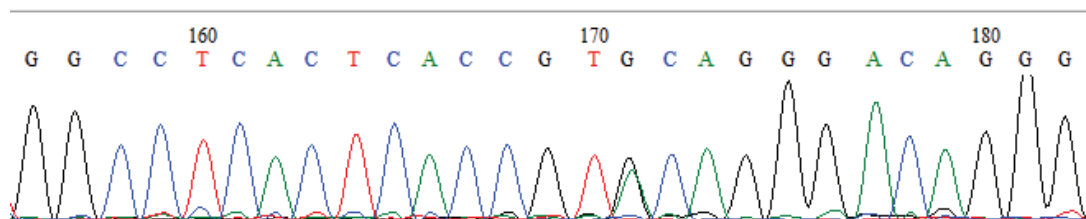




**Universidad  
Zaragoza**

## Trabajo Fin de Grado

### Marcadores moleculares aplicados a la identificación genética en animales



Autor:

**Javier Sánchez Marco**

Directores:

**Clementina Rodellar Penella  
Arianne Sanz Fernández**

Facultad de Ciencias/Grado Biotecnología  
Departamento de Anatomía, Embriología y Genética Animal de la Facultad de  
Veterinaria  
Año 2015

**Clementina Rodellar Penella**, Profesora Titular y **Arianne Sanz Fernández**, Personal Investigador del Departamento de Anatomía, Embriología y Genética Animal de la Facultad de Veterinaria de la Universidad de Zaragoza.

CERTIFICAN QUE:

El trabajo que se presenta bajo el título “Marcadores moleculares aplicados a la identificación genética en animales” que se recoge en la presente Memoria, del que es autor **Javier Sánchez Marco**, ha sido realizado en el Departamento de Anatomía, Embriología y Genética Animal de la Facultad de Veterinaria de la Universidad de Zaragoza bajo nuestra dirección.

Zaragoza, 14 de septiembre de 2015



Fdo. Clementina Rodellar Penella



Fdo. Arianne Sanz Fernández

# **Índice**

1. ABREVIATURAS .....	1
2. RESUMEN .....	2
3. ANTECEDENTES .....	3
3.1. Marcadores moleculares aplicados a la identificación animal .....	3
3.1.1. Microsatélites .....	3
3.1.2. SNPs .....	4
3.2. Microsatélites aplicados a la verificación de las relaciones de parentesco en la especie equina .....	4
3.2.1. Estandarización de marcadores por la International Society on Animal Genetics (ISAG).....	4
3.3. Enfermedad autosómica recesiva poliquística renal (ARPKD) .....	5
3.3.1. Fibrosis hepática congénita y renal en humanos .....	5
3.3.2. Fibrosis hepática congénita en caballo .....	5
3.4. Gen de la enfermedad poliquística hepática y renal (PKHD1) .....	6
4. OBJETIVOS .....	7
5. MATERIALES Y MÉTODOS .....	8
5.1. Extracción y cuantificación de DNA.....	8
5.1.1. Obtención de DNA a partir de sangre. ....	8
5.1.2. Purificación de DNA a partir de hígado .....	8
5.1.3. Cuantificación DNA.....	8
5.2. Verificación de parentesco mediante el uso de marcadores microsatélites.....	9
5.2.1. Amplificación de microsatélites mediante multiplex-PCR .....	9
5.2.2. Análisis de los fragmentos y lectura de los genotipos.....	9
5.2.3. Exclusión y asignación de paternidades .....	9
5.3.1. Búsqueda del gen PKHD1 en caballos.....	10
5.3.1.1. Diseño de cebadores.....	10
5.3.1.2. Amplificación mediante PCR.....	10
5.3.1.3. Purificación de los productos de PCR.....	11
5.3.2. Reacción de secuenciación.....	11
5.3.2.1. Purificación de la reacción de secuenciación y electroforesis capilar .....	11
5.4. Búsqueda de SNPs .....	12
5.4.1. Alineamientos mediante ClustalW .....	12
5.4.2. Estimación de las frecuencias alélicas y genotípicas. ....	12
6. RESULTADOS Y DISCUSIÓN.....	13
6.1. Cantidad y calidad del DNA extraído .....	13
6.2. Verificación de las relaciones de parentesco.....	13
6.3. Aislamiento del gen PKHD1 en la especie equina.....	13
6.4. Análisis de las secuencias obtenidas. ....	15
6.4.1. Análisis de SNPs asociados con ARPKD .....	15
6.4.2. Análisis de nuevos SNPs.....	18
6.4.3. Análisis de las frecuencias alélicas y genotípicas .....	20
7. CONCLUSIONES .....	22
8. ANEXO.....	25

## 1. ABREVIATURAS

- ADPKD: Autosomal Dominant Polycystic Kidney Disease
- ARPKD: Autosomal Recessive Polycystic Kidney Disease
- Chr: Cromosoma
- DNA: ácido desoxirribonucleico
- dNTP's: Desoxinucleótidos trifosfatos
- EDTA: ethylenediaminetetraacetic acid/ ácido etildiaminotetraacético
- ISAG: International Society on Animal Genetics/ Sociedad Internacional de Genética Animal
- NCBI: National Center for Biotechnology Information/ Centro Nacional de Información Biotecnológica
- ORF: Open Reading Frame/ Marco Abierto de Lectura
- Pb: pares de bases
- PCR: Polymerase Chain Reaction/ Reacción en Cadena de Polimerasa
- PE1: probabilidad de exclusión asumiendo un progenitor conocido
- PE2: probabilidad de exclusión asumiendo dos progenitores conocidos
- *PKD1*: gen Polycystic Kidney Disease 1
- *PKD2*: gen Polycystic Kidney Disease 2
- *PKHD1*: gen Polycystic Kidney and Hepatic Disease 1
- SNP: Single Nucleotide Polymorphism
- SSR: Simple Sequence Repeat
- STR: Short Tandem Repeat

## **2. RESUMEN**

La fibrosis congénita hepática ha sido descrita en la raza de caballos Franches Montagnes como una enfermedad letal en los primeros estadios de vida y de herencia monogénica autosómica recesiva. La fibrosis congénita hepática en caballos se ha asociado con el gen *PKHD1*, cuyas variantes causan fibrosis congénita hepática y renal en la especie humana (ARPKD).

En el Departamento de Patología Animal de la Universidad de Zaragoza se ha diagnosticado recientemente el primer caso de fibrosis congénita hepática en un potro Pura Raza Español. El objetivo de este proyecto fue el estudio del gen *PKHD1* y su asociación con la fibrosis congénita hepática, para ello se analizaron dos polimorfismos asociados fuertemente con la enfermedad en la raza de caballos Franches-Montagnes, g.49,597,760A>T y g.49,630,834G>A. Estos polimorfismos causan cambios en la proteína fibrocistina, codificada por el gen *PKHD1*, en una región conservada en varias especies, siendo la pérdida de función de dicha proteína una posible causa de la enfermedad

Se llevó a cabo el análisis de estos SNPs en el potro afectado y en sus parientes cercanos. También se hizo un estudio de la frecuencia de los polimorfismos en animales no emparentados de Pura Raza Español. Se han detectado cuatro nuevos SNPs en los fragmentos analizados correspondientes al gen *PKHD1*. Los resultados obtenidos permiten afirmar que no existe una asociación entre los polimorfismos analizados con la fibrosis hepática congénita en la raza Pura Raza Española a diferencia de lo descrito en la raza Franches-Montagnes.

## **ABSTRACT**

Congenital hepatic fibrosis has been described in the Franches-Montagnes horse breed as a lethal disease in the early stages of life, with monogenic autosomal recessive inheritance. Congenital hepatic fibrosis in horses has been associated with the *PKHD1* gene, whose variants cause congenital hepatic and renal disease in the human species (ARPKD).

The first case of congenital hepatic fibrosis has been recently diagnosed in a Purebred Spanish Horse foal by the Department of Animal Pathologic of the University of Zaragoza. The objective of this project was the study of *PKHD1* gene and his association with the congenital hepatic fibrosis. For that purpose, we analyzed two polymorphisms which were strongly associated with the congenital hepatic fibrosis in the Franches-Montagnes horse breed, g.49,597,760A>T y g.49,630,834G>A. These polymorphisms cause changes into the protein encoded by the *PKHD1*, the protein fibrocystin, in a conserved region in many species. The lack of function of this protein is a possible disease cause.

Both SNPs of the foal affected and their closely related family were analyzed. Also a frequency study was performed in unrelated Purebred Spanish Horse animals. Our results show there is no association between the analyzed polymorphism and the congenital hepatic fibrosis in Purebred Spanish Horse. These results disagree with the described association in the Franches-Montagnes horse breed.

### **3. ANTECEDENTES**

#### **3.1. Marcadores moleculares aplicados a la identificación animal**

Los marcadores moleculares son fragmentos de DNA que sirven de referencia para seguir la transmisión de un segmento de cromosoma de una generación a otra, siendo, en un sentido restringido, una entidad genética que manifiesta polimorfismo y se hereda de forma mendeliana. Su uso ha demostrado ser una herramienta muy útil en la identificación individual (Herráeza et al., 2005; Pakstis et al., 2010).

Los marcadores de DNA necesitan reunir una serie de características para ser utilizables en identificación genética (Herráeza et al. 2005; Liu et al. 2004):

- Que sean muy polimórficos
- Que estén distribuidos por todos los cromosomas de una especie para obtener información representativa de la posible variación existente.
- Que sean independientes del estado físico y desarrollo del individuo.
- Que tengan un bajo coste tanto en la obtención como en la aplicación
- Que sean públicos y no estén sometidos a ningún tipo de patente.
- Que su interpretación sea sencilla y objetiva.
- Que requieran poca cantidad y calidad de muestra.
- Que sean reproducibles en cualquier experimento de laboratorio
- Que sean estables para que las mutaciones no puedan influir en los resultados.
- Que tengan preferiblemente herencia codominante.
- Que sean discriminantes, o sea, capaces de detectar diferencias entre individuos estrechamente emparentados
- Que no estén sujetos a influencias ambientales.
- Que sean neutros

Para que se den muchas de estas características, los marcadores deben ser susceptibles de ser analizados mediante la Reacción en Cadena de Polimerasa (PCR; Mullis and Faloona, 1987), que ha impulsado enormemente el desarrollo de la genética molecular gracias a su gran fiabilidad, rapidez y repetitividad.

##### **3.1.1. Microsatélites**

Los microsatélites, también conocidos como SSR (Simple Sequence Repeat) o STR (Short Tandem Repeat) consisten en repeticiones de secuencias cortas (2-8pb) que han sido descritos en la mayoría de las especies y poseen una tasa de mutación que oscila entre  $10^{-3}$  y  $10^{-5}$ . Presentan herencia codominante y son fáciles de detectar metodológicamente (Chistiakov et al., 2006; Forbes et al., 1995).

Los microsatélites han demostrado ser marcadores muy informativos para estudios poblacionales, con un alto grado de polimorfismo. Actualmente se utilizan también en análisis forenses, pruebas de paternidad, estudios de biodiversidad y análisis filogenéticos (Chistiakov et al., 2006; Holt et al., 2000).

### **3.1.2. SNPs**

SNPs (Single Nucleotide Polymorphisms) o polimorfismos de una base consisten en la sustitución de un nucleótido por otro. Se encuentran distribuidos por todo el genoma y pueden encontrarse tanto en regiones codificantes como no codificantes. Son los polimorfismos más comunes en la especie humana, y se estima que existe un SNP por cada 1000 pares de bases en la especie humana (Brookes, 1999; Wang et al., 1998).

En este tipo de polimorfismos pueden estar presentes cualquiera de los cuatro nucleótidos en cualquier posición en el genoma, sin embargo, en la práctica la mayoría de los SNPs tienen solamente dos variantes, la secuencia original y la versión mutada. Esto se debe a la vía en que estos aparecen y se distribuyen en una población. Un SNP se origina cuando ocurre una mutación puntual en el genoma. Cuando esta mutación ocurre en un gameto, puede heredarse y al cabo de varias generaciones establecerse en la población. Para que se dé una tercera variante en un determinado SNP tendría que producirse los mismos pasos, por lo que es poco probable; siendo la mayoría de los SNPs bialélicos (Brookes, 1999).

El método más directo para detectarlos es la secuenciación de segmentos de DNA previamente amplificados por PCR. Una vez localizado el SNP se pueden utilizar diversas técnicas para su análisis, como son: extensión del primer o snapshot (Pastinen et al., 1996), secuenciación Sanger (Sanger et al., 1977), pirosecuenciación (Ronaghi et al., 1998) o microarrays o microchips de DNA (Schena et al., 1998).

Debido a las ventajas que ofrecen este tipo de marcadores y a la diversidad de métodos disponibles para su estudio son los utilizados mayoritariamente en diversas especies en estudios de identificación y paternidad (Pakstis et al., 2010), caracterización de poblaciones (Inagaki et al., 2002), estudios de biodiversidad (Twito et al., 2007) y estudio de enfermedades genéticas (Brookes, 1999) entre otras.

## **3.2. Microsatélites aplicados a la verificación de las relaciones de parentesco en la especie equina**

Una de las aplicaciones de los microsatélites son los test de exclusión de paternidad. Se ha obtenido un porcentaje de exclusión mayor al 99,99% usando 20 marcadores microsatélites, todos ellos con un grado de polimorfismo alto. Este alto porcentaje de exclusión los sitúa como marcadores moleculares de elección a la hora de establecer parentesco con gran fiabilidad (Choi et al., 2008).

### **3.2.1. Estandarización de marcadores por la International Society on Animal Genetics (ISAG)**

Actualmente no existe obligatoriedad impuesta por ningún organismo oficial que indique los marcadores genéticos que deben utilizarse para una correcta identificación de los animales y productos derivados. Sin embargo, la Sociedad Internacional de Genética Animal (ISAG; <http://www.isag.us>) favorece el intercambio a nivel internacional de información y reactivos entre sus miembros, regulando la estandarización de los marcadores y técnicas utilizadas internacionalmente para la identificación animal en las distintas especies domésticas.

La ISAG recomienda a nivel internacional un grupo de 17 microsatélites (ver Tabla S1 en Anexo) para llevar a cabo pruebas de identificación y de filiación, que garanticen una

correcta identificación y permitan una estandarización a nivel internacional (van de Goor et al., 2010).

### **3.3. Enfermedad autosómica recesiva poliquística renal (ARPKD)**

Existe un gran número de desordenes heredables que causan desarrollo de quistes renales. La forma más común, enfermedad poliquística hepática autosómica dominante (ADPKD) suele ser diagnosticada en adultos y es causada por mutaciones en el gen *PKD1* o *PKD2*. ARPKD es la forma neonatal de la enfermedad y está asociada con agrandamiento de hígado, de herencia autosómica recesiva y asociada con mutaciones en el gen *PKHD1* (Harris and Torres, 2009).

#### **3.3.1. Fibrosis hepática congénita y renal en humanos**

ARPKD (Autosomal recessive polycystic kidney disease) tiene una incidencia en humanos de 1:20000. En más de un 30% causa muerte neonatal, el resto sufren hipertensión e insuficiencia renal. Otro signo fenotípicamente característico de la enfermedad es una disgenesia biliar resultante en una fibrosis hepática congénita, además de la dilatación del conducto biliar intrahepático. La enfermedad se puede manifestar no sólo en el periodo neonatal, pudiendo ser diagnosticada tanto en los primeros años de vida como en la adolescencia o en edad adulta (Harris and Torres, 2009).

#### **3.3.2. Fibrosis hepática congénita en caballo**

La fibrosis congénita hepática de carácter dominante o recesivo ha sido descrita en humanos, gatos, perros y cerdos. En caballos, se identificaron 30 potros con lesiones en el hígado que podrían ser debidas a esta fibrosis hepática congénita (Haechler et al., 2000).

El examen patológico de estos potros mostró hígados gravemente alargados con quistes en las paredes delgadas. Histológicamente, se encontró severa fibrosis difusa en puente porto portal con varios quistes pequeños e irregulares, encontrándose también quistes en el conducto biliar. El análisis del pedigrí permitió rastrear la enfermedad hasta llegar a un único semental. Con todos estos resultados, se sugirió que muy posiblemente esta fibrosis hepática congénita sea un enfermedad de herencia recesiva (Haechler et al., 2000).

En un estudio posterior, se estudiaron 5 potros de raza Franches-Montagnes de entre tres semanas y doce meses con signos de daño hepático. Se pudieron rastrear los parientes de estos 5 potros, encontrando un antecesor común, coincidiendo con el estudio de Haechler et al. (2000), en que la enfermedad es de herencia recesiva.(Drögemüller et al., 2014)

Se genotiparon 65.157 SNPs en los 5 potros con la enfermedad y en 12 caballos no afectados de la misma raza, Franches-Montagnes. Se detectó una posible asociación en una región del cromosoma 20 con la enfermedad. En esta región, entre otros genes, se encuentra el gen *PKHD1*, que se ha postulado como un excelente gen candidato de la enfermedad en caballos (Drögemüller et al., 2014).

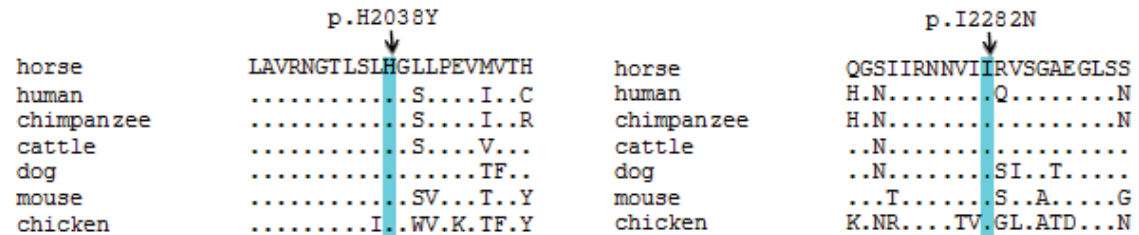
De todos los SNPs detectados, dos de ellos, g.49,597,760A>T y g.49,630,834G>A y producían cambios en la proteína p.I2282N y p.H2038Y , respectivamente. En ambos casos se comprobó que los potros afectados presentaban homocigosis mientras que los padres eran heterocigotos (Drögemüller et al., 2014).

Se observó una fuerte asociación, pero no perfecta, entre el SNP g.49,597,760A>T y la fibrosis congénita hepática. Se especuló que en el caso del SNP g.49,630,834G>A se obtendrían



resultados de asociación similares, puesto que no se consiguió secuenciar (Drögemüller et al., 2014).

Se comprobó alineando las proteínas fibrocistina de varias especies que en las dos variantes, los cambios de base producidos, cambiaban un aminoácido conservado en la secuencia de varias especies. Se sugiere que si estos dos aminoácidos están conservados puede ser debido a que tengan una importancia en la proteína y un cambio produzca la pérdida de función de la proteína que ocasione la enfermedad. Como se observa en la Figura 1 se conservan el aminoácido histidina en la posición 2038 y el aminoácido isoleucina en la posición 2282 en la proteína codificada por el gen *PKHD1* (Drögemüller et al., 2014).



**Figura 1:** Alineamiento de la proteína codificada por el gen *PKHD1* en diferentes especies. Las secuencias han sido obtenidas de las siguientes códigos de acceso: *H. sapiens* NP\_619639.3, *P. troglodytes* XP\_518534.2, *B. taurus* XP\_002697357.1, *C. lupus* XP\_532169.2, *M. musculus* NP\_694819.2, and *G. gallus* XP\_420050.3.

### 3.4. Gen de la enfermedad poliquística hepática y renal (*PKHD1*)

*PKHD1* (polycystic kidney and hepatic disease 1) es el gen causante de la ARPKD (Autosomal recessive polycystic kidney disease). Está localizado en la región 21 del cromosoma 6 en humanos. El gen *PKHD1* contiene 67 exones y cubre una amplia región de 470kb. El mayor marco abierto de lectura (ORF) era de 12,222 pares de bases, pero puede haber otros productos debido al splicing alternativo (Harris and Torres, 2009; Onuchic et al., 2002).

El gen *PKHD1* codifica una proteína conocida como fibrocistina o poliductina, formada por 4074 aminoácidos. La fibrocistina tiene un rol fundamental en células de la cilia del epitelio tubular del riñón y en células hepáticas. Es una proteína transmembrana, estando parte en el interior celular y parte en el exterior celular, participando como mecanosensor y regulando la entrada de calcio. Es posible que interaccione con la policistina 2 (*PKD2*) y que intervenga en funciones de adhesión, repulsión y proliferación celular; sin embargo la función precisa y el rol de las diferentes isoformas es desconocido (Wang et al., 2007; Zhang et al., 2004).

## **4. OBJETIVOS**

En el Departamento de Patología Animal de la Universidad de Zaragoza se ha diagnosticado recientemente el primer caso de fibrosis congénita hepática en un potro de un mes de Pura Raza español. Tras la muerte del animal se observó en la necropsia un hígado hipertrofiado y fino, observándose en el estudio histológico una severa fibrosis difusa en puente porto portal y múltiples quistes en hígado y riñón.

Los resultados histológicos y clínicos fueron consistentes con una fibrosis congénita hepática por lo que basándonos en el estudio de Drögemüller et al. (2014) llevados a cabo en la raza de caballos “Frenches-Montagnes”, nos propusimos como objetivo general de este trabajo el estudio de los polimorfismos g.49,597,760A>T y g.49,630,834G>A, localizados en los exones 43 y 37 del gen *PKHD1*, en el potro Pura Raza Español afectado así como en animales emparentados y otros animales no emparentados de la misma raza .

Para conseguir el objetivo general del presente trabajo nos propusimos los siguientes objetos específicos:

1. Aislamiento y secuenciación de fragmentos que incluyan los SNPs g.49,597,760A>T y g.49,630,834G>A.
2. Búsqueda de polimorfismos en los fragmentos aislados.
3. Análisis de los polimorfismos en potro con fibrosis congénita hepática y animales emparentados.
4. Análisis de los polimorfismos en individuos no emparentados para el cálculo de frecuencias alélicas y genotípicas.
5. Comparación entre los resultados obtenidos en Pura Raza Español y los obtenidos en otras razas equinas.

## **5. MATERIALES Y MÉTODOS**

### **5.1. Extracción y cuantificación de DNA**

Para todos nuestros experimentos posteriores usaremos DNA obtenido a partir de muestras de sangre o de hígado. La extracción de DNA de calidad y concentración suficiente condicionará los posteriores experimentos.

#### **5.1.1. Obtención de DNA a partir de sangre.**

Todas las muestras de DNA se obtuvieron a partir de sangre, salvo la del potro afectado. Se utilizó un kit comercial de purificación mediante columnas con membrana de sílice de la casa GE Healthcare Life Sciences (Illustra Blood GenomicPrep Mini Spin Kit)

En primer lugar se incubaron las muestras de sangre con proteinasa K y un buffer óptimo para que la enzima pueda producir la lisis celular, durante 10'. Las muestras se colocaron en las columnas y se centrifugaron de forma que el DNA genómico se una selectivamente a las membranas de sílice.

Posteriormente se añadió de nuevo buffer de lisis y se volvió a centrifugar. A continuación se adicionó buffer de lavado y se hizo una centrifugación más larga, consiguiendo que solo quede DNA en la columna.

Por último, se añadió el buffer de elución precalentado a 70°C y por centrifugación se consiguió la separación del DNA genómico de la membrana.

#### **5.1.2. Purificación de DNA a partir de hígado**

La muestra de DNA del potro afectado se obtuvo de hígado proveniente de la necropsia y conservado en congelación.

Para la extracción de DNA se añadió nitrógeno líquido para disgregar el tejido lo máximo posible. Con estas muestras se realizó una incubación con proteinasa K en un buffer de lisis que disponía de las condiciones óptimas para la acción de la enzima. El DNA se obtuvo mediante precipitación con etanol absoluto.

#### **5.1.3. Cuantificación DNA**

Para la cuantificación y la detección de la calidad del DNA obtenido se utilizó un espectrofotómetro NanoDrop D-1000.

El espectrofotómetro Nanodrop utilizado mide la absorbancia de la muestra de DNA en tres puntos, a 230nm, 260nm y 280nm. El máximo de absorbancia del DNA es a 260 y con esa medida se calcula la concentración de DNA en ng/μl. La absorbancia a 230nm y 280nm son usadas para hacer las ratios de calidad  $A_{260nm}/A_{230nm}$  y  $A_{260nm}/A_{280nm}$ . La ratio de DNA puro  $A_{260nm}/A_{230nm}$  está en torno a 2, ratios menores indican una posible contaminación con EDTA, carbohidratos o urea. La ratio de DNA puro  $A_{260nm}/A_{280nm}$  se sitúa alrededor de 1.8, ratios menores pueden estar causadas por una contaminación con proteínas o grupos fenol.

El programa asociado ND-1000 DataViewer v3.3.0 permite ver los resultados y crear tablas con las concentraciones y los parámetros de calidad.

## **5.2. Verificación de parentesco mediante el uso de marcadores microsatélites**

Se usaron microsatélites multialélicos estandarizados por la ISAG para poder confirmar las relaciones de parentesco del potro con fibrosis congénita hepática.

### **5.2.1. Amplificación de microsatélites mediante multiplex-PCR**

Se realizaron 3 reacciones de amplificación múltiple que llamaremos M1, M2 y M3 para la amplificación de los 10 microsatélites incluidos en el estudio. En la reacción M1 se amplificaron los microsatélites AHT 4, HMS3, HMS6, HMS7, HTG4 y VHL20; en la reacción M2 se amplificaron los microsatélites HTG10 y HMS2; y en la reacción M3 se amplificaron los microsatélites ASB17 y ASB23.

Se utilizó el kit de amplificación múltiple, Quiagen Multiplex PCR. Los cebadores utilizados se marcaron con diferentes fluorocromos que permitieron su posterior visualización. En las Tablas S2, S3 y S4 del Anexo se muestran los componentes de la reacción M1, M2 y M3 y sus características.

Para la reacción de amplificación se utilizó el termociclador 2720 Thermal Cycler (Applied Biosystems) en el programa [95°C, 15' - 35x(95°C, 30'' - 58°C, 1' - 72°C, 1') - 72°C, 30' - 4°C, ∞] para M1, el programa [95°C, 15' - 35x(95°C, 30'' - 55°C, 1' - 72°C, 1') - 72°C, 30' - 4°C, ∞] para M2 y el programa [95°C, 15' - 35x(95°C, 30'' - 60°C, 1' - 72°C, 1') - 72°C, 30' - 4°C, ∞] para M3.

### **5.2.2 Analisis de los fragmentos y lectura de los genotipos**

El producto de PCR final se diluyó en 25µl de agua miliQ. Se tomó 1µl del producto diluido y se adicionaron 12µl de formamida Hi-Di y 0,3µl de Std 500 Liz (Applied Biosystems) como marcador interno de talla, finalmente se desnaturalizó 5 minutos a 95°C.

La electroforesis capilar se realizó usando como soporte polímero POP7 y el secuenciador ABI Prism 3130 (Applied Biosystem). Los genotipos fueron interpretados usando el programa GeneMapper v3.7 ambos desarrollados por Applied Biosystems.

### **5.2.3. Exclusión y asignación de paternidades**

Con los genotipos obtenidos, se realizaron las exclusiones manualmente y se confirmaron los parentescos, basándonos en la herencia codominante de todas los microsatélites analizados.

## **5.3. Aislamiento y secuenciación fragmentos del gen PKHD1**

Los polimorfismos g.49,597,760A>T y g.49,630,834G>A están localizados en el exón 43 y 37 del gen PKHD1 respectivamente. En adelante llamaremos a los fragmentos que contienen los polimorfismos fragmentos 43 y 37, por los exones que contienen.

El aislamiento y la posterior secuenciación de los fragmentos 43 y 37 amplificados nos permitió detectar los polimorfismos en dichas secuencias del gen *PKHD1* del potro afectado, animales emparentados y no emparentados en relación a la secuencia de referencia.

### **5.3.1 Búsqueda del gen PKHD1 en caballos**

Los SNPs descritos por Drögemüller et al. (2014) están anotados en una posición del cromosoma y no en la posición del gen *PKHD1*. Se localizó la secuencia del cromosoma 20 en las base de datos NCBI (<http://www.ncbi.nlm.nih.gov/>).

A partir de la secuencia, sabiendo la posición exacta del SNP en el cromosoma 20 (g.49,597,760 y g.49,630,834) se seleccionaron dos grandes regiones que serán usadas para el posterior diseño de cebadores que servirán para la amplificación de los fragmentos 43 y 37.

#### **5.3.1.1. Diseño de cebadores**

Los cebadores que diseñamos nos permitieron amplificar los fragmentos 43 y 37 que contienen los polimorfismos asociados a la fibrosis hepática congénita de acuerdo a la bibliografía existente (Drögemüller et al., 2014). Para el diseño de cebadores se utilizó el programa Primer Express Software v2.0 (Applied Biosystems) que nos permitió, agregando la secuencia, encontrar los cebadores que nos permitieran una óptima amplificación.

Las características óptimas para el diseño de unos cebadores cuyo objetivo final es la secuenciación del fragmento amplificado serán las siguientes:

- Longitud de las parejas de oligonucleótidos en torno a 20pb (pares de bases)
- Temperatura de fusión (melting) próxima a 60°C
- Evitar la complementariedad entre los cebadores directo y reverso
- Longitud de los fragmentos a amplificar alrededor de 400pb, dejando lo mas centrados posible los SNPs que queremos estudiar, ya que la secuenciación Sanger tiende a producir errores en los extremos

#### **5.3.1.2. Amplificación mediante PCR**

Se realizaron para cada muestra de DNA dos amplificaciones, una para amplificar el fragmento 43 y una segunda para amplificar el fragmento 37. Se prepararon dos mezclas de reacción, cuya única diferencia fueron los cebadores usados y cuyos componentes se muestran en la Tabla 1.

**Tabla 1:** Componentes de la reacción de amplificación de los fragmentos 43 y 37 para su posterior secuenciación.

Componentes	Volumen (µL)	Características
H <sub>2</sub> O	11,2	Agua purificada
PCR Buffer (10X)	2	[200 mM Tris HCl (pH 8.4), 500 mM KCl]
MgCl <sub>2</sub>	1	Iones Magnesio y Cloro [100mM]
dNTP's	2,5	Dideoxinucleótidos trifosfatos [10mM]
Cebador directo	0,5	[10mM]
Cebador inverso	0,5	[10mM]
Taq polimerasa	0,3	[5unidades/µl]
DNA	2	DNA purificado obtenido en el punto 4.1
Volumen final	20	

Para comprobar las amplificaciones se utilizaron 3 controles positivos de muestras de caballos del banco de DNA de LAGENBIO (Q-449, Q-450 y Q-451). La reacción de amplificación se llevo a cabo en un termociclador 2720 Thermal Cycler (Applied Biosystems®) utilizando el programa: [95°C, 5' - 30X ( 95°C,30''-60°C,30''-72°C, 30'')-4°C, ∞].

Tras la reacción de amplificación se comprobó si se había producido amplificación mediante una electroforesis en gel de agarosa. Se preparó un gel de Agarosa al 1% y se utilizó un marcador de talla de 100pb para comprobar que la talla de los fragmentos amplificados era la esperada.

### **5.3.1.3 Purificación de los productos de PCR**

Una vez comprobada la amplificación se realizó una purificación del producto de amplificación necesaria para que la posterior reacción de secuenciación se realice perfectamente. En 10µl de producto de PCR se añadió 1µL del producto ExoSAP (USB) que contiene fosfatasa alcalina y exonucleasa para degradar los cebadores y desfosforilar los dNTP's. Siguiendo el protocolo se incubó 45' a 37°C, con una posterior inactivación a 80°C durante 15'.

### **5.3.2 Reacción de secuenciación**

Para la reacción de secuenciación usaremos el kit comercial BigDye® Terminator v3.1 CycleSequencing Kit (AppliedBiosystems).

La reacción se llevó a cabo en un termociclador 2720 Thermal Cycler (Applied Biosystems®) en un programa [96°C. 1' - 25X(96°C,10''-50°C,5''- 60°C, 4'')-4°C, ∞]. Para cada producto de amplificación se realizó una reacción cuyos componentes están detallados en la Tabla 2.

**Tabla 2:** Componentes de la reacción de secuenciación de los fragmentos 43 y 37

Componentes	Volumen (µL)
<b>BigDye v3.1 ReactionMix</b>	11,2
<b>Cebador [10mM]</b>	0,5
<b>H<sub>2</sub>O miliQ</b>	4,5
<b>Producto de PCR</b>	3
<b>Volumen final</b>	10

#### **5.3.2.1. Purificación de la reacción de secuenciación y electroforesis capilar**

Se llevó a cabo una purificación con el kit de purificación BigDye XTerminator Purification (Applied Biosystems) siguiendo el protocolo del fabricante. La electroforesis capilar se llevó a cabo empleando el soporte polímero POP7 y como secuenciador ABI Prism 3130 (Applied Biosystem).

Esta información es transformada mediante un programa informático (Sequencing Analysis 5.2) que crea archivos .AB1 en los que se pueden observar los picos de fluorescencia producidos correspondientes a cada fluorocromo junto con la longitud del fragmento.

## 5.4. Búsqueda de SNPs

Se revisaron cada una de las secuencias para la detección de la variabilidad utilizando el software BioEdit que permite analizar los archivos .AB1.

### 5.4.1. Alineamientos mediante ClustalW

Se alinearon las secuencias con la herramienta ClustalW Multiple Alignment del programa BioEdit. Con esta herramienta se pueden detectar más fácilmente los polimorfismos, y saber a qué posición del gen corresponde así como comparar con la secuencia encontrada en la base de datos NCBI.

### 5.4.2. Estimación de las frecuencias alélicas y genotípicas.

Para poder estimar la proporción de diferentes polimorfismos en la población, se realizó el cálculo de frecuencias alélicas y genotípicas.

La frecuencia alélica representa el porcentaje de cada alelo en la población y se calcula usando la Fórmula 1. Las frecuencias genotípicas son el porcentaje de cada genotipo en la población y se calculan usando la Fórmula 2. Las frecuencias alélicas miden la cantidad de variabilidad genética en una población mientras que las frecuencias genotípicas muestran de qué manera se distribuye entre sus miembros (Sadava, 2009).

$$F. \text{alélica } (A) = \frac{2 \cdot N(AA) + N(Aa)}{2N} \quad F. \text{alélica } (a) = \frac{2 \cdot N(aa) + N(Aa)}{2N}$$

**Fórmula 1:** Fórmulas para el cálculo de frecuencias alélicas

$$F. \text{genotípicas } (AA) = \frac{N(AA)}{N} \quad F. \text{genotípicas } (Aa) = \frac{N(Aa)}{N} \quad F. \text{genotípicas } (aa) = \frac{N(aa)}{N}$$

**Fórmula 2:** Fórmulas para el cálculo de frecuencias genotípicas

## 6. RESULTADOS Y DISCUSIÓN

### 6.1 Cantidad y calidad del DNA extraído

En la Tabla S5 del Anexo se muestra la tabla que incluye los resultados de las concentraciones y las ratios  $A_{260nm}/A_{230nm}$  y  $A_{260nm}/A_{280nm}$ . Para el potro problema obtuvimos una concentración de 155ng/ $\mu$ l y para el resto de caballos cuyo DNA se extrajo en columnas, una concentración media de 15,9ng/ $\mu$ l.

### 6.2. Verificación de las relaciones de parentesco

En la Tabla 3 se muestran los resultados obtenidos para los diferentes microsatélites utilizados para la verificación de las relaciones de parentesco.

**Tabla 3:** Genotipos para los 10 microsatélites analizado: AHT4, ASB2, ASB17, ASB23, HMS2, HMS3, HMS6, HMS7, HTG4 y VHL20.

	AHT4		ASB2		ASB17		ASB23		HMS2	
Potro afectado	152	154	236	244	103	107	204	208	H	K
Madre(Luna)	152	152	236	246	103	111	184	204	K	R
Padre (Mas)	154	154	244	252	107	107	204	208	H	H
Abuela (Lozana)	152	154	236	246	103	117	184	186	K	K
Hermana de la madre	152	154	240	246	103	117	184	184	K	K
Potro hermano	152	154	236	252	107	111	204	208	H	R
	HMS3		HMS6		HMS7		HTG4		VHL20	
Potro afectado	166	166	159	161	181	181	131	133	97	103
Madre (Luna)	162	166	159	169	181	181	131	131	97	101
Padre (Mas)	160	166	161	169	181	183	133	133	103	107
Abuela (Lozana)	166	166	159	169	179	181	131	133	97	99
Hermana de la madre	166	166	159	169	179	181	131	133	99	99
Potro hermano	160	162	169	169	181	183	131	133	97	103

A partir de los genotipos obtenidos para los 10 microsatélites se compararan las relaciones de parentesco. Con 10 microsatélites realizados, hay un PE1 (probabilidad de exclusión asumiendo un progenitor conocido) mayor de 0.99 y un PE2 (probabilidad de exclusión asumiendo dos progenitores conocidos) mayor del 0.999. Podemos confirmar con buena fiabilidad las relaciones de parentesco (van de Goor et al., 2011).

### 6.3. Aislamiento del gen PKHD1 en la especie equina

Se aislaron los fragmentos 43 y 37 que contienen los SNPs descritos por Drögemüller et al. (2014) y asociados a la fibrosis congénita hepática en la raza “Franches-Montagnes” (SNP g.49.597.760 A>T y SNP g.49.630.834 G>A).

Para el aislamiento y posterior análisis de la presencia de variaciones en los animales afectados se partió de la secuencia del cromosoma 20 depositada en la página NCBI (Accession Number: NW\_001867389).



Una vez localizados los dos SNPs descritos por Drögemüller et al. (2014) se seleccionaron dos regiones de 800pb que contienen dichos SNPs en la parte central de la misma, que serán usadas para el diseño de cebadores:

- SNP g.49.597.760 A>T (exón 43): región Chr 20: (49.598.160- 49.597.360)
- SNP g.49.630.834 G>A (exón 37): región Chr 20: (49.631.234- 49.630.434)

Utilizando el programa Primer Express Software v2.0 (Applied Biosystems) se introdujeron estas dos regiones de 800 pb en FASTA. Se seleccionaron dos cebadores para la amplificación de cada uno de los fragmentos que mejor se ajustaban a las características óptimas de los mismos (Ver Apartado 4.2.2). En la Tabla 4 se muestran los cebadores diseñados para la amplificación de los fragmentos 43 y 37.

**Tabla 4:** Características y secuencias de los cebadores usados para la amplificación por PCR de los fragmentos 43 y 37

Nombre fragmento	Secuencia del cebador	Longitud del cebador (pb)	Tª de fusión (°C)	Longitud fragmento (pb)	Localización en el cromosoma
Fragmento 43	Directo 5'- CACTGAAGCCTCCACTCCAAA- 3'	21	59	403	Chr20: 49597586- 49597988
	Reverso 5'- GGCCTGACAGCAGAATGTTAGAT-3'	23	59		
Fragmento 37	Directo 5'- CTCTGCCACGGGAATTACAAC- 3'	21	59	351	Chr20: 49630643- 49630993
	Reverso 5'- TCCATCCTTGTTTCCCATGG-3'	20	59		

Las secuencias correspondientes a los fragmentos 43 y 37 del gen *PKHD1* que fueron amplificados se muestran a continuación.

```

1  cactgaagcc tccactccaa acaacttccg gagtggacat tatgccctct caactctgac
61  ctccctggca gagctgcaga ccttacctcc acaacattgg tggggttgcg gatgtagatg
121 ccagatggtg tcagcacttc cggactggag agtcctcag caccggaaac acggAtgatac
181 acattgtttc ttattatact tccctgttct gactcgtcta attgaagaaa agatacatca
241 aacagagatg atttacaac atagttaaaa tttatacaac tcaacactct tctttctgct
301 aatacacatt acatttcaca taatgcacat atgtgtatgt atatatataa aatatttagt
361 atgttcaaca aatctgagat atctaacatt ctgctgtcag gcc

```

**Secuencia 1:** Fragmento 43, correspondiente a la región 49597586-49597988 del cromosoma 20 equino, código de acceso NCBI: NW\_001867389.1. En mayúscula, negrita y subrayado la posición del SNP g.49,597,760A>T

```

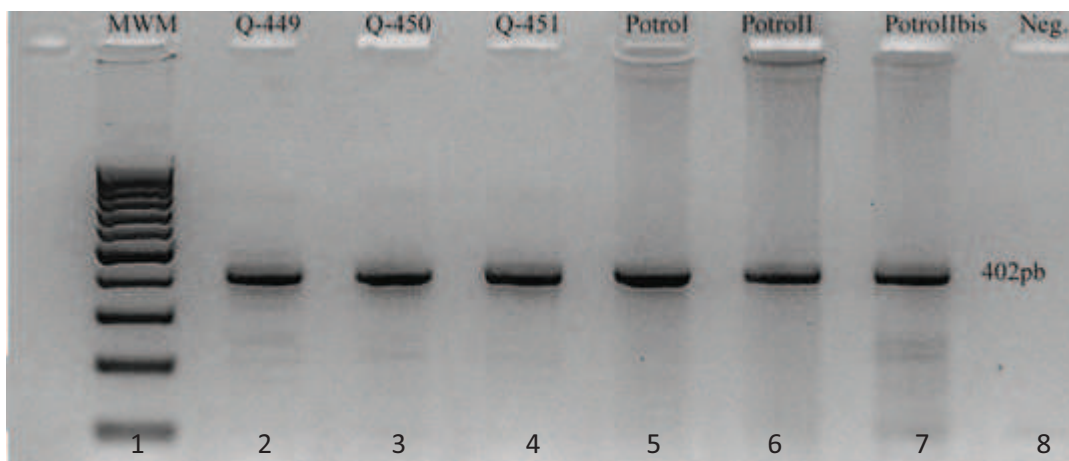
1  ctctgccacg ggaattacaa caattaatat ttctttttta aggaaatagt gcctgtgctc
61  tgataaactt tgttgcctca acaactaaca gccaaagagt ttttaagtagt tttgccattt
121 tccccggttg ccacacattg gctagacaca gctctcccc aactcctctg gccacggcc
181 tactcaccg tGcagggaca gggttccatt cctcacagcc aggaacttga caccatatgg
241 aaagaagggg gtagagtggg aactcccgtg gagtttgatc tcagctttgc cttggaaggg
301 cttgtccttg gacccaatcc ggagctctcc accatgggaa acaaggatgg a

```

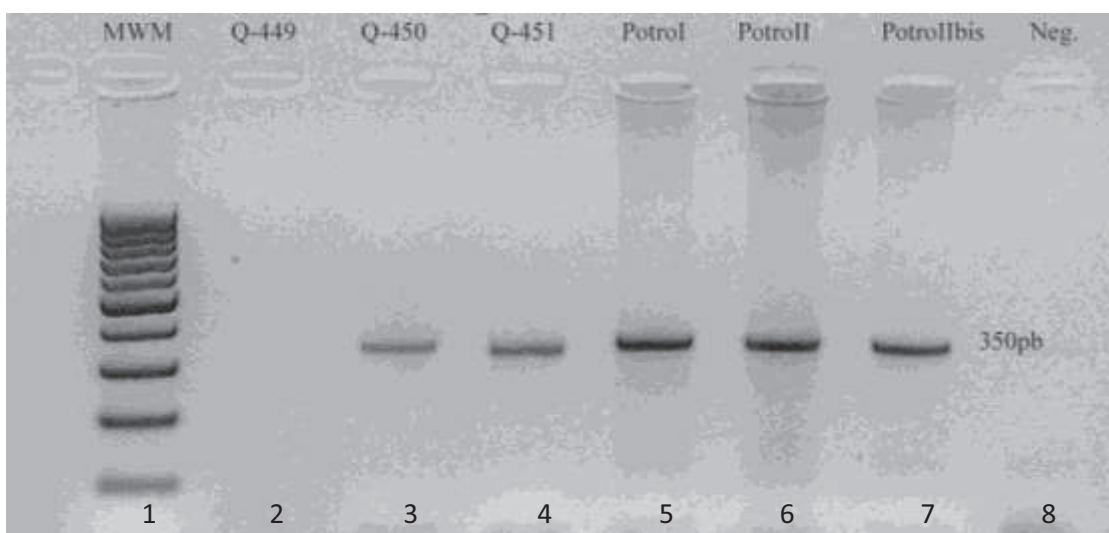
**Secuencia 2:** Fragmento 37, correspondiente a la región 49630643-49630993 del cromosoma 20 equino, código de acceso NCBI: NW\_001867389.1. En mayúscula, negrita y subrayado la posición del SNP g.49,630,834G>A

Se comprobó mediante gel de agarosa al 1% que la amplificación de ambos fragmentos se correspondía con el tamaño esperado. A modo de ejemplo se muestran a continuación (Figura 2

y 3) los resultados obtenidos a partir de las muestras de animales utilizados como control positivo y del potro afectado.



**Figura 2:** Electroforesis en gel de agarosa de la amplificación del Fragmento 43 (1: MWM 100pb; 2, 3 y 4: caballos Q-449, Q-450 y Q-451 como controles positivos; 5, 6, 7: potro afectado; 8: control negativo).



**Figura 3:** Electroforesis en gel de agarosa de la amplificación del Fragmento 37 (1: MWM 100pb; 2, 3 y 4: caballos Q-449, Q-450 y Q-451 como controles positivo; 5, 6, 7: potro afectado; 8: control negativo).

## 6.4. Análisis de las secuencias obtenidas.

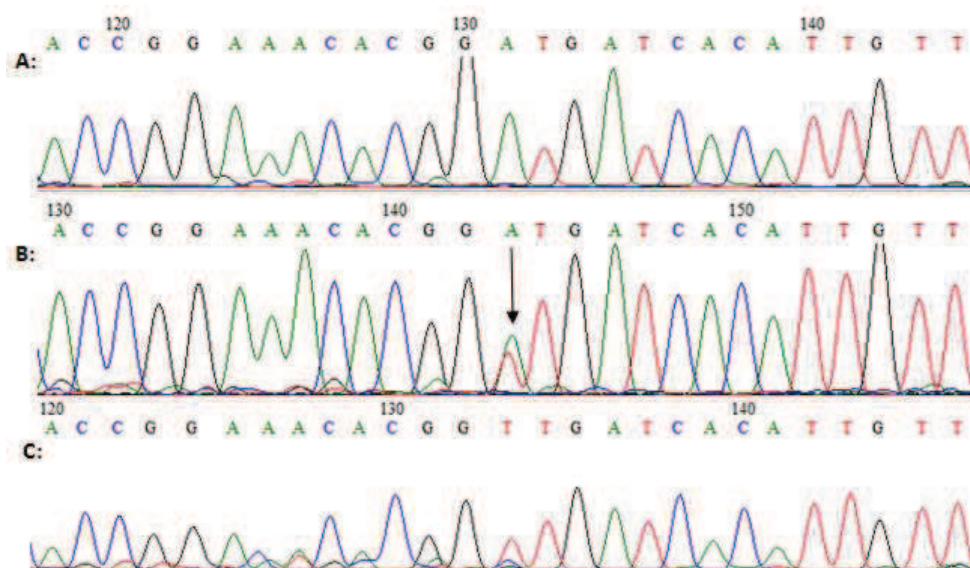
Se utilizó el software Bioedit para el análisis de secuencias.

### 6.4.1. Análisis de SNPs asociados con ARPKD

Se analizó la posición correspondiente al SNP g.49,597,760 A>T en el Fragmento 43. Los resultados se muestran en la Figura 2, procedente de la secuencia del potro afectado. Se observó que era heterocigoto A/T para el SNP g.49,597,760 A>T.

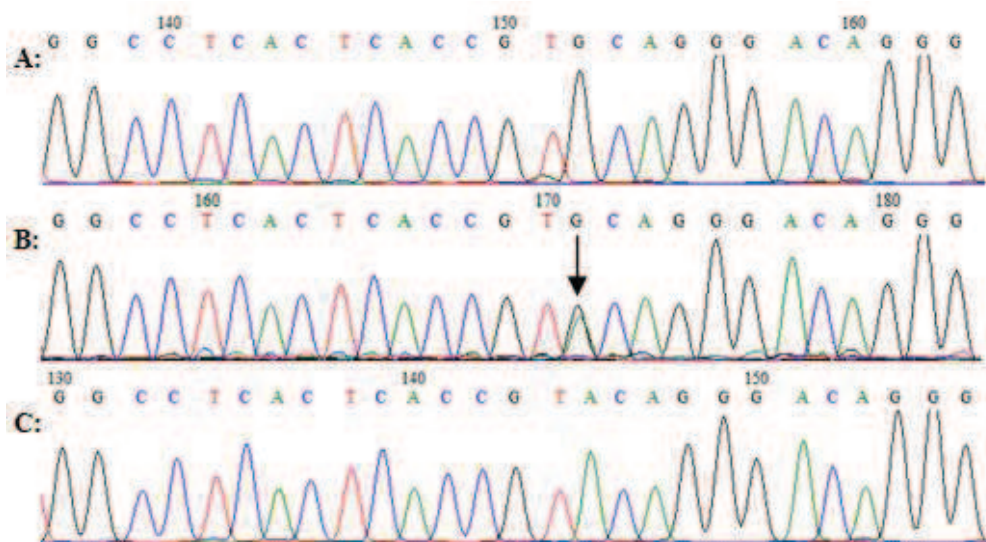
En las Figuras 4 y 5 se muestran ejemplos de homocigosis y heterocigosis correspondientes a los SNPs analizados. Los animales homocigotos para cada una de las bases

mostrarían solo uno de los dos picos mientras que los animales heterocigotos muestran dos picos superpuestos.



**Figura 4:** Parte de la secuencia del Fragmento 43, se muestra el SNP g.49,597,760A>T en los tres posibles casos, **A:** homocigosis AA, **B:** heterocigosis AT, **C:** homocigosis TT . La flecha señala los dos picos correspondientes a la heterocigosis en el SNP g.49,597,760A>T.

Se analizó la posición correspondiente al SNP g.49,630,834 G>A en el Fragmento 37. Los resultados se muestran en la Figura 3, procedente de la secuencia del potro afectado. Se observó que era heterocigoto G/A para el SNP g.49,630,834 G>A .



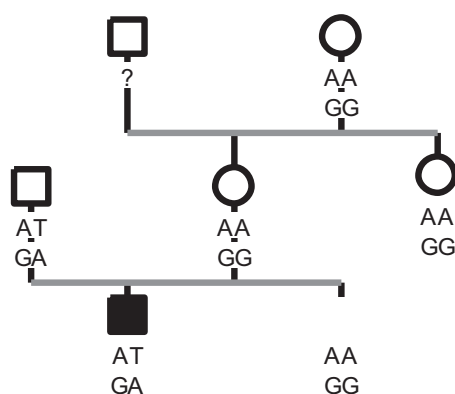
**Figura 5:** Parte de la secuencia del Fragmento 37, se muestra el SNP g.49,630,834 G>A en los tres posibles casos, **A:** homocigosis GG, **B:** heterocigosis GA, **C:** homocigosis AA .La flecha señala los dos picos correspondientes a la heterocigosis en el SNP g.49,630,834 G>A.

Los resultados obtenidos para el resto de caballos analizados, tanto parientes del potro afectado, como caballos no emparentados Pura Raza español se muestran en la Tabla 5.

**Tabla 5:** Genotipos de los SNPs g.49,597,760A>T y g.49,630,834G>A del potro afectado y parientes y de otros caballos de la misma raza no emparentados

SNPs:			SNPs:		
Nombre	g.49,597,760 A>T	g.49,630,834 G>A	Nombre	g.49,597,760 A>T	g.49,630,834 G>A
Sec. referencia	AA	GG	Q-776	AT	GA
Potro problema	AT	GA	Q-777	AA	GG
Madre Juna	AA	GG	Q-778	AA	GG
Abuela	AA	--	Q-779	AT	GA
Padre nas	AT	GA	Q-780	AA	GG
Hermana madre	AA	GG	Q-781	AA	GG
Potro hermano	AA	GG	Q-782	AA	GG
Q-758	AA	GG	Q-783	AA	--
Q-760	TT	--	Q-784	AA	GG
Q-761	AA	AA	Q-785	AA	GG
Q-762	TT	--	Q-786	AA	GG
Q-763	AA	--	Q-787	TT	AA
Q-764	AA	GG	Q-788	AT	GA
Q-765	TA	--	Q-789	AA	GG
Q-766	AT	--	Q-790	AT	GA
Q-767	AA	GG	Q-791	AT	GA
Q-768	AT	--	Q-792	AT	GA
Q-769	AT	GA	Q-793	TT	GG
Q-770	AA	--	Q-794	AT	GA
Q-771	AT	GA	Q-795	--	GA
Q-772	AT	GA	Q-796	--	GA
Q-773	AA	GG	Q-797	AA	GG
Q-774	AA	GG	Q-798	TA	AA
Q-775	AT	GA			

A partir de los genotipos de los individuos emparentados y sabiendo que las relaciones de parentesco son fiables se estableció el pedigrí. (Ver figura 6)

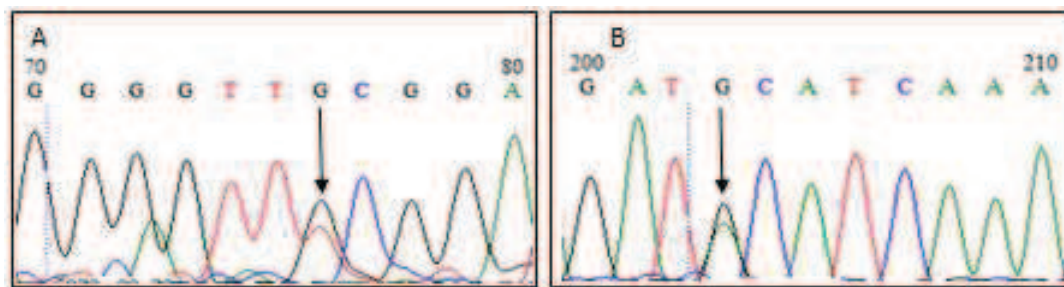


**Figura 6:** Pedigrí del potro afectado con sus parientes. Machos representados como cuadrados y hembras como círculos. El cuadrado negro representa al potro problema. Los genotipos correspondientes al SNP g.49,597,760 A>T están en la primera línea y los genotipos del SNP g.49,630,834 G>A en la segunda línea.



#### 6.4.2. Análisis de nuevos SNPs

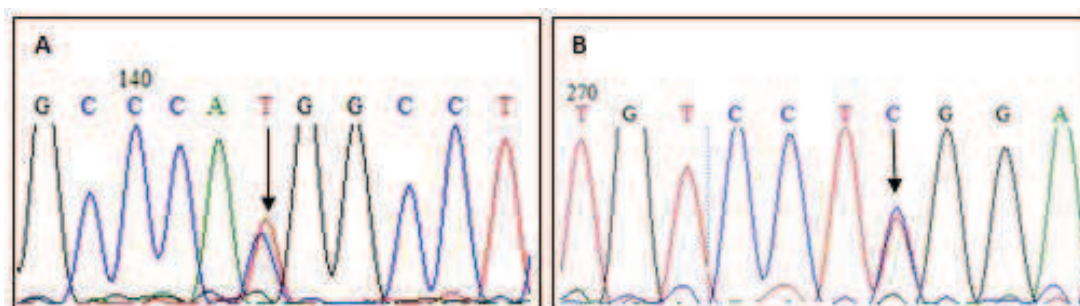
En los Fragmentos 43 y 37 se analizó la presencia de otros posibles SNPs. En el fragmento 43 se detectaron dos nuevos polimorfismos no descritos hasta el momento, Figuras 7 y 8.



**Figura 7:** *A: Parte de la secuencia del fragmento 43 del potro afectado. La flecha señala una heterocigosis para el nuevo SNP g.49.597.693G>T. B: Parte de la secuencia del fragmento 43 del potro afectado. La flecha señala una heterocigosis para el nuevo SNP g.49.597.820A>G.*

Los nuevos SNPs se anotaron en comparación con la secuencia de referencia del NCBI NW\_001867389.1. Ambos polimorfismos, g.49.597.693G>T y g.49.597.820A>G, se encontraron en heterocigosis en el potro afectado.

En el Fragmento 37 se detectaron 2 nuevos polimorfismos no descritos hasta el momento, que se muestran en las Figura 8.



**Figura 8:** *A: Parte de la secuencia del fragmento 37 del potro afectado. La flecha señala una heterocigosis para el nuevo SNP g.49.630.818C>T. B: Parte de la secuencia del fragmento 37 del potro afectado. La flecha señala una heterocigosis para el nuevo SNP g.49.630.951T>C.*

Los nuevos SNPs se anotaron en comparación con la secuencia de referencia del NCBI NW\_001867389.1. Ambos polimorfismos, g.49.630.818C>T y g.49.630.951T>C, se encontraron en heterocigosis en el potro afectado.

Los cuatro nuevos polimorfismos detectados fueron analizados en todos los animales incluidos en el estudio observándose para ellos los genotipos que se muestran a continuación.

**Tabla 6:** Genotipos de los cuatro polimorfismos nuevos, g.49.597.693G>T, g.49.597.820 A>G, g.49.630.818C>T y g.49.630.951T>C del potro afectados y familiares y de otros caballos de la misma raza no emparentados

Nombre	SNPs Fragmento 43:		SNPs Fragmento 37:	
	g.49.597.693G>T	g.49.597.820 A>G	g.49.630.818C>T	g.49.630.951T>C
Sec. referencia	GG	AA	CC	TT
Potro problema	GT	AG	TT	CC
Madre Juna	GT	AG	CT	TC
Abuela	GG	AA	CT	TC
Padre Nas	GG	AA	TT	CC
Hermana madre	GG	AA	TT	CC
Potro hermano	GG	AA	CT	TC
Q-758	GG	AA	CC	TT
Q-760	GG	AA	-	-
Q-761	GG	AA	TT	CC
Q-762	GG	AA	-	-
Q-763	GG	AA	-	-
Q-764	GG	AA	CT	TC
Q-765	GG	AA	-	-
Q-766	GG	AA	-	-
Q-767	GG	AA	CT	TC
Q-768	GG	AA	-	-
Q-769	GG	AA	CT	TC
Q-770	GG	AA	-	-
Q-771	GG	AA	CT	TC
Q-772	GG	AA	CT	TC
Q-773	GG	AA	TT	CC
Q-774	GG	AA	CT	TC
Q-775	GG	AA	CT	TC
Q-776	GG	AA	CT	TC
Q-777	GG	AA	TT	CC
Q-778	GG	AA	CC	TT
Q-779	GG	AA	CT	TC
Q-780	GG	AA	CT	TC
Q-781	GG	AA	CT	TC
Q-782	GG	AA	CT	TC
Q-783	GT	GA	-	-
Q-784	GT	GA	CT	TC
Q-785	GT	GA	TT	CC
Q-786	GG	AA	CC	TT
Q-787	GG	AA	TT	CC
Q-788	GG	AA	CT	TC
Q-789	GG	AA	CT	TC
Q-790	GT	GA	TT	CC
Q-791	GT	GA	TT	CC
Q-792	GG	AA	CT	CT
Q-793	GG	AA	TT	CC
Q-794	GT	GA	TT	CC
Q-795	-	-	CT	CT
Q-796	-	-	CT	TC
Q-797	GG	AA	CC	TT
Q-798	GT	GA	TT	CC

### 6.4.3. Análisis de las frecuencias alélicas y genotípicas

Se calcularon las frecuencias alélicas y genotípicas de los SNPs g.49.597.760 A>T y g.49.630.834 G>A para la Pura Raza Español, la raza Franches-Montagnes y una población de varias razas, que nos servirán para contrastar los resultados.

Para el cálculo de frecuencias de Pura Raza Español se usaron los genotipos de los animales no emparentados (Tabla 5) y se compararon con las frecuencias de la raza Franches-Montagnes y una población de varias razas una población de caballos genotipada en el artículo de Drögemüller et al. (2014)

**Tabla 7:** Frecuencias genotípicas y alélicas del polimorfismo g,49,597,760T>A para la raza Franches Montagnes, Pura Raza Español y caballos de otras razas.

SNP g,49,597,760A>T					
Raza	Frecuencias genotípicas			Frecuencias alélicas	
	AA	AT	TT	A	T
Franches Montagnes	0,834	0,166	$3,93 \times 10^{-4}$	0,917	0,083
Caballos otras razas	0,803	0,192	0,004	0,899	0,1
Pura Raza Española	0,5	0,395	0,105	0,697	0,303

En la Tabla 7 se observa que la frecuencia del alelo T es tres veces superior en la raza Pura Raza Español en comparación con el resto de razas y una presencia en la población de 10% de animales homocigotos (TT) que según el estudio realizado por Drögemüller et al. (2014) estaría fuertemente relacionado con la fibrosis hepática congénita. La presencia de este genotipo resulta ser aparentemente 250 veces superior en la raza Pura Raza Español en comparación con la raza Franches-Montagnes.

**Tabla 8:** Frecuencias genotípicas y alélicas del polimorfismo g,49,597,760T>A para la raza Franches Montagnes, Pura Raza Español y caballos de otras razas.

SNP g,49,630,834G>A					
Raza	Frecuencias genotípicas			Frecuencias alélicas	
	GG	GA	AA	G	A
Franches Montagnes	0,843	0,145	0,012	0,916	0,084
Caballos otras razas	0,779	0,206	0,015	0,882	0,118
Pura raza española	0,5	0,406	0,094	0,703	0,296

En la Tabla 8 se observan diferencias en las frecuencias genotípicas y alélicas, la frecuencia del alelo A en la raza Pura Raza Español está mucho más presente en la población que en las otras razas. Casi un 10% de la población es homocigota (AA), que según el estudio realizado por Drögemüller et al. (2014) estaría fuertemente asociado con la fibrosis hepática congénita.

A pesar de que la presencia de homocigotos para el alelo asociado a la enfermedad en ambos polimorfismos es del 10%, cabe destacar que el caso detectado con fibrosis hepática que es objeto de estudio del presente trabajo es el primer caso de esta enfermedad descrito en la Raza Pura Sangre Español. Este hecho podría sugerir que estos SNPs no están asociados a esta

enfermedad en la raza Pura Raza Español, a diferencia de lo que ocurre en la raza Franches-Montagnes

Aunque en la raza Franches-Montagnes ambos polimorfismos estén fuertemente asociados con la enfermedad, el hecho de que en nuestra raza haya un 10% de homocigotos para el alelo relacionado con la enfermedad y que sólo se haya diagnosticado un caso de fibrosis congénita hepática nos hace sospechar que estos cambios en la proteína (p.H2038Y y p.I2282N) codificada por el gen *PKHD1* puedan no ser la causa de la enfermedad.

Los resultados obtenidos en el potro afectado podrían indicar que los cambios aminoacídicos generados por los dos polimorfismos no originan una pérdida de función de la proteína fibrocistina que sea la causante de la enfermedad, al contrario de lo sugerido por Drögemüller et al. (2014).

Los resultados obtenidos en el presente Trabajo Fin de Grado junto con los realizados por el Departamento de Patología Animal han sido presentados al 33rd Annual Meeting of The European Society of Pathology & 26th Annual Meeting of The European College of Veterinary Pathologists (ESVP-ECVP annual meeting 2015) celebrado en Helsinki del 2 al 5 de septiembre obteniendo el premio al mejor poster (ver Abstract, Poster y Certificado en Anexo)



## **7. CONCLUSIONES**

1. El potro afectado es heterocigoto para los SNPs g.49.597.760 A>T y g.49.630.834 G>A.
2. La frecuencia de los alelos asociados con la enfermedad para los SNPs g.49.597.760 A>T y g.49.630.834 G>A es tres veces mayor para la población de caballos Pura Raza Española que en la raza Franches-Montagnes.
3. La presencia del genotipo TT (SNP g.49.597.760 A>T) es aproximadamente 250 veces superior en la raza Pura Raza Español que en la raza Franches-Montagnes.
4. No se encontró una asociación entre los polimorfismos analizados y la fibrosis hepática congénita en la raza Pura Raza Español.
5. La presencia de un 10% de homocigotos para los polimorfismos asociados con la enfermedad en la población sugiere que los cambios proteicos p.H2038Y y p.I2282N en la proteína fibrocistina, causados por los SNPs g.49.597.760 A>T y g.49.630.834 G>A no sean la causa de la enfermedad.
6. Se han descrito cuatro nuevos polimorfismos en el caballo Pura Raza Español que no habían sido descritos hasta el momento en las razas analizadas.

## **Conclusions**

1. The affected foal is heterozygote for both SNPs g.49.597.760 A>T and g.49.630.834 G>A.
2. The frequency for the SNP's alleles associated with the disease (g.49.597.760 A>T and g.49.630.834 G>A) is three times higher in the Purebred Spanish Horse population than the Franches-Montagnes horse breed.
3. The frequency of the TT genotype (SNP g.49.597.760 A>T) is approximately 250 times higher in the Purebred Spanish Horse breed than the Franches-Montagnes breed.
4. No evidence of association was found between the analyzed polymorphisms and the congenital hepatic fibrosis in the Purebred Spanish Horse.
5. There is a 10% of homozygous in the population for the polymorphisms associated with the disease. This result suggests that the protein changes in the fibrocystin protein, p.H2038Y and p.I2282N, caused by the SNPs g.49.597.760 A> T and g.49.630.83 G>A are not the cause of the disease.
6. Four novel polymorphisms have been described in the Purebred Spanish Horse. These polymorphisms hadn't been described in the analyzed breeds so far.

## **8. Bibliografía**

- Brookes, A.J. (1999). The essence of SNPs. *Gene* 234, 177–186.
- Chistiakov, D.A., Hellemans, B., and Volckaert, F.A.M. (2006). Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* 255, 1–29.
- Choi, S.-K., Cho, C.-Y., Yeon, S.-H., Cho, B.-W., and Cho, G.-J. (2008). Genetic characterization and polymorphisms for parentage testing of the Jeju horse using 20 microsatellite loci. *J. Vet. Med. Sci.* 70, 1111–1115.
- Drögemüller, M., Jagannathan, V., Welle, M.M., Graubner, C., Straub, R., Gerber, V., Burger, D., Signer-Hasler, H., Poncet, P.-A., Klopfenstein, S., et al. (2014). Congenital hepatic fibrosis in the Franches-Montagnes horse is associated with the polycystic kidney and hepatic disease 1 (PKHD1) gene. *PLoS One* 9, e110125.
- Forbes, S.H., Hogg, J.T., Buchanan, F.C., Crawford, A.M., and Allendorf, F.W. (1995). Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Mol. Biol. Evol.* 12, 1106–1113.
- Van de Goor, L.H.P., Panneman, H., and van Haeringen, W.A. (2010). A proposal for standardization in forensic equine DNA typing: allele nomenclature for 17 equine-specific STR loci. *Anim. Genet.* 41, 122–127.
- Van de Goor, L.H.P., van Haeringen, W.A., and Lenstra, J.A. (2011). Population studies of 17 equine STR for forensic and phylogenetic analysis. *Anim. Genet.* 42, 627–633.
- Haechler, S., Van den Ingh, T.S., Rogivue, C., Ehrensperger, F., and Welle, M. (2000). Congenital hepatic fibrosis and cystic bile duct formation in Swiss Freiberger horses. *Vet. Pathol.* 37, 669–671.
- Harris, P.C., and Torres, V.E. (2009). Polycystic kidney disease. *Annu. Rev. Med.* 60, 321–337.
- Herráeza, D.L., Schäfer, H., Mosner, J., Fries, H.-R., and Wink, M. (2005). Comparison of microsatellite and single nucleotide polymorphism markers for the genetic analysis of a Galloway cattle population. *Z. Naturforsch. C.* 60, 637–643.
- Holt, C.L., Stauffer, C., Wallin, J.M., Lazaruk, K.D., Nguyen, T., Budowle, B., and Walsh, P.S. (2000). Practical applications of genotypic surveys for forensic STR testing. *Forensic Sci. Int.* 112, 91–109.
- Inagaki, S., Yamamoto, Y., Doi, Y., Takata, T., Ishikawa, T., Yoshitome, K., Miyaishi, S., and Ishizu, H. (2002). Typing of Y chromosome single nucleotide polymorphisms in a Japanese population by a multiplexed single nucleotide primer extension reaction. *Leg. Med. (Tokyo)* 4, 202–206.
- Liu, Z.J., and Cordes, J.F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238, 1–37.
- Mullis, K.B., and Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155, 335–350.

- Onuchic, L.F., Furu, L., Nagasawa, Y., Hou, X., Eggermann, T., Ren, Z., Bergmann, C., Senderek, J., Esquivel, E., Zeltner, R., et al. (2002). PKHD1, the polycystic kidney and hepatic disease 1 gene, encodes a novel large protein containing multiple immunoglobulin-like plexin-transcription-factor domains and parallel beta-helix 1 repeats. *Am. J. Hum. Genet.* *70*, 1305–1317.
- Pakstis, A.J., Speed, W.C., Fang, R., Hyland, F.C.L., Furtado, M.R., Kidd, J.R., and Kidd, K.K. (2010). SNPs for a universal individual identification panel. *Hum. Genet.* *127*, 315–324.
- Pastinen, T., Partanen, J., and Syvänen, A.C. (1996). Multiplex, fluorescent, solid-phase minisequencing for efficient screening of DNA sequence variation. *Clin. Chem.* *42*, 1391–1397.
- Ronaghi, M., Uhlén, M., and Nyrén, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* *281*, 363, 365.
- Sadava, D. (2009). Vida (Ed. Médica Panamericana).
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* *74*, 5463–5467.
- Schena, M., Heller, R.A., Thériault, T.P., Konrad, K., Lachenmeier, E., and Davis, R.W. (1998). Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol.* *16*, 301–306.
- Twito, T., Weigend, S., Blum, S., Granevitze, Z., Feldman, M.W., Perl-Treves, R., Lavi, U., and Hillel, J. (2007). Biodiversity of 20 chicken breeds assessed by SNPs located in gene regions. *Cytogenet. Genome Res.* *117*, 319–326.
- Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., et al. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* *280*, 1077–1082.
- Wang, S., Zhang, J., Nauli, S.M., Li, X., Starremans, P.G., Luo, Y., Roberts, K.A., and Zhou, J. (2007). Fibrocystin/polyductin, found in the same protein complex with polycystin-2, regulates calcium responses in kidney epithelia. *Mol. Cell. Biol.* *27*, 3241–3252.
- Zhang, M.-Z., Mai, W., Li, C., Cho, S., Hao, C., Moeckel, G., Zhao, R., Kim, I., Wang, J., Xiong, H., et al. (2004). PKHD1 protein encoded by the gene for autosomal recessive polycystic kidney disease associates with basal bodies and primary cilia in renal epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 2311–2316.