



# CARACTERIZACIÓN GENÉTICA DE LA RAZA CHATO MURCIANO

---

UNIVERSIDAD DE MURCIA  
FACULTAD DE VETERINARIA

**Juan Manuel Herrero Medrano**

**2012**



D. GUILLERMO RAMIS VIDAL, Profesor Contratado Doctor del Área de PRODUCCIÓN ANIMAL y Presidente de la Comisión Académica del Programa de Doctorado Porcinocultura Profesional y Científica, INFORMA:

Que la Tesis Doctoral titulada “CARACTERIZACIÓN GENÉTICA DE LA RAZA CHATO MURCIANO”, ha sido realizada por D JUAN MANUEL HERRERO MEDRANO, bajo la inmediata dirección y supervisión del Dr. Guillermo Ramis Vidal, el Dr. Antonio Muñoz Luna y el Dr. Richard Crooijman y que la Comisión Académica ha dado su conformidad para que sea presentada ante la Comisión de Doctorado.

En Murcia, a 29 de mayo de 2012

Fdo:

Prof. Dr. Guillermo Ramis Vidal

**Facultad de Veterinaria**

Campus de Espinardo 30100 Murcia

T. 868 883 905 – F. 868 884 147– [www.um.es /veterina/](http://www.um.es/veterina/)



D. GUILLERMO RAMIS VIDAL, Profesor Contratado Doctor del Área de  
PRODUCCIÓN ANIMAL AUTORIZA:

La presentación de la Tesis Doctoral titulada “CARACTERIZACIÓN GENÉTICA DE  
LA RAZA CHATO MURCIANO”, realizada por D JUAN MANUEL HERRERO MEDRANO,  
bajo mi inmediata dirección y supervisión, y que presenta para la obtención del grado  
de Doctor por la Universidad de Murcia.

En Murcia, a 29 de mayo de 2012

Fdo:

Prof. Dr. Guillermo Ramis Vidal



D. ANTONIO MUÑOZ LUNA, Catedrático de Universidad del Área de  
PRODUCCIÓN ANIMAL AUTORIZA:

La presentación de la Tesis Doctoral titulada “CARACTERIZACIÓN GENÉTICA DE  
LA RAZA CHATO MURCIANO”, realizada por D JUAN MANUEL HERRERO MEDRANO,  
bajo mi inmediata dirección y supervisión, y que presenta para la obtención del grado  
de Doctor por la Universidad de Murcia.

En Murcia, a 29 de mayo de 2012

Fdo:

Prof. Dr. Antonio Muñoz Luna

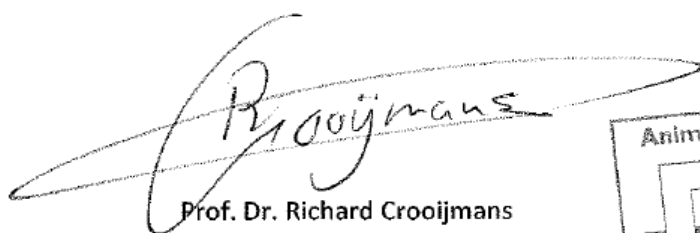


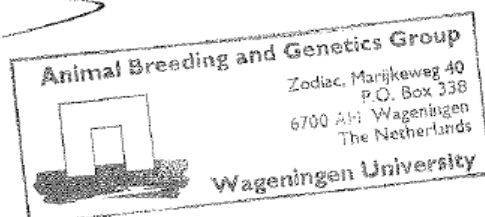
D. RICHARD CROOIJMANS, Profesor asistente del Departamento de Animal Breeding and Genomics Centre de la Universidad de Wageningen AUTORIZA:

Que la Tesis Doctoral titulada "CARACTERIZACIÓN GENÉTICA DE LA RAZA CHATO MURCIANO", realizada por D JUAN MANUEL HERRERO MEDRANO, bajo mi inmediata dirección y supervisión, y que presenta para la obtención del grado de Doctor por la Universidad de Murcia.

En Wageningen, a 29 de mayo de 2012

Fdo:

  
Prof. Dr. Richard Crooijmans



## **AGRADECIMIENTOS**

*A Guillermo Ramis y Antonio Muñoz, cuya dirección y supervisión ha sido fundamental para la elaboración de este trabajo.*

*A Richard Crooijmans y Hendrik-Jan Megens de la Universidad de Wageningen (Holanda), por su supervisión y cesión de muestras utilizadas en el trabajo.*

*Al Servicio de Formación y Transferencia Tecnológica de la Dirección General de Modernización de Explotaciones y Transferencia Tecnológica de la Consejería de Agricultura y Agua de la Comunidad Autónoma de la Región de Murcia que subvencionó la mayor parte de este trabajo a través del contrato de investigación 12603, especialmente a Juan Antonio Mora por su implicación en este trabajo.*

*A mis compañeros de Departamento, tanto a los que están (Aida, Obdulio, Juani) como a los que se fueron (Juanjo, Livia), cuya ayuda me dio el tiempo necesario para realizar este trabajo y cuya amistad me permite el lujo de ir contento a trabajar cada mañana. Mención especial a mi amiga Juani, que quizás no sea consciente del peso que ha tenido en esta tesis.*

*A mis amigos por darme el apoyo y el consejo que de vez en cuando uno necesita.*

*A mi madre, siempre está pase lo que pase.*

---

## ÍNDICE

<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>OBJETIVOS .....</b>	<b>7</b>
<b>REVISIÓN BIBLIOGRÁFICA.....</b>	<b>11</b>
1. Origen de la raza Chato Murciano .....	13
2. Marcadores genéticos utilizados en especies de animales domésticos .....	17
<b>Microsatélites.....</b>	<b>18</b>
<b>Genotipado de alta densidad (SNP) .....</b>	<b>20</b>
<b>ADN mitocondrial.....</b>	<b>22</b>
3. Diversidad genética en animales domésticos. Estudios previos en el cerdo Chato Murciano .....	24
4. Diferenciación de razas de animales domésticos en base a marcadores moleculares. Pureza de una raza y trazabilidad.....	28
5. Desequilibrio de Ligamiento.....	31
<b>MATERIAL Y MÉTODOS .....</b>	<b>35</b>
1. Animales y muestras .....	37
2. Recogida de las muestras .....	41

---

3. Extracción de ácidos nucleicos .....	41
4. Amplificación de los fragmentos microsatélites .....	42
5. Genotipado de alta densidad .....	43
6. Secuenciación del fragmento de ADN mitocondrial D-loop .....	44
7. Análisis de datos .....	45
<b>Microsatélites</b> .....	45
<b>Genotipado de alta densidad</b> .....	47
<b>ADN mitocondrial</b> .....	50
8. Análisis estadístico de marcadores genéticos.....	50
<b>RESULTADOS</b> .....	57
1. Estudio de la diversidad y subestructura genética de la raza Chato Murciano .....	59
2. Estudio de la relación genética del Chato Murciano con otras razas .....	67
<b>Genotipado de alta densidad</b> .....	67
<b>ADN mitocondrial</b> .....	80
3. Desequilibrio de Ligamiento.....	85
4. Marcadores genéticos distintivos de la raza Chato Murciano .....	89
<b>Marcadores SNP distintivos de la raza Chato Murciano</b> .....	89
<b>Frecuencias alélicas de los loci microsatélites</b> .....	95



<b>DISCUSIÓN</b> .....	99
Diversidad y subestructura genética de la raza Chato Murciano .....	102
Relación genética del Chato Murciano con otras razas.....	109
Desequilibrio de ligamiento en el cerdo Chato Murciano.....	115
Identificación genética de la raza Chato Murciano. Trazabilidad y pureza.....	118
<b>CONCLUSIONES</b> .....	125
<b>RESUMEN</b> .....	131
<b>SUMMARY</b> .....	137
Introduction.....	139
Materials and Methods.....	141
Results.....	145
Discussion.....	156
<b>ANEXOS</b> .....	167
<b>BIBLIOGRAFÍA</b> .....	173



## **LISTA DE ABREVIATURAS**

ADN: ácido desoxirribonucleico.

ADNmt: ADN mitocondrial.

BIC: Bayesian Information Criterion

BK: Berkshire

CM: Chato Murciano

CP: Coeficiente de pertenencia

$D_A$ : Distancia genética de Nei

DO: Denominación de Origen

DU: Duroc

F: Frecuencia alélica

He: Heterocigosidad esperada

Ho: Heterocigosidad observada

IB: Ibérico

IBS: Identity by State

IGP: Indicación Geográfica Protegida

LD: Desequilibrio de ligamiento

LR: Landrace

LW: Large White

Mb: Megabase

MDS: Escalamiento multidimensional

MH: Meishan

NA: Número de alelos.

NI: Negro Ibérico.

NJ: Neighbor-Joining

NMA: Número medio de alelos por locus

Pdb: Pares de bases

RE: Retinto

SNP: single nucleotide polymorphisms

TA: Tamworth

# INTRODUCCIÓN



## **INTRODUCCIÓN**

Las razas locales son de gran importancia en el mantenimiento de la diversidad genética de una especie, además de constituir en muchos casos un elemento cultural propio de comunidades locales o nacionales. La FAO (Food and Agricultural Organization) ha advertido de la progresiva erosión de los recursos genéticos, estableciendo que aproximadamente el 20% de las razas de animales domésticos están en peligro de extinción (Informes de la FAO Roma, 2006 y Interlaken, 2007). Una de estas razas locales categorizada como en peligro de extinción por el Sistema de Información de la Diversidad de los Animales Domésticos (DAD-IS, <http://dad.fao.org/>) es la raza Chato Murciano.

Se denomina Chato Murciano a una raza porcina autóctona de la Región de Murcia considerada dentro del grupo de razas locales españolas. El Chato Murciano comparte las características propias de dicho grupo de razas, es decir, gran rusticidad y adaptación al medio, baja productividad y explotación en extensivo o semiextensivo, aunque en los últimos años se ha orientado hacia la producción intensiva. La raza Chato Murciano ha sido calificada por el Ministerio de Agricultura de España como una raza de Protección Especial (Real Decreto

1682/1997 del 7 de noviembre), e incluida en diferentes programas de recuperación a lo largo de la última década. Este esfuerzo llevado a cabo por las autoridades y los productores refleja la preocupación actual de la sociedad ante la posible pérdida definitiva de una raza autóctona de la Región de Murcia que constituye un patrimonio cultural irremplazable. Otro de los factores que están ayudando a la progresiva recuperación de la raza es el creciente interés que han cobrado sus productos cárnicos, comercializados como originarios de la Región de Murcia y erigiéndose como diferenciados respecto a productos de cruces comerciales de razas precoces. Aún así, el reducido número de animales, los cruces con otras razas y la ausencia de un programa de conservación amenazan el futuro de una raza porcina con un impacto socio-económico importante en la Región de Murcia.

El primer paso para el diseño de un programa de conservación de una población es el estudio detallado de su diversidad genética y estructura de la población. El estudio de marcadores genéticos como los microsatélites, la secuencia de ADN mitocondrial (ADNmt) y los polimorfismos de un solo nucleótido (SNPs, por sus siglas en inglés single nucleotide polymorphisms) han demostrado su utilidad para evaluar tanto la diversidad y estructura genética de una raza, como



también las relaciones con otras poblaciones de la misma especie (Kim et al., 2002; Negrini et al., 2008a; Megens et al., 2008; Nidup y Moran, 2011). Además, la utilización conjunta de dichos marcadores moleculares permite realizar estudios más comprensibles y detallados (Ramírez et al., 2009; Luetkemeier et al. 2010). Otro aspecto importante en la conservación de una raza de producción animal es poder dotar a la industria alimentaria de herramientas que permitan diferenciar productos derivados de diferentes razas y aplicarlas en sus programas de trazabilidad. De esta forma se puede asegurar al consumidor el origen del producto cárnico y diferenciarlo de los productos derivados de otras razas. Las nuevas tecnologías de secuenciación aportan herramientas muy eficaces para la realización de test genéticos en los protocolos de trazabilidad de productos de origen animal (Ramos et al., 2011).

En el presente trabajo se estudiaron 34 marcadores microsatélites en la práctica totalidad de la población de Chato Murciano. Además, se estudiaron los marcadores moleculares ADNmt y genotipado de alta densidad de SNPs por primera vez en el Chato Murciano. Finalmente, se analizaron animales de aquellas razas porcinas que pudieron haber participado en la composición genética actual de la raza Chato Murciano.



# OBJETIVOS



## OBJETIVOS

Los objetivos principales del presente trabajo fueron:

- Evaluar la diversidad genética de la raza Chato Murciano.
- Estudiar la diferenciación genética del cerdo Chato Murciano con respecto a otras razas porcinas así como la influencia de las mismas en el Chato Murciano actual.
- Identificar marcadores genéticos que diferencien a la raza Chato Murciano de otras razas porcinas.

Los objetivos específicos fueron:

- Evaluar la diversidad genética actual de la raza Chato Murciano, tanto en su conjunto como en cada granja por separado.
- Investigar la subestructura genética de la población.
- Investigar la contribución genética de otras razas porcinas sobre la composición genética de la raza objeto de estudio, tanto aquellas descritas por la documentación histórica como las debidas a cruces recientes con razas porcinas comerciales.
- Obtener resultados del patrón de desequilibrio de

ligamiento en la raza Chato Murciano.

- Evaluar diferentes estrategias para determinar la pureza de la raza desde un punto de vista genético.
- Diseñar un panel de marcadores genéticos característicos de la raza Chato Murciano que permita la trazabilidad genética de productos cárnicos derivados de la misma.

REVISIÓN

BIBLIOGRÁFICA





## REVISIÓN BIBLIOGRÁFICA

### 1. Origen de la raza Chato Murciano

El origen de la raza porcina Chato Murciano se estima entre finales del siglo XIX y principios del siglo XX a partir del ya desaparecido Negro Mediterráneo, perteneciente al tronco ibérico (Lobera, 1998). El cerdo Negro Mediterráneo se localizaba en el levante español, principalmente en la Región de Murcia, por lo que también se le ha llamado cerdo Murciano primitivo. Se distinguieron dos variedades de cerdo Murciano primitivo conocidas como Gabana y Pintada (Herranz, 1987). Este cerdo era un animal muy rústico, que se alimentaba de subproductos de la producción agrícola o simplemente de los frutos caídos de los árboles. Un inconveniente importante del cerdo Murciano primitivo era su crecimiento extremadamente lento (el engorde podía durar hasta dos años) y el tamaño pequeño de las camadas (Lobera, 1998).

La tendencia a la intensificación de las producciones ganaderas de finales del siglo XIX y principios del siglo XX, dio lugar a una serie de cambios importantes en los sistemas de producción ganadera que afectaron a un gran número de razas locales, entre ellas el cerdo Murciano primitivo. Estos cambios llevaron al abandono y casi

extinción de muchas razas locales mientras que otras fueron cruzadas con razas foráneas con el fin de mejorar sus parámetros productivos. En este último grupo se incluyó el cerdo Murciano primitivo que se cruzó con reproductores seleccionados de las razas Berkshire, Large White, Craonés, Alderney, Tamworth y Colorado Extremeño (Lobera, 1998). Como resultado de estos cruces se describió la raza porcina denominada Chato Murciano en 1913. A partir de los años cincuenta hubo un cambio de mentalidad en los mercados de productos cárnicos, en los que comenzaron a primar las razas con altas producciones y carnes más magras, depreciándose de esta forma las carnes grasas. Esta nueva situación llevó al cerdo Chato Murciano a un declive de su población por la tendencia de los productores a criar razas foráneas y cruces comerciales.

El número de reproductores de la raza se estimó en unos 30 animales a principios de los años noventa, estando la raza en un alto riesgo de extinción (Martínez et al., 1998). Ante esta situación, siendo el cerdo Chato Murciano una raza autóctona de la Región de Murcia y que forma parte de la historia y cultura de la misma, el Ministerio de Agricultura y Pesca de España catalogó al Chato Murciano como una raza de Protección Especial (Real Decreto 1682/1997). A partir de entonces se incluyó en diferentes programas de recuperación a lo largo

de la última década, tanto regionales como nacionales, aumentando su población a un número cercano a 300 reproductores censados en el año 2009, repartidos entre granjas de pequeño y mediano tamaño (**Tabla 1**) en las que se explota principalmente en sistema intensivo o semi-intensivo. Otro acontecimiento que demostró el interés por la recuperación de esta raza porcina fue la creación en el año 2005 de la Asociación de Criadores para la conservación y recuperación del Chato Murciano (ACHAMUR) con la finalidad de recuperar y fomentar los productos de la raza.

Las características morfológicas del cerdo Chato Murciano han sido descritas en detalle en el trabajo de Peinado et al. (2006). En líneas generales el cerdo Chato Murciano se describe como un animal subhipermétrico, cóncavo y longilíneo, orejas erectas o con tendencia a la verticalidad y hocico corto y achatado. El color de la capa es generalmente negro uniforme, si bien se han observado animales con pelos blancos en regiones periféricas, supuestamente como reminiscencia del uso de verracos Berkshire en la formación de la raza.

**Tabla 1.** Censo de reproductores de Chato Murciano registrados en el año 2009, por productor ordenado alfabéticamente

<b>N</b>	<b>GANADERO</b>	<b>HEMBRAS</b>	<b>VERRACOS</b>
1	Antonio Cánovas Balsas	2	0
2	Cárnicas La Noria	6	0
3	Granja La Pilica	2	0
4	Jysaps	45	4
5	Alfredo López Gómez	5	2
6	Salvador Marín López	6	0
7	Juan Martínez Antolinos	13	4
8	Antonio Martínez Ros	1	0
9	Antonio Nietos Giménez	1	1
10	José Reverte Navarro	80	4
11	Salvador Sánchez Espín	70	3
12	José Sánchez Sánchez	2	1
13	Ángel Serrano Belluga	2	0
14	Miguel Serrano González	20	2
15	María Victoria Torrecillas	1	0
16	Macario Zabala Díaz	5	1
	<b>TOTAL</b>	<b>261</b>	<b>22</b>

## 2. Marcadores genéticos utilizados en especies de animales domésticos

Los marcadores genéticos son regiones de ADN variables que se heredan de una generación a la siguiente. Para que sean útiles deben ser polimórficos, es decir, en el marcador genético deben existir dos o más variantes en la población estudiada. El uso de marcadores genéticos ha sido de gran utilidad para estudios de diversidad genética (ej. San Cristobal et al., 2006) y para evaluar las relaciones genéticas entre razas y poblaciones (ej. Pariset et al., 2010). Además, el uso de estos marcadores ha permitido diferenciar razas de una misma especie, siendo posible asignar individuos a su raza de origen en base a estas diferencias genéticas (Blott et al., 1999; Ramos et al., 2011; Pant et al., 2012).

Entre los marcadores genéticos disponibles en la actualidad para el estudio de poblaciones, los microsatélites y el estudio de determinadas regiones del ADN mitocondrial han sido los más utilizados en la última década (ej. Larson et al., 2005; Megens et al., 2008). Recientemente han aparecido nuevas técnicas de genotipado de alta densidad capaces de estudiar miles de SNPs, erigiéndose como un

marcador molecular de gran potencia para estudios presentes y futuros.

### **Microsatélites**

Los microsatélites o STR (por sus siglas en inglés, Short Tandem Repeat) son secuencias cortas de ácido desoxirribonucleico (ADN), entre 2 y 6 pares de bases (pdb), que se repiten en tándem y que están amplia y aleatoriamente distribuidas por todo el genoma, aunque usualmente se localizan en regiones no codificantes. Se pueden clasificar según el tamaño de su secuencia en dinucleótidos (2 pdb), trinucleótidos (3 pdb), tetranucleótidos (4 pdb), etc. Las diferencias en el número de repeticiones de estas secuencias de ADN van a definir los diferentes alelos de cada microsatélite. El elevado polimorfismo de estas secuencias de ADN se puede explicar por la alta tasa de mutación observada en las mismas (Ellegren, 2002; Whittaker et al., 2003).

Debido al reducido tamaño de los microsatélites y a que las regiones que los flanquean están altamente conservadas, se pueden amplificar por medio de PCR a partir de cualquier muestra biológica susceptible de contener células nucleadas como por ejemplo sangre, pelo, piel, productos cárnicos, etc. Posteriormente el producto de PCR

se visualiza mediante electroforesis en gel o electroforesis capilar en un secuenciador determinándose el tamaño del fragmento microsatélite.

Las características que hacen de los microsatélites un marcador molecular tan utilizado, pueden resumirse en los siguientes puntos:

- Elevado grado de polimorfismo.
- Herencia mendeliana simple.
- Herencia codominante
- Análisis rápido y reproducible.

Estas características han propiciado su uso por un gran número de investigadores para estudiar la diversidad genética de poblaciones de diferentes especies, así como las relaciones genéticas entre ellas (Paszek et al., 1998; Laval et al., 2000; Yang et al., 2003; Manea et al., 2009, etc.), relaciones de parentesco e identificación individual (Jobling y Gill, 2004) y asignación de individuos a sus razas de origen (Ciampolini et al., 2006).

## **Genotipado de alta densidad (SNP)**

Los SNPs son variaciones en la secuencia de ADN en un solo nucleótido que se localizan tanto en regiones codificantes como en no codificantes del genoma. Constituyen la forma de variación genética más común en el genoma de los vertebrados (Wang et al., 1998). Al ser un marcador bialélico es menos informativo que otros marcadores más polimórficos como los microsatélites (Blott et al., 1999). Se ha estimado que por cada microsatélite son necesarios tres SNPs para obtener la misma cantidad de información (Schopen et al., 2008). Dicha desventaja se compensa por el elevado número de SNPs distribuidos a lo largo de todo el genoma, permitiendo estudiar la variabilidad genética en miles de regiones y por ende, realizar estudios genéticos de alta densidad que aportan más información que otros marcadores genéticos más polimórficos. Otra ventaja de los marcadores SNP es que son más estables de generación en generación que los microsatélites (Krawczak, 1999). Actualmente existen en el mercado varias tecnologías que han permitido identificar un número de SNPs muy alto como las desarrolladas por las plataformas Illumina Genome Analyzer (Bennett, 2004), Roche 454 (Margulies et al., 2005), AB SOLiD (Shendure et al., 2005) y Helicos (Milos, 2008).



La aparición reciente de paneles comerciales de alta densidad de SNPs, que permiten estudiar simultáneamente más de 50.000 SNPs en distintas especies de animales domésticos (Matukumalli et al., 2009; Ramos et al., 2009; Groenen et al., 2011), ha hecho posible la aplicación de estos marcadores en múltiples estudios (Mckay et al., 2007; de Roos et al., 2008; Gorbach et al., 2010). En el caso concreto de la especie porcina, el PorcineSNP60 BeadChip comercializado por la empresa Illumina permite genotipar 62612 SNPs (Ramos et al., 2009). Estos marcadores se distribuyen en todos los cromosomas, si bien algunos de los SNPs incluidos ocupan una posición del genoma desconocida. El panel PorcineSNP60 se ha utilizado principalmente por sus aplicaciones en el mapeo de QTL (por sus siglas en inglés, Quantitative Trait Loci), selección asistida por marcadores genéticos (MAS por sus siglas en inglés, Marker assisted selection,) y en estudios de asociación del genoma completo (GWAS, por sus siglas en inglés, Genome-wide association studies) (Duijvesteijn et al., 2010; Ponsuksili et al., 2011; Onteru et al., 2012; Fu et al., 2012) aunque también supone una herramienta para diferenciar y/o comparar poblaciones en base a sus diferencias genéticas (Uimari y Tapio, 2010; Ramos et al., 2011; Badke et al., 2012).

## **ADN mitocondrial**

El ADNmt es un fragmento de ADN circular localizado en las mitocondrias de las células eucariotas. Ha sido ampliamente utilizado para estudiar las relaciones filogenéticas entre poblaciones porcinas, permitiendo obtener gran cantidad de información de los procesos que llevaron a la domesticación de la especie y al origen y distribución de las razas modernas. Así el estudio de las diferencias de la secuencia de ADNmt entre poblaciones porcinas de todo el mundo ha llevado a sugerir que existieron varios centros de domesticación independientes en Asia y Europa (Giuffra et al., 2000; Larson et al., 2005; Larson et al., 2007; Larson et al., 2010). También se ha observado la presencia de haplotipos de ADNmt propios de razas asiáticas en razas europeas como consecuencia de los acontecimientos históricos de los siglos XVIII y XIX cuando se llevaron a cabo cruces entre cerdos asiáticos y cerdos ingleses con el fin de mejorar los parámetros productivos de las razas europeas (Giuffra et al., 2000; Kim et al., 2002; Larson et al., 2005; Wu et al., 2007).

Existen diversas razones que han hecho del ADNmt una herramienta de gran valor en los estudios de poblaciones. El ADNmt tiene una elevada capacidad de mutación, mucho más rápido que la del

ADN nuclear, que ha generado un elevado polimorfismo genético posibilitando su uso en estudios evolutivos. Se hereda vía materna puesto que se localiza en las mitocondrias y no en el núcleo, siendo los procesos de recombinación inexistentes o muy raros (Brown et al., 1979; Wolstenholme, 1992; Savolainen et al., 2002). Esta característica permite trazar líneas genéticas que se remontan mucho tiempo atrás.

No obstante hay autores que consideran que el uso del ADNmt como marcador genético presenta ciertas desventajas. Este marcador representa a una parte del genoma muy pequeña y se considera como un locus único (Pamilo y Nei, 1988) por lo que es un mal indicador de la diversidad genética. Además, al ser heredado por vía materna exclusivamente, el estudio de este marcador no tiene en cuenta la información genética heredada del macho, el cual tiene un papel muy importante en la producción porcina (Luetkemeier et al., 2010).

3. Diversidad genética en animales domésticos. Estudios previos en el cerdo Chato Murciano.

La biodiversidad ha sido definida en términos generales como el grado de variedad de la naturaleza (McNeeley, 1988). La FAO, define la biodiversidad como “la variedad y la variabilidad de los ecosistemas, los animales, las plantas y los microorganismos a nivel genético, de especie y de ecosistema, necesarias para sostener la vida humana y las funciones clave de los ecosistemas” siendo necesario, en referencia a los animales domésticos, “el mantenimiento de la diversidad genética para asegurar la producción de alimentos en el futuro”.

Se han descrito una serie de factores genéticos implicados en el riesgo de extinción de las poblaciones y por tanto importantes en el manejo genético de las mismas (Frankham et al., 2002):

- Los efectos negativos de la endogamia sobre la reproducción y supervivencia.
- Pérdida de la diversidad genética y de la habilidad para adaptarse a los cambios ambientales.
- Fragmentación de la población y reducción del flujo de genes.
- Efecto predominante de la deriva genética sobre la selección natural.

- Acumulación de mutaciones deletéreas.
- Adaptación genética a la vida en cautividad y sus efectos negativos en el éxito de reintroducción a la vida salvaje.
- Efectos deletéreos en cruzamientos entre diferentes poblaciones

Entre los factores mencionados, uno de los más importantes a los que se enfrentan las poblaciones de pequeño tamaño es la deriva genética (Lacy, 1987) que incrementa el efecto de la endogamia y la fragmentación de las poblaciones (Rocha y Gasca, 2007). Así, la deriva genética hace que, en ausencia de procesos de selección, migración y mutación, las poblaciones modifiquen sus frecuencias alélicas por un proceso aleatorio de herencia de genes de generación en generación. Si no se establecen mecanismos para incrementar la variabilidad genética, las poblaciones tienden a perder los alelos menos frecuentes y a fijar los más frecuentes, disminuyendo la variabilidad genética (Nei, 2000) y aumentando la endogamia con los consecuentes efectos negativos sobre el crecimiento y la supervivencia de la población (Falconer, 1981).

El estudio de la diversidad genética y la estructura de las poblaciones es importante para evaluar cómo han influido los cambios ambientales en las mismas, aportando información muy valiosa para

proteger y preservar esas poblaciones. Entre los marcadores moleculares disponibles en la actualidad, los microsatélites han sido los más utilizados para evaluar la diversidad de poblaciones, constituyendo una herramienta para el diseño de estrategias de conservación de las razas porcinas (Nidup y Moran, 2011). Los parámetros más utilizados para medir la diversidad genética de las poblaciones son las heterocigosidades observada y esperada y el número medio de alelos por locus (Allendorf, 1986).

Zhang y Plastow (2011) realizaron una revisión bibliográfica sobre la diversidad genética de las razas porcinas a nivel mundial mostrando que la diversidad genética en términos de heterocigosidad observada y esperada es superior en las razas porcinas asiáticas que en las europeas. Concretamente, la heterocigosidad esperada media en las razas asiáticas fue de 0,752 mientras que en las razas europeas fue de 0,570. En el caso concreto del cerdo Chato Murciano, los trabajos realizados previamente mostraron valores de diversidad genética inferiores a la media de las razas europeas. Calvo et al. (2000) realizaron el primer estudio de la diversidad genética de la raza en base al análisis de ocho marcadores microsatélites analizados en 43 cerdos de la raza Chato Murciano, si bien solamente 27 fueron

catalogados como puros por los autores. Según este estudio, la diversidad genética de la población total de cerdos Chato Murciano, medida en términos de heterocigosidad esperada, fue de 0,397, aunque solamente de 0,29 para los animales considerados como puros mientras que el número medio de alelos fue 3,8. Posteriormente, Vega-Pla et al. (2004) realizaron un estudio más amplio, tanto en número de marcadores (26 microsatélites) como en número de animales, al estudiar 24 cerdos Chato Murciano y 59 animales pertenecientes a la F2, F3, F4 y F5 de un cruce inicial de macho Chato Murciano y hembra Large White. Estos últimos autores determinaron una heterocigosidad esperada de 0,53 y un número medio de alelos de 3,92 lo que indica una mayor diversidad genética que en el estudio previo realizado por Calvo et al. (2000). En ambos trabajos se observó que los microsatélites con frecuencias alélicas más elevadas fueron: S0215, SW951, S0227 y S0355. Resulta interesante que solamente se detectara fijación alélica de los loci microsatélite SW951 y S0215 en los cerdos Chato Murciano considerados como puros. Finalmente, los mismos autores sugirieron la posibilidad de cruzamientos controlados con otras razas comerciales ya que observaron distancias genéticas pequeñas entre los descendientes F3, F4 y F5 de animales Chato

Murciano cruzados con Large White, y los Chato Murciano considerados como puros.

No existen hasta la fecha trabajos en los que se haya estudiado la relación genética del cerdo Chato Murciano con otras razas porcinas utilizando el ADN mitocondrial y marcadores SNP como marcadores genéticos.

#### 4. Diferenciación de razas de animales domésticos en base a marcadores moleculares. Pureza de una raza y trazabilidad.

El efecto de diferentes fuerzas evolutivas sobre las especies de animales domésticos, como son la deriva genética y la selección natural y artificial, ha llevado a la modificación de razas de animales e incluso a la creación de otras nuevas. Diversos estudios han demostrado que las razas porcinas son genéticamente diferenciables mediante el análisis de microsatélites (Megens et al., 2008; Boitard et al., 2010) y marcadores SNP (Rohrer et al., 2007; Ramos et al., 2011; Matsumoto et al., 2012), siendo incluso posible diferenciar poblaciones de la misma raza localizadas en diferentes países (Boitard et al., 2010). Los marcadores genéticos más adecuados para diferenciar razas son aquellos en los que existe fijación de alelos diferentes en las razas



comparadas (Blott et al., 1999). Bajo dicha premisa se han utilizado técnicas de AFLP (por sus siglas en inglés amplified fragment length polymorphism; Negrini et al., 2007) y marcadores microsatélites en distintas especies (Koskinen, 2003; Dalvit et al., 2008) en la asignación de individuos a su raza de origen, si bien se estima que dichos marcadores son menos fiables que los paneles de alta densidad de SNPs para este tipo de estudios (Ramos et al., 2011).

El término trazabilidad alimentaria hace referencia al conjunto de procedimientos destinados a mantener la identificación de un animal o producto animal a lo largo de la cadena productiva. Esta práctica es de gran importancia para garantizar la seguridad alimentaria, ya que, hace posible la rápida detección y aislamiento del origen de eventuales crisis alimentarias. Debido a que el ADN es inalterable, estable a los tratamientos tecnológicos de productos cárnicos e identificativo de cada individuo, se ha estudiado la utilización de marcadores genéticos para diferenciar individuos o razas de animales domésticos con el fin de asegurar la correcta trazabilidad de productos cárnicos (Dalvit et al., 2007). La utilización de marcadores genéticos en los protocolos de trazabilidad suele limitarse a la verificación de que los métodos físicos se están realizando correctamente (Goffaux et al., 2005).

La trazabilidad también adquiere relevancia en las razas animales con un valor comercial superior a los productos de razas híbridas o comerciales, como son por ejemplo los productos con Denominación de Origen (DO) y con Indicación Geográfica Protegida (IGP). En muchos casos estos productos proceden de una raza local y el consumidor los asocia con una mayor calidad. Ejemplos de este tipo de productos son la IGP de la carne de Avila y la carne de Cantabria, ambas de vacuno, y las DO Jamón de Guijuelo y Jamón de Jabugo (Real Decreto 1335/2011, de 3 de octubre). Todos estos productos tienden a comercializarse a precios más elevados lo que los hace susceptibles al fraude.

Por otro lado, los marcadores genéticos han permitido la detección de introgresión genética o mezcla entre razas (Bradley et al., 1994; Clop et al., 2004), si bien el número de marcadores genéticos necesarios para identificar individuos híbridos es mucho mayor que el necesario para identificar individuos puros (Blott et al., 1999). La identificación de individuos híbridos o mezclados adquiere una gran importancia en los programas de conservación de razas. En efecto, en los programas de conservación es necesario identificar a los animales

con alta pureza racial para priorizar su uso y así mantener la identidad e integridad genética de la raza.

#### 5. Desequilibrio de Ligamiento.

El desequilibrio de ligamiento (LD por sus siglas en inglés, Linkage Disequilibrium) es una propiedad que hace referencia a la tendencia que tienen dos alelos situados en loci diferentes a heredarse conjuntamente en lugar de hacerlo al azar a lo largo de sucesivas generaciones. Debido a esta propiedad, es posible encontrar combinaciones de alelos o de marcadores genéticos con una frecuencia mayor o menor a lo esperado en condiciones de formación de haplotipos aleatorias. El grado de desequilibrio dependerá de la diferencia entre las frecuencias alélicas observadas y las esperadas.

La disponibilidad de un gran número de SNPs ha incrementado el interés del estudio de los patrones de LD en especies de producción, debido a su aplicación en estudios de mapeo genético de genes asociados a enfermedades (Pritchard y Donnelly, 2001) o de genes importantes en mejora genética animal (Chao et al., 2012). La mayor densidad de marcadores aumenta la probabilidad de que un SNP esté próximo al gen responsable del fenotipo buscado, lo que dará lugar a

un valor de LD elevado. De esta forma, el LD es un aspecto clave de los estudios de asociación del genoma completo y selección genómica (van der Steen et al., 2005; Hayes et al., 2009). Por otro lado, el LD está influido por la historia y demografía de las poblaciones, por lo que su estudio permite obtener datos sobre el origen y desarrollo de razas o poblaciones humanas (Tenesa et al., 2007; Teo et al., 2009), porcinas (Zhang y Plastow, 2011), ovinas (Meadows et al., 2008), etc.

El grado de LD está afectado por factores numerosos y complejos y muchas veces no comprendidos en su totalidad (Ardlie et al., 2002). Además de depender de la mutación y recombinación, el LD lo hace de factores que afectan a las poblaciones como la deriva genética, la estructura y subdivisión de la población, la mezcla entre poblaciones y la selección natural y artificial (Ardlie et al., 2002; Du et al., 2007). Así, el LD desaparece gradualmente por efecto de la recombinación y mutación mientras que se va regenerando por efecto de la deriva genética (Abecasis et al., 2005). Por ejemplo, se ha observado que el LD se extiende hasta mayores distancias genéticas en poblaciones con un número efectivo de población pequeño o en aquellas que se han mezclado recientemente con otras (Zhu et al., 2005).

El primer estadístico para medir el LD fue el propuesto por Lewontin (1964) que cuantificó el LD como la diferencia entre la frecuencia observada de un haplotipo formado por dos loci y la frecuencia esperada en el caso de que los dos alelos se segregaran al azar. A este estadístico se le llamó coeficiente de disequilibrio de ligamiento,  $D$ . El valor  $D$ , aún siendo muy intuitivo, depende de las frecuencias de los alelos de ambos loci, lo que limita su uso para medir el grado de LD. Actualmente existen otros estadísticos para medir el LD basados en el descrito previamente, de los cuales  $D'$  y  $r^2$  son los más utilizados. El valor  $D'$  se obtiene de la división de  $D$  entre su máximo valor posible. Adquiere valores entre 0 y 1. Dos SNPs tendrán un valor de LD de  $D'=1$  solamente si dichos SNPs no han sido separados por recombinación o mutación, estando así en LD completo. Un inconveniente importante de este estadístico es que su valor está muy influido por el tamaño muestral, además, los valores  $D'$  intermedios son difíciles de interpretar (Ardlie et al., 2002). Un estadístico para medir el LD que supera los inconvenientes mencionados es el coeficiente de correlación  $r^2$  entre SNPs (Hill y Robertson, 1968). El estadístico varía entre 0 y 1. El valor  $r^2=1$  se obtiene si los dos marcadores no se han separado por recombinación y, además, tienen las mismas frecuencias alélicas, lo que se denomina disequilibrio

perfecto. La existencia de LD perfecto implica que la observación de un marcador da una información completa del otro, por lo que el genotipado de ambos marcadores resultaría ser redundante.

# MATERIAL Y MÉTODOS





## MATERIAL Y MÉTODOS

### 1. Animales y muestras

Se tomaron muestras de un total de 196 reproductores de la raza Chato Murciano (CM). Tras eliminar aquellas muestras que no cumplían con los mínimos de calidad de genotipado requeridos, se utilizaron para los análisis 175 animales, lo que supone aproximadamente el 70% de la población total. Para localizar a los productores que criaban reproductores CM se utilizó el censo aportado por la Consejería de Agricultura, Agua y Medioambiente, así como los datos aportados por ACHAMUR referentes al año 2009 (**Tabla 1**), descartándose aquellas granjas con un número de reproductores inferior a tres. De esta forma se tomaron finalmente muestras de 8 granjas pertenecientes a 6 productores (las granjas R1, R2 y R3 pertenecen al mismo productor) situadas en los términos municipales de Bullas, Lorca, Baños y Mendigo, Puerto Lumbreras, Alhama de Murcia y Murcia (**Tabla 2**). Además, se incluyeron en el estudio 194 muestras de las razas porcinas que, atendiendo a la información histórica disponible, eran susceptibles de haber contribuido en mayor medida en la composición genética actual de la raza CM, tanto en el pasado como recientemente. Concretamente se analizaron animales de

las razas Large White (LW, n=53), Landrace (LR, n=29), Duroc (DU, n=42), Tamworth (TA, n=15), Berkshire (BK, n=19) y cerdo Ibérico (IB, n=26). Entre estos últimos, 11 fueron cerdos retintos (RE) y 15 cerdos ibéricos de capa negra (NI). Finalmente se analizaron 10 cerdos de la raza China Meishan (MH), como raza control que no ha tenido relación con el cerdo CM.

Para la elección de estos animales de otras razas se priorizó a aquellos que fueran representativos de la raza y no emparentados entre sí. En la **Tabla 3** se detalla el número de animales usados así como los análisis realizados en cada raza.

**Tabla 2.** Relación por granjas de CM muestreadas en el estudio

GRANJAS	NÚMERO DE ANIMALES MUESTREADOS		% ANIMALES MUESTREADOS/TOTAL		NÚMERO DE ANIMALES UTILIZADOS*
	<i>Hembras</i>	<i>Machos</i>	<i>Hembras</i>	<i>Machos</i>	
AI	13	3	100	100	11
JJ	41	3	100	100	36
R1	16	2	100	100	16
R2	19	3	100	100	21
R3	30	2	100	100	32
MS	8	1	100	100	7
SS	42	2	100	100	42
MC	9	2	100	100	10
<b>TOTAL</b>	<b>178</b>	<b>18</b>			<b>175</b>

\*Animales usados para los análisis tras descartar aquellos que presentaban datos perdidos derivados de errores en el genotipado.

**Tabla 3.** Análisis realizados en cada población

<b>POBLACIÓN</b>	<b>PAÍS</b>	<b>N</b>	<b>ADNmt</b>	<b>Porcine SNP60</b>	<b>MICROSATÉLITES</b>
<b>Chato Murciano</b>	España	175	152	39	175
<i>Granja JJ</i>	España	36	32	8	36
<i>Granja R1</i>	España	16	14	5	16
<i>Granja R2</i>	España	21	18	5	21
<i>Granja R3</i>	España	32	26	6	32
<i>Granja MS</i>	España	7	7	3	7
<i>Granja SS</i>	España	42	35	8	42
<i>Granja AI</i>	España	11	10	2	11
<i>Granja MC</i>	España	10	10	2	10
<b>Large White</b>	EEUU, Países Bajos, Dinamarca	53	48	53	0
<b>Landrace</b>	EEUU, Países Bajos, Dinamarca	29	34	29	0
<b>Duroc</b>	EEUU, Países Bajos, Dinamarca	42	42	42	0
<b>Tamworth</b>	Reino Unido	15	14	15	0
<b>Berkshire</b>	Reino Unido	19	19	19	0
<b>Cerdo Ibérico</b>	España	26	38	26	0
<b>Meishan</b>	China	10	10	10	0

## 2. Recogida de las muestras.

Para los análisis de laboratorio se utilizaron muestras de sangre recogidas mediante venopunción yugular en tubos con EDTA (Venoject®, Becton Dickinson, Reino Unido) y muestras de pelo en los casos en los que la extracción de sangre resultó imposible de realizar. Las muestras de los cerdos con los que se comparó el cerdo CM fueron amablemente donadas por el Grupo de Investigación Animal Breeding and Genomics (Universidad de Wageningen, Holanda).

## 3. Extracción de ácidos nucleicos.

Se extrajo ADN de las muestras de sangre y folículo piloso mediante los kit comerciales para la extracción de ácidos nucleicos denominados SSS y Danapure Spin Kit respectivamente (Genedan S.L., España). La cantidad de ADN obtenida se cuantificó mediante método espectrofotométrico usando un Nanodrop 1000 (Thermo, Reino Unido).

#### 4. Amplificación de los fragmentos microsatélites.

Se analizaron un total de 34 marcadores microsatélite, 24 pertenecientes al panel recomendado por el Grupo Asesor sobre Diversidad Genética Animal (ISAG-FAO) (Hoffmann et al. 2004) y diez al panel propuesto por el Instituto Roslin (Russell et al. 2003). Los 34 loci microsatélites se eligieron en base a la ausencia de alelos nulos, la facilidad para la identificación de los picos de fluorescencia y a la posibilidad de agruparse en sets para ser amplificados mediante PCR multiplex, evitando el solapamiento de los picos. Los grupos o sets de primers diseñados para la amplificación por PCR multiplex se detallan en el **Anexo I**. La reacción tuvo lugar en un volumen final de 15µl, conteniendo: 50-80 ng de ADN genómico, desoxinucleótidos trifosfato (dntp's) 0,33 mM, tampón para PCR, MgCl<sub>2</sub> 1,5 mM, y 1,35 U de Taq Polimerasa (5U/µl) (Genecraft, Alemania). Se optimizaron las concentraciones para cada primer, estando las cantidades comprendidas entre 80-600 nmol. La amplificación se realizó en un termociclador ABI 2720 (Applied Biosystems, EEUU). El protocolo de amplificación requirió de un calentamiento previo de 95°C, seguido de 27-30 ciclos según set, consistiendo cada ciclo de 3 etapas: desnaturalización a 95°C (10 minutos), hibridación a 55-60°C según

set (30 segundos), y polimerización a 72°C (1 minuto). El protocolo finalizó con la etapa de extensión final a 72°C durante 30 minutos.

El producto de PCR se sometió a un análisis de fragmentos mediante un secuenciador ABI 3100® (Applied Biosystems, EEUU). Este análisis se realizó en el Laboratorio de Biología Molecular del Servicio de Apoyo a la Investigación de la Universidad de Murcia. Se utilizó como estándar de tamaño el GS LIZ 500® (Applied Biosystems, EEUU). Los picos resultantes se analizaron mediante el programa informático GeneMapper v3.7® (Applied Biosystems, EEUU). Para la normalización del tamaño de los alelos se utilizaron las muestras de referencia F9110010 y F9119912 que fueron amablemente proporcionadas por el INRA (Institut National de la Recherche Agronomique, Francia).

## 5. Genotipado de alta densidad

El genotipado de alta densidad se realizó mediante el Porcine SNP60 BeadChip (Illumina Inc, Estados Unidos) diseñado para el genotipado de 62163 SNPs en el laboratorio certificado GeneSeek (Lincoln, Nebraska, EEUU). En el presente estudio se utilizaron los SNPs localizados en cromosomas autosómicos, excluyéndose los

marcadores SNPs localizados en los cromosomas sexuales y aquellos con más de un 5% de datos perdidos. El número final de SNPs utilizado para los análisis fue de 46887.

#### 6. Secuenciación del fragmento de ADN mitocondrial D-loop.

Se amplificó un fragmento de la región D-loop de ADNmt mediante PCR. La secuencia de los primers utilizados fue: Forward 5'CTCCGCCATCAGCACCCAAAG3'; reverse 5'GCACCTTGTTTGGATTRTCG3', los cuales fueron utilizados previamente por Luetkemeier et al. (2010). La reacción se llevó a cabo en un volumen final de 12 µl conteniendo PCR master mix (PCR master mix 1,5 mM MgCl<sub>2</sub>, Thermo Fisher Scientific), 5 pmol de cada primer y 40 ng de ADN aproximadamente. El protocolo de amplificación requirió de un calentamiento previo de 95°C durante 2 minutos, seguido de 35 ciclos consistiendo cada ciclo de 3 etapas, primero desnaturalización a 95°C, hibridación a 55°C y finalmente polimerización a 72°C. La presencia del amplicón resultante fue determinada por electroforesis en gel de agarosa 1,5% teñido con bromuro de etidio y visualizado usando un transiluminador ultravioleta (Consort N.V., Bélgica). El producto de PCR restante se



purificó mediante el kit comercial Dana Clean Spin Kit (Genedan S.L., España). Finalmente se realizaron las reacciones de secuenciación de forma bidireccional usando Big Dye Terminator Cycle Sequencing Ready Reaction kit v3.1 (Applied Biosystem, EEUU) para ser analizados en el secuenciador ABI 3100® (Applied Biosystems, EEUU). Estos análisis se realizaron en el Laboratorio de Biología Molecular del Servicio de Apoyo a la Investigación de la Universidad de Murcia y en el Departamento Animal Breeding and Genomics Centre de la Universidad de Wageningen (Países Bajos). Debido a que no todas las muestras analizadas alcanzaron el tamaño completo de la secuencia de ADNmt (722 pb), se utilizó un fragmento de 624 pb común a la mayoría de muestras para la realización de los análisis. Los números de acceso asignadas por Genbank a las secuencias utilizadas en este trabajo fueron los comprendidos entre JQ238237-JQ238602.

## 7. Análisis de datos

### **Microsatélites**

Para estudiar la diversidad genética de la población de CM se calcularon el número de alelos y las frecuencias alélicas así como los valores de heterocigosidad observada y esperada, tanto en la población

total como en cada granja por separado utilizando el programa GenALEX V. 6.3 (Peakall y Smouse, 2006). Se analizó la diferenciación genética entre granjas utilizando el estadístico  $F_{st}$  de Wright (1969) mediante el método de Weir y Cockerham (1984) implementado por el programa GENEPOP 4.0.10 (Rousset et al., 2008), el cual incorpora el tamaño muestral y el número de subpoblaciones. Se utilizó el programa GenALEX V. 6.3 para calcular la matriz de distancias genéticas (Nei, 1972) entre pares de granjas que crían cerdos CM. Dicha matriz se utilizó para la construcción de un dendograma por medio del método Neighbor-joining (NJ) con el programa Mega 5.03 (Tamura et al. 2011).

El estudio de la subestructura de población CM en base a los datos microsatélite se realizó mediante el programa Structure v2.3 (Pritchard et al., 2000). Se aplicó el modelo de agrupación bayesiano implementado por el programa para diferentes valores de K (o número de grupos fijados previamente por el investigador), estudiándose las variaciones en los agrupamientos realizados y su correspondencia con las granjas incluidas en el estudio. En todos los análisis realizados con el programa Structure se ajustaron los parámetros a 10.000 iteraciones de Burn-in y con número de iteraciones de Cadenas de Markov de Monte Carlo de 10.000, suponiendo frecuencias alélicas no

correlacionadas entre poblaciones y sin aportar información previa del origen de los animales. La representación gráfica de los resultados obtenidos con el programa Structure se realizó con el programa Distruct 1.1 (Rosenberg 2004). Los gráficos resultantes muestran a cada individuo representado por una línea vertical que se divide en un número “K” de colores (siendo K el número de poblaciones fijado por el investigador) mientras que las líneas verticales de color negro separan las diferentes poblaciones designadas previamente.

### **Genotipado de alta densidad**

El filtrado de los SNPs con más de un 5% de datos perdidos se realizó con el programa Plink (Purcell et al, 2007). Este programa se utilizó también para calcular la matriz IBS (Identity by State) para todas las parejas de individuos basándose en la similitud genómica entre los mismos. A partir de dicha matriz se realizó el análisis de escalamiento multidimensional (MDS) implementado por el programa Plink, siendo finalmente visualizado gráficamente con el paquete estadístico R v. 2.12.1. (<http://www.r-project.org/>). Al igual que con los marcadores microsatélite, se utilizó el programa Structure para analizar los SNPs. De esta forma se estudiaron las relaciones entre el

cerdo CM y otras razas porcinas. Para la determinación de valor de K se utilizó el método BIC (Bayesian Information Criterion) utilizando el paquete Adegenet incluido en el paquete estadístico R (Jombart, 2008). Además, se compararon los resultados de subestructura de la población de cerdo CM obtenida en base a los marcadores microsatélite con la obtenida en base a los SNPs. Los parámetros utilizados para ejecutar los análisis fueron los mismos que los previamente descritos para los datos microsatélites. Las distancias genéticas (Nei, 1972) y los valores  $F_{st}$  entre los pares de poblaciones se calcularon con el programa Power marker (Liu y Muse 2005).

Los marcadores con desviación significativa del equilibrio de Hardy-Weinberg ( $P < 0,001$ ) y con la frecuencia del alelo menor (MAF) por debajo de 0,05 fueron excluidos del análisis para eliminar los efectos de las frecuencias alélicas en el análisis de LD utilizando el programa de Plink (Purcell et al., 2007). El LD se midió mediante el estadístico  $r^2$  entre las parejas de SNPs separadas por una distancia inferior a 3Mb, en cada cromosoma y población por separado utilizando el programa Haploview 4.2 (Barrett et al. 2005). Las representaciones gráficas del valor de  $r^2$  vs. Distancia y los valores medios de  $r^2$  por cromosoma y por raza se hicieron utilizando paquete estadístico R v. 2.12.1.

La selección de los marcadores SNP distintivos de la raza CM se realizó en base a las frecuencias alélicas calculadas con el programa Power marker (Liu y Muse 2005). Los criterios utilizados para dicha selección fueron: 1. SNPs con frecuencias alélicas muy altas en la raza CM y muy bajas en el resto de poblaciones; 2. SNPs con frecuencias alélicas muy bajas en la raza CM y muy altas en el resto de poblaciones; 3. Selección de un número máximo de 10 SNPs. Con el fin de evaluar en qué medida la utilización de un panel que incluyera pocos marcadores SNPs podría diferenciar a los animales de la raza CM del resto de razas, se realizaron los mismos análisis de estructuración de la población realizados previamente con los SNPs seleccionados. Para estos análisis se excluyeron aquellos cerdos CM que mostraron un grado alto de mezcla en los análisis previos con el panel PorcineSNP60. Finalmente se compararon, en cada cerdo CM independientemente, los valores de coeficiente de pertenencia obtenidos con el programa Structure en base al panel PorcineSNP60 con los obtenidos en base a los SNPs seleccionados.

## **ADN mitocondrial**

Las secuencias de ADNmt fueron revisadas y corregidas con el programa GAP4 (Bonfield et al., 1995) utilizándose la secuencia porcina GenBank AJ002189 (Ursing y Arnason 1998) como referencia. Las secuencias fueron alineadas con el programa Clustal X v.2 (Larkin et al., 2007) para posteriormente ser agrupadas en haplotipos con el programa ALTER (Glez-Peña et al., 2010). Para la creación de los dendogramas de poblaciones se utilizó el programa MEGA 5 usando el análisis NJ y el método de máxima probabilidad (maximum likelihood).

### 8. Análisis estadístico de marcadores genéticos.

Los polimorfismos genéticos observados tras analizar los distintos marcadores moleculares fueron estudiados mediante técnicas matemáticas y estadísticas con el fin de alcanzar unos resultados informativos acerca de factores como la diversidad genética, estructura de la población o las relaciones genéticas entre poblaciones. Los estadísticos más importantes utilizados en el presente trabajo fueron:

**Frecuencia alélica (F):** Proporción observada de un alelo específico respecto al total de alelos que pueden ocupar un locus en una población.

$$F_x = \frac{2N_{xx} + N_{xy}}{2N}$$

Donde  $N_{xx}$  es el número de homocigotos para el alelo "x" y  $N_{xy}$  es el número de heterocigotos que contienen el alelo "x" (siendo "y" el otro alelo).  $N$  es el número de muestras analizadas.

**Heterocigosidad observada (Ho):** Proporción de individuos heterocigotos observados con respecto al número total de individuos estudiados por locus.

$$H_o = \frac{N^{\circ} \text{ de heterocigotos}}{N}$$

Donde  $N$  es el número de muestras analizadas.

**Heterocigosidad esperada (He):** Proporción de individuos heterocigotos esperados por locus bajo condiciones de apareamiento al azar:

$$He = 1 - \sum p_i^2$$

Donde  $p_i$  es la frecuencia alélica de cada alelo ( $i$ )

El valor promedio de la He en todos los loci permite estimar la variabilidad genética de la población:

$$He = 1 - \sum_{i=1}^h p_{Ti}^2$$

Donde  $p_{Ti}$  es la frecuencia del alelo  $i$  en el total de la población y  $h$  es el número máximo de alelos.

**Número medio de alelos por locus (NMA):** Suma de todos los alelos detectados en todos los loci dividido por el número total de loci.



**Estadísticos F de Wright:** La subdivisión de una población lleva consigo un efecto de deriva genética que actúa de forma similar a la endogamia por su tendencia al aumento de homocigotos. Los estadísticos F de Wright tratan de medir este efecto en términos de reducción en la proporción de genotipos heterocigotos.

**F<sub>st</sub>:** Mide la reducción de heterocigosidad originada por la subdivisión en subpoblaciones y la deriva genética, por lo que se utiliza como medida de diferenciación genética entre subpoblaciones. Su valor teórico está comprendido entre 0 y 1. Cuando las poblaciones tienen frecuencias similares el valor F<sub>st</sub> es pequeño mientras que si se alcanza la fijación la diferenciación es grande, y por tanto adquiere el valor de 1.

$$F_{ST} = \frac{H_e - \bar{H}_e}{H_e}$$

Donde  $H_e$  es el valor promedio de la heterocigosidad esperada en todos los loci y  $\hat{H}_e$  es el promedio de heterocigosidad esperada en las subpoblaciones ( $k$ ):

$$\bar{H}_e = \sum_{l=1}^k \frac{H_e}{k}$$

Wright propuso interpretar los valores de  $F_{st}$  del siguiente modo:

- 0 - 0,05 = muy poca diferenciación genética
- 0,05 - 0,15 = poca a moderada diferenciación genética
- 0,15 - 0,25 = alta diferenciación genética
- 0,25 = diferenciación genética muy grande

**Distancia genética:** es una medida cuantitativa de la divergencia genética entre individuos o poblaciones. La distancia genética de Nei (Nei, 1972) es la más utilizada. Varía entre 0 (poblaciones con frecuencias alélicas idénticas) e infinito (poblaciones con ningún alelo en común). Su valor tiene en cuenta los cambios en las frecuencias alélicas derivados de mutaciones y de los efectos de la deriva genética.

$$D_{ij} = -\ln \left( \frac{\sum_{k=1}^q |x_{kt} x_{kj}|}{\sqrt{\sum_{k=1}^q x_{kt}^2 x_{kj}^2}} \right)$$

Donde  $D_{ij}$  es la distancia genética de Nei (Nei, 1972) entre las poblaciones  $i$  y  $j$ ;  $x_{ki}$  y  $x_{kj}$  es la frecuencia del alelo  $k$  en las poblaciones  $i$  y  $j$  respectivamente;  $q$  es el número máximo de alelos para el locus  $x$ .

**Estructura genética de la población:** El programa Structure (Pritchard et al., 2000) utiliza un algoritmo de agrupamiento Bayesiano que asume que un individuo puede ser descendiente de otras poblaciones. Para deducir la estructura de la población el programa asume que hay un número  $K$  de grupos o clusters (siendo  $K$  un valor desconocido y asignado por el investigador) en el que cada grupo se caracteriza por el conjunto de frecuencias alélicas de cada locus, asignando a los individuos a su población o poblaciones más probables en base a sus frecuencias alélicas. El programa genera coeficientes de pertenencia a un cluster o grupo, tanto para cada individuo como para el conjunto de cada población estudiada, que varía entre 0 y 1. Este análisis puede realizarse a partir de marcadores microsatélite y SNP y permite estudiar la estructura de las poblaciones, asignar a cada individuo a su población así como identificar individuos mezclados o cruzados.

**Identity by state (IBS).** Método para estudiar la estructura de la población. A diferencia del anterior, se basa en los alelos compartidos por dos individuos en cada SNP (Roberson y Pevsner, 2009). De esta forma se estima la similitud genotípica entre dos individuos considerando que, para cada SNP, se pueden dar tres situaciones: 1. Los dos individuos tienen el mismo genotipo; 2. Solamente un alelo es compartido; 3. Ningún alelo es compartido por ambos individuos. La matriz de datos resultante puede representarse gráficamente mediante un análisis de **escalamiento multidimensional (MDS)** que es una técnica multivariante de interdependencia que trata de representar en un espacio geométrico de pocas dimensiones las proximidades existentes entre un conjunto de objetos.

**Análisis neighbour-joining (NJ):** Algoritmo utilizado para la construcción de dendogramas en base a una matriz de distancias genéticas (Saitou y Nei, 1987).



# RESULTADOS



---

## RESULTADOS

### 1. Estudio de la diversidad y subestructura genética de la raza Chato Murciano.

El estudio de la diversidad genética de la población de CM en su conjunto mostró un valor medio de  $He=0,54$  y de  $Ho=0,52$  (**Tabla 4A**). Todos los marcadores microsatélite fueron polimórficos siendo el número medio de alelos por loci (NMA) 6,39, variando entre 2 (SW796) y 12 (S0005) alelos (**Anexo II**). El estudio de los parámetros de diversidad genética de cada granja por separado se resume en la **Tabla 4B**. La diversidad genética fue inferior a 0,5 en todas las granjas menos en la granja SS, la cual mostró los valores más elevados de variabilidad genética en términos de  $He$  (0,56) y de NMA (4,73). Por el contrario, las granjas AI y MC fueron las que presentaron los valores más bajos de  $He$  (0,32 y 0,26) y de NMA (2,26 y 1,18) respectivamente (**Figura 1**), además de un número de marcadores microsatélites monomórficos más elevado que el observado en las otras granjas de cría de cerdo CM como se detallará posteriormente (**Tabla 13**).

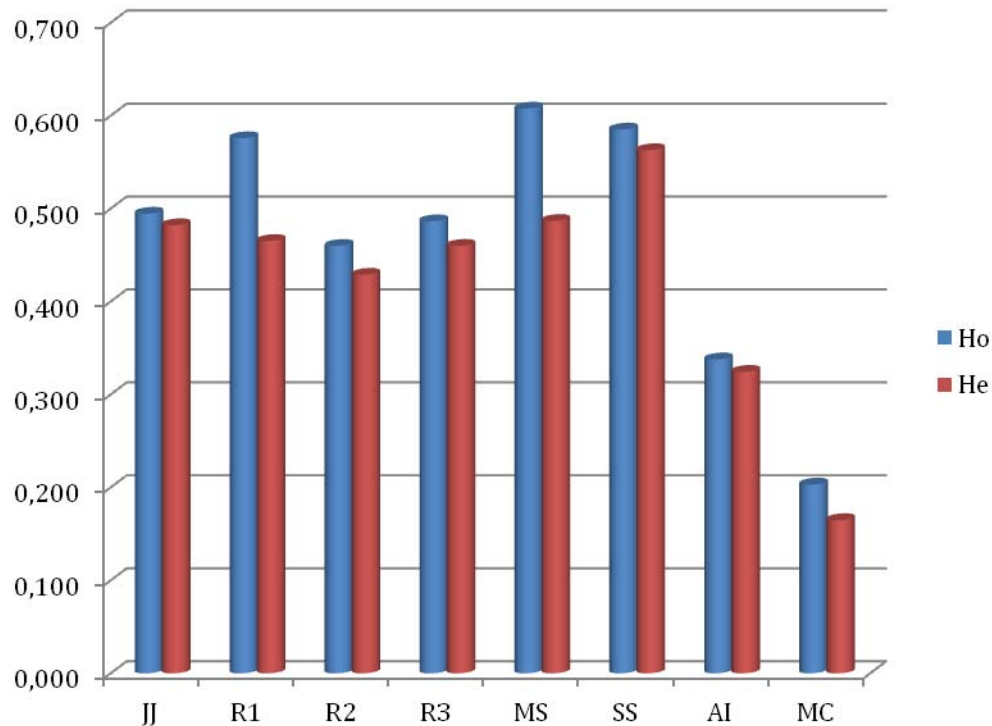


**Tabla 4A.** Número de alelos ( $N_a$ ) y  $H_e$ - $H_o$  en toda la población de CM

POBLACIÓN	N	$N_a$	NMA	$H_o$	$H_e$
CM	175	217	6,39	0,52	0,54

**Tabla 4B.** Número de alelos y  $H_e$ - $H_o$  en las granjas de CM

POBLACIÓN	N	$N_A$	NMA	$H_o$	$H_e$
JJ	36	122	3,50	0,49	0,48
R1	16	114	3,35	0,58	0,46
R2	21	116	3,41	0,46	0,43
R3	32	138	4,06	0,49	0,46
MS	7	95	2,79	0,61	0,49
SS	42	161	4,73	0,58	0,56
AI	11	77	2,26	0,34	0,32
MC	10	40	1,18	0,32	0,26

**Figura 1.** Gráfico de las He y Ho en cada población

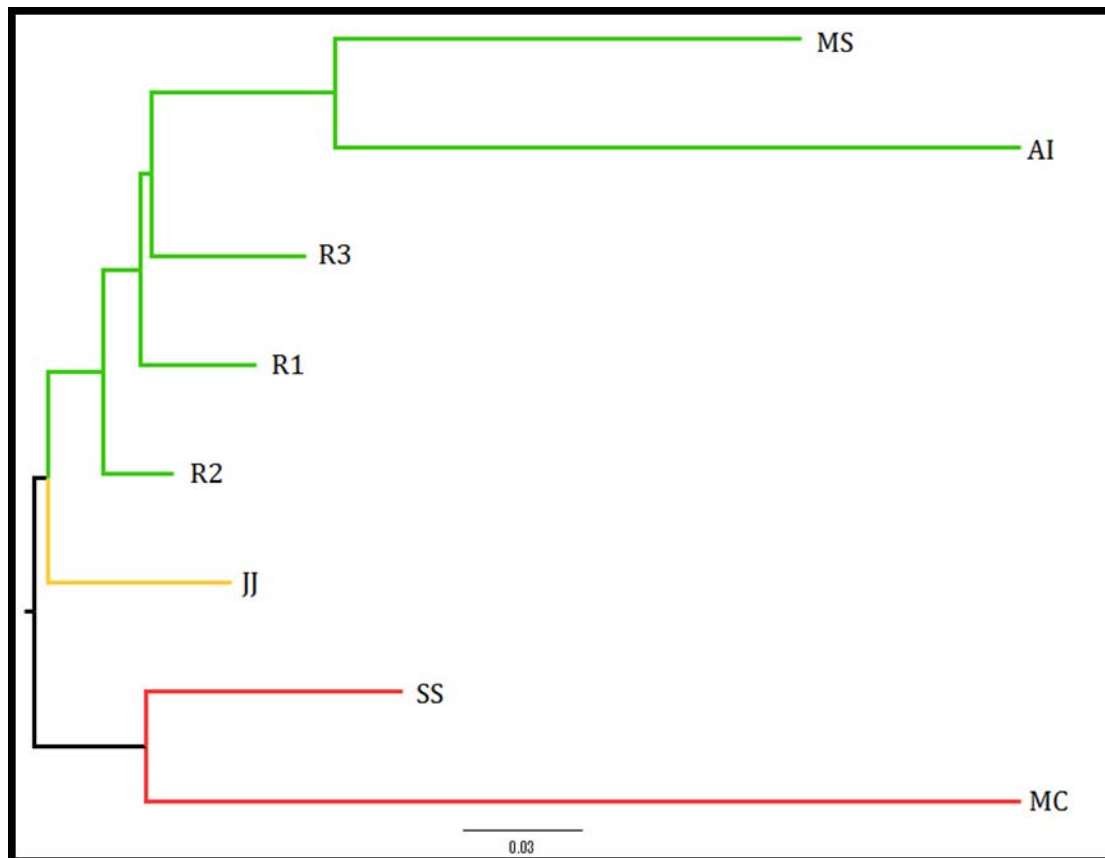
El valor medio del estadístico  $F_{st}$  de Wright para toda la población de CM y loci microsatélites fue  $F_{st}=0,11$ , indicando que el 11% de la variación genética existente en la población se explica por las diferencias genéticas entre granjas. En la **Tabla 5** se muestran los valores entre pares de granjas para los estadísticos  $F_{st}$  y distancias genéticas de Nei ( $D_A$ ) (Nei, 1972). Las tres granjas R1, R2 y R3, presentaron entre sí una diferenciación genética baja o moderada mientras que los valores  $D_A$  entre estas granjas fueron los más bajos

(<0,07). Las otras dos explotaciones consideradas grandes por el número de reproductores de la raza CM criados en ellas, granjas JJ y SS, presentaron una diferenciación genética moderada entre ellas y con las granjas R1-3. Las granjas de menor tamaño (granjas MS, AI y MC) fueron las que mostraron una mayor diferenciación genética al compararse con el resto de granjas ( $F_{st}>0,1$ ), además de presentar valores  $D_A$  más elevados entre pares de granjas. Concretamente, la granja MC fue la que presentó los valores  $F_{st}$  y  $D_A$  más elevados y por tanto fue la población de cerdos CM más divergente. La representación gráfica en base a las distancias genéticas de Nei (**Figura 2**) permitió diferenciar dos grupos principales (granjas R1-3, MS y AI en verde y granjas SS y MC en rojo). La granja JJ (en amarillo) se encuadró en una rama del árbol diferente, próximo al grupo coloreado en verde.

**Tabla 5.** Valores  $F_{st}$  (cuadrante inferior izquierdo) y  $D_A$  (cuadrante superior derecho) entre pares de granjas

	JJ	R1	R2	R3	MS	SS	AI	MC
JJ	0	0,106	0,095	0,118	0,216	0,124	0,253	0,320
R1	0,051	0	0,052	0,069	0,208	0,175	0,245	0,286
R2	0,049	0,029	0	0,064	0,241	0,142	0,225	0,239
R3	0,073	0,052	0,052	0	0,217	0,183	0,246	0,324
MS	0,105	0,103	0,124	0,121	0	0,194	0,292	0,474
SS	0,050	0,072	0,065	0,085	0,085	0	0,359	0,287
AI	0,141	0,139	0,135	0,156	0,170	0,161	0	0,557
MC	0,204	0,193	0,181	0,222	0,281	0,175	0,351	0

**Figura 2.** Árbol NJ de las 8 granjas de Chato Murciano en base a  $D_A$  obtenidas del análisis de 34 marcadores microsatélites.



Con el objetivo de determinar la subestructura genética de la raza CM, se utilizó el algoritmo de agrupamiento implementado por el programa Structure en base a los datos microsatélite y a los datos Porcine SNP60 en cada una de las granjas estudiadas (**Figura 3**). Se testaron valores de K entre dos y cuatro. Cuando K=2, los resultados fueron consistentes entre los dos tipos de marcadores utilizados, microsatélites y SNP, y también con el dendograma basado en las distancias genéticas (**Figura 2**). De esta forma se detectaron dos subpoblaciones genéticas: la primera representada gráficamente en color verde y que incluyó a las granjas R1-3, MS y AI; la segunda representada en color rojo que incluyó a las granjas SS y MC. La granja JJ incluyó animales de ambas subpoblaciones. Sin embargo, cuando K = 3 la mayoría de los animales de la granja JJ formaron un grupo independiente (de color amarillo) al considerar los marcadores microsatélites, aunque no se detectó tal diferenciación genética al usar los marcadores SNP. Los resultados para K=3 y K=4 fueron idénticos entre ellos por lo que la existencia de una cuarta subpoblación fue descartada.

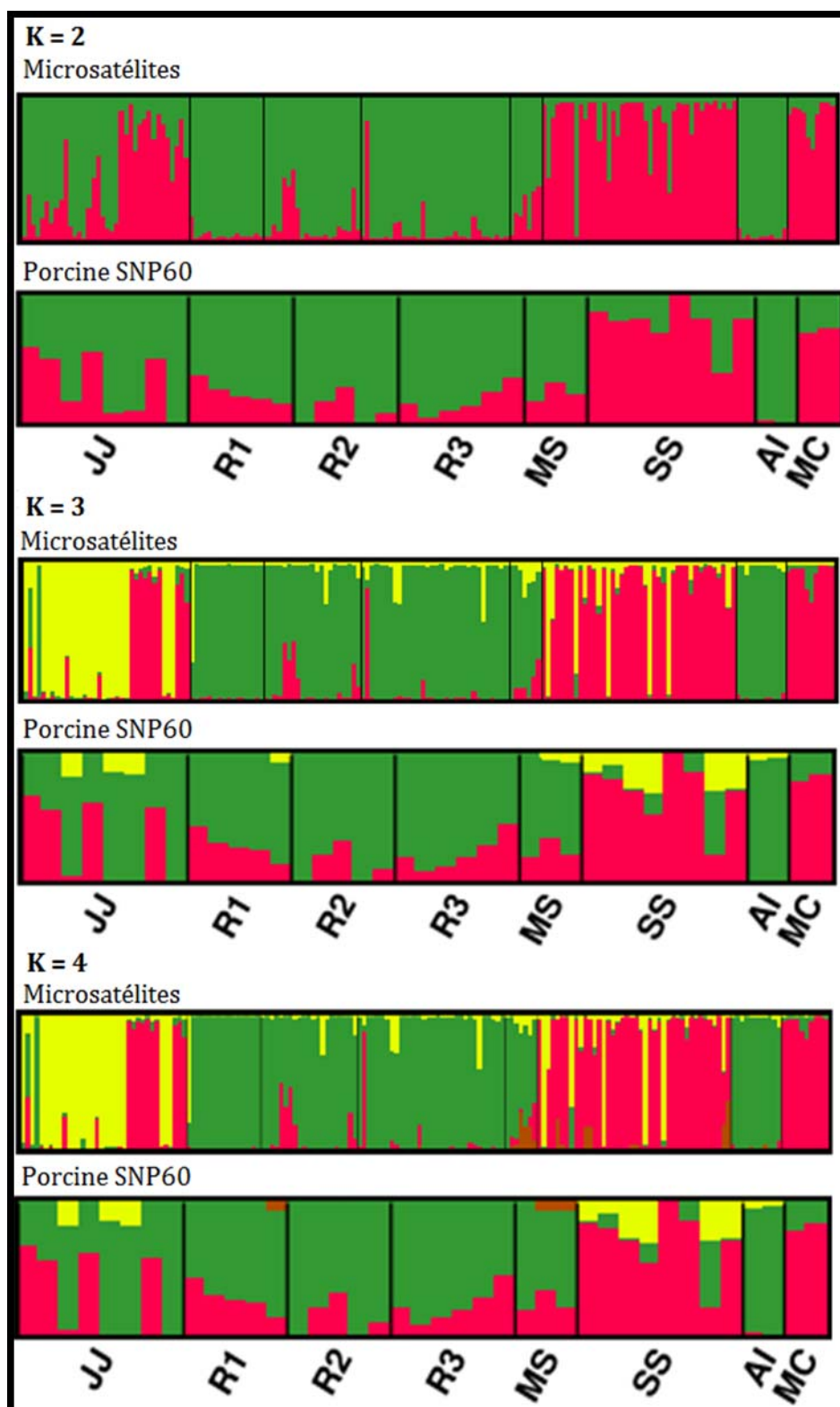
El estudio detallado de los coeficientes de pertenencia (CP) determinados por el programa Structure mostró que las granjas R1-R3, AI y MC se colorearon de forma muy homogénea como resultado de

valores elevados de CP obtenidos en las mismas, especialmente en las granjas AI y MC (0,96 y 0,93 respectivamente) lo que indica una diferenciación genética elevada. Sin embargo, la otra granja considerada de menor tamaño, granja MS, se agrupó con los valores CP más bajos en subpoblación marcada en verde (0,71). La granja JJ mostró siempre los valores CP más bajos ( $CP < 0,68$ ) independientemente del marcador genético utilizado y del valor de K (Tabla 6).

**Tabla 6.** Coeficientes de pertenencia para  $K = 3$  en base a los análisis de marcadores microsatélites.

Grupos inferidos por el programa (K = 3)				
Grupos predefinidos		1	2	3
	JJ	0,643	0,283	0,074
	R1	0,054	0,008	0,938
	R2	0,032	0,083	0,884
	R3	0,042	0,044	0,914
	MS	0,181	0,113	0,706
	SS	0,162	0,812	0,025
	AI	0,024	0,014	0,962
	MC	0,009	0,926	0,066

**Figura 3.** Representación gráfica del coeficiente de pertenencia para valores  $K=2-4$ . Para cada valor de  $K$ , se detallan los resultados en base a los datos microsatélite y al panel Porcine SNP60.



2. Estudio de la relación genética del Chato Murciano con otras razas.

**Genotipado de alta densidad**

El análisis de escalamiento multidimensional (MDS) basado en los datos Porcine SNP60 permitió diferenciar tres grupos principales (**Figura 4**). El primero, y también el más divergente, fue el grupo “Asiático” representado exclusivamente por la raza MH. Tomando como referencia el grupo asiático, el segundo grupo se situó en el extremo opuesto de la primera componente y estuvo representado por la raza DU. Finalmente, el tercer grupo fue el que incluyó a un mayor número de individuos pertenecientes a las razas europeas LW, LR, TA, BK, IB y CM. Dentro de este grupo de razas europeas, la raza IB fue la más diferenciada situándose en el extremo próximo a la raza DU. En el extremo opuesto del grupo europeo, y por tanto más próximo al asiático, se situaron las razas LW y LR.

Debido a que una gran parte de la variación observada en la **Figura 4** se debió a las razas MH y DU, estas dos razas se eliminaron del análisis MDS para así realizar un estudio más detallado del grupo europeo (**Figura 5A**). La gráfica resultante mostró que todos los animales se agruparon según su raza de origen. A su vez, todas las

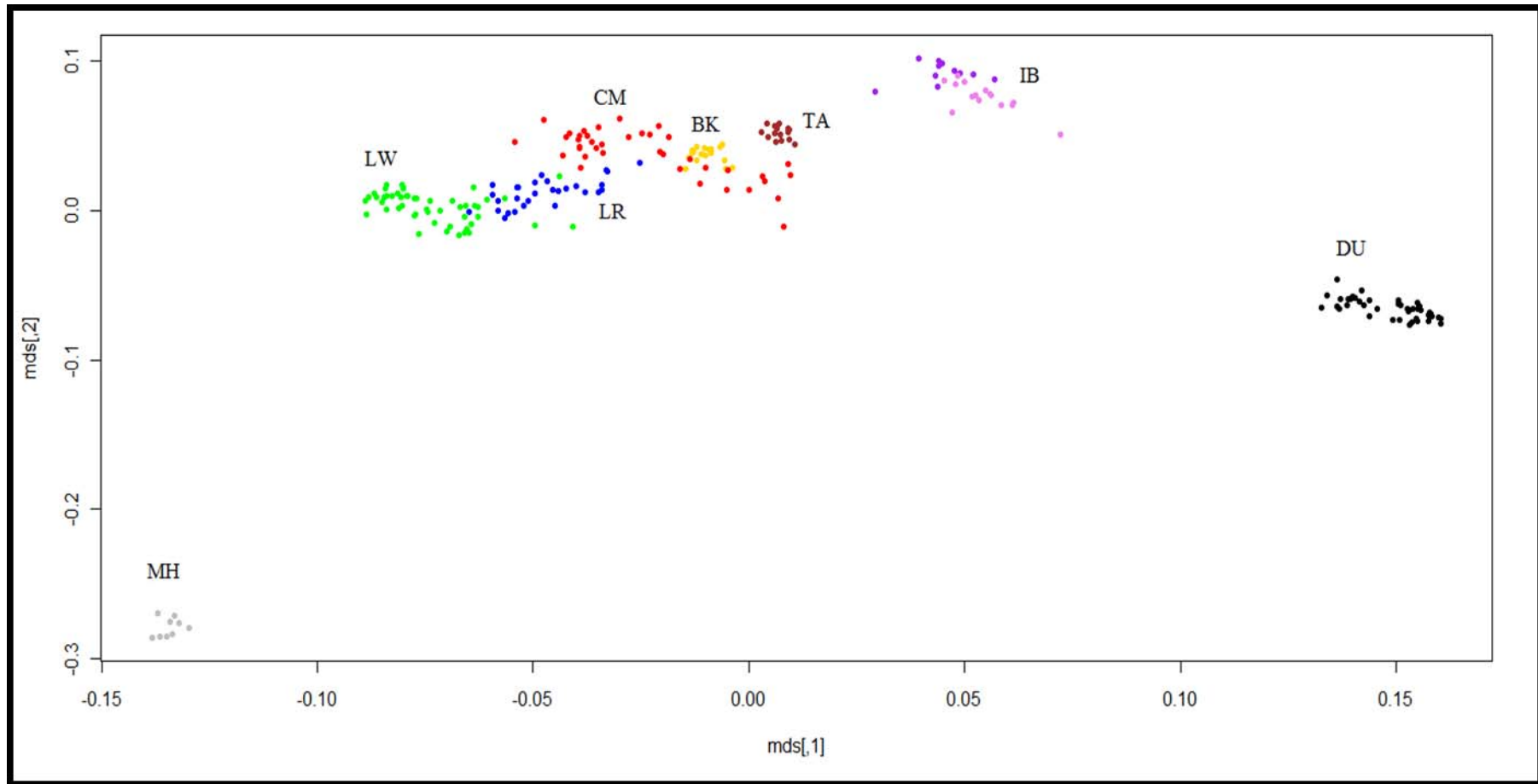


razas porcinas fueron claramente diferenciadas entre sí. Las variedades RE, en color violeta, y NI, en color púrpura, de Ibérico no se diferenciaron entre ellas, si bien hay que tener en cuenta que ambas constituyen dos variedades de la misma raza (IB). Las razas LW y LR se diferenciaron poco entre ellas pero claramente del resto de razas europeas.

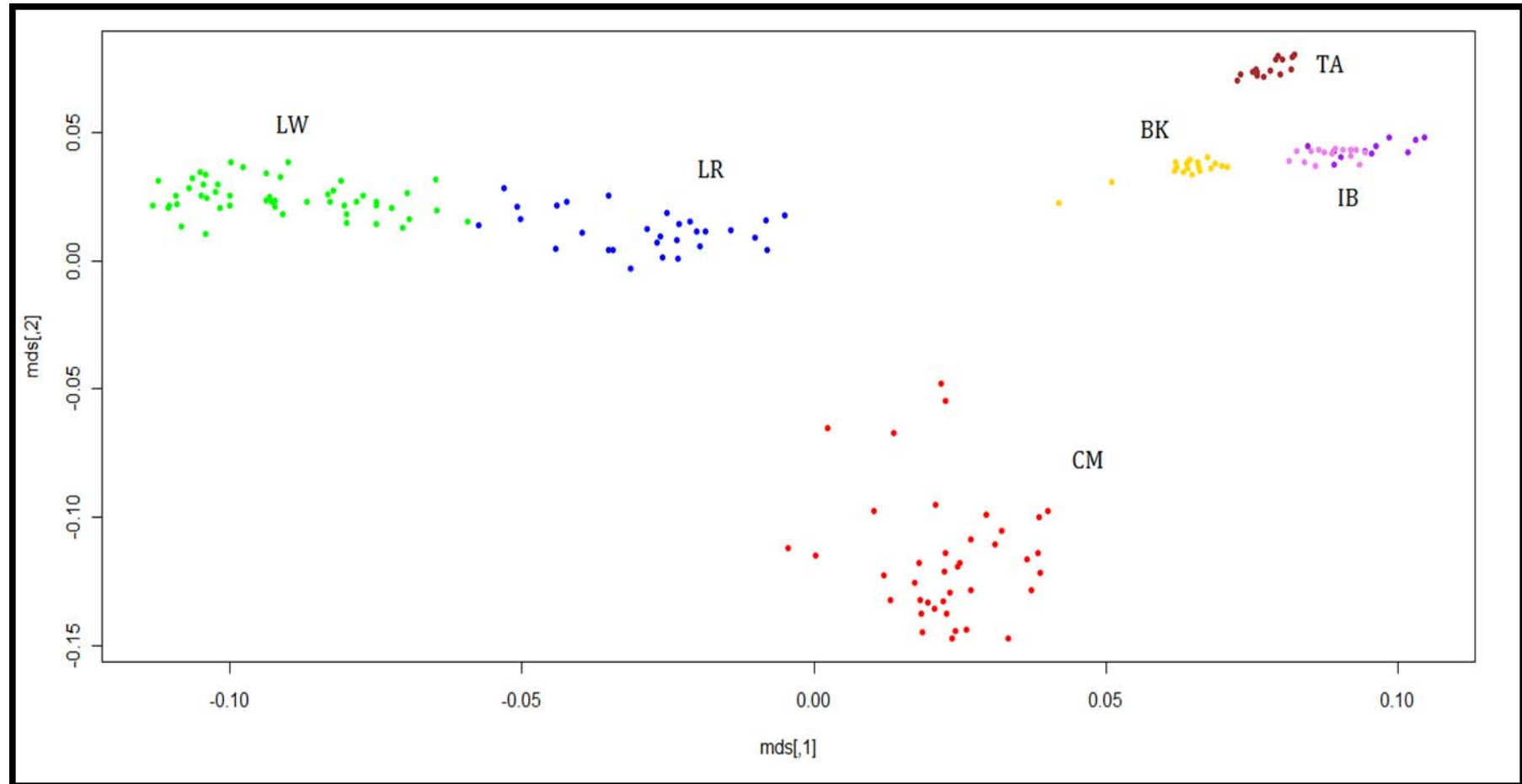
Todos los animales pertenecientes a la raza CM se agruparon en la misma área del gráfico, diferenciándose inequívocamente de las otras razas estudiadas. Cuando se consideró el primer componente de la gráfica (eje mds, 1) la población CM se situó en una posición intermedia entre las razas LW-LR y la raza BK. El segundo componente (eje mds, 2) explicó las diferencias entre el grupo CM y el resto de razas al quedar encuadrado en la región inferior del gráfico. Aunque todos los individuos de la raza CM se agruparon en el mismo área del gráfico, cuatro cerdos CM pertenecientes a la granja SS se separaron claramente del grupo principal en dirección a las otras razas europeas (**Figura 5B**).



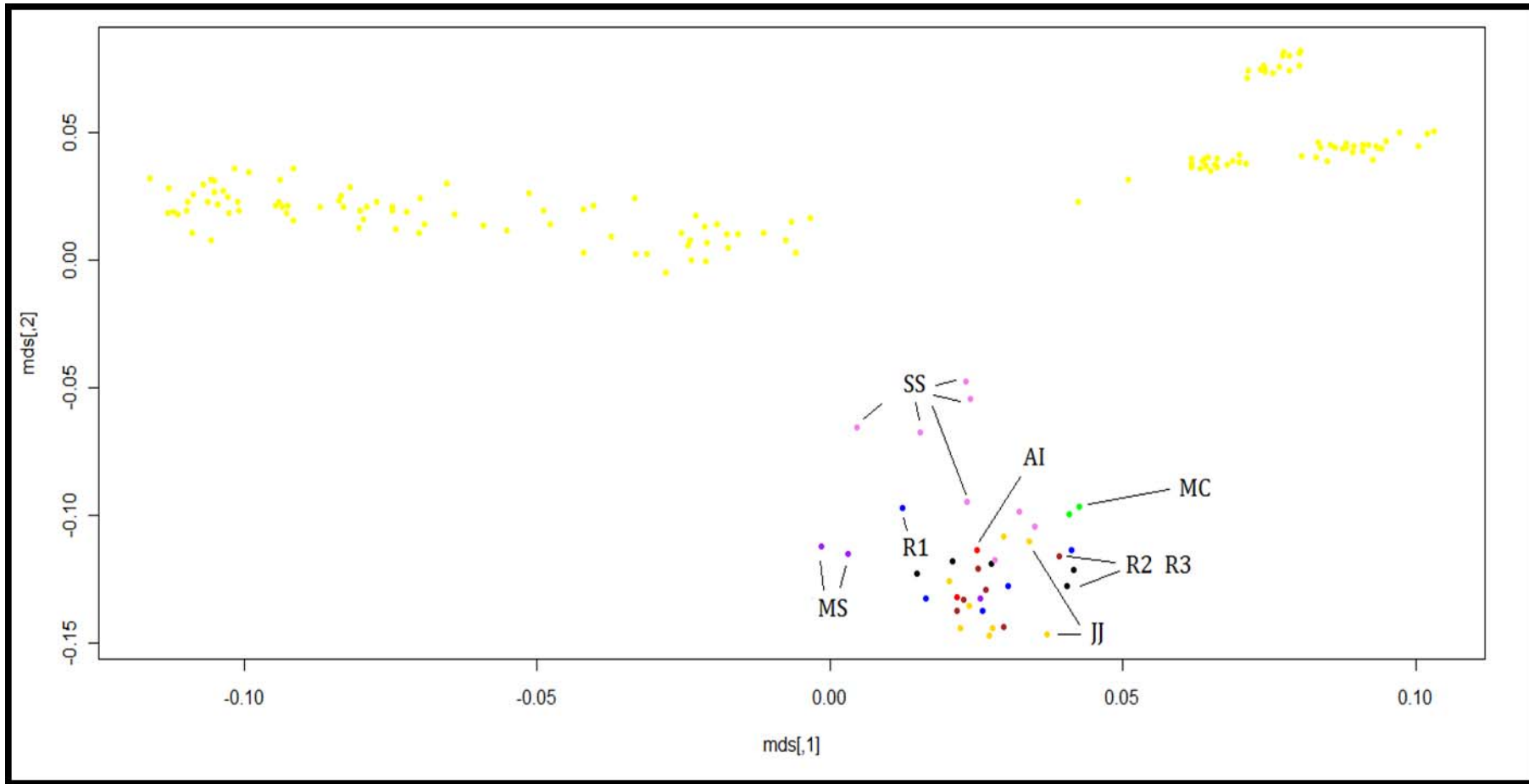
**Figura 4.** Escalamiento multidimensional en base en base a los datos Porcine SNP60. Se representan todas las razas incluidas en el estudio



**Figura 5A.** Escalamiento multidimensional en base en base a los datos Porcine SNP60. Detalle de las razas europeas.



**Figura 5B.** Escalamiento multidimensional en base en base a los datos Porcine SNP60. Diferenciación entre los animales pertenecientes a cada granja.



Para determinar si la raza CM es una población genéticamente diferenciada y detectar cerdos CM mezclados o cruzados con otras razas, se utilizó el modelo bayesiano de agrupamiento implementado por el programa Structure en base a la información aportada por los marcadores SNPs. Debido a que se incluyeron un total de nueve poblaciones en el estudio, se testaron valores de K entre K= y K=9 (**Figura 6A**). K=8 resultó ser el valor óptimo de K según el test BIC (**Figura 7**). Para un valor K=8 todos los individuos fueron inequívocamente asignados a su población de origen, maximizándose el valor del coeficiente de pertenencia en las nueve poblaciones estudiadas. Las razas LW y LR se agruparon conjuntamente para valores K=6 y 7 (grupo verde) lo que sugiere una mayor proximidad genética entre las mismas. Las razas BK y TA solo se agruparon conjuntamente cuando K=6. El resto de poblaciones se diferenciaron para todos los valores de K testados con la excepción de las variedades de cerdo Ibérico RE y NI que se agruparon en el mismo grupo (color marrón) para todos los valores de K testados. El coeficiente de pertenencia más elevado se observó en la raza MH con un valor CP=1 que implica una correspondencia del 100% entre la población asumida por el investigador y la población asignada por el programa Structure y que se muestra gráficamente por un único color violeta en la región

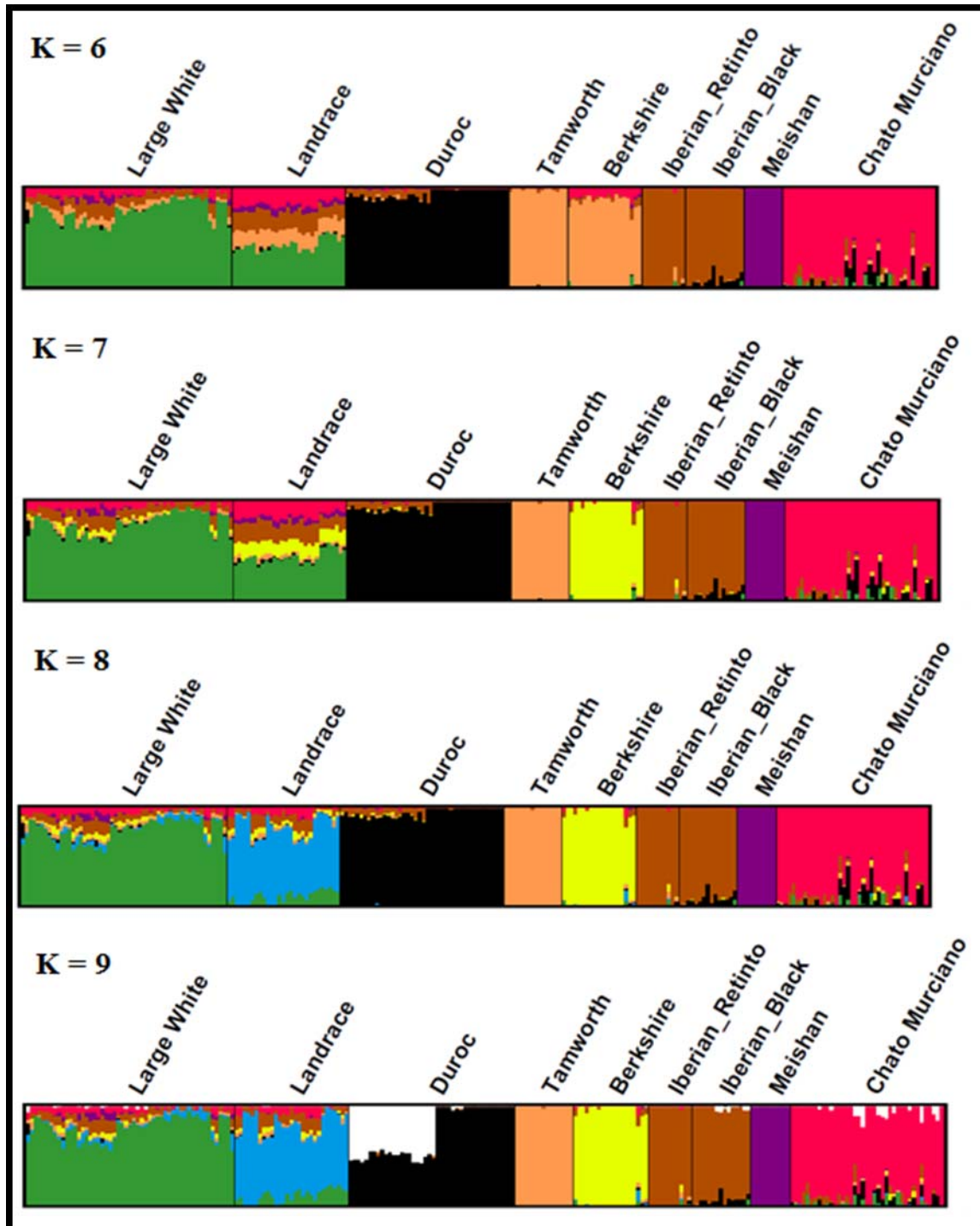
asignada a los individuos de dicha raza. Por el contrario las razas LW y LR presentaron los valores de CP más bajos (0,77 y 0,71 respectivamente). El resto de poblaciones se agruparon con valores superiores a 0,85 para todos los valores de K testados. Concretamente, las razas DU, BK, TA y IB tuvieron valores CP superiores a 0,93 mientras que la raza CM lo hizo con un valor CP=0,85 (**Tabla 7**).

Un estudio detallado de la raza CM (**Figura 6B**) mostró la contribución del resto de razas estudiadas a la genética actual del cerdo CM. Todos los individuos de la raza CM analizados se agruparon con CP superiores a 0,75 con la excepción de cuatro animales que lo hicieron con valores mucho menores (CP≈0,5) (**Anexo III**). Dichos animales pertenecían a la granja SS. Por otro lado, los resultados obtenidos indicaron que las razas DU e IB fueron las que contribuyeron en mayor medida a la genética actual del CM con un 6 y 4% respectivamente, seguidos de LW (2%) y BK (1%). La raza MH, utilizada como control de raza que no ha tenido relación con el cerdo CM, fue la única raza con una participación del 0% confirmándose así la falta de relación entre ambas razas. Con el fin de eliminar el efecto que los cuatro animales considerados como mezclados (valores CP≈0,5) tuvieron sobre la media total del valor de CP en la raza CM, se realizó un segundo análisis sin considerar a dichos animales (**Tabla 8**).

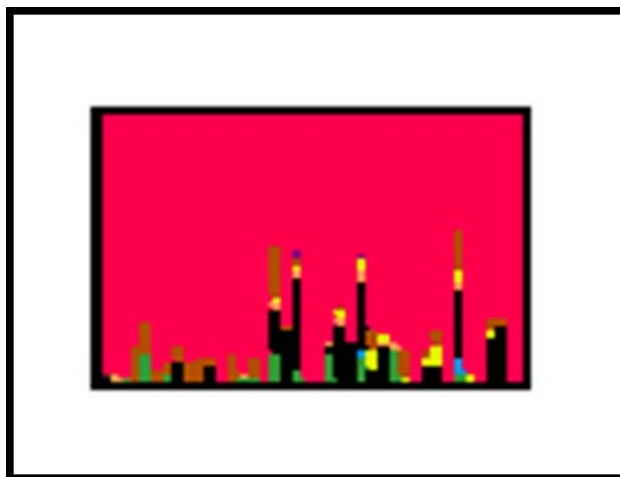
El CP total de la raza CM aumentó de 0,85 a 0,899 disminuyendo principalmente la contribución de la raza DU que pasó de un 6% a un 4% mientras que la contribución del resto de razas fue similar a la obtenida al considerar a todos los cerdos CM.



**Figura 6A.** Estructuración de la población para K=6 a K=9. Existe una sobreestimación del número de poblaciones que se observa como zona blanca en la región del grafico designado para la raza DU.



**Figura 6B.** Detalle de la estructuración de la población de la raza Chato Murciano. Cuatro animales presentan valores  $CP \approx 0,5$ .



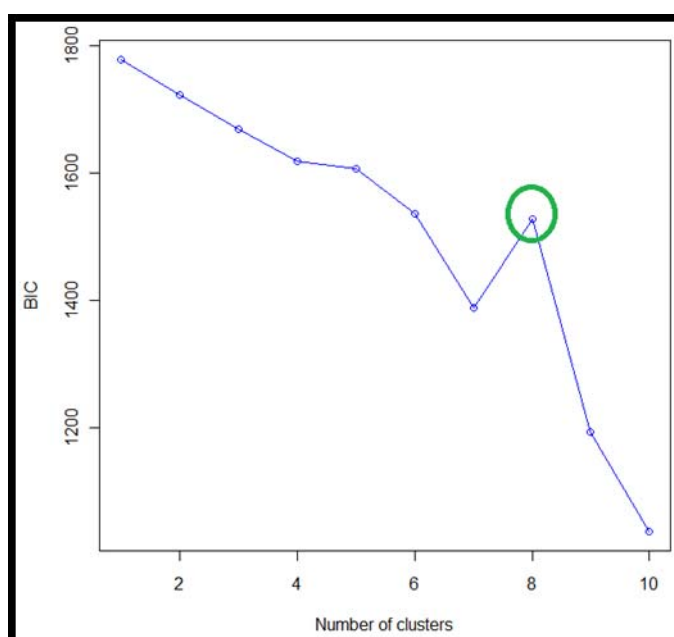
**Tabla 7.** Proporción de la contribución de los ancestros inferidos por el programa Structure (K=8) a las razas estudiadas.

		GRUPOS INFERIDOS POR EL PROGRAMA (K = 8)								
		1_DU	2_BK	3_MH	4_LR	5_LW	6_TA	7_CM	8_IB	N
GRUPOS PREDEFINIDOS	LW	0,011	0,025	0,017	0,036	0,768	0,030	0,045	0,069	49
	LR	0,005	0,036	0,002	0,710	0,087	0,025	0,048	0,087	29
	DU	0,957	0,005	0,001	0,001	0,000	0,003	0,001	0,033	42
	TA	0,001	0,000	0,000	0,000	0,000	0,998	0,000	0,001	15
	BK	0,003	0,950	0,001	0,007	0,005	0,006	0,008	0,020	19
	IB_RE	0,004	0,008	0,000	0,002	0,004	0,010	0,002	0,969	11
	IB_NI	0,063	0,002	0,000	0,000	0,004	0,001	0,001	0,929	15
	MH	0,000	0,000	1,000	0,000	0,000	0,000	0,000	0,000	10
	CM	0,064	0,011	0,000	0,003	0,019	0,007	0,854	0,041	39

**Tabla 8.** Proporción de la contribución de los ancestros inferidos por el programa Structure (K=8) a las razas estudiadas sin considerar a los cerdos Chato Murciano con valores  $CP \approx 0,5$ .

		GRUPOS INFERIDOS POR EL PROGRAMA (K = 8)								
		1_CM	2_LW	3_MH	4_IB	5_DU	6_TA	7_BK	8_LR	N
GRUPOS PREDEFINIDOS	LW	0,050	0,780	0,017	0,092	0,008	0,011	0,019	0,024	49
	LR	0,052	0,095	0,001	0,103	0,002	0,016	0,026	0,705	29
	DU	0,003	0,000	0,001	0,040	0,954	0,002	0,001	0,000	42
	TA	0,000	0,000	0,000	0,002	0,000	0,998	0,000	0,000	15
	BK	0,009	0,005	0,001	0,022	0,003	0,005	0,950	0,006	19
	IB_RE	0,004	0,006	0,000	0,970	0,003	0,009	0,007	0,000	11
	IB_NI	0,001	0,004	0,000	0,934	0,060	0,000	0,000	0,000	15
	MH	0,000	0,000	1,000	0,000	0,000	0,000	0,000	0,000	10
	CM	0,899	0,015	0,000	0,036	0,041	0,002	0,007	0,000	35

**Figura 7.** Valor BIC (Bayesian Information Criterion) Vs. Número de grupos (K) para la determinación del valor de K más probable. Se observa el pico para el valor K = 8

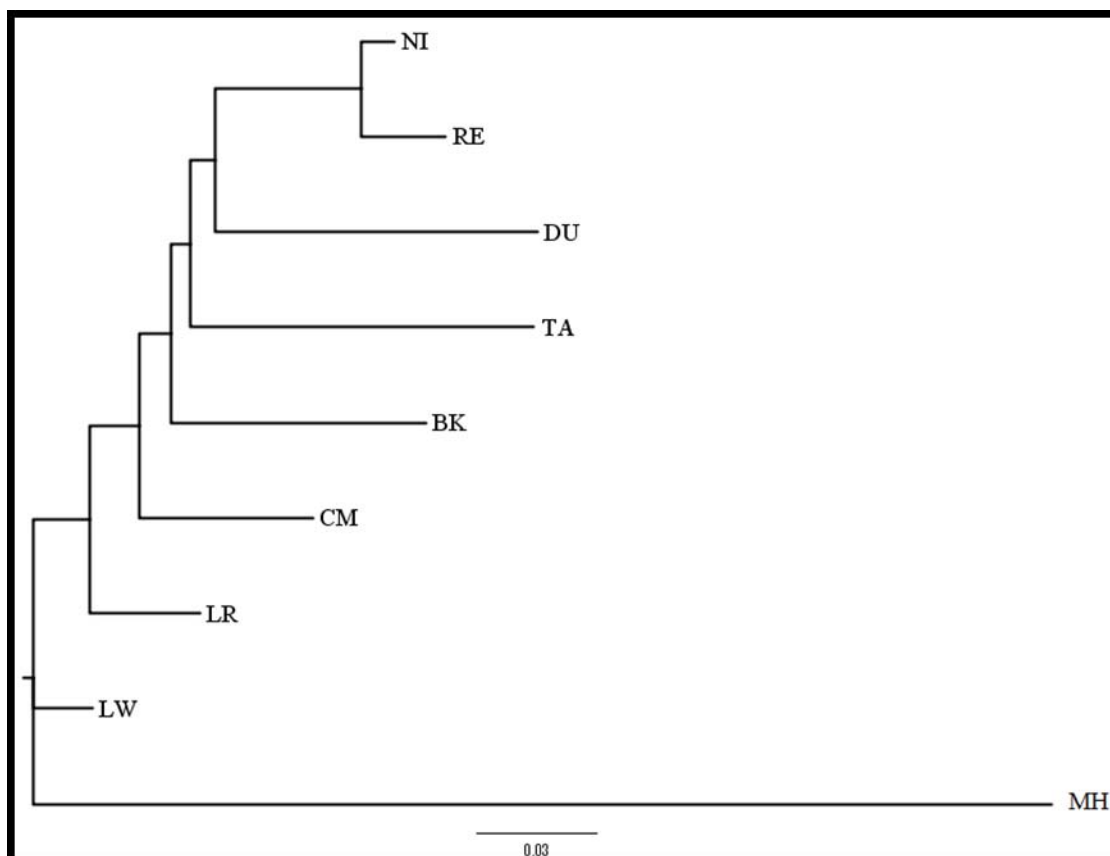


Los valores  $F_{st}$  entre pares de poblaciones (**Tabla 9**) fueron más altos entre la raza MH y el resto de razas estudiadas ( $>0,45$ ) mientras que dicho estadístico registró el valor más bajo al comparar las variedades RE y NI (0,068). La raza CM tuvo los valores  $F_{st}$  más bajos cuando se comparó con las razas LW y LR (0,15). Resultados similares se obtuvieron cuando se estudiaron las distancias genéticas (Nei, 1972), representadas gráficamente en la **Figura 8**. La raza MH fue de nuevo la más divergente mientras que las variedades RE y NI de cerdo ibérico presentaron el valor de distancia genética más bajo. Las razas que presentaron menores distancias genéticas en relación al CM fueron LR y LW (0,08) seguidas de IB ( $\approx 0,11$ ) y BK (0,12).

**Tabla 9.** Distancias genéticas de Nei (bajo la diagonal) y valores  $F_{st}$  entre pares de poblaciones (arriba de la diagonal).

	BK	CM	DU	LR	LW	MH	NI_IB	RE_IB	TA	
	<i>Valores <math>F_{st}</math> entre pares de poblaciones</i>									
<b>BK</b>	0,000	0,248	0,308	0,218	0,261	0,521	0,256	0,289	0,330	
<b>CM</b>	0,120	0,000	0,247	0,149	0,156	0,517	0,251	0,315	0,340	
<b>DU</b>	0,162	0,135	0,000	0,224	0,237	0,540	0,271	0,351	0,383	
<b>LR</b>	0,112	0,089	0,139	0,000	0,113	0,463	0,227	0,285	0,299	
<b>LW</b>	0,114	0,084	0,139	0,067	0,000	0,517	0,289	0,375	0,358	
<b>MH</b>	0,346	0,322	0,346	0,290	0,273	0,000	0,558	0,587	0,581	
<b>NI_IB</b>	0,123	0,109	0,121	0,106	0,107	0,404	0,000	0,068	0,306	
<b>RE_IB</b>	0,131	0,118	0,146	0,115	0,117	0,427	0,030	0,000	0,329	
<b>TA</b>	0,154	0,149	0,181	0,140	0,136	0,389	0,137	0,143	0,000	

**Figura 8.** Árbol NJ construido en base a las distancias genéticas de Nei (1972) utilizando los datos Porcine SNP60.



### **ADN mitocondrial**

Se secuenció la región D-loop de ADNmt con el fin de estudiar el origen y las relaciones genéticas entre el cerdo CM y otras razas porcinas. La **Figura 9A** muestra el dendograma construido por el método NJ en base a las distancias genéticas entre las secuencias de la región D-loop. Para cada rama del árbol se detalló el número del haplotipo, seguido del código de la raza y del porcentaje de individuos de la raza incluidos en ese haplotipo con respecto al total de cerdos de dicha raza. En el dendograma se pudieron diferenciar dos zonas o nodos principales: El nodo superior, que incluyó exclusivamente razas europeas y comerciales; la zona inferior o nodo asiático, que incluyó a todos los animales de la raza china MH así como algunos animales pertenecientes a razas comerciales y británicas.

Se detectaron un total de 43 haplotipos, de los cuales 15 se agruparon en el nodo asiático y 28 en el europeo (**Tabla 10**). Se estudió detalladamente la presencia o ausencia de haplotipos asiáticos en las razas porcinas europeas. La frecuencia de haplotipos asiáticos en razas británicas fue muy elevada, estando presentes en el 100% de los individuos de la raza BK y en el 93% de la raza TA. Las razas comerciales mostraron un número de haplotipos muy elevado,

concretamente 29, lo que representa el 67% del total de haplotipos observados. El 35% de los cerdos LW, el 38% de los cerdos LR y el 7% de los cerdos DU portaron haplotipos asiáticos. De todas las razas europeas estudiadas, solamente dos portaron exclusivamente haplotipos europeos: las razas IB y CM. Se detectaron un total de 10 haplotipos en cerdos IB, cinco de los cuales fueron exclusivos de dicha población. En lo que respecta al CM, se observaron un total de cinco haplotipos (HP1-HP5). El 77% de los cerdos CM portaron haplotipos muy frecuentes en razas porcinas comerciales (HP2, 3, 4); el 7% presentaron el HP1 que se compartió exclusivamente con el cerdo IB. Finalmente, se observó un haplotipo exclusivo del CM (HP5) el cual apareció en la población con una frecuencia del 14%.

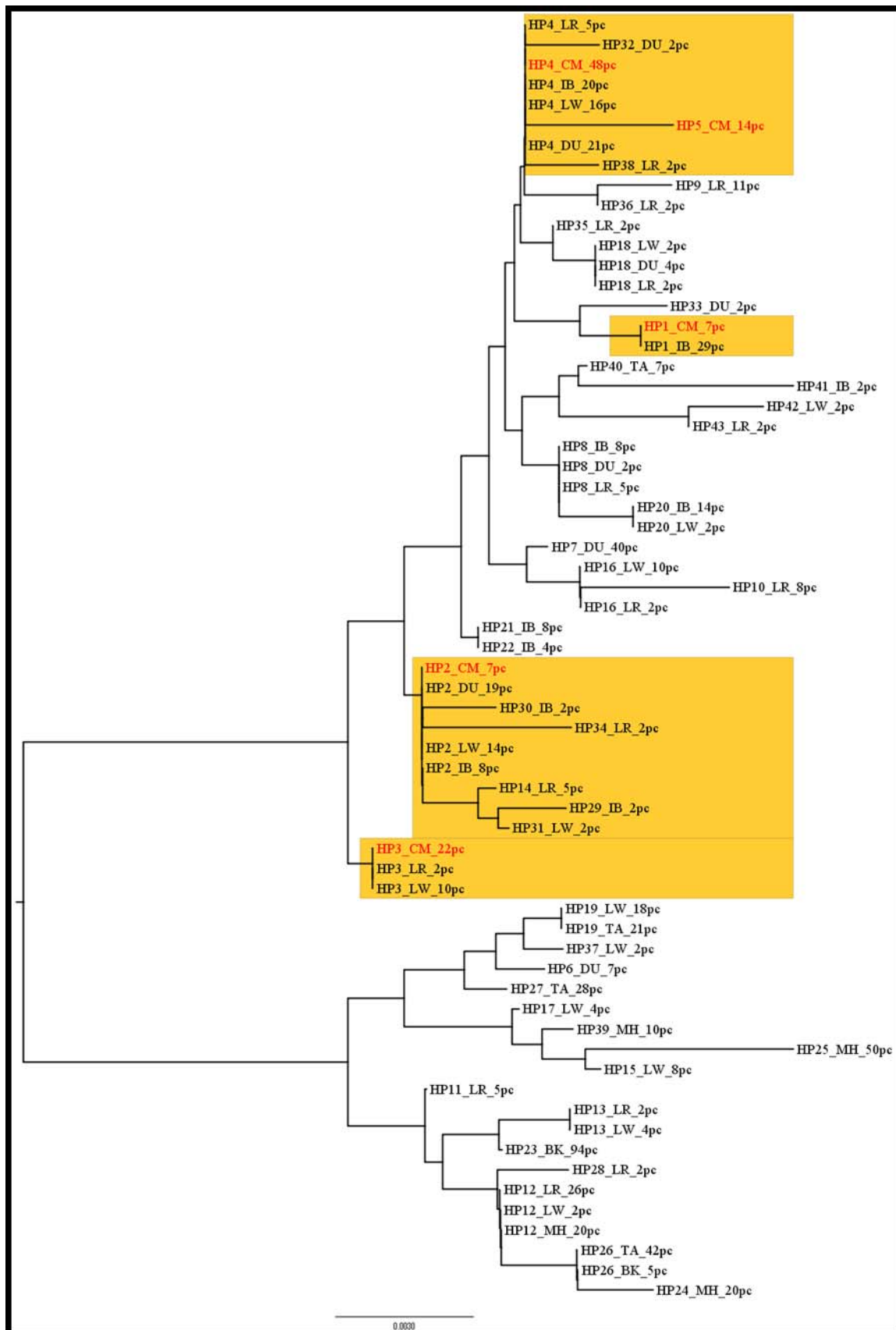
Para un estudio más detallado de la raza CM se realizó un dendograma utilizando el mismo método que el descrito para la **Figura 9A** pero detallándose la granja de procedencia de cada uno de los cerdos CM, además del haplotipo y porcentaje con respecto al número total de cerdos analizados en cada granja (**Figura 9B**). El HP1 agrupó a un 29% de los cerdos IB analizados y a un 7% de cerdos CM. Debe mencionarse que todos los cerdos CM portadores del HP1 pertenecieron a la granja AI. Los haplotipos HP4 y el HP2 fueron comunes a varias razas (DU, LW, IB y CM) siendo el HP4 el más

frecuente en la población de cerdos CM (48%). El haplotipo HP5, exclusivo de la raza CM, se encontró en las granjas MS (77%) y R1, R2 y R3 ( $\approx 20\%$ ).

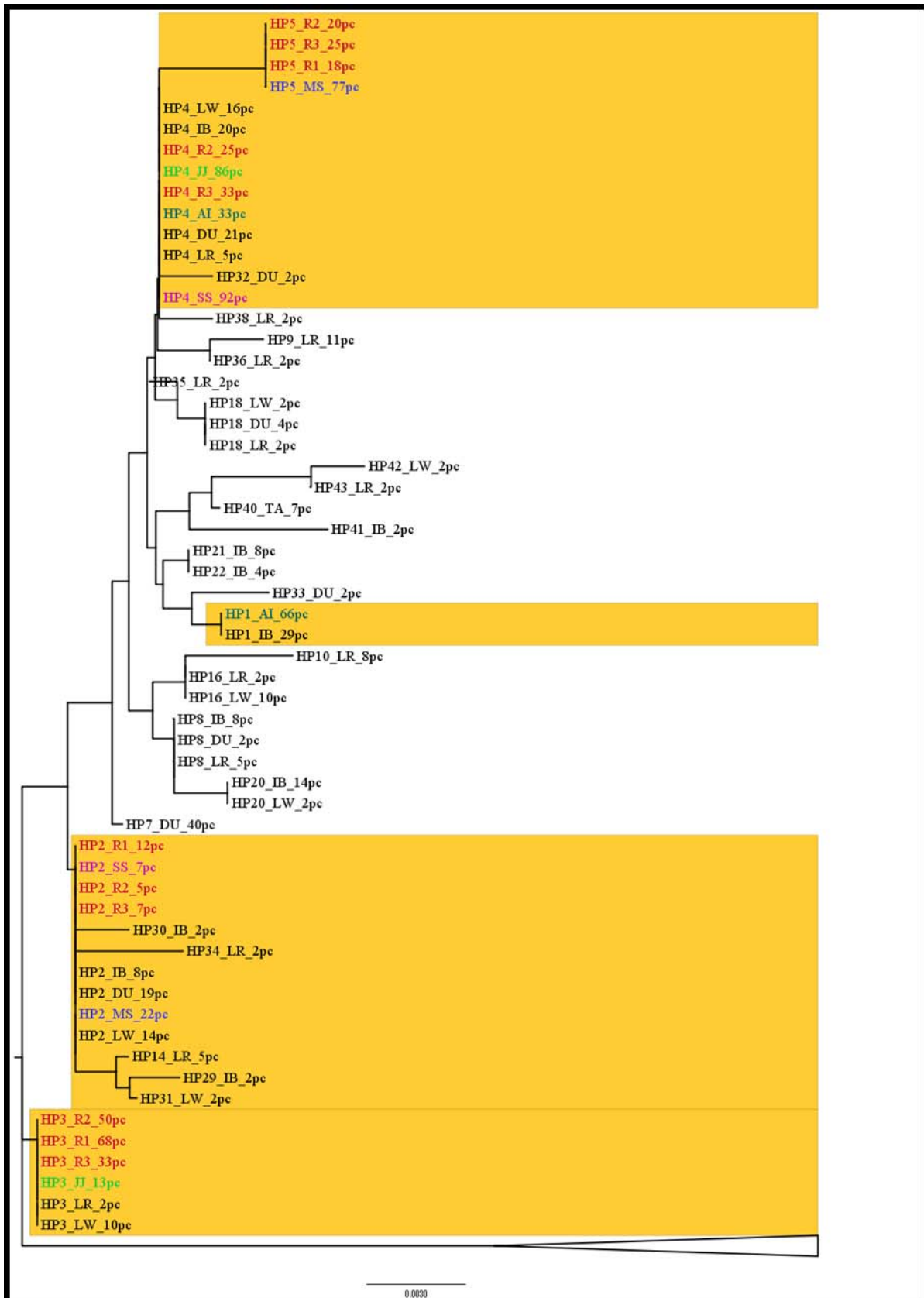
**Tabla 10.** Número total de haplotipos (HP) detectados en cada raza porcina. Se especifica el número de haplotipos asiáticos y europeos en cada raza.

POBLACIÓN	NÚMERO TOTAL HP	HP ASIA	HP EUROPA
CM	5	0	5
LW	14	6	8
LR	17	4	13
DU	8	1	7
TA	4	3	1
BK	2	2	0
IB	10	0	10
MH	4	4	0





**Figura 9A.** Dendrograma usando el método NJ. Se muestra el haplotipo (HP), el código de la raza (Chato Murciano, CM; Large White, LW; Landrace, LR; Duroc, DU; Berkshire, BK; Tamworth, TA; Ibérico, IB; Meishan, MS) y porcentaje de animales analizados con respecto al total de animales analizados en cada raza.



**Figura 9B.** Dendrograma usando el método NJ. Se ha colapsado el nodo asiático y detallado la distribución de los animales CM por granjas.

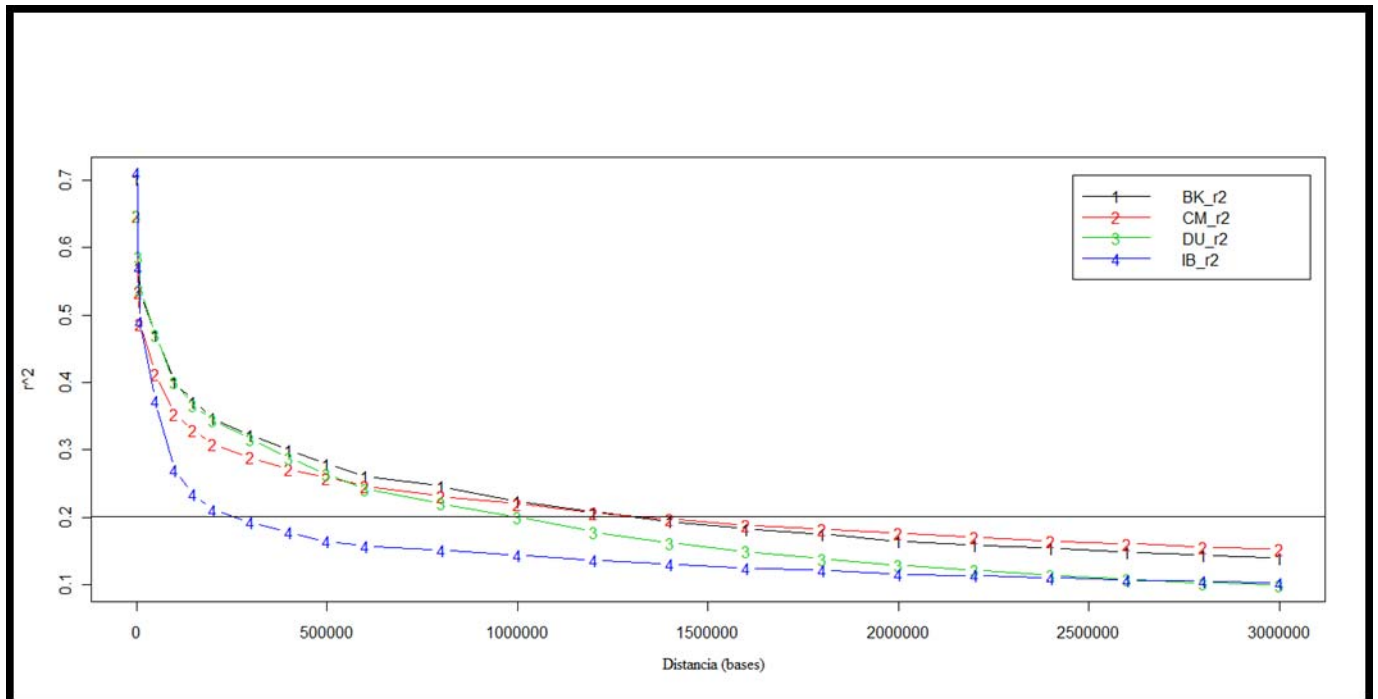
### 3. Desequilibrio de Ligamiento.

El desequilibrio de ligamiento fue estimado para todas las parejas de SNPs separadas por una distancia inferior a 3Mb en cada cromosoma utilizando el coeficiente de correlación entre SNPs ( $r^2$ ). Dichos análisis se realizaron en las dos razas porcinas locales de la península Ibérica, CM y IB, en la raza Inglesa BK y en la raza comercial raza DU. Se calcularon los valores medios de la variación de  $r^2$  conforme aumenta la distancia genética entre marcadores en los 18 cromosomas para cada población y se representó gráficamente (**Figura 10**). El valor  $r^2$  disminuyó conforme se incrementó la distancia genética entre las parejas de marcadores en todas las poblaciones. Sin embargo, la persistencia y la intensidad del LD variaron significativamente entre las diferentes poblaciones y cromosomas. Considerando un valor umbral de  $r^2=0,2$  (Línea negra, **Figura 10**) para determinar el grado y extensión del LD, éste descendió más rápidamente en la raza IB extendiéndose hasta 0,25 Mb. El LD persistió hasta una distancia superior a la observada en el cerdo IB en la raza comercial DU (1Mb) y en las razas minoritarias CM y BK (1,2 Mb). Se observaron patrones similares al calcular la media total de LD en todos los cromosomas en cada población (**Tabla 11**). De este modo, la raza IB presentó el valor promedio de  $r^2$  más bajo ( $r^2=0,14$ ) mientras que en

las razas CM y BK los valores fueron más elevados ( $r^2 \approx 0,21$ ). La raza DU mostró un valor intermedio ( $r^2 = 0,18$ ).

El LD varió significativamente entre los diferentes cromosomas en todas las poblaciones estudiadas (**Figura 11**). El cromosoma SSC14 presentó el valor más alto de LD mientras que los cromosomas SSC18 y SSC10 mostraron los valores menores en todas las poblaciones estudiadas. En la población CM los cromosomas que presentaron mayor LD fueron SSC14, SSC13, SSC1, SSC15, SSC4 y SSC9. Estos resultados fueron similares a otras razas con la excepción del cromosoma SSC4 que tuvo valores de LD bajos en el resto de poblaciones con la excepción de la raza BK. Los cromosomas que presentaron menor valor de LD en la raza CM fueron SSC18, SSC11, SSC10 y SSC12, los cuales también se agruparon entre los cromosomas con bajo LD en el resto de poblaciones. Por tanto, el patrón de LD en los 18 cromosomas observado en el cerdo CM fue similar al observado en las otras razas porcinas.

**Figura 10.** Media de LD en todos los cromosomas medido con el estadístico  $r^2$  en cada población estudiada según aumenta la distancia (en bases) entre marcadores. La línea negra marca el umbral para

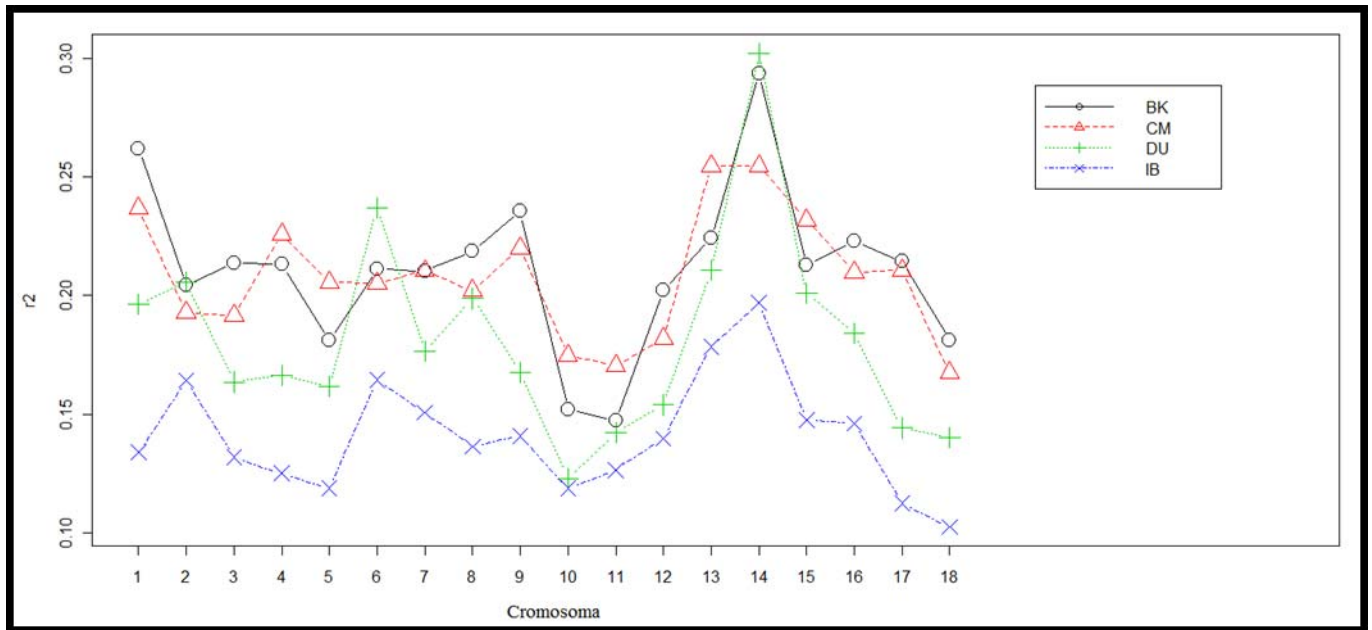


determinar la extensión del LD ( $r^2=0,2$ )

**Tabla 11.** Desequilibrio de Ligamiento por cromosoma y raza. Media de todos los cromosomas en cada raza.

	CM	IB	BK	DU
<i>Cromosoma</i>	<i>r<sup>2</sup></i>			
1	0,237	0,134	0,262	0,196
2	0,193	0,164	0,204	0,206
3	0,192	0,132	0,214	0,164
4	0,226	0,125	0,213	0,166
5	0,206	0,119	0,181	0,162
6	0,205	0,164	0,211	0,237
7	0,210	0,151	0,210	0,176
8	0,202	0,137	0,219	0,199
9	0,220	0,141	0,236	0,168
10	0,175	0,119	0,152	0,123
11	0,171	0,126	0,147	0,142
12	0,182	0,140	0,202	0,154
13	0,254	0,178	0,224	0,211
14	0,255	0,197	0,293	0,302
15	0,232	0,148	0,213	0,201
16	0,210	0,146	0,223	0,184
17	0,211	0,113	0,214	0,144
18	0,167	0,103	0,181	0,140
<b>Media</b>	0,208	0,141	0,211	0,182

**Figura 11.** Representación gráfica de las medias de LD medidas con el estadístico  $r^2$  en cada cromosoma y en todas las poblaciones.



Las líneas que unen los puntos no tienen significado estadístico, pero ayudan a comparar las tendencias de LD en las diferentes razas.

#### 4. Marcadores genéticos distintivos de la raza Chato Murciano

##### **Marcadores SNP distintivos de la raza Chato Murciano.**

Para seleccionar los marcadores SNP que influyen en mayor medida en la diferenciación entre el cerdo CM y las otras razas incluidas en este trabajo se estudiaron las frecuencias alélicas (F) de los 46887 SNPs analizados. Se detectaron tres grupos de SNPs con frecuencias alélicas muy diferentes entre el CM y las otras razas estudiadas (**Tabla 12**). El primer grupo incluyó tres SNPs (SNP1, SNP2,

SNP3) con  $F \geq 0,82$  en el cerdo CM pero inferiores a 0,20 en el resto de razas, con la excepción de las razas MH y TA cuyas  $F=0,40$  para el SNP2. Dentro de este grupo es destacable que los SNP1 Y SNP2 mostraron una  $F=0,00$  en las razas con mayor contribución a la genética del cerdo CM como son las razas IB y DU y una  $F < 0,1$  en la raza LW. En el segundo grupo se incluyeron aquellos SNPs con  $F \geq 0,95$  en el CM pero con  $F < 0,40$  en las razas DU, LW, LR y IB (SNP4, SNP5, SNP6, SNP7). Estas razas adquieren especial importancia al compartir mercado comercial que los productos cárnicos del cerdo CM y habitar en muchas ocasiones las mismas granjas. El tercer grupo incluyó un solo marcador (SNP8) con una frecuencia alélica muy baja en el cerdo CM ( $F=0,08$ ) pero superior a 0,70 en las otras razas con la excepción de la raza IB ( $F \approx 0,40$ ).



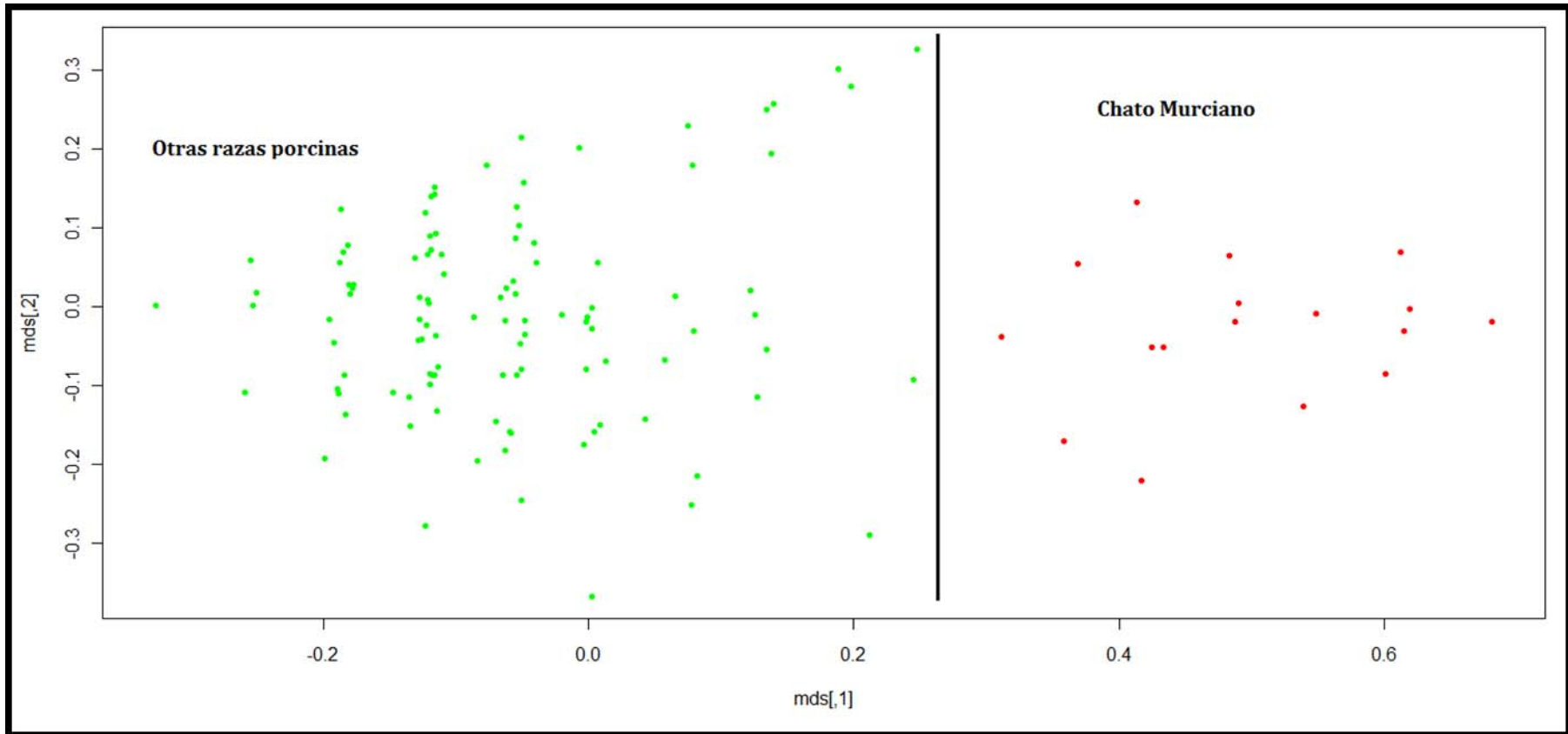
**Tabla 12.** SNPs cuyas frecuencias alélicas fueron significativamente diferentes a las de las otras razas porcinas incluidas en el presente trabajo.

SNP	GRUPO	ALELO	BK	CM	DU	LR	LW	MH	NI	RE	TA
SNP1	I	A	0,03	0,82	0,00	0,16	0,05	0,00	0,00	0,00	0,00
SNP2	I	A	0,13	0,83	0,00	0,16	0,08	0,40	0,07	0,18	0,40
SNP3	I	G	0,03	0,82	0,00	0,16	0,05	0,00	0,00	0,00	0,00
SNP4	II	T	0,50	0,95	0,15	0,40	0,38	0,05	0,33	0,09	0,00
SNP5	II	G	0,68	0,95	0,01	0,14	0,38	0,10	0,20	0,00	1,00
SNP6	II	C	0,16	0,97	0,38	0,17	0,36	1,00	0,08	0,18	0,90
SNP7	II	A	0,37	0,95	0,23	0,29	0,40	0,00	0,27	0,09	1,00
SNP8	III	G	0,95	0,08	0,82	0,97	0,91	1,00	0,40	0,36	0,70

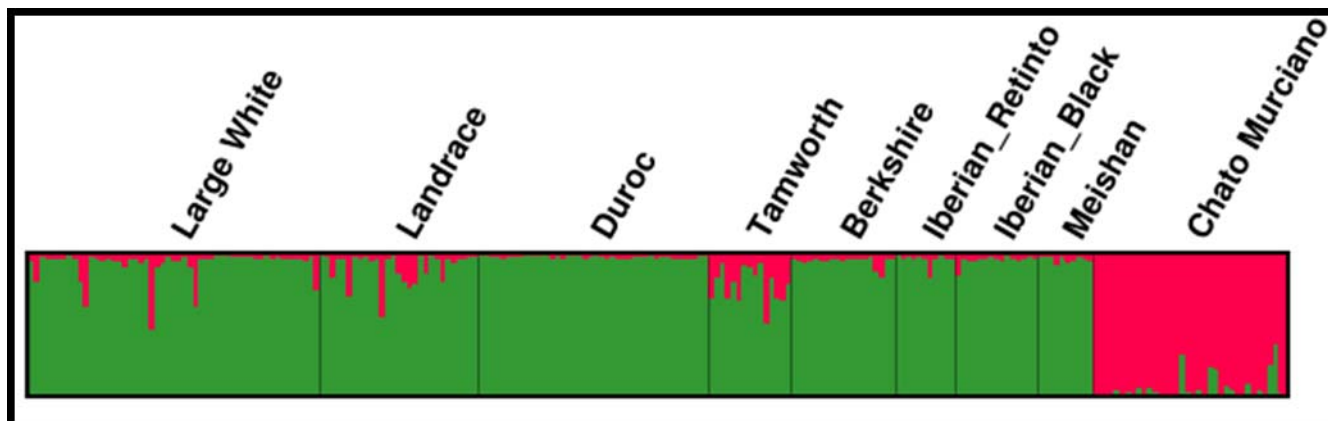
El análisis de escalamiento multidimensional (MDS) utilizando los ocho SNPs detallados en la **Tabla 12** mostró que todos los cerdos CM analizados se diferenciaron de los cerdos no pertenecientes a dicha raza (**Figura 12**). De forma similar, el modelo de agrupación implementado por el programa Structure en base a los ocho SNPs mencionados para un valor de K=2 permitió diferenciar a la población CM de los individuos de otras razas. El CM se agrupó con un coeficiente de pertenencia de 0,95 (**Tabla 13**) en el grupo coloreado de rojo

(Figura 13) mientras que el resto de razas se agruparon en el segundo grupo, representado de color verde, con valores de  $CP > 0,92$  en todas las razas con la excepción de la raza TA cuyo  $CP = 0,79$ . El estudio de los CP por animal demostró que 28 de los 35 cerdos de la raza CM incluidos en el estudio se agruparon con  $CP \geq 0,95$  mientras que cinco lo hicieron con  $CP < 0,82$ , los cuales también habían mostrado valores CP más bajos cuando se analizaron utilizando el panel completo de 46887 SNPs. Solamente un cerdo CM mostró un valor inferior a 0,77 (animal CM14,  $CP = 0,63$ ). Por lo tanto, el estudio de los 8 marcadores SNP distintivos de la raza CM estuvo en consonancia con los resultados obtenidos al estudiar el panel completo de 46887 marcadores SNP. En cuanto a los cerdos pertenecientes a otras razas, de los 194 animales utilizados 185 tuvieron una contribución al grupo del cerdo CM inferior a 0,2 y solamente dos animales, uno de la raza LW y otro de la raza TA se agruparon con valores  $CP \approx 0,5$ .

**Figura 12.** Escalamiento multidimensional en base en base a los 8 SNPs característicos de la raza Chato Murciano



**Figura 13.** Asignación individual mediante el programa Structure para diferenciar los cerdos CM del resto de poblaciones



utilizando ocho SNPs seleccionados en base a sus F.

**Tabla 13.** Coeficientes de pertenencia inferidos por el programa Structure (K=2) a las razas estudiadas en base a los datos de ocho SNPs seleccionados.

		GRUPOS INFERIDOS POR EL PROGRAMA (K = 2)	
		1_CM	2_Otras razas
GRUPOS PREDEFINIDOS	LW	0,054	0,946
	LR	0,077	0,923
	DU	0,011	0,989
	TA	0,205	0,795
	BK	0,035	0,965
	IB_RE	0,026	0,974
	IB_NI	0,028	0,972
	MH	0,025	0,975
	CM	0,947	0,053

### Frecuencias alélicas de los loci microsatélites

De los 34 loci microsatélites estudiados cinco presentaron alelos con frecuencias  $>0,75$  (**Tabla 14**). Además, el estudio por granjas mostró que todas ellas tuvieron frecuencias alélicas elevadas en los cinco alelos referidos en la **Tabla 14** excepto la granja AI para alelo 245 del locus S0355. La frecuencia del alelo 123 del loci SW951 fue superior a 0,8 en todas las granjas menos en la granja SS (0,54).

**Tabla 14.** Frecuencias alélicas de los cinco alelos con  $F > 0,75$  en cada granja y en el total de la población de CM.

	LOCI	S0215	SW 951	S0227	SW769	S0355
<i>Granja</i>	<i>Alelo</i>	<b>155</b>	<b>123</b>	<b>230</b>	<b>105</b>	<b>245</b>
<b>JJ</b>		0,81	0,93	0,92	1,00	1,00
<b>R1</b>		0,57	0,94	1,00	1,00	0,93
<b>R2</b>		0,90	0,95	1,00	1,00	0,98
<b>R3</b>		0,94	0,88	1,00	1,00	0,97
<b>MS</b>		1,00	0,80	1,00	1,00	0,57
<b>SS</b>		0,94	0,54	0,80	0,93	0,98
<b>AI</b>		1,00	1,00	1,00	1,00	0,09
<b>MC</b>		1,00	1,00	1,00	1,00	1,00
<b>Total CM</b>		0,89	0,75	0,93	0,98	0,90

Para estudiar la posible relación entre pureza racial, los loci microsatélite con  $F > 0,75$  y la información obtenida del análisis de los ocho SNPs seleccionados, se comparó a los animales heterocigotos para alguno de esos cinco loci microsatélites con sus valores de CP determinados por el programa Structure en base a los datos PorcineSNP60 y a los ocho SNPs seleccionados previamente (**Tabla 15**). Así, de los 39 cerdos CM genotipados con el panel de alta densidad de SNPs, 34 fueron homocigotos para los cinco loci microsatélites con  $F > 0,75$  y presentaron los valores CP más elevados, tanto al estudiar el panel completo de SNPs como al estudiar los ocho SNPs seleccionados. En cuanto a los cinco cerdos CM restantes (CM112, CM113, CM116, CM118 y CM14), éstos fueron justamente los únicos cerdos CM con valores  $CP < 0,63$  al estudiar los ocho alelos SNPs específicos de la raza CM. Finalmente, remarcar que cuatro de ellos fueron aquéllos que previamente mostraron valores  $CP \approx 0,5$ .

**Tabla 15.** Comparación entre los valores de coeficiente de pertenencia asignados por el programa Structure y los animales heterocigotos para algunos de los cinco microsatélites con frecuencias alélicas altas en la raza.

ID	GRANJA	8 SNPS	PORCINE SNP60	S0215	SW951	S0227	SW769	S0355
CM112	SS	0,097	0,428	Homocigoto	Heterocigoto	Homocigoto	Heterocigoto	Homocigoto
CM113	SS	0,453	0,491	Heterocigoto	Homocigoto	Heterocigoto	Homocigoto	Homocigoto
CH116	SS	0,527	0,501	Homocigoto	Homocigoto	Heterocigoto	Homocigoto	Homocigoto
CM118	SS	0,525	0,504	Homocigoto	Heterocigoto	Homocigoto	Homocigoto	Homocigoto
CH14	JJ	0,625	0,759	Heterocigoto	Homocigoto	Homocigoto	Homocigoto	Homocigoto





# DISCUSIÓN



## **DISCUSIÓN**

En los programas de conservación genética de razas de animales domésticos deben considerarse dos amenazas principales. En primer lugar, la pérdida progresiva de diversidad genética en las poblaciones (Lacy, 1987), especialmente en aquéllas de pequeño tamaño, que amenaza la supervivencia de la raza a largo plazo debido a los efectos negativos de la endogamia. En segundo lugar, el cruzamiento con otras razas que puede amenazar a la identidad genética de una raza ya que la variación genética o fenotípica resultante conllevaría la pérdida del reservorio genético característico de la misma. De este modo, el cruce no controlado con otras razas podría reducir el valor de la raza en términos de herencia cultural e incluso culinaria. El CM se cría actualmente en 15 granjas de las cuales solamente ocho tienen más de 10 reproductores. El escaso número de integrantes de la población, unido a los pocos productores y a la limitada distribución geográfica de los mismos, facilita la realización de un estudio detallado de cómo la pérdida de diversidad genética y el cruce con otras razas porcinas están afectando a la composición genética actual de la raza porcina CM.

### Diversidad y subestructura genética del Chato Murciano

Los resultados de este estudio muestran que la diversidad genética del CM está en el mismo rango que otras razas porcinas criadas en Europa, siendo la  $H_e$  del CM inferior a la media de las razas europeas ( $H_e = 0,57$ , Zhang y Plastow, 2011) y similar o sensiblemente inferior a otras razas locales. Así, Vicente et al. (2008) y Druml et al. (2012) estudiaron razas locales portuguesas y balcánicas respectivamente obteniendo valores similares de  $H_e$  (0,54-0,58) y número de alelos medio por locus (3,8-6,3) a los observados en la raza CM ( $H_e=0,54$ ;  $NMA=6,4$ ), aunque se debe mencionar que el tamaño muestral y el número de microsatélites utilizados fueron inferiores a los del presente estudio. El CM presentó valores de diversidad genética ligeramente inferiores al IB (Alves et al., 2006) y la raza húngara Mangalitsa (Manea et al., 2009). En lo referente a trabajos previos en la raza CM, los valores de diversidad genética obtenidos en este trabajo fueron superiores a los registrados por Calvo et al. (2000) ( $H_e=0,397$ ) y muy similares a los observados por Vega-Pla et al. (2004) ( $H_e=0,53$ ). Estos datos indican que los programas de recuperación de la raza CM tuvieron buenos resultados hasta el año 2004, produciéndose un aumento del tamaño de la población y de la diversidad genética de la

misma. Sin embargo, la similitud entre los resultados del presente trabajo y los obtenidos por Vega-Pla et al. (2004) probablemente no indiquen un mantenimiento de la diversidad genética entre el año 2004 y 2012 ya que estos últimos autores utilizaron un número muy inferior de animales (83 frente a 175). Además, estos autores no especificaron las granjas incluidas en su estudio dificultando la comparación de los resultados como consecuencia a las variaciones existentes entre las granjas.

Los valores de diversidad genética obtenidos en la población global de CM deben valorarse con precaución debido a que, cuando una población está subdividida, la diversidad genética de la misma está sobreestimada y los análisis deben hacerse en cada subpoblación por separado (Ji et al., 2011). La población de CM está subdividida en subpoblaciones muy diferenciadas genéticamente siendo muy probable que la diversidad genética observada en el total de la población de CM esté sobreestimada. Por esta razón se estudió la variabilidad genética de cada granja por separado.

A diferencia de los resultados descritos para el conjunto de la población, el análisis de las granjas que crían cerdos CM por separado permitió detectar valores de diversidad genética inferiores a la media

de la población en todas las granjas con la excepción de la granja SS (0,56). No obstante, la granja SS presentó indicios de mezcla con otras razas porcinas que se discutirán más adelante. Se observaron valores de diversidad genética muy bajos en granjas de pequeño tamaño como la granja AI y la MC ( $He \leq 0,32$ ) mostrando las características propias de granjas pequeñas y aisladas con elevada consanguinidad y mayor efecto de la deriva genética. Así por ejemplo, la tendencia a la pérdida del alelo 245 del locus S0355, el cual es muy frecuente en la raza CM, sugiere un aislamiento reproductivo de la granja AI. Por otro lado, ambas granjas fueron las que presentaron el mayor número de loci microsatélites fijados, indicando un efecto más intenso de la deriva genética. La deriva genética es un proceso natural por el cual poblaciones aisladas modifican sus frecuencias alélicas de generación en generación, dando lugar tanto a la disminución de la variabilidad genética por fijación o pérdida de variantes alélicas (Nei 2000), como a la progresiva diferenciación genética de las poblaciones. Se ha observado que el efecto de la deriva genética es más intenso en poblaciones aisladas y con un número de reproductores pequeño (Holsinger y Weir, 2009).

La raza CM mostró un alto grado de subestructuración correlacionada con las granjas en las que se crían cerdos CM, existiendo al menos dos subpoblaciones genéticas. La primera subpoblación, representada en color verde la **Figura 3**, incluyó a las granjas R1, R2, R3, MS y AI. Las granjas R1-3 mostraron una mayor semejanza genética entre ellas que sugiere un flujo de reproductores entre las mismas y refleja el hecho de que las granjas R1-3 son propiedad del mismo productor. La granja MS mostró una influencia elevada en su composición genética de la segunda subpoblación, que unido a la alta diversidad genética detectada en la granja MS a pesar del bajo número de reproductores ( $N=11$ ) y a que la  $H_o$  fue mucho mayor que la  $H_e$ , indica que la granja MS se formó a partir de la mezcla de reproductores de las dos subpoblaciones inferidas. De este modo, la presencia de un haplotipo mitocondrial exclusivo de las granjas R1-3 y MS sugiere un origen común de las hembras CM de esas granjas o que ha habido flujo de animales entre las mismas. La segunda subpoblación, que se marcó en rojo, estuvo representada por las granjas SS y MC mientras que la granja JJ pareció formar por si sola una tercera subpoblación cuando se consideraron los datos microsatélites, pero no fue así cuando se utilizaron los datos SNPs, donde los animales se asignaron a alguna de las dos subpoblaciones principales. Teniendo

en cuenta: la posición ocupada por la granja JJ en el árbol de distancias genéticas (**Figura 2**), que los animales de dicha granja se asignaron a las dos subpoblaciones inferidas por el programa Structure en base a los datos microsatélites cuando  $K < 3$  y para cualquier valor de  $K$  en base a los datos SNPs y que los dos haplotipos de ADN mitocondrial observados en la granja JJ fueron altamente frecuentes en las granjas SS y R1-3, la granja JJ representaría realmente una mezcla de animales de las dos subpoblaciones principales en lugar de formar realmente una tercera subpoblación.

La existencia de subpoblaciones en la raza CM puede ser consecuencia de diversos factores que actúan conjuntamente. El aislamiento reproductivo y el tamaño reducido de algunas granjas acentuaría el efecto de la deriva genética tal y como demuestra el hecho de que las tres granjas de menor tamaño (granjas MS, AI y MC) fueron las más diferenciadas genéticamente. Otro factor que parece haber aumentado la diferenciación es la implementación por parte de los granjeros de sus propias estrategias de manejo reproductivo. Los resultados obtenidos sugieren que tales estrategias no sólo se limitaron al intercambio de productores entre las granjas R1-3 ya que también se detectaron signos de introgresión genética con otras razas porcinas, principalmente en las granjas AI y SS. El 60% de los cerdos



CM criados en la granja AI presentaron un haplotipo exclusivo de la raza IB (HP1, **Figura 9B**). Estos datos son consistentes con el hecho de que en dicha granja se crían conjuntamente ambas razas porcinas. Es posible que el ganadero responsable decidiera modificar los parámetros de calidad de producto de su granja mediante cruzamientos con IB. En cuanto a la granja SS, se observaron varios indicadores de introgresión genética del CM con otras razas. En primer lugar, dicha granja mostró las frecuencias más bajas de los alelos de dos loci microsatélites descritos como característicos de cerdos CM puros (Vega-Pla et al., 2004; Calvo et al., 2000). En segundo lugar, animales de la granja SS se separaron claramente del resto de cerdos CM en el análisis MDS basado en los datos SNPs (ver **Figura 5B**). Finalmente, cuatro cerdos de la granja SS presentaron valores  $CP \approx 0,5$ , mientras que el resto de cerdos CM tuvieron valores superiores a 0,7, sugiriendo por tanto un elevado grado de mezcla genética. La existencia de cerdos CM mezclados es un hecho reseñable debido a que los animales con elevado grado de mezcla se excluirían de futuros programas de conservación a fin de preservar la identidad genética de la raza. No obstante, esta práctica es comprensible en explotaciones comerciales que intentan adaptar y diferenciar su producción.

El estudio de la subestructura genética de la raza CM combinado con los datos de diversidad genética de cada granja por separado proporciona una información muy valiosa para el diseño de un futuro programa de conservación de la raza CM que permitiera preservar su identidad y a la vez evitar la pérdida de diversidad genética de la misma. Teniendo en cuenta el bajo número de reproductores CM, las granjas pequeñas y genéticamente aisladas constituyen una parte importante de la diversidad total de la población. Se ha sugerido que la creación de subgrupos de cría para un posterior intercambio de reproductores entre los mismos es una medida efectiva para el mantenimiento de la diversidad genética en una población aislada (Lacy, 1987). De esta forma, las granjas pequeñas y muy diferenciadas genéticamente se pueden considerar como subgrupos de la población total de cerdo CM y podrían utilizarse para aumentar la diversidad genética de la raza en su conjunto mediante intercambio de reproductores, al mismo tiempo que disminuirían los efectos negativos derivados de la consanguinidad. Considerando la baja productividad de las razas locales en general y del cerdo CM en particular, es de vital importancia evitar que la consanguinidad disminuya todavía más los parámetros productivos. La creación de un centro de verracos común que dispensara semen a todas las granjas de cría de CM, observando

esta circunstancia, sería una medida recomendable. Otra opción no desdeñable sería la creación de un banco de embriones congelados de cara a mantener una reserva de individuo de genotipo conocido.

#### Relación genética del Chato Murciano con otras razas.

Todos los animales analizados fueron agrupados inequívocamente a sus razas de origen con los modelos de estratificación utilizados en base al panel Porcine SNP60. Ramos et al. (2011) también observaron que el genotipado de alta densidad de SNPs es una herramienta molecular adecuada para diferenciar poblaciones y asignar a los individuos a su raza de origen. La raza MH fue la población más divergente mientras que las variedades NI y RE de Ibérico por un lado y las razas LW y LR por otro fueron las menos diferenciadas como se observó en estudios previos (Martínez et al., 2000; Alves et al., 2006; Megens et al., 2008). Todos los animales pertenecientes a la raza CM se agruparon inequívocamente, sin que ningún individuo analizado se asignara a otras poblaciones, lo cual indica que el CM es genéticamente diferenciable, tanto de las razas próximas geográficamente como de las utilizadas históricamente en cruces con la raza objeto de estudio. El análisis de la estructura de las

poblaciones realizado con el programa Structure mostró un valor de coeficiente de pertenencia en la raza CM inferior al de otras razas como BK, TA, DU y IB (0,85 Vs.>0,90). El menor coeficiente de pertenencia en la raza CM se debió en gran medida a la presencia de un reducido número de cerdos CM mezclados que disminuyeron la media total de la población, tal y como se demostró al eliminar del análisis a los cuatro cerdos CM con  $CP \approx 0,5$  obteniéndose entonces un valor de coeficiente de pertenencia para el cerdo CM más elevado (0,899).

Por otro lado, estos resultados evidenciaron una contribución genética de otras poblaciones porcinas en la raza CM. Considerando el análisis realizado tras eliminar a los cerdos CM mezclados por considerarse más representativo de la raza CM, el valor  $CP=0,899$  implicaría una contribución de al menos un 10% de otras razas porcinas. Los resultados indican que las razas que han contribuido en la composición genética de la raza CM fueron DU (4%), IB (3,6%), LW (1,5%) BK (0,7%) y TA (0,02%). Las razas LW, BK y IB, además de contribuir significativamente a la genética del CM actual, también presentaron menores distancias genéticas con el cerdo CM en comparación con la raza TA, lo que concuerda con la documentación histórica que indica que esta raza inglesa fue menos utilizada que las razas LW, BK y RE en los cruces realizados con el cerdo Murciano

primitivo a finales del siglo XIX y principios del siglo XX y que dieron lugar al actual CM (Lobera, 1998). La contribución del IB fue mucho más alta que la de otras razas relacionadas históricamente con el cerdo CM. Además la raza DU tuvo una gran influencia en la composición genética del actual CM, a pesar de que no se encontró en la bibliografía consultada ningún indicio de la utilización de cerdos DU en la formación de la raza CM. Estos datos sugieren que tal introgresión genética tiene un origen reciente y no histórico, ya que se localizó principalmente en animales pertenecientes a la granja SS y en menor medida a la granja JJ (**Anexo III**), en las cuales se crían conjuntamente animales CM, IB, DU además de cruces IB x DU. Es probable que estos ganaderos hayan buscado mejorar los parámetros productivos de sus animales mediante cruces con cerdos DU o con cruces IB x DU siguiendo el ejemplo del IB, en el cual los cruces con DU están aceptados y legislados.

El estudio de la secuencia de ADNmt ha determinado una elevada distancia filogenética entre razas porcinas asiáticas y europeas, habiéndose incluso determinado que dicha divergencia ocurrió hace aproximadamente 760.000 años (Fernández et al., 2011). Esta divergencia en la secuencia de ADNmt permitió estudiar los fenómenos de introgresión genética de cerdos asiáticos en razas europeas como

consecuencia del uso de cerdos procedentes de Asia para mejorar las razas inglesas en los siglos XVIII y principios del XIX (Jones, 1998). Los haplotipos mitocondriales asiáticos son especialmente frecuentes en razas comerciales como por ejemplo LW o LR y en razas inglesas (Giuffra et al., 2000; Kim et al., 2002; Wu et al. 2007) tal y como se observó en el presente trabajo. A pesar de la alta frecuencia de haplotipos asiáticos en razas que estuvieron involucradas en la creación del actual CM, no se observó ningún haplotipo asiático en esta raza española. Cabría esperar sin embargo que durante los cruces realizados entre razas inglesas y el cerdo CM a principios de siglo XX, reproductores portadores de haplotipos asiáticos hubieran introducido dichas secuencias de ADNmt en la raza CM, tal y como ha ocurrido en otras razas locales como el Manchado de Jabugo y el cerdo Negro Canario en las que se han observado haplotipos asiáticos como resultado de cruces con razas inglesas (Alves et al., 2003; Clop et al., 2004). La ausencia de haplotipos asiáticos en la raza CM concuerda con los documentos históricos que establecen que solamente se importaron machos de razas foráneas a la Región de Murcia para ser cruzados con hembras CM (Lobera 1998). Por lo tanto, es posible que nunca haya existido introgresión de ADNmt en la raza o que ésta fuera débil y desapareciera durante el cuello de botella sufrido por la raza

CM en la segunda parte del siglo XX. Tampoco se detectó introgresión asiática en los cerdos de la raza IB como previamente observaron Alves et al. (2003).

El análisis del fragmento D-loop de la secuencia de ADNmt realizado en 175 animales de la raza CM mostró un total de cinco haplotipos mitocondriales ocupando diferentes regiones del dendograma (**Figura 5**). Esta diferenciación en la secuencia de los haplotipos del cerdo CM contrasta con los resultados obtenidos por Wang et al. (2010) al estudiar la raza minoritaria china Pudong White, que observaron que todos los animales analizados ocuparon el mismo área del dendograma, además de ser sus haplotipos muy diferentes a los de otras razas chinas de la misma región geográfica. A raíz de esos resultados, Wang et al. (2010) concluyeron que no se habían producidos cruzamientos recientes entre el Pudong White y otras razas. En cambio, el 55% de cerdos CM presentaron haplotipos mitocondriales iguales a los de otras razas europeas. Este hecho puede deberse tanto a cruces pasados como recientes entre el cerdo CM y otras razas porcinas, aunque también es posible que todos o algunos de esos haplotipos presentaran una amplia distribución geográfica en la época en que apareció la raza CM. El estudio por granjas reveló que no existieron haplotipos propios de una sola granja, con la excepción

del HP1 que solamente se detectó en animales de la granja AI y que fue compartido con la raza IB como se discutió previamente. Por otro lado, es especialmente interesante la detección de un haplotipo (HP5) que se observó en el 14% de cerdos CM pero no en cerdos de otras razas, siendo por tanto un haplotipo exclusivo de la raza CM. Ese mismo haplotipo tampoco se detectó en más de 3000 secuencias de cerdos pertenecientes a más de 200 poblaciones porcinas de todo el mundo (Herrero-Medrano et al., artículo en preparación). Es posible que este haplotipo exclusivo de la raza CM fuera heredado directamente del antiguo cerdo Murciano primitivo, aunque no se puede descartar una mutación reciente en la secuencia D-loop de ADNmt que se hubiera expandido de forma importante por la población de CM tras el reciente cuello de botella sufrido por la raza. Desafortunadamente, no existen restos biológicos del antiguo cerdo Murciano primitivo para realizar una comparación directa.



### Desequilibrio de ligamiento en el cerdo Chato Murciano.

El panel PorcineSNP60 contiene un elevado número de marcadores repartidos por todo el genoma porcino y separados por una distancia genética estimada en 30-40 Kb de media (Ramos et al., 2009), lo que permite utilizar este panel para realizar estudios de LD en razas porcinas (Uimari y Tapio, 2010; Badke et al., 2012). En este trabajo se utilizó dicho panel de genotipado de alta densidad para realizar por primera vez un estudio del LD en la raza CM. Además, aun no siendo el objetivo de este trabajo estudiar el LD en otras razas, se consideró comparar los resultados en la raza CM con otras razas minoritarias (IB y BK) y con una raza comercial (DU). El estudio de las diferencias genéticas entre poblaciones es importante debido a que la fuerte estructuración de las mismas es una de las razones más importantes en la aparición de falsas asociaciones entre fenotipos y marcadores en los estudios de asociación y mapeo de QTLs (Zhao et al., 2007). Por ello, se considera que un aspecto clave en la utilidad de la selección asistida por marcadores es la identificación de efectos de LD consistentes entre poblaciones diferentes (Short et al., 1997) al facilitar la aplicación de dichos estudios de asociación y aumentar la confianza de los mismos (Zhang y Plastow, 2011).

En estudios previos en los que se ha estudiado el LD en razas porcinas finlandesas (Uimari y Tapio, 2010) y americanas (Badke et al., 2012) utilizando el panel *PorcineSNP60*, el valor de  $r^2$  para SNPs adyacentes (< 100Kb) mostró valores entre 0,36 y 0,48 según la población. Todas las razas estudiadas se incluyeron en estos rangos con la excepción de la raza IB que presentó un valor muy inferior (0,27). Diversos autores han observado que eventos que afectan a la historia de las poblaciones tales como la deriva genética, estructura poblacional, la mezcla entre poblaciones, la selección natural y artificial y los cuellos de botella influyen en el grado de LD (Ardlie et al, 2002; Du et al., 2007). Los eventos históricos acontecidos en la raza IB explicarían los valores más bajos de LD obtenidos en la misma, así la selección y cruzamiento son factores que tienden a elevar el LD (Du et al., 2007; Ardlie et al., 2002; Badke et al., 2012). Hasta tiempos recientes, la raza de cerdo IB no estuvo sometida a una selección intensa. Además los animales utilizados en este estudio se eligieron como representantes de IB con un elevado estándar de pureza, al menos fenotípicamente. En cambio, el cerdo CM sí ha estado sujeto al cruce reciente con otras razas, presentando además una diversidad genética relativamente baja y un tamaño de la población menor, lo cual explicaría que el grado y la extensión de LD fuera mayor que en el

cerdo IB. De las razas incluidas en el presente estudio, el LD solamente se había estudiado previamente con el mismo panel de SNPs en la raza DU. Badke et al. (2012) obtuvo valores de LD en esta raza prácticamente idénticos a los observados en este estudio.

Las razas minoritarias están siendo objeto de un interés creciente, no solo por su importancia para el mantenimiento de la diversidad genética, sino también por el hecho de que podrían jugar un papel importante en la ganadería del futuro como se infiere de la reciente creación del proyecto “Low Input Breeds” ([www.lowinputbreeds.org](http://www.lowinputbreeds.org)). La progresiva reducción del precio de genotipado de alta densidad podría hacer que la selección asistida por marcadores y/o la selección genómica aplicada en razas locales fuera una realidad en un futuro próximo.

El grado de LD determina la densidad de marcadores requerida por los estudios de asociación del genoma completo. Nuestros resultados indicarían que la raza IB necesitaría un mapeo de marcadores fino debido a la menor extensión de LD (0,25 Mb), mientras que el número de marcadores necesarios en estudios de asociación del genoma completo sería menor en la raza CM (extensión de LD de 1,2 Mb). La evaluación conjunta del grado de LD y de la

estructura de las poblaciones proporciona información muy importante para el diseño de estudios de asociación del genoma completo (Chao et al., 2010). Si bien los resultados expuestos suponen un punto de partida, son necesarios más estudios para la implementación de estudios de asociación del genoma completo en la raza.

#### Identificación genética de la raza Chato Murciano. Trazabilidad y pureza.

El rápido desarrollo de nuevas tecnologías de secuenciación acontecido en los últimos años ha puesto a disposición de los investigadores herramientas cada vez más potentes y económicas para la realización de estudios en especies animales de importancia económica como es el caso de la especie porcina (Ramos et al., 2009) o bovina (Matukumalli et al. 2009). Entre sus múltiples aplicaciones, las técnicas de genotipado de alta densidad generan un número suficiente de marcadores SNPs para diferenciar razas (Ramos et al., 2011; Pant et al., 2012) y por tanto se erigen como una herramienta de gran fiabilidad aplicables en la trazabilidad de productos diferenciados de origen animal, a veces relacionados con un valor diferente en el

mercado. Los productos cárnicos derivados de la raza CM se comercializan con un precio de mercado más elevado que los productos derivados de cruces comerciales de razas precoces. Las razones principales son el crecimiento lento de esta raza, que alarga el periodo de engorde (Galián et al., 2009) y el elevado valor tecnológico de los productos cárnicos, superiores a los productos de otras razas porcinas precoces (Peinado et al., 2004). Los resultados obtenidos al seleccionar ocho marcadores SNPs extraídos del panel PorcineSNP60 en base sus las frecuencias alélicas fueron equivalentes a los resultados observados al analizar 46887 (**Tabla 15**) en términos de diferenciación genética entre los cerdos CM y el resto de animales de otras razas. Ramos et al. (2011) seleccionaron un panel de SNP específicos de raza utilizando el panel de SNPs PorcineSNP60 con un número muy superior de SNPs que en el presente trabajo (193 SNPs). Hay que considerar que estos autores diseñaron un panel para identificar no una sola raza porcina, sino cinco razas comerciales, lo cual justifica la necesidad de utilizar un número más elevado de marcadores SNP.

Además de las diferencias en el número de marcadores observadas al comparar este trabajo con otros realizados previamente, también se siguieron premisas diferentes para la selección de los SNPs

característicos de las poblaciones. Así, tanto Ramos et al. (2011) como Pant et al. (2012) consideraron como SNPs específicos de raza aquellos alelos cuyas frecuencias fueran elevadas (cercanas a uno) en una raza y a la vez cercanas a cero en las otras razas comparadas, pero en el presente trabajo no se encontró ningún marcador que cumpliera ambas condiciones. El panel de ocho SNPs distintivos de la raza CM, a pesar de que también se basó en las frecuencias alélicas, utilizó criterios menos estrictos que en los trabajos mencionados. Hay que considerar que Ramos et al. (2011) y Pant et al. (2012) utilizaron animales puros mientras que en este estudio la única información previa de pureza de los animales era la resultante de los análisis con el panel PorcineSNP60 completo. Cuando se testó la eficacia del panel de ocho marcadores SNP para diferenciar cerdos CM mediante dos métodos de agrupamiento diferentes, siendo uno de ellos igual al utilizado por Ramos et al. (2011) (modelo bayesiano de agrupamiento implementado por Structure), se observó que la eficacia de asignación individual de cerdos CM fue similar a la obtenida en ambos trabajos previos, al diferenciarse inequívocamente a todos los cerdos CM de los cerdos de otras razas.

Por otro lado, el estudio de los ocho marcadores SNP seleccionados también mostró una sensibilidad igual o mayor que el panel completo *PorcineSNP60* para diferenciar los cerdos CM con alto grado de mezcla de los cerdos CM con signos de pureza racial. El mantenimiento de subpoblaciones de reproductores de raza pura es fundamental en los programas de conservación genética de las razas domésticas (Bradley et al., 1994). La posibilidad de identificar a los animales puros de una raza por medio de marcadores genéticos resulta de especial interés en razas en las que se carece de libro genealógico como es el caso del cerdo CM. Existen pocos trabajos previos en los que se hayan identificado marcadores genéticos que pudieran ser utilizados para identificar a aquellos reproductores CM puros. Vega-Pla et al. (2004) observaron que el alelo 123 del locus SW951 estaba presente en los cerdos CM definidos como puros por ese estudio. De forma similar Calvo et al. (2000) solamente observó fijación alélica en el loci S0227 en la población de CM considerados como puros por dicho autor. Los resultados confirmaron que dichos loci tienden a la fijación alélica en la raza CM, existiendo además otros loci con frecuencias alélicas próximas a uno (S0215, SW769 y S0355). En este trabajo, los animales con elevados signos de mezcla al estudiar los marcadores SNP fueron siempre heterocigotos para uno o varios de los

loci SW951, S0227 y S0215, mientras que los animales sin signos de cruzamiento con otras razas en base a los marcadores SNP, fueron homocigotos para los cinco loci microsatélites con frecuencias alélicas superiores a 0,75 en raza CM. Estos datos sugieren que la homocigosis de estos loci microsatélite constituye un parámetro a tener en cuenta en la selección de reproductores puros. No obstante, la fiabilidad del uso de los datos SNP puede ser considerada superior a la de los datos microsatélite. En efecto, aunque los individuos heterocigotos para los loci microsatélites SW951, S0227 y S0215 serían candidatos a ser excluidos de un programa de selección de reproductores CM puros, animales homocigotos para dichos marcadores podrían no ser animales puros puesto que se ha observado que aquellos alelos descritos como altamente frecuentes en el cerdo CM no son exclusivos de esta raza y por tanto, animales no puros podrían ser también homocigotos para estos loci. Además, trabajos previos consideran que el uso marcadores SNP arroja mejores resultados que los marcadores microsatélite siempre y cuando el número de SNPs sea muy elevado (Rohrer et al., 2007; Leroy et al., 2012). Concretamente se ha sugerido que por cada microsatélite son necesarios tres SNPs para obtener una cantidad de información similar (Schopen et al., 2008). El hecho de que en el presente trabajo se utilizaran 1383 SNPs por cada marcador



microsatélite parece indicar claramente la mayor potencia del panel PorcineSNP60 en comparación con el panel de 34 microsatélites.

Las implicaciones prácticas de estos resultados son importantes. El uso de sólo ocho SNPs distintivos de la raza CM abre la posibilidad de una herramienta genética para la implementación de sistemas rutinarios de trazabilidad y detección de animales con alto grado de mezcla que sería económica y rápida de analizar. Por ejemplo, el análisis de estos SNPs permitiría diferenciar los productos cárnicos derivados de la raza CM de cerdos o productos cárnicos que procedieran de otras razas porcinas con menor valor en el mercado.

La certificación de origen de productos con alto valor comercial procedentes de una raza autóctona de la Región de Murcia motivaría a ganaderos, comercios y al gobierno de la Región de Murcia a la conservación de la raza CM e incluso podría facilitar el desarrollo de una denominación de origen de productos de la raza CM.



# CONCLUSIONES



## **CONCLUSIONS**

1. Despite historical records and genetic observations that support past and present introgression, the Chato Murciano breed remains genetically distinct from breeds that likely contributed genetically to the current CM gene pool.
2. The Chato Murciano breed shows a high degree of substructure that is correlated with the farms they are reared on. Consequently, although the study found levels of genetic diversity in the breed to be within the range of other European local breeds, a farm-by-farm analysis revealed marked differences between farms in the genetic diversity, with particularly farms displaying the characteristics of isolated small farms with high inbreeding and genetic drift.

3. Evidence of genetic contributions from Duroc, Iberian, Large White and Berkshire into CM breed was found. Specifically, the breeds Duroc and Iberian appear to have had large, recent and localized influence on the CM gene pool as a result of independent breeding strategies developed by the farmers.
  
4. Asian haplotypes were not found in the Chato Murciano breed which is however consistent with historical records that report that only males were imported to Murcia Region. It is noticeable the existence of one haplotype exclusive of the Chato Murciano breed.
  
5. According to the results of LD, the Chato Murciano would need a lower number of markers for genome-wide association studies than Iberian pig but a similar number to Duroc.

6. Eight SNP markers selected from the PorcineSNP60 BeadChip resulted to distinguish the CM pigs from other breeds, facilitating design of traceability assays as well as purebred tests at low cost.





# RESUMEN



## **RESUMEN**

El Chato Murciano es una raza local porcina autóctona del sureste de España y en peligro de extinción. Esta raza, al igual que otras muchas razas locales, está amenazada por los efectos de la endogamia y el cruce con otras razas porcinas. En el presente trabajo se utilizaron conjuntamente tres marcadores genéticos, concretamente un panel de 34 microsatélites, un panel de alta densidad de SNPs y un fragmento de ADN mitocondrial que se analizaron en la práctica totalidad de la población de Chato Murciano. Además, se analizó el fragmento de ADN mitocondrial y el panel de SNPs en cerdos de aquellas razas que con mayor probabilidad han contribuido a la composición genética actual de la raza Chato Murciano.

La diversidad genética de la raza fue similar a la de otras razas locales europeas, sin embargo, cuando las granjas que crían al menos 10 reproductores de la raza fueron analizadas por separado, se observó que la diversidad genética en algunas de ellas era preocupantemente baja. Los análisis de los marcadores moleculares ADN mitocondrial y el panel de alta densidad de SNPs indicaron la existencia de contribución genética de las razas de cerdo Ibérico, Large White y Berkshire en la raza Chato Murciano, además de signos de

reciente introgresión genética de las razas Duroc y cerdo Ibérico en tres de las granjas que crían cerdos Chato Murciano, lo que sugiere que algunos productores han implementando sus propias estrategias de manejo reproductivo. Por otro lado, se demostró que la raza Chato Murciano es diferenciable genéticamente, tanto de las razas que influyeron en su creación como de aquellas razas que se crían en mismas granjas o mismo entorno. Dicha diferenciación genética fue evidente incluso utilizando un número muy reducido de marcadores SNPs seleccionados en base a sus frecuencias alélicas.

La ausencia de un libro genealógico de la raza Chato Murciano supone un importante inconveniente para el desarrollo de un programa de conservación eficaz. Los resultados de diversidad genética, de subestructura de la población y de los cruces con otras razas obtenidos permitieron describir la situación actual de la raza Chato Murciano, no solo a nivel global sino también en cada granja. Estos datos serán muy valiosos para el diseño de estrategias de conservación de la raza en el futuro. Además, la posibilidad de diferenciar genéticamente al cerdo Chato Murciano de otras razas mediante un reducido número de marcadores genéticos que pudieran ser genotipados a bajo coste supondría una herramienta muy interesante para trazar los productos cárnicos procedentes del cerdo

Chato Murciano, y también para detectar a aquellos animales con signos de mezcla con otras razas que serían candidatos a ser excluidos de los programas de conservación de la raza.



# SUMMARY





## **Introduction**

Local breeds often are regarded as part of the cultural heritage of local and national communities. One of those breeds categorized as endangered on the List for Domestic Animal Diversity Information System (DAD-IS, <http://dad.fao.org/>) is the Chato Murciano (CM) pig breed.

The CM breed is autochthonous to the Region of Murcia (Spain). It has similar characteristics to other Spanish local breeds such as rusticity, high adaptation to the environment and low productivity compared with commercial porcine breeds. The ancestor of the breed is a rustic black Mediterranean pig, also referred to as the primitive Murcia pig that lived in the South-East of Spain at least 150 years ago (Lobera, 1998). farmers designed crosses with breeds such as the Berkshire, Large White, Retinto Iberian pig and Tamworth in the late 19<sup>th</sup> century and early 20<sup>th</sup> century (Poto et al. 2000) in order to improve production parameters. As a result, a phenotypically differentiated breed was described as “CM”. The depreciation of animal fat by the consumer in the second part of the 20<sup>th</sup> century lead the CM breed to the edge of extinction with less than 40 breeders in 1997. However, since the inclusion of the CM pig in the group of Spanish

breeds under special protection and consequently, the population number rose to 287 individuals in 2009. This effort made by the authorities and producers reflects the current concerns of society to an imminent loss of this pig that remains to have an important socio-economic influence in the Region of Murcia. The favourable appreciation of CM meat products by local customers (Peinado et al. 2004) is another important factor for the future of the breed. Despite these factors, several concerns remain regarding the long-term survival of the breed. Nowadays, breeding stocks tend to be isolated by farm, and the farmers each tend to apply their own breeding strategies, which in some cases may include crossing with other breeds. For the smallest farms in particular, there is a concern of high inbreeding levels and related risks of inbreeding depression. Further loss of breeding stock may result from farmers being unable to cope with the economic consequences of the low productivity of the CM breed. In order to design an effective conservation program, a detailed genetic study of the CM breed diversity is a necessary first step.

To assess the genetic threats to the integrity and survival of the CM breed, and to aid in designing a conservation program, three genetic marker systems – microsatellites, SNPs and mtDNA– were applied across the majority of the total breeding stock. In addition,

mtDNA and SNPs were genotyped in breeds that likely contributed genetically to the current CM gene pool.

## **Materials and Methods**

### 1. Animals and sampling

Genomic DNA was extracted from blood and hair using the Genra PureGene Blood kit (Qiagen) and the Danapure Spin kit (Genedan SL, Spain) respectively according to the manufacturer's protocol. Although CM pigs are currently being bred and reared on approximately 15 farms, the study included 175 CM pigs collected from the eight registered farms that bred at least 10 breeders, representing 70% of the current breeding population. In addition, 194 domestic pigs from the breeds Large White (LW, n=53), Landrace (LR, n=29), Duroc (DU, n=42), Tamworth (TA, n=15), Berkshire (BK, n=19), Iberian pig (IP, n=26) and Meishan (MH, n=10) were included. The samples used for each molecular marker are detailed in the **Table 1**.

**Table 1.** Sampling information and analysis performed in each pig population.

Breed	Country	N	mtDNA	Porcine SNP60	Microsatellites
<b>Chato Murciano</b>					
<i>farm JJ</i>	Spain	36	32	8	36
<i>farm R1</i>	Spain	16	14	5	16
<i>farm R2</i>	Spain	21	18	5	21
<i>farm R3</i>	Spain	32	26	6	32
<i>farm MS</i>	Spain	7	7	3	7
<i>farm SS</i>	Spain	42	35	8	42
<i>farm AI</i>	Spain	11	10	2	11
<i>farm MC</i>	Spain	10	10	2	10
Total	Spain	175	152	39	175
<b>Large White</b>	Commercial	53	48	53	0
<b>Landrace</b>	Commercial	29	34	29	0
<b>Duroc</b>	Commercial	42	42	42	0
<b>Tamworth</b>	United Kingdom	15	14	15	0
<b>Berkshire</b>	United Kingdom	19	19	19	0
<b>Iberian Pig</b>	Spain	26	38	26	0
<b>Meishan</b>	China	10	10	10	0

## 2. Microsatellite genotyping

A total of 34 autosomal markers were analyzed. Microsatellites were chosen based on their absence of null alleles, sharpness of peaks and possibility of being grouped into six multiplex PCR. Amplified products were electrophoresed in an ABI 3130® (Applied Biosystems, USA). The French Pig Map reference samples DNAs F9110010 and

F9119912 (courtesy of INRA, <http://www.toulouse.inra.fr>) were used to normalize the allele sizes.

### 3. SNP genotyping

High density SNP genotyping was performed using the PorcineSNP60 BeadChip (IlluminaInc, USA; Ramos et al. 2009) according to manufacturer's protocol. For this study, only SNPs mapped to one of the 18 autosomes on *Sus scrofa* build 10.2 were included in the analysis and SNP markers with more than 5% missing genotypes were excluded by using Plink software (Purcell et al. 2007). Finally, 46887 SNPs of the 62163 potential SNPs were used for the analysis.

### 4. Mitochondrial sequencing

The primers previously used in Luetkemeier et al. (2010) were used to amplify the D-loop region by PCR. As not all samples yielded the complete fragment (722 bp), a 624 bp fragment common to most samples was finally used for the analysis.

## 5. Data analysis

The number of alleles per marker, the allele frequencies as well as observed and expected heterozygosities were calculated with Genalex V.6.3 (Peakall and Smouse, 2006). Wright's F-statistics (Wright1969) were used as implemented by GENEPOP 4.0.10 software (Rousset et al., 2008).

The Plink software was used to compute the Identity by State (IBS) matrix for all pairs of individuals based on genome similarity derived from the Porcine SNP60 data and used to perform multidimensional scaling analysis (MDS). The Structure software version 2.0 (Pritchard et al. 2000) was used to examine relatedness among breeds and also population stratification between farms using both SNP and microsatellite data.

The individual SNP marker allele frequencies were calculated with Power marker (Liu y Muse 2005) and compared across breeds in order to identify a maximum of ten SNP markers with high allele frequencies in CM pig and low in the other breeds as well as SNPs with low allele frequencies in CM pig and high in the other breeds were selected.

To calculate LD, pairwise  $r^2$  values were estimated for adjacent SNP that were less than 3 Mb apart for each autosomal chromosome and for each population using Haploview 4.2 (Barrett et al. 2005).

D-loop region for each individual relative to pig mtDNA sequence GenBank ID AJ00218 were aligned by ClustalX V.2 (Larkin et al. 2007) and grouped into haplotypes using the program ALTER (Glez-Peña et al. 2010). Phylogenetic relationships among the haplotypes were determined with Mega 5.03 using the NJ method based on the maximum composite likelihood. The Genbank accession numbers of the sequences are JQ238237-JQ238602.

## **Results**

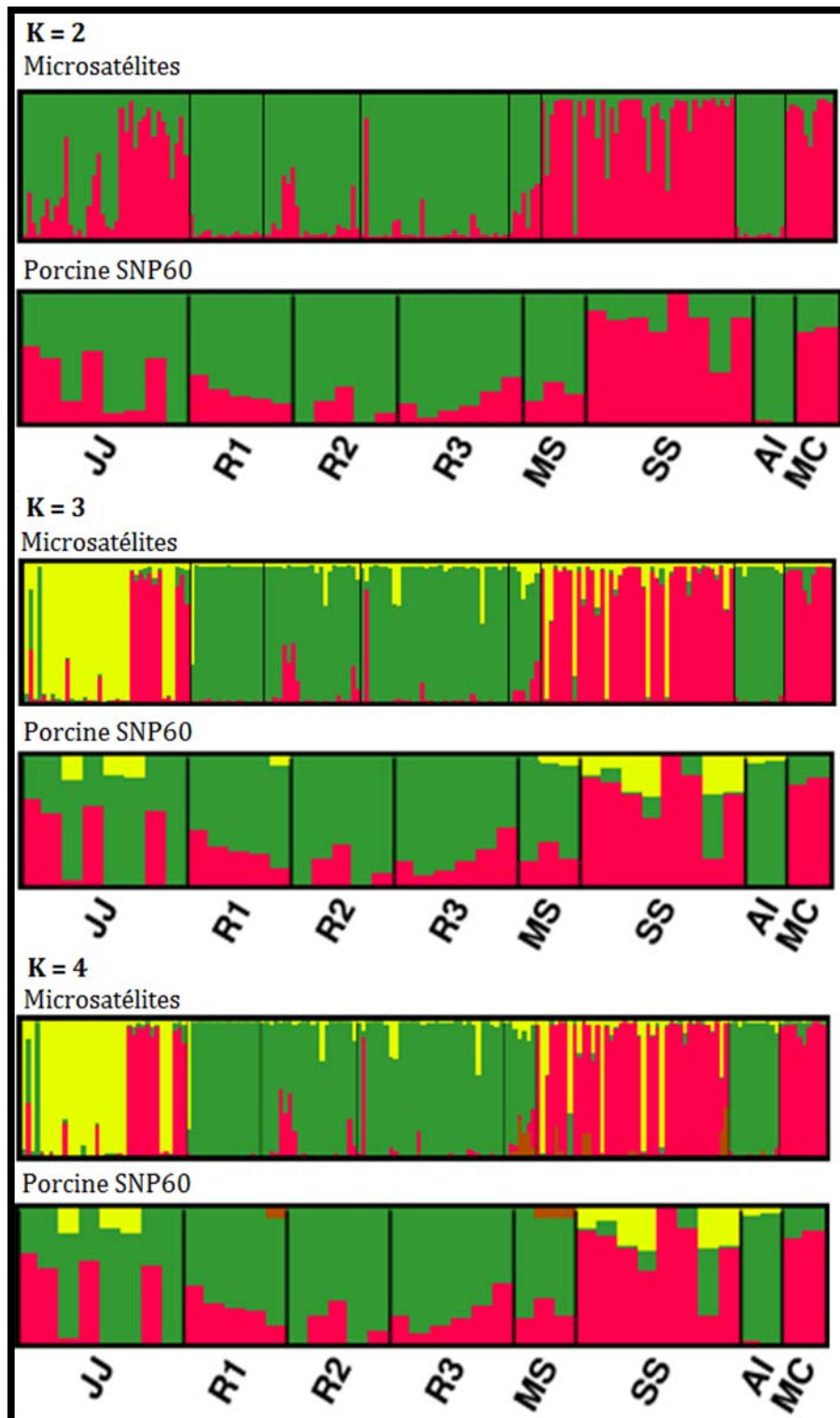
### Genetic diversity and population substructure of CM pig

All tested microsatellite markers were polymorphic for all the CM pigs with an average number of detected alleles of 6.4 ranging from 2 (SW796) to 12 (S0005). The expected heterozygosity ( $H_e$ ) was 0.539 and the observed heterozygosity ( $H_o$ ) was 0.516. The analysis performed by farms showed that farm SS had the highest genetic diversity ( $H_e=0.563$ ) and also the highest mean number of alleles per

locus (NMA=4.735), while farms AI and MC had the lowest values of both, He (0.324 y 0.262) and NMA (2.265 y 1.176) respectively, although they represent the smallest number of animals in the CM population. The average value of  $F_{st}$  across all loci was 0.114 indicating that 11% of genetic variation was explained by differences between farms.

To estimate the existence of different genetic clusters in the CM population, the model implemented by Structure software was used for the microsatellites and SNP genotyping data separately (**Figure 1**). For K=2 consistent results were obtained across microsatellites and SNPs data and also with the tree of genetic distances. farms R1-3 y MS and AI (green), and farms SS and MC (red) each appeared to represent different gene pools, while pigs from farm JJ were assigned to each of these gene pools. However, for K=3 a third cluster represented by farm JJ (yellow) appeared with microsatellite data but not with SNPs.



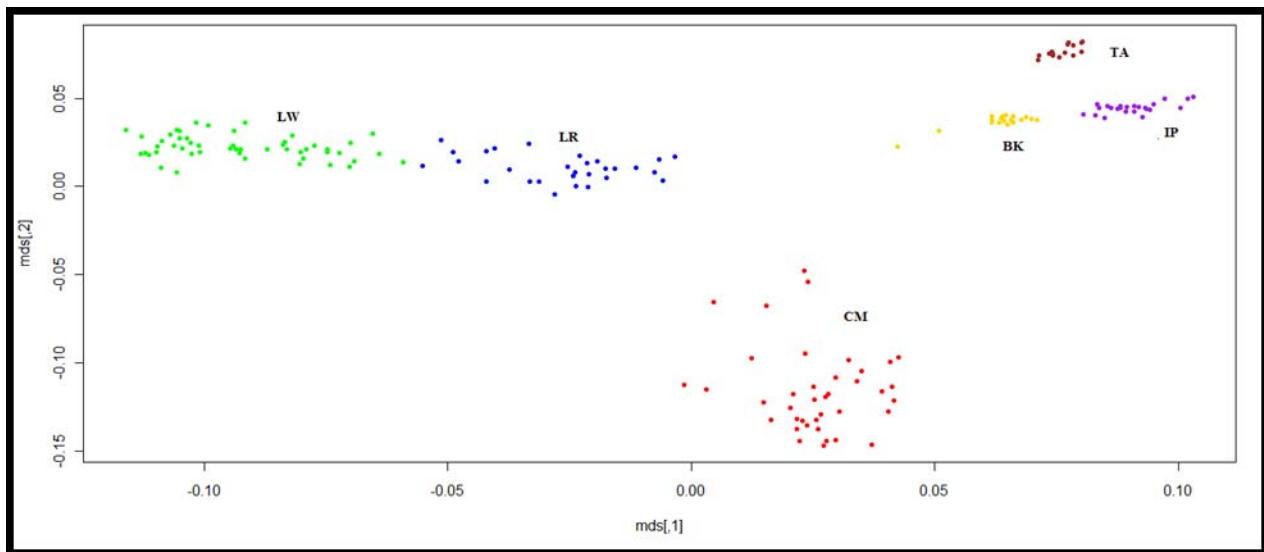


**Figure 1.** Graphic representation of estimated membership coefficients for each individual from  $K = 2$  to  $K = 4$ . Results with microsatellites data; results with SNPs data. Each color represents the proportion of the genome assigned to each assumed cluster.

### Relationships of CM with other domestic breeds.

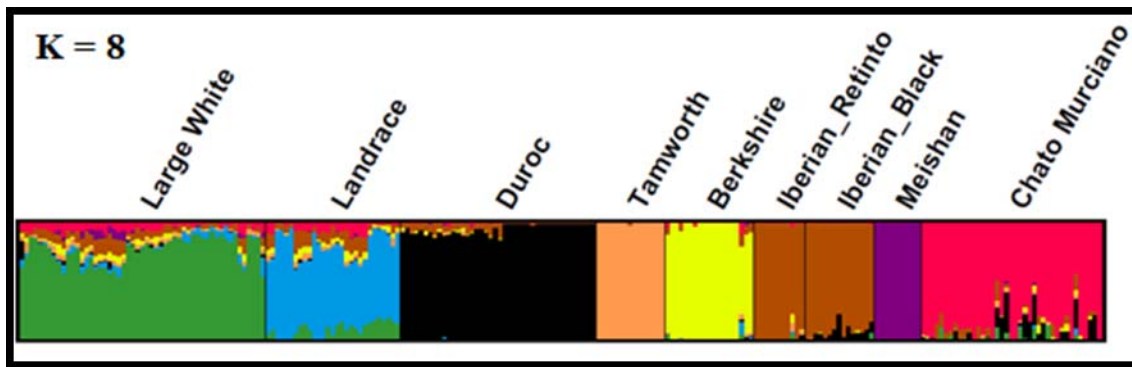
Based on the Porcine SNP60 data three main clusters were identified by the MDS analysis. The first one was the Asian cluster represented by the Meishan breed which was the most divergent population. The second cluster was in the opposite end of the first component relative to the Asian cluster, and was represented by the Duroc breed. Finally, a third cluster that included two commercial breeds, Large White and Landrace, English breeds Tamworth and Berkshire, and the two Spanish breeds: the Iberian and the CM.

Since the Meishan breed and the Duroc breed accounted for a large part of the variation observed, these two breeds were removed subsequently to analyze the European cluster (**Figure 2**). All the breeds formed discrete clusters with the exception of Retinto and Negro Ibérico that are actually two varieties of Iberian Pig. CM occupied an intermediate position, in between the commercial breeds and Tamworth, Berkshire, and Iberian Pig, in the first component, but the second component explained the difference between CM and the other breeds as shown in **Figure 2**. All the CM pigs were grouped in the same area of the plot, although seven individuals separated slightly from the main CM cluster. Four of those animals belonged to farm SS.



**Figure 2:** MDS analyses. European breeds.

K values (number of assumed clusters) from six to nine were tested using the Porcine SNP60 data, with K=8 (**Figure 3**) determined to be the optimum value (Bayesian Information Criterion, BIC test). Large White and Landrace were the breeds with lower membership coefficient (0.77 and 0.71 respectively) while Meishan showed the highest (1.00). For the rest of the breeds the average membership coefficient was over 0.93 with the exception of CM, for which it was 0.85. A detailed study of the individual membership coefficient in CM was made. All the CM pigs clustered together with a membership coefficient over 0.75 except for four pigs that had a membership coefficient of around 0.5 which belonged to farm SS.



**Figure 3.** Graphic representation of the membership coefficients estimated for each individual. Predefined populations are separated by black lines. Each color represents the proportion of the genome assigned to each assumed cluster.

The Control Region or D-loop fragment of the mtDNA was used to study the genetic origin and relationships within and between pig breeds. There were two main clades: the first one included exclusively European and commercial breeds and the second one represented an Asian mitochondrial clade, and included all pigs of the Chinese Meishan breed, and some commercial and English pigs. The Asian mtDNA haplotypes were highly frequent in the British pigs. On the other hand, the Iberian and CM pigs carried exclusively European haplotypes. Three common haplotypes in the CM (HP2-4, together 77% occurrence) were shared with the commercial breeds. One haplotype was shared with the Iberian Pig (HP1). All CM pigs carrying HP1

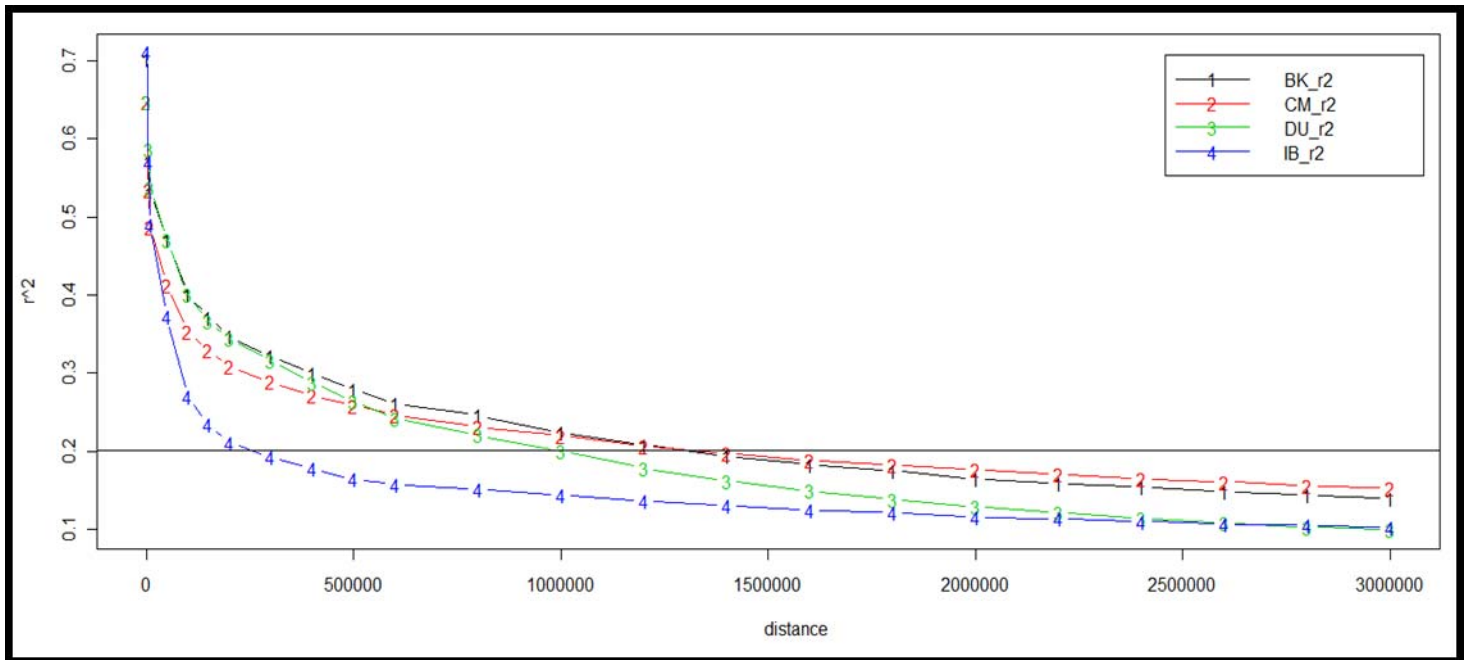
belonged to farm AI both rearing the Iberian and CM pigs. One haplotype was unique to the CM breed (HP5) and occurred in 14% of the CM pigs from farms R1-3 and MS.

### Linkage disequilibrium

To perform a linkage disequilibrium (LD) analysis four populations were chosen: the CM and Iberian breeds from Spain, the Berkshire breed from England and the commercial Duroc breed. LD was estimated for all marker pairs less than 3Mb apart in each autosomal chromosome independently using the measure  $r^2$ . The mean values  $r^2$  decayed with increasing distance between markers in all populations and chromosomes, although the persistence as the distance increased as well as the strength of LD varied greatly among both, populations and chromosomes.

The decay of LD as a function of the marker distance was faster in the Iberian pig than in the other breeds since it decayed below the level of  $r^2 < 0.2$  –black line, **Figure 4**– within 0.25 Mb. In contrast, LD persisted over longer distances in Duroc (1 Mb), CM and Berkshire (1,2Mb). When the average extent of LD for each population was

compared, Iberian pig had the lowest average LD ( $r^2=0.14$ ) while the rest of domestic breeds showed higher values ( $r^2=0.18-0.21$ ).



**Figure 4.** Decline of LD over marker distance (in bases) in each population. The threshold, settled in  $r^2 = 0.2$ , is marked as a black line.

### Distinctive genetic markers set for the CM breed

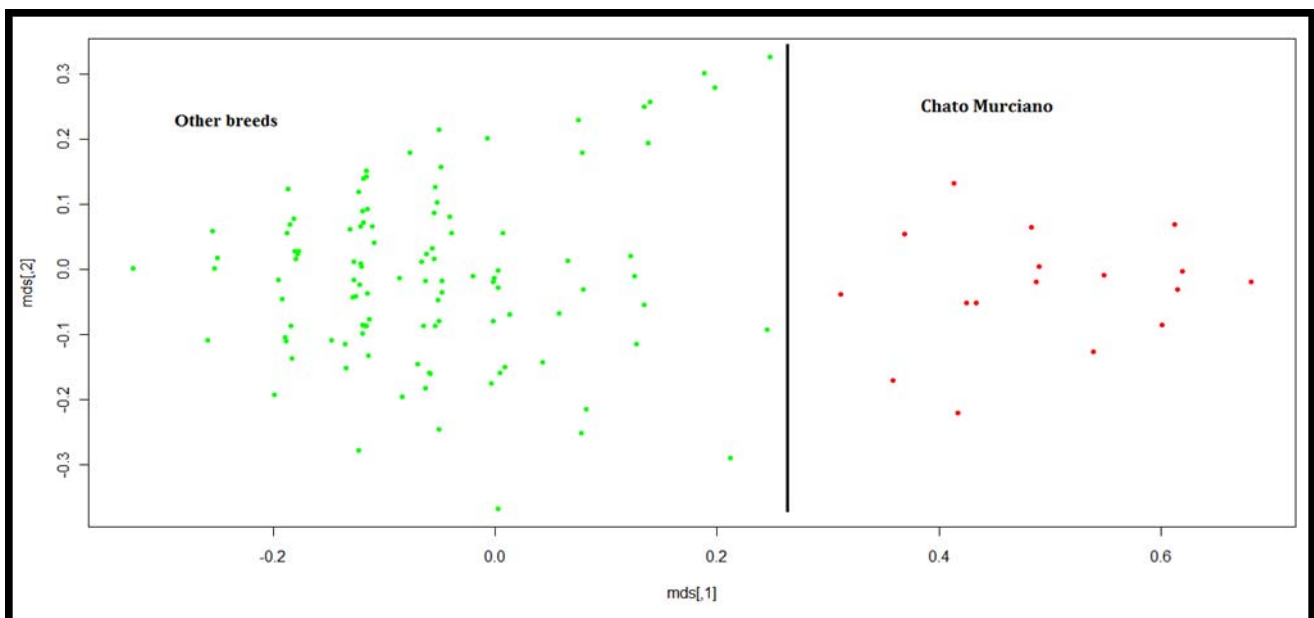
Allele frequencies (F) of the 46887 SNPs used in the study were compared across populations in order to identify the breed specific SNPs in the CM breed.

In total 8 SNP markers (SNP1-SNP8) did have an allele frequency difference between the CM breed and the other breeds studied. The 8 SNPs were classified into three sets according with the screening process.

The first set included the three SNP markers (SNP1, SNP2 and SNP3) that showed an  $F \geq 0.82$  in the CM pig and an  $F < 0.20$  in the other breeds with the exception of SNP2 ( $F=0.40$ ) in the Meishan and the Tamworth. The second set included the SNPs with an  $F \geq 0.95$  in the CM pig and  $F < 0.40$  in the breeds Iberian, Duroc, Large White and Landrace (SNP4, SNP5 and SNP6). The third set included SNP8 that showed an allele frequency in the CM pig of 0.08 and  $F > 0.70$  in the other breeds with the exception of Iberian pig ( $F \approx 0.40$ ).

CM pigs were clustered together and clearly differentiated of the other breeds by MDS analysis (**Figure 5**) and Structure analysis ( $K = 2$ ). The CM breed showed a high membership coefficient ( $CP=0.95$ )

while the other breeds were grouped in the second group with membership coefficient values higher than 0.92 except for the Tamworth breed (CP=0.79). The study of each animal independently showed that 28 out of 35 CM pigs were grouped with  $CP \geq 0.95$  and only one of them had a membership coefficient value lower than 0.77 (animal CM14, CP=0.63).



**Figure 5:** MDS analysis based on 8 distinctive SNPs of CM breed. CM pigs are represented in red dots. The other breeds studied are represented in green dots. Animals with the same genotype showed superimposed dots.



Regarding microsatellite data, five out of the 34 loci microsatellites studied showed allele frequencies  $>0.75$  (**Table 2**). Interestingly, those CM pigs that showed the lowest membership coefficients in the Structure analysis based on PorcineSNP60 data ( $CP \approx 0.5$ ), were heterozygote to any of the five loci microsatellite with high allele frequency. On the contrary, all the animals showing high membership coefficients based on PorcineSNP60 ( $CP > 0.75$ ) were homozygote to the five loci detailed in **Table 2**, with the exception of the pig CM14, although this pig showed the lowest membership coefficient value when the eight distinctive SNPs of CM breed were tested ( $CP = 0.63$ ), and was heterozygote to the locus S0215.

**Table 2.** The allele frequency of the most frequent alleles in the Chato Murciano breed of 5 microsatellite markers in each farm separately and the total population.

	LOCI	S0215	SW 951	S0227	SW769	S0355
<i>Farm</i>	<i>Allele</i>	<b>155</b>	<b>123</b>	<b>230</b>	<b>105</b>	<b>245</b>
<b>JJ</b>		0.81	0.93	0.92	1.00	1.00
<b>R1</b>		0.57	0.94	1.00	1.00	0.93
<b>R2</b>		0.90	0.95	1.00	1.00	0.98
<b>R3</b>		0.94	0.88	1.00	1.00	0.97
<b>MS</b>		1.00	0.80	1.00	1.00	0.57
<b>SS</b>		0.94	0.54	0.80	0.93	0.98
<b>AI</b>		1.00	1.00	1.00	1.00	0.09
<b>MC</b>		1.00	1.00	1.00	1.00	1.00
<b>Total CM</b>		0.89	0.75	0.93	0.98	0.90

## Discussion

In conservation genetics of specific populations or breeds of domestic animals there are two major threats to consider: inbreeding and crossbreeding. Inbreeding may erode genetic diversity and threaten long-term survival because of reduced productivity or inbreeding depression. Crossbreeding may threaten the historical significance of the population or breed by losing what is perceived to be the 'genetic integrity'. In addition, crossbreeding may reduce the perceived value of a breed for being a representative for cultural or culinary heritage. Because the CM breed is currently almost exclusively confined to a single region in Spain, where it is being bred in small numbers on a small number of farms, together with the high propensity of being crossbred with other pig breeds, it is feasible to study how both threats have affected the CM breed.

### Genetic diversity and population substructure of CM.

The levels of genetic diversity in the CM breed detected within this study ( $He=0.54$ ) is slightly lower than the average of European pig breeds ( $He=0.57$ , Zhang y Plastow, 2011) and within the range of other

European local breeds such as Iberian pig (Alves et al. 2006), Mangalitsa (Manea et al. 2009), Portuguese breeds (Vicente et al. 2008) and Balkan breeds (Druml et al., 2012). The genetic diversity of the CM breed in 2012 was higher than the observed by Calvo et al. (2000) ( $He=0.39$ ) but similar to the genetic diversity showed by Vega-Pla et al. (2004) ( $He=0.53$ ). Taking into account the farm history, Vega-Pla et al. (2004) did not indicate in their study which farms were included, therefore making it difficult to compare the results of both studies. If a population is subdivided all analysis should be performed in each population (Ji et al., 2011). A farm-by-farm analysis revealed marked differences between farms, with particularly low genetic diversity in farms AI and MC ( $He\leq 0.32$ ) displaying the characteristics of isolated small farms with high inbreeding and genetic drift.

The CM breed shows a high degree of substructure that is correlated with the farms they are reared on. Specifically, there appear to be two distinct gene pools. The first one includes farms R1-3, MS and AI, and within these farms, farms R1, R2, R3 displaying an even closer relationship ( $D_A < 0.07$ ). The three farms belong to the same farmer and therefore the data suggested that an increased rate of swapping of breeding stock between farms enhances gene flow between them, explaining the high similarity. Farms SS and MC represented the

second gene pool while farm JJ seemed to constitute a third distinct group (e.g. when  $K \geq 3$  using microsatellite data). However, taking into account the fact that Farm JJ had animals assigned to both clusters inferred by Structure using SNPs data and that all mtDNA haplotypes found in farm JJ were highly frequent in pigs of either the R1-3 or SS farms, it is likely that farm JJ represents a mix of animals from both clusters instead of forming a distinct third group.

The results demonstrated that all farmers appear to have their own breeding strategies. It is clear that breeding strategy is not limited to effective population size on farms and exchange of breeding stock of the CM pigs between farms, however farms SS and AI showed signs of genetic introgression with other breeds. In farm AI, 60% of the pigs presented a mtDNA haplotype found only in Iberian pigs (IP, haplotype HP1). Considering that farm AI reared both CM and Iberian pigs, and its high level of inbreeding, it is likely that the farmer preferred to cross animals of those breeds to improve production parameters. Similarly, for farm SS, several indications of genetic introgression were found. Firstly, farm SS showed the lowest allele frequency in those loci microsatellite previously described as exclusive of CM pure animals (Vega-Pla et al. 2004). Secondly, animals belonging to farm SS were clearly separated from the CM cluster when the IBS matrix was

computed. In addition, four pigs from farm SS were assigned to CM breed with membership coefficients below 0.5 indicating a high level of genetic admixture. It seems, therefore, that breeding practices of farms SS and AI include, or included, crossbreeding strategies. This is noteworthy since breeders with high level of admixture should be excluded from a future conservation program in order to preserve the genetic identity of the CM breed.

#### Relationships of CM breed with other domestic breeds.

All the pig breeds, including the CM, were unambiguously identified by the stratification models used, showing that high density SNP genotyping data provided a powerful tool for assessing genetic differentiation between populations and breed assignment (Ramos et al. 2011). Structure analysis showed a lower membership coefficient in the CM (0.85) than in the Berkshire, Tamworth, Duroc and Iberian pig breeds (all above 0.90), but this was mainly due to the existence of a limited number of admixed animals within the CM pigs tested. There was evidence of genetic contributions from other pig populations into the CM. Specifically, the breeds Large White, Berkshire and Iberian pig appear to have had large influence on the CM gene pool while

Tamworth had little influence, which is in agreement with the recorded history of the CM (Lobera, 1998). Interestingly, the Duroc breed seemingly had a large contribution to the CM gene pool although no historical records exist that support such crossbreeding. Crossbreeding with Duroc and/or Iberian X Duroc crosses appears to have been recent and localized as it is mostly evident in the CM pigs from farm SS and farm JJ. These farms reared both Duroc and Iberian x Duroc crosses, so farmers may have tried to improve production parameters of the CM by crossing CM pigs with Duroc or Iberian x Duroc crosses pigs.

The analysis of mtDNA has been an important tool to identify genetic introgression of Asian pigs into European breeds (e.g. Giuffra et al. 2000). Occurrence of Asian haplotypes can be particularly high in commercial and English breeds (Giuffra et al., 2000; Kim et al., 2002; Wu et al. 2007), as confirmed by this study. However, despite the very high occurrence of Asian haplotypes in English breeds such as Berkshire, which are thought to have been used in the past for crossbreeding with the CM pigs, no Asian haplotypes were found in this Spanish breed. Because of recorded breed history, the presence of Asian haplotypes in CM might be expected since English pig breeds could have carried Asian haplotypes into CM as happened with other

local breeds such as Jabugo Spotted (Alves et al. 2003). The absence of Asian haplotypes is however consistent with historical records that report that only males were imported to Murcia Region (Lobera 1998). Most of the CM (55%) carried haplotypes that are common in other European pigs. Although this may reflect past and present day crossbreeding with commercial pig breeds, these common haplotypes may have had a wide distribution area even at the time the first CM pigs were historically described. It is noticeable that one haplotype (HP5), which was present in 14% of the CM pigs analyzed, was not detected in any other breed. It is possible that HP5 had been inherited from the primitive Murcia pig, or, alternatively, a more recent mutation may have gained prominence during recent bottlenecks because of drift.

#### Linkage disequilibrium in CM pig.

The usefulness of this SNP panel to study the LD can be considered reliable due to the large number of markers utilized (Uimari and Topai, 2010; Badke et al., 2012). This is the first study in which the extent of LD in the CM pigs was assessed. Studying differences between populations is important since the strong

structure of some domestic populations is one of the major reasons for false associations between phenotypes and genetic markers in association mapping studies (Zhao et al., 2007). Therefore, a key aspect of the usefulness of marker assisted selection is to identify common patterns of LD among different populations (Short et al., 1997) in order to increase the confidence in the associations and make their application possible (Zhang and Plastow, 2011). The extent of LD is determined by important features in the population history such as genetic drift, population structure, admixture, artificial and natural selection, population bottlenecks, and also by recombination and mutation rates and gene conversion (Ardlie et al., 2002; Du et al., 2007).

Uimari and Topai (2010) and Badke et al. (2012) noticed that  $r^2$  ranged from 0.36 to 0.48 for markers less than 100kB apart in the pig breeds studied using the Porcine 60SNP panel which was consistent with the results obtained in CM pig, Berkshire and Duroc but not in the Iberian pig (0.27). Iberian pig has been included in selection programs recently and the samples that were used in this study belong to pure Iberian pigs, which may explain the small extent of LD in this breed. On



contrary, the CM pig was crossed with other domestic pigs and shows a relatively low genetic diversity. These events tend to increase the LD (Ardlie et al., 2002; Du et al., 2007; Badke et al., 2012).

There is an increasing interest in preserving local breeds, not only because they are important for the maintenance of the genetic diversity, but also because they could develop an important role in the future of animal breeding. Specifically, the implementation of the so-called Low Input Breeds project ([www.lowinputbreeds.org](http://www.lowinputbreeds.org)) could represent an economically viable strategy for small farmers to increase revenue because of the higher value of their products. Furthermore, the cost of high density genotyping is decreasing gradually. Both factors could make it possible to develop tailor-made marker assisted selection for local breeds in a near future. The extent of LD determines the marker density required for genome-wide association studies (GWAS). According to the results provided, Iberian pig would need a higher number of markers due to shorter extent of LD than CM. More studies are necessary to confirm these results.

Distinctive genetic markers of CM breed for the identification of admixed animals and traceability purposes.

The development of new sequencing technologies has generated a high number of SNP markers that can be used for the identification of breeds in livestock species, and therefore to ensure the traceability of animal products (Ramos et al., 2011, Pant et al., 2012). CM pig products have a higher technological quality (Peinado et al., 2004) but they require longer rearing period (Galián et al., 2009) than commercial breeds, and thus the CM products are sold at higher prices. Due to the differences in terms of quality and price, tools to enforce product identity certification are needed to prevent fraud and assure the customer confidence.

Eight SNP markers resulted to distinguish the CM from other breeds. This distinctness offers the opportunity to design a genetic marker set using only eight SNP markers that could unambiguously trace products sold to be from CM pig back to the origin and at a low cost. On the other hand, the maintenance of purebred animals is essential to develop a conservation program in domestic breeds (Bradley et al., 1994). Therefore, the possibility of identifying the pure breed pigs through genetic markers is particularly interesting. The results showed that the level of admixture obtained by analyzing eight SNPs selected based on their allele frequencies were consistent with those obtained by analyzing 46887 SNPs. Interestingly, those animals

with higher levels of admixture when SNP markers were studied, were also heterozygote for those microsatellite markers described by Vega-Pla et al. (2004) as being monomorphic in purebred CM pigs, whilst those pigs with low level of admixture using PorcineSNP60K data were homozygote for those microsatellite markers. Despite the fact that the heterozygosity of those microsatellite markers may suggest admixture, high density SNP data seems to be a more reliable genetic marker to identify purebred individuals. Indeed, the microsatellite data cannot guarantee low level of admixture since those alleles can also be found in commercial crosses and the Iberian pig.

The distinctness of CM pigs showed by genotyping the selected eight SNP markers would provide valuable information for both, traceability purposes and detecting admixed animals. The use of these markers to enforce product identity certification may stimulate local farmers, local butchers and the local government to maintain their breeding stock sustainably. Moreover, the identification of breeders with high level of admixture is important to preserve the identity of the CM.



# ANEXOS



**Anexo I.** Panel de microsatélites utilizado. Se indica el set, número de cromosoma y temperatura de hibridación.

SET	NOMBRE	SECUENCIA DE OLIGONUCLEÓTIDOS (5' → 3')	CROMOSOMA	TEMPERATURA DE HIBRIDACIÓN
1	SW24	F-CTTTGGGTGGAGTGTGTGC R-ATCCAAATGCTGCAAGCG	17	58°C
	S0068	F-AGTGGTCTCTCCTCTTGCT R-CCTTCAACCTTTGAGCAAGAAC	13	58°C
	SW72	F- ATCAGAACAGTGCGCCGT R- TTTGAAAATGGGGTGTFTCC	3	58°C
	S0355	F-TCTGGCTCCTACACTCCTTCTTGATG R-TTGGGTGGGTGCTGAAAAATAGGA	15	58°C
	S0215	F-TAGGCTCAGACCTGCTGCAT R- TGGGAGGCTGAAGGATTGGGT	13	58°C
	SW936	F- TCTGGAGCTAGCATAAGTGCC R- GTGCAAGTACACATGCAGGG	15	58°C
	SW911	F-CTCAGTTCTTTGGGACTGAACC R-CATCTGTGGAAAAAAAAAGCC	9	58°C
2	S0005	F- <a href="https://www-&lt;br/&gt;lgc.toulouse.inra.fr/pig/panel/panel2004.htm">https://www- lgc.toulouse.inra.fr/pig/panel/panel2004.htm</a> - DTCCTTCCTCCTGGTAACTA R-GCACTTCCTGATTCTGGGTA	5	60°C
	S0090	F-CCAAGACTGCCTTGTAGGTGAATA R-GCTATCAAGTATTGTACCATTAGG	12	60°C
	S0026	F-AACCTTCCTTCCCAATCAC R-CACAGACTGCTTTTTACTCC	16	60°C
	S0101	F-GAATGCAAAGAGTTCAGTGTAGG R-GTCTCCCTCACACTTACCGCAG	7	60°C
	S0155	F-TGTTCTCTGTTTCTCCTCTGTTTG R-AAAGTGAAAGAGTCAATGGCTAT	1	60°C
	SW240	F-AGAAATTAGTGCCTCAAATTGG R-AAACCATTAAGTCCCTAGCAAA	2	60°C
	SW951	F- TTTCACAACCTCTGGCACCAG R- GATCGTGCCCAAATGGAC	10	60°C
	IGF1	F-GCTTGGATGGACCATGTTG R-CATATTTTTCTGCATAACTTGAACCT	5	60°C
	SW857	F-TGAGAGGTCAGTTACAGAAGACC R-GATCCTCCTCCAAATCCCAT	14	60°C
3	S0226	F-GCACTTTTAACTTTCATGATACTCC R-GGTAAACTTTTNCCCAATACA	2	60°C
	S0227	F- GATCCATTTATAATTTTAGCACAAAGT R- GCATGGTGTGATGCTATGTCAAGC	4	60°C

	SW122	F-CAAAAAAGGCAAAAGATTGACA R-TTGTCTTTTTATTTTGCTTTTGG	6	60°C
	S0225	F-GCTAATGCCAGAGAAATGCAGA R-CAGGTGGAAAGAATGGAATGAA	8	60°C
	S0228	F-GGCATAGGCTGGCAGCAACA R-AGCCACCTCATCTTATCTACACT	6	60°C
4	SW174	F-GCCAAAATAGCTATTGGACAGC R-TCATGCTATTTTGTTCCAGATG	9	55°C
	SW787	F-CTGGAGCAGGAGAAAAGTAAGTTC R-GGACAGTTACAGACAGAAGAAGG	18	55°C
	S0097	F-GACCTATCTAATGTCATTATAGT R-TTCCTCCTAGAGTTGACAAACTT	4	55°C
	SW902	F-ATCAGTTGGAAATGATGGCC R-CTTGCCTCAAAGAGTTGTAAGG	3	55°C
	S0070	F- GCGAGCATTTTCATTACAG R- GAGCAAACAGCATCGTGAGC	10	55°C
	SW632	F-TGGGTTGAAAGATTTCCCAA R-GGAGTCAGTACTTTGGCTTGA	7	55°C
5	SWR1004	F- TGGGAACACCTGCTTCATTC R- TCCATATGCCCCAAGTGTG	17	60°C
	SW1111	F- AGGTCCTACTGTCCATCACAGG R-GAAGCAGAGTTGGCTTACAGTG	15	60°C
	SW1023	F-AACCTGCTGAGCCACAGTG R-GCAAGTACCCAATCTTTTTTCC	18	60°C
6	SW769	F-GGTATGACCAAAAGTCCTGGG R-TCTGCTATGTGGGAAGAATGC	13	55°C
	SW830	F-AAGTACCATGGAGAGGGAAATG R-ACATGGTTCCAAAGACCTGTG	10	55°C
	S0143	F-ACTCACAGCTTGTCTGGGTGT R-CAGTCAGCAGGCTGACAAAAAC	12	55°C
	SW2410	F-ATTTGCCCCCAAGGTATTTTC R-CAGGGTGTGGAGGGTAGAAG	8	55°C



**Anexo II. Número de alelos (Na), Heterocigosidad observada (Ho) y esperada (He) por loci en la población Chato Murciano.**

<b>Loci</b>	<b>Na</b>	<b>Ho</b>	<b>He</b>
<b>S101a</b>	5	0,655	0,691
<b>IGF1a</b>	4	0,545	0,568
<b>S0005a</b>	12	0,722	0,763
<b>S0026a</b>	4	0,451	0,457
<b>S0068a</b>	9	0,551	0,628
<b>S0090a</b>	6	0,619	0,638
<b>S0215a</b>	5	0,230	0,210
<b>S0155a</b>	5	0,477	0,516
<b>S0225a</b>	6	0,529	0,544
<b>S0226b</b>	9	0,485	0,645
<b>S0227a</b>	3	0,138	0,125
<b>S0355a</b>	4	0,145	0,175
<b>S0228a</b>	6	0,700	0,707
<b>SW 951a</b>	7	0,162	0,410
<b>SW24a</b>	9	0,478	0,566
<b>SW240a</b>	8	0,416	0,491
<b>SW122a</b>	10	0,725	0,784
<b>SW72a</b>	6	0,577	0,605
<b>SW911a</b>	6	0,672	0,689
<b>SW936a</b>	9	0,747	0,701
<b>SW857a</b>	6	0,653	0,615
<b>SW632a</b>	6	0,704	0,664
<b>S0097a</b>	5	0,500	0,508
<b>SW787a</b>	7	0,804	0,722
<b>SW902a</b>	3	0,509	0,584
<b>S0070a</b>	11	0,733	0,757
<b>SW174a</b>	4	0,758	0,668
<b>SW1023a</b>	10	0,601	0,645
<b>SWR1004a</b>	6	0,589	0,649
<b>SW1111a</b>	8	0,431	0,423
<b>SW769a</b>	2	0,041	0,035
<b>SW830a</b>	6	0,459	0,425
<b>SW2410a</b>	3	0,161	0,150

<b>S0143a</b>	6	0,509	0,572
---------------	---	-------	-------

**Anexo III.** Coeficientes de pertenencia/membership coefficients (CP) individuales para los cerdos Chato Murciado en base a los datos Porcine SNP60.  
Valor K = 8.

GRANJA	ID individual	MS	BK	LR	CM	TA	LW	IB	DU
R1	R1_1	0,000	0,000	0,000	0,965	0,007	0,000	0,004	0,024
R1	R1_2	0,001	0,000	0,000	0,978	0,021	0,000	0,000	0,000
R1	R1_3	0,001	0,005	0,002	0,990	0,000	0,000	0,002	0,000
R1	R1_4	0,000	0,000	0,000	0,862	0,000	0,000	0,137	0,000
R1	R1_5	0,000	0,000	0,000	0,771	0,000	0,108	0,120	0,000
R2	R2_6	0,000	0,000	0,000	0,942	0,000	0,000	0,058	0,000
R2	R2_7	0,000	0,000	0,000	0,925	0,000	0,031	0,044	0,000
R2	R2_8	0,000	0,000	0,000	0,865	0,000	0,000	0,056	0,079
R2	R2_9	0,000	0,000	0,000	0,915	0,000	0,000	0,085	0,000
R3	R3_1	0,001	0,000	0,000	0,903	0,000	0,016	0,080	0,000
R3	R3_2	0,000	0,000	0,000	0,909	0,000	0,000	0,033	0,058
R3	R3_3	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000
R3	R3_4	0,000	0,000	0,000	0,887	0,000	0,000	0,112	0,000
R3	R3_5	0,001	0,000	0,000	0,961	0,021	0,016	0,000	0,000
R3	R3_6	0,000	0,000	0,000	0,900	0,000	0,016	0,083	0,000
MS	CM90	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000
SS	CM113	0,001	0,020	0,000	0,491	0,035	0,108	0,183	0,163
JJ	CM159	0,001	0,000	0,000	0,790	0,000	0,000	0,017	0,192
SS	CM116	0,027	0,003	0,000	0,501	0,025	0,058	0,042	0,344
R2	CM88	0,000	0,000	0,000	0,999	0,000	0,000	0,000	0,000
JJ	CM168	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000
MS	CM91	0,001	0,000	0,000	0,849	0,000	0,121	0,001	0,028
SS	CM128	0,002	0,016	0,000	0,710	0,026	0,021	0,016	0,209
JJ	CM160	0,001	0,000	0,000	0,847	0,000	0,000	0,000	0,152
SS	CM118	0,027	0,035	0,024	0,504	0,026	0,102	0,019	0,261
MC	CM194	0,001	0,069	0,000	0,794	0,000	0,000	0,087	0,049
SS	CM134	0,001	0,038	0,000	0,814	0,008	0,000	0,000	0,139
MS	CM92	0,001	0,000	0,000	0,850	0,018	0,132	0,000	0,000
AI	CM143	0,001	0,010	0,000	0,873	0,000	0,006	0,110	0,000
JJ	CM161	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000
SS	CM130	0,001	0,025	0,004	0,880	0,000	0,001	0,030	0,059
MC	CM195	0,001	0,066	0,000	0,795	0,000	0,000	0,066	0,073
JJ	CM183	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000
SS	CM112	0,003	0,040	0,047	0,428	0,034	0,048	0,154	0,246
AI	CM154	0,001	0,026	0,000	0,973	0,000	0,000	0,000	0,000
JJ	CM166	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000

SS	CM135	0,001	0,029	0,000	0,754	0,000	0,001	0,050	0,164
JJ	CM14	0,000	0,000	0,000	0,759	0,000	0,000	0,032	0,208
JJ	CM184	0,000	0,000	0,000	1,000	0,000	0,000	0,000	0,000

# BIBLIOGRAFÍA



## **BIBLIOGRAFÍA**

Abecasis GR, Ghosh D, Nichols TE. Linkage disequilibrium: ancient history drives the new genetics. 2005. *Hum Hered.*, 59:118-24.

Allendorf FW. Genetic drift and the loss of alleles versus heterozygosity. 1986. *Zoo Biol.*, 5:181-190.

Alves E, Ovilo C, Rodríguez MC, Silió L. Mitochondrial DNA sequence variation and phylogenetic relationships among Iberian pigs and other domestic and wild pig populations. 2003. *Anim Genet.*, 34:319-24.

Alves E, Fernández AI, Barragán C, Ovilo C, Rodríguez C, Silió L. Inference of hidden population substructure of the Iberian pig breed using multilocus microsatellite data. 2006. *Span. J. Agric. Res.*, 4:37-46.

Andersson P, Tong SY, Bell JM, Turnidge JD, Giffard PM. Minim Typing - A Rapid and Low Cost MLST Based Typing Tool for *Klebsiella pneumoniae*. 2012. *PLoS One.*, 7:e33530.

Ardlie KG, Kruglyak L, Seielstad M. Patterns of linkage disequilibrium in the human genome. 2002. *Nat Rev Genet.* 3:299-309.

Badke YM, Bates RO, Ernst CW, Schwab C, Steibel JP. Estimation of linkage disequilibrium in four US pig breeds. 2012. *BMC Genomics.* 13:24.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. 2005. *Bioinformatics*. 21:263-5.

Bennett S. Solexa Ltd. 2004. *Pharmacogenomics*, 5:433-8.

Blott SC, Williams JL, Haley CS. Discriminating among cattle breeds using genetic markers. 1999. *Heredity*, 82:613-9.

Boitard S, Chevalet C, Mercat MJ, Meriaux JC, Sanchez A, Tibau J, Sancristobal M. Genetic variability, structure and assignment of Spanish and French pig populations based on a large sampling. 2010. *Anim Genet.*, 41:608-18.

Bonfield JK, Smith K, Staden R. A new DNA sequence assembly program. 1995. *Nucleic Acids Res.*, 23:4992-9.

Bradley DG, MacHugh DE, Loftus RT, Sow RS, Hoste CH, Cunningham EP. Zebu-aurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in west African trypanotolerant cattle populations. 1994. *Anim Genet.*, 25:7-12.

Brown WM, George M Jr, Wilson AC. Rapid evolution of animal mitochondrial DNA. 1979. *Proc Natl Acad Sci USA.*, 76:1967-71.

Calvo JH, Lobera J, Osta R, Zaragoza P. Genetic diversity of Chato Murciano pig breed. 2000. *Arch. Zootec.*, 49: 53-58.

Chao S, Dubcovsky J, Dvorak J, Luo MC, Baenziger SP, Matnyazov R, Clark DR, Talbert LE, Anderson JA, Dreisigacker S, Glover K, Chen J, Campbell K, Bruckner PL, Rudd JC, Haley S, Carver BF, Perry S, Sorrells ME, Akhunov ED. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). 2010. *BMC Genomics*. 11:727.

Chao Z, Wang F, Deng CY, Wei LM, Sun RP, Liu HL, Liu QW, Zheng XL. Distribution and linkage disequilibrium analysis of polymorphisms of MC4R, LEP, H-FABP genes in the different populations of pigs, associated with economic traits in DIV(2) line. 2012. *Mol Biol Rep*. Jan 31.

Ciampolini R, Cetica V, Ciani E, Mazzanti E, Fosella X, Marroni F, Biagetti M, Sebastiani C, Papa P, Filippini G, Cianci D, Presciuttini S. Statistical analysis of individual assignment tests among four cattle breeds using fifteen STR loci. 2006. *J Anim Sci*. 84:11-9.

Clop A, Amills M, Noguera JL, Fernández A, Capote J, Ramón MM, Kelly L, Kijas JM, Andersson L, Sánchez A. Estimating the frequency of Asian cytochrome B haplotypes in standard European and local Spanish pig breeds. 2004. *Genet Sel Evol*. 36:97-104.

Dalvit C, De Marchi M, Cassandro M. Genetic traceability of livestock products: A review. 2007. *Meat Sci*, 77:437-49.

Dalvit C, De Marchi M, Targhetta C, Gervaso M, Cassandro M. Genetic traceability of meat using microsatellite markers. 2008. *Food Res Int*, 41:301-7.



de Roos AP, Hayes BJ, Spelman RJ, Goddard ME. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. 2008. *Genetics.*, 179:1503-12.

Druml T, Salajpal K, Dikic M, Urosevic M, Grilz-Seger G, Baumung R. Genetic diversity, population structure and subdivision of local Balkan pig breeds in Austria, Croatia, Serbia and Bosnia-Herzegovina and its practical value in conservation programs. 2012. *Genet Sel Evol.*, 44:5.

Du FX, Clutter AC, Lohuis MM. Characterizing linkage disequilibrium in pig populations. 2007. *Int J Biol Sci.* 3:166-78.

Duijvesteijn N, Knol EF, Merks JW, Crooijmans RP, Groenen MA, Bovenhuis H, Harlizius B. A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6. 2010. *BMC Genet.*, 11:42.

Ellegren H. Heterogeneous mutation processes in human microsatellite DNA sequences. 2000. *Nat Genet.*, 24:400-2.

Falconer DS. *Introduction to quantitative genetics*, 2nd. edition. 340pp. London, New York: Longmans., 1981.

Fernández AI, Alves E, Óvilo C, Rodríguez MC, Silió L. Divergence time estimates of East Asian and European pigs based on multiple near complete mitochondrial DNA sequences. 2011. *Anim Genet.*, 42:86-8.

Frankham R, Ballou JD, Briscoe DA. Introduction to Conservation Genetics. Cambridge University Press, Chapter 1. 2002.

Fu WX, Liu Y, Lu X, Niu XY, Ding XD, Liu JF, Zhang Q. A Genome-Wide Association Study Identifies Two Novel Promising Candidate Genes Affecting Escherichia coli F4ab/F4ac Susceptibility in Swine. 2012. PLoS One. 7:e32127.

Galián M, Poto A, Peinado B. Carcass and meat quality traits of the Chato Murciano pig slaughtered at different weights. 2009. Livest Sci., 124:314-320.

Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, Andersson L. The origin of the domestic pig: independent domestication and subsequent introgression. 2000. Genetics. 154:1785-91.

Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D. ALTER: program-oriented format conversion of DNA and protein alignments. 2010. Nucleic Acids Res., 38: 14-8.

Goffaux F, China B, Dams L, Clinquart A, Daube G. Development of genetic traceability test in pig based on single nucleotide polymorphism detection. 2005. Forensic Sci Int., 151:239-47.

Gorbach DM, Makgahlela ML, Reecy JM, Kemp SJ, Baltenweck I, Ouma R, Mwai O, Marshall K, Murdoch B, Moore S, Rothschild MF. Use of SNP genotyping to determine pedigree and breed composition of dairy cattle in Kenya. 2010. J Anim Breed Genet., 127:348-51.

Groenen MA, Megens HJ, Zare Y, Warren WC, Hillier LW, Crooijmans RP, Vereijken A, Okimoto R, Muir WM, Cheng HH. The development and characterization of a 60K SNP chip for chicken. 2011. *BMC Genomics.*, 12:274.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: Genomic selection in dairy cattle: progress and challenges. 2009. *J Dairy Sci.*, 92:433-43.

Herranz J. Elogio y Reivindicación del Cerdo. 1987. Ed. Cajamurcia.

Hill WG, Robertson A. Linkage disequilibrium in finite populations. 1968. *Theor Appl Genet*, 38: 226-231.

Hoffmann I, AjmoneMarsan P, Barker JSF, Cothran EG, Hanotte O, Lenstra JA, Milan D, Weigend S, Simianer H. New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/FAO working group. 2004. In: *Proceedings of 29th International Conference on Animal Genetics*, p. 123, Tokyo, Japan.

Holsinger KE, Weir BS. Genetics in geographically structured populations: defining, estimating and interpreting  $F_{ST}$ . 2009. *Nat Rev Genet.*, 10:639-50.

Ji YQ, Wu DD, Wu GS, Wang GD, Zhang YP. Multi-locus analysis reveals a different pattern of genetic diversity for mitochondrial and nuclear DNA between wild and domestic pigs in East Asia. 2011. *PLoS One.* 6:e26416.

Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. 2004. *Nat Rev Genet.*, 10:739-751.

Jombart T. Adegenet: a R package for the multivariate analysis of genetic markers. 2008. *Bioinformatics*, 24:1403-1405.

Jones GF. Genetic aspects of domestication, common breeds and their origin. 1998. *The Genetics of the Pig*. Edited by: A. Ruvinsky and M. F. Rothschild. CPB International, Oxon, UK. pp. 17-50.

Krawczak M. Informativity assessment for biallelic single nucleotide polymorphisms. 1999. *Electrophoresis*. 20:1676-81.

Kim KI, Lee JH, Li K, Zhang YP, Lee SS, Gongora J, Moran C. Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. 2002. *Anim Genet* 33:19-25.

Koskinen MT. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. 2003. *Animal Genetics*, 34:297-301.

Lacy RC. Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. 1987. *Conserv Biol.*, 1:143-158.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. 2007. *Bioinformatics*, 23:2947-48.

Larson G, Dobney K, Albarella U, Fang M, Matisoo-Smith E, Robins J, Lowden S, Finlayson H, Brand T, Willerslev E, Rowley-Conwy P, Andersson L, Cooper A. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. 2005. *Science*, 307:1618-21.

Larson G, Cucchi T, Fujita M, Matisoo-Smith E, Robins J, Anderson A, Rolett B, Spriggs M, Dolman G, Kim TH, Thuy NT, Randi E, Doherty M, Due RA, Bollt R, Djubiantono T, Griffin B, Intoh M, Keane E, Kirch P, Li KT, Morwood M, Pedriña LM, Piper PJ, Rabett RJ, Shooter P, Van den Bergh G, West E, Wickler S, Yuan J, Cooper A, Dobney K. Phylogeny and ancient DNA of *Sus* provides insights into neolithic expansion in Island Southeast Asia and Oceania. 2007. *Proc Natl Acad Sci U S A*. 104:4834-9.

Larson G, Liu R, Zhao X, Yuan J, Fuller D, Barton L, Dobney K, Fan Q, Gu Z, Liu XH, Luo Y, Lv P, Andersson L, Li N. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. 2010. *Proc Natl Acad Sci U S A*. 107:7686-91.

Laval G, Iannuccelli N, Legault C, Milan D, Groenen MA, Giuffra E, Andersson L, Nissen PH, Jørgensen CB, Beeckmann P, Geldermann H, Foulley JL, Chevalet C, Ollivier L. Genetic diversity of eleven European pig breeds. 2000. *Genet Sel Evol*. 32:187-203.

Leroy G, Danchin-Burge C, Palhiere I, Baumung R, Fritz S, Mériaux JC, Gautier M. An ABC estimate of pedigree error rate: application in dog, sheep and cattle breeds. 2012. *Anim Genet.*, 43:309-314.

Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. 1964. *Genetics*, 49:49-67.

Liu K, Muse SV. Power Marker: an integrated analysis environment for genetic marker analysis. 2005 *Bioinformatics*, 21:2128-9.

Lobera JB. El cerdo Chato Murciano: Orígenes e Historia. Murcia (España): Consejería de Medio Ambiente, Agricultura y Agua. 1988 ISBN 84-95531-19-4. pp.58.

Luetkemeier ES, Sodhi M, Schook LB, Malhi RS. Multiple Asian pig origins revealed through genomic analyses. 2010. *Mol Phylogenet Evol.*, 54:680-6.

Manea MA, Georgescu SE, Kevorkian S, Costache M. Genetic diversity analyses of seven Romanian pig populations based on 10 microsatellites. 2009. *Romanian Biotechnological Letters*, 14: 4827-34.

Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. 2005. *Nature*, 437:376-80.

Martínez M, Peinado B, Martín J, Lobera J, Barba C, Poto A. El Chato Murciano, la raza autóctona de la Región. Situación actual desde el punto de vista genético. 1998. *Adea-Asaja*, 8:24-26.

Martínez AM, Delgado JV, Rodero A, Vega-Pla JL. Genetic structure of the Iberian pig breed using microsatellites. 2000. *Anim Genet.* 31:295-301.

Matsumoto T, Okumura N, Uenishi H, Hayashi T, Hamasima N, Awata T. Population structure of pigs determined by single nucleotide polymorphisms observed in assembled expressed sequence tags. 2012. *Anim Sci J.*, 83:14-22.

Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TP, Sonstegard TS, Van Tassell CP. Development and characterization of a high density SNP genotyping assay for cattle. 2009. *PLoS One.* 4:e5350.

McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J, Coppeters W, Crews D, Dias Neto E, Gill CA, Gao C, Mannen H, Stothard P, Wang Z, Van Tassell CP, Williams JL, Taylor JF, Moore SS. Whole genome linkage disequilibrium maps in cattle. 2007. *BMC Genet.*, 8:74.

McNeeley JA. Economics and biological diversity. International Union for Conservation of Nature and Natural Resources. 1988. Gland, Switzerland, 236p

Meadows JR, Chan EK, Kijas JW. Linkage disequilibrium compared between five populations of domestic sheep. 2008. *BMC Genet.*, 30;9:61.

Megens HJ, Crooijmans RP, San Cristobal M, Hui X, Li N, Groenen MA. Biodiversity of pig breeds from China and Europe estimated from pooled DNA samples: differences in microsatellite variation between two areas of domestication. 2008. *Genet Sel Evol.*, 40:103-28.

Milos P. Helicos bioSciences. 2008. *Pharmacogenomics*, 9:477-80.

Montgomery JL, Sanford LN, Wittwer CT. High-resolution DNA melting analysis in clinical research and diagnostics. 2010. *Expert. Rev. Mol. Diagn.*, 10:219-240

Negrini R, Milanesi E, Colli L, Pellecchia M, Nicoloso L, Crepaldi P, Lenstra JA, Ajmone-Marsan P. Breed assignment of Italian cattle using biallelic AFLP markers. 2007. *Anim Genet.* 38:147-53.

Negrini R, Nicoloso L, Crepaldi P, Milanesi E, Colli L., Chegdani F, Pariset L, Dunner S, Leveziel H, Williams JL, Ajmone Marsan P. Assessing SNP markers for assigning individuals to cattle populations. 2008a. *Anim Genet.*, 40:18-26.

Negrini R., Nicoloso L., Crepaldi P. et al. Traceability of four European Protected Geographic Indication (PGI) beef products using Single Nucleotide Polymorphisms (SNP) and Bayesian statistics. 2008b. *Meat Science*, 80: 1212-7.

Nei M. Genetic distance between populations. 1972. *Am. Nat.*, 106: 283-292.



Nei M, Kumar S. Molecular evolution and phylogenetics. 2000. Oxford University Press, New York (USA), pp. 333.

Nidup K, Moran C. Genetic diversity of domestic pigs as revealed by microsatellites: a mini review. 2011. *Genomics and Quantitative Genetics*, 2:5-18.

Onteru SK, Fan B, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF. A whole-genome association study for pig reproductive traits. 2012. *Anim Genet.* 43:18-26.

Pamilo P, Nei M. Relationships between gene trees and species trees. 1988. *Mol Biol Evol.*, 5:568-83.

Pant SD, Schenkel FS, Verschoor CP, Karrow NA. Use of breed-specific single nucleotide polymorphisms to discriminate between holstein and jersey dairy cattle breeds. 2012. *Anim Biotechnol.*, 23:1-10.

Pariset L, Mariotti M, Nardone A, Soysal MI, Ozkan E, Williams JL, Dunner S, Leveziel H, Maróti-Agóts A, Bodò I, Valentini A. Relationships between Podolic cattle breeds assessed by single nucleotide polymorphisms (SNPs) genotyping. 2010. *J Anim Breed Genet.*, 127:481-8

Paszek AA, Flickinger GH, Fontanesi L, Beattie CW, Rohrer GA, Alexander L, Schook LB. Evaluating evolutionary divergence with microsatellites. 1998. *J Mol Evol.*, 46:121-6.

Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. 2006. *Mol Ecol Notes*, 6:288-95.

Peinado B., Poto A., Gilb F. López G. Characteristics of the carcass and meat of the Chato Murciano pig. 2004. *Livest Prod Sci.*, 90:285-92.

Peinado B, Vega-Pla JL, Martinez MA, Galian M, Barba C, Delgado JV, Poto A. Chato Murciano pig breed: genetic and ethnozoological characterization. 2006. *Animal Genetic Resources Information*, 38:77-86.

Ponsuksili S, Murani E, Brand B, Schwerin M, Wimmers K. Integrating expression profiling and whole-genome association for dissection of fat traits in a porcine model. 2011. *J Lipid Res.*, 52:668-78.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. 2000. *Genetics*. 155:945-59.

Pritchard JK, Donnelly P. Case-control studies of association in structured or admixed populations. 2001. *Theor Popul Biol.* 60:227-37.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. 2007. *Am J Hum Genet.* 81:559-75.

Ramírez O, Ojeda A, Tomàs A, Gallardo D, Huang LS, Folch JM, Clop A, Sánchez A, Badaoui B, Hanotte O, Galman-Omitogun O, Makuza SM, Soto H, Cadillo J, Kelly L, Cho IC, Yeghoyan S, Pérez-Enciso M, Amills M. Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig breeds. 2009. *Mol Biol Evol.* 26:2061-72.

Ramos A.M., Crooijmans R.P., Affara N.A., Amaral A.J., Archibald A.L., Beever J.E., Bendixen C., Churcher C., Clark R., Dehais P., et al. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. 2009 *PLoS One*4., e6524.

Ramos AM, Megens HJ, Crooijmans RP, Schook LB, Groenen MA. Identification of high utility SNPs for population assignment and traceability purposes in the pig using high-throughput sequencing. 2011. *Anim Genet.*, 42:613-20.

Roberson ED, Pevsner J. Visualization of shared genomic regions and meiotic recombination in high-density SNP data. 2009. *PLoS One.*, 4:e6711.

Rocha M. y Gasca J. 2007. Ecología molecular de la conservación. En: Eguiarte L., Souza V. y Aguirre X. Eds. *Ecología Molecular*, pp.251-272. SEMARNAT-INE-UNAM-CONABIO, México.

Rohrer GA, Freking BA, Nonneman D. Single nucleotide polymorphisms for pig identification and parentage exclusion. 2007. *Anim Genet.*, 38:253-8.

Rosenberg N.A. Distruct: a program for the graphical display of population structure. 2004. *Mol Ecol Notes*, 4:137-8.

Rousset F. GenePop007: A Complete Reimplementation of the GENEPOP Software for Windows and Linux. 2008. *Mol Ecol. Resour.* 8:103-6.

Russell GA, Archibald AL, Haley CS, Law AS. The pig genetic diversity database and the WWW. 2003. *Arch. Zootec.*, 52, 165-72.

San Cristobal M, Chevalet C, Haley CS, Joosten R, Rattink AP, Harlizius B, Groenen MA, Amigues Y, Boscher MY, Russell G, Law A, Davoli R, Russo V, Désautés C, Alderson L, Fimland E, Bagga M, Delgado JV, Vega-Pla JL, Martinez AM, Ramos M, Glodek P, Meyer JN, Gandini GC, Matassino D, Plastow GS, Siggins KW, Laval G, Archibald AL, Milan D, Hammond K, Cardellino R. Genetic diversity within and between European pig breeds using microsatellite markers. 2006. *Anim Genet.*, 37:189-98.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. 1987. *Mol Biol Evol.* 4:406-25.

Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. Genetic evidence for an East Asian origin of domestic dogs. 2002. *Science*, 298:1610-1613.

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, Wang MD, Zhang K, Mitra RD, Church GM. Accurate

multiplex polony sequencing of an evolved bacterial genome. 2005. *Science* 309:1728–32.

Short TH, Rothschild MF, Southwood OI, McLaren DG, de Vries A, van der Steen H, Eckardt GR, Tuggle CK, Helm J, Vaske DA, Mileham AJ, Plastow GS. 1997. *J Anim Sci.*, 75:3138-42.

Silio L. Iberian pig breeding programme. S. Galal, J. Boyazouglu, K. Hammond (Eds.), *Developing Breeding Strategies for Lower Input Animal Production Environments*, International Committee for Animal Recording, Roma. 2000. pp. 511–520

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. 2011. *Mol Biol Evol.*, 28:2731-9.

Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, Goddard ME, Visscher PM. Recent human effective population size estimated from linkage disequilibrium. 2007. *Genome Res.*, 17:520-6.

Teo YY, Small KS, Fry AE, Wu Y, Kwiatkowski DP, Clark TG. Power consequences of linkage disequilibrium variation between populations. 2009. *Genet Epidemiol.*, 33:128-35.

Uimari P, Tapio M. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. 2011. *J Anim Sci.*, 89:609-14.

Ursing BM, Arnason U. The complete mitochondrial DNA sequence of the pig (*Sus scrofa*). 1998. *J Mol Evol.*, 47:302-6.

van der Steen HAM, Prall GFW, Plastow GS. Application of genomics to the pork industry. 2005. *J. Anim. Sci.*, 83:E1E8.

Vega-Pla JL, Martínez AM, Peinado B, Poto A, Delgado JV. The use of Molecular Techniques in the Support to the Conservation of the Chato Marciano Breed of Pig. 2004. *Arch Latinoam Prod Anim.* 12:45-48.

Vicente AA, Carolino MI, Sousa MC, Ginja C, Silva FS, Martinez AM, Vega-Pla JL, Carolino N, Gama LT. Genetic diversity in native and commercial breeds of pigs in Portugal assessed by microsatellites. 2008. *J Anim Sci.*, 86:2496-507.

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES. Large-scale identification, mapping and genotyping of single-nucleotide polymorphisms in the human genome. 1998. *Science*, 280:1077-82.

Wang M, Sheng Z, Wu X, Wang Q, Pan Y. Mitochondrial DNA-based genetic structure analysis of Pudong White pigs. 2010. *Biochem Genet.*, 48:924-37.

Weir BS, Cockerham CC. Estimating F-Statistics for the analysis of population structure. 1984. *Evolution*, 38: 1358–1370.

Whittaker JC, Harbord RM, Boxall N, Mackay I, Dawson G, Sibly RM. Likelihood-based estimation of microsatellite mutation rates. 2003. *Genetics*, 164:781-7.

Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. 1992. *Int Rev Cytol.*, 141:173–216.

Wright S. *Evolution and the Genetics of Population*. 1969. Vol. IV. Variability Within and Among Natural Populations. University of Chicago Press. Chicago.

Wu GS, Yao YG, Qu KX, Ding ZL, Li H, Palanichamy MG, Duan ZY, Li N, Chen YS, Zhang YP. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. 2007. *Genome Biol.*, 8:R245.

Yang SL, Wang ZG, Liu B, Zhang GX, Zhao SH, Yu M, Fan B, Li MH, Xiong TA, Li K. Genetic variation and relationships of eighteen Chinese indigenous pig breeds. 2003. *Genet Sel Evol.*, 35:657-71.

Zhang C, Plastow G. Genomic Diversity in Pig (*Sus scrofa*) and its Comparison with Human and other Livestock. 2011. *Curr Genomics.*, 12:138-46.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M. An Arabidopsis example of association mapping in structured samples. 2007. PLoS Genet., 3:e4.

Zhou ZW, Yan JB, Li H, Ren ZR. Application of high-resolution melting for genotyping bovine mitochondrial DNA. 2011. Biotechnol Prog., 27:592-5.

Zhu X, Luke A, Cooper RS, Quertermous T, Hanis C, Mosley T, Gu CC, Tang H, Rao DC, Risch N, Weder A. Admixture mapping for hypertension loci with genome-scan markers. 2005. Nat Genet., 37:177-81.



