cm.

Colorante DYE

- 0,025 g de Bromophenol Blue.
- 4 g de sacarosa.
- Completar hasta un volumen final de 100 ml con agua destilada.

Revelado de los geles

1. Se colocan 5 μ l de Bromuro de etidio en 5 ml de agua destilada.
2. Sumergir el gel en la solución por unos segundos.
3. Visualizar en un transiluminador de luz UV.

Anexo 4. Chequeo de los amplificadores en geles de poliacrilamida

-Agua destilada 9 ml

-3 ml TBE 5X (Tris Borato EDTA)

-3 ml Acrilamida/Bisacrilamida 19:1

- 5 μ l T M D

- 5 μ l APS

Procedimiento:

1. Limpiar los vidrios con alcohol para desengrasarlos.
2. Colocar los separadores (0,375 mm) sobre el vidrio entero y apoyar encima, con mucho cuidado, el vidrio calado. Sujetar con ganchos o clips.
3. Agregar el TEMED (Tetramethylethylenediamina) y el APS (Persulfato de amonio) a la mezcla de poliacrilamida, mezclar bien y proceder (desde el extremo calado) a cargar el gel con una pipeta de vidrio, evitando que se formen burbujas de aire.
4. Una vez que se llenó el molde con el gel, colocar el peine con cuidado.
5. Esperar que polimerice y montar en la cuba electroforética con TBE 1X. No dejar el gel polimerizado mucho tiempo sin montar para evitar la deshidratación.
6. Sembrar las muestras y correr a 170 watts durante el tiempo necesario.

TBE 5X:

-54 g de Tris base.

-27,5 g de Ácido Bórico.

-20 ml de EDTA 0,5M (pH=8).

-Completar hasta 1000 ml con agua destilada.

Preparación de las muestras para su siembra

La muestra se siembra en el gel con colorante DYE, permitiendo así su visualización durante la corrida electroforética. Se toman μ l del ADN genómico y se le agregan μ l de DYE los μ l resultantes se colocan en los pocillos del gel.

La corrida electroforética se realiza a 170 voltios durante 1 hora y cuarto.

Colorante DYE

- 0,025 g de Bromophenol Blue.

- 4 g de sacarosa.

- Completar hasta un volumen final de 100 ml con agua destilada.

Revelado de los geles

1. Se colocan 5 μ l de Bromuro de etidio en 5 ml de agua destilada.

2. Sumergir el gel en la solución por unos segundos.

3. Visualizar el ADN en un transiluminador de luz UV.

Anexo 5. Secuenciación

5.1- Purificación de las reacciones de PCR por precipitación con polietilenglicol (PEG) para secuenciación

este protocolo está ajustado para reacciones de PCR de 5 μ l. Si el volumen de la PCR es diferente, modificar la escala manteniendo las proporciones.

1. Correr 5 μ l del producto de PCR en un gel de agarosa para verificar la amplificación.

2. Al resto de la reacción de PCR agregarle 5 μ l de una solución de PEG - 2,5 MNaCl y mezclar bien.

3. Incubar a 37°C por 15 minutos.

4. Centrifugar a \sim 15.000 x g durante 15 minutos a temperatura ambiente.

5. Retirar el sobrenadante con pipeta y descartarlo. El pellet es incoloro y queda adherido a la pared del tubo.

6. Agregar 5 μ l de etanol 70%. Si se agrega el etanol en el fondo del tubo centrifugar 2 minutos. Si se agrega el etanol suavemente por la pared del tubo, dejar reposar un minuto. Eliminar el sobrenadante y descartar. Eliminar la mayor cantidad posible del etanol.

7. Secar el pellet a 37°C durante 10-15 minutos. Asegurarse de que no queden restos de etanol y que el pellet esté seco. Un exceso en el secado del pellet dificulta la resuspensión.

8. Disolver el producto de PCR en 5 μ l de agua bidestilada estéril. Resuspender completamente el pellet (con pipeta o vortex). Ayuda incubar por unos minutos a temperatura ambiente o a 37°C.

Solución 20% PEG - 2,5 M NaCl

Para 50 ml mezclar:

- 10,0 g polietilenglicol 8000 (funciona igualmente con el PEG 6000 u 8000).
- 7,3 g de NaCl.
- Agregar 45 ml de agua bidestilada. Agitar y dejar disolver el PEG durante al menos 20 minutos. Se puede colocar la solución en un agitador a 37°C.

Cuando se disolvió el PEG (la solución se vuelve incolora), completar el volumen hasta 50 ml con agua bidestilada. Conservar a temperatura ambiente.

5.2- Protocolo a seguir para realizar las reacciones de secuenciación

Una reacción estándar se realiza en μl finales de acuerdo al siguiente protocolo:

- ADN (120- n g) + agua 5 μl
- primer (5 uM = 5 p ol / μl) μl
- premix (DYEnamic ET Dye Terminator) 4 μl
- Total μl

5.3- Protocolo Precipitación con Acetato de Amonio

- Reacción de secuencia μl
- Acetato de amonio 7.5 M μl
- Etanol absoluto 30 μl

1. Centrifugar a 15.000 rpm durante 15 minutos.
2. Descartar el sobrenadante.

. Lavar con μl de etanol 7 μl agregando el alcohol con suavidad para no despegar el pellet.

4. Dejar reposar 1 minuto y eliminar tanto etanol como sea posible (si quedaran gotas en las paredes, centrifugar brevemente y eliminarlas con la pipeta). Esto es importante ya que los restos de etanol producen artefactos en la corrida del secuenciador, denominados dye blobs.

5. Secar el etanol a 37°C durante aproximadamente 10 minutos. No secar excesivamente, ya que dificulta la resuspensión del ADN.

. Disolver el pellet en μl de MegaBAC Loading Solution e clar vigorosamente con un vortex durante 20-30 segundos para asegurar una resuspensión completa.

Anexo 6. Pirosecuenciación

6.1- Protocolos de amplificación: (volumen final 50 μ l)

Agua Destilada μ l

Buffer 5 μ l

dNTP's μ l

Mg μ l

Primer – F (5 μ M) 5 μ l

Primer – R – M (5 μ M) μ l

Primer – 13 – Biotin. μ l

Taq. Poli er asa 5 μ l

ADN μ l

6.2- Preparado de muestras para Pirosecuenciación

Luego de obtener el producto de PCR biotinilados se procedió a purificarlos utilizando el siguiente protocolo:

A) Preparación de Placa para purificación

1. Preparar la siguiente mezcla multiplicando por el número de muestras a pirosecuenciar:

X 1 muestra

Buffer de “binding” μ l

Perlas de Sefarosa 3 μ l

H₂O 17 μ l

Volu en f inal μ l

2. Mezclar muy bien antes de distribuir en la placa de purificación. Evitar la decantación de las perlas de sefarosa. Distribuir en cada fosa de la placa de Purificación:

Mezcla de binding μ l

Producto de PCR μ l

3. Tapar la placa con film. Incubar la mezcla por 5-10 minutos a temperatura ambiente mientras se agita a 1400 rpm.

4. Durante la incubación preparar la Placa para Pirosecuenciación.

B) Preparación de la Placa para Pirosecuenciación

1. Preparar la siguiente mezcla multiplicando por el número de muestras a pirosecuenciar:

X 1 muestra

Buffer de “annealing” 5μ l

Primer interno de secuenciación 5μ l

. Distribuir 5μ l en cada fosa de la placa para Pirosecuenciación (PSQ).

C) Purificación

1. Con las perlas de sefarosa en suspensión, aspirar la solución usando la estación de purificación (Vacuum Prep Tool TM). De esta forma las perlas de sefarosa se adhieren al filtro de las puntas del soporte de aspiración.

Mantener las puntas dentro de la placa de PCR hasta verificar que toda la solución haya sido aspirada.

2. Transferir el soporte de aspiración a la primera cuba y aspirar el etanol 70 % por 5 segundos.
3. Transferir el soporte de aspiración a la segunda cuba y aspirar NaOH 0,2 M por 5 segundos para desnaturalizar el ADN.
4. Transferir el soporte de aspiración a la tercera cuba y aspirar el buffer de lavado (10 a 20 ml por placa) por 5 segundos.
5. Levantar el soporte de aspiración a una posición vertical para permitir que el líquido drene de las puntas. Retornar a la posición horizontal. Colocar el soporte de aspiración sobre la placa de Pirosecuenciación (introduciendo las puntas en los pocillos sin tocar el líquido), apagar la bomba de aspiración y agitar suavemente para permitir que las perlas caigan en la placa PSQ.
6. Calentar la placa a 80° C por dos minutos para permitir el annealing del primer de secuenciación. Enfriar a temperatura ambiente.
7. Limpiar las puntas del soporte de aspiración agitándolo en la cuarta cuba que contiene agua Milli Q, con el fin de eliminar cualquier perla remanente.
8. Cuando todas las placas han sido preparadas, transferir el soporte de aspiración a la quinta cuba y aspirar el agua Milli Q por 20-30 seg. para lavar el soporte de aspiración.