



Facultad de Veterinaria
Universidad Zaragoza



Trabajo Fin de Máster en

Calidad, Seguridad y Tecnología de los Alimentos

*Estudio de marcadores genéticos relacionados con la calidad de la
leche y sanidad de la ubre*

Study of genetic markers related to milk quality and udder health traits

Autor/es

María Piquer Sabanza

Director/es

Jorge Hugo Calvo

Facultad de Veterinaria

2020

AGRADECIMIENTOS

Este trabajo ha sido financiado a través de dos proyectos FITE (Fondo de Inversiones de Teruel) denominados “Nuevas estrategias aplicadas a la mejora del proceso productivo del Queso de Teruel (ASCYTEQ)” y “Cadena de valor en la producción de leche y queso de Teruel. Hacia una economía circular (LACTOCYNARA)”, y de un convenio con el INIA y la Asociación Turolense de Productores de Leche y Queso titulado “Proceso integrado de mejora de la calidad del queso de Teruel”, así como fondos del Grupo reconocido del Gobierno de Aragón “Sistemas agroganaderos alimentarios sostenibles (SAGAS; A14_17R)”.

Quisiera expresar mi agradecimiento tanto a las instituciones como a las personas que me han ayudado en la realización de este Trabajo de Fin de Máster.

Al Centro de Investigación y Tecnología Agroalimentario de Aragón (CITA), en cuyas instalaciones se ha llevado a cabo el presente trabajo.

A mi director de tesis Jorge Hugo Calvo, por confiar en mí y por haberme dado la oportunidad de desarrollarme en el ámbito de la biología molecular, bioinformática y bioestadística, así como de acompañarme en este proceso de formación. Además de ser una persona cercana y comprensiva, dándome muy buenos consejos siempre que ha sido necesario.

A mis compañeras de laboratorio, Pilar Sarto, Laura Iguacel y Kenza Lakhssassi, por ser unas personas maravillosas. Por su dedicación, su disciplina y los ánimos que han hecho de este aprendizaje, una vivencia gratificante que me ha formado tanto profesional como personalmente.

A mis compañeros de máster, en especial David, Josune, Luis, Mónica y Víctor, que gracias a ellos he vivido una maravillosa experiencia, en la Facultad de Veterinaria de Zaragoza. Por darnos apoyo en los estudios y prácticas siempre que lo hemos necesitado.

A todos mis amigos y familia de Caspe, en especial a Marta Piquer, Carlos Marqués y a mis padres Eva Sabanza y Antonio Piquer por escucharme y apoyarme en todo, viendo siempre el lado positivo en cada momento.

ABREVIATURAS

C

CL: Contenido en Lactosa

D

DEL: Días en Lactación

E

E: Edad

ES: Extracto Seco

ETLs: Caracteres de Interés Económico; del inglés “Economic Trait Loci”

F

FDR: Tasa de descubrimientos falsos; del inglés “False Discovery Rate”

FITE: Fondo de Inversiones de Teruel

FRET: Transferencia de energía por resonancia de fluorescencia; del inglés “Fluorescence resonance energy transfer”

G

G: Genotipo

GAS: Selección Asistida por Genes; del inglés “Gene Assisted Selection”

GS: Selección genómica; del inglés “Genomic selection”

GWAS: Estudios de asociación del genoma completo; del inglés “Genome Wide Association Studies”

I

IBS: Idéntico por estado; del inglés “Identity By State”

ICO: Índice de mérito genético combinado

Indels: Polimorfismos de inserción y delección en loci específicos

ISGC: Consorcio Internacional de Genómica Ovina; del inglés “International Sheep Genomics Consortium”

K

KASP: PCR competitiva específica de alelo; del inglés “Kompetitive Allele Specific PCR”

L

LD: Desequilibrio de ligamiento; del inglés “Linkage Disequilibrium”

LSmeans: Medias mínimo cuadráticas; del inglés “Least Squares Means”

M

MAF: Frecuencia del alelo menos frecuente; del inglés “Minor Allele Frequency”

Mb: Un millón de pares de bases

MDS: Escalado Multidimensional; del inglés “Multidimensional Scaling”

MLMA: Modelo Lineal Mixto; del inglés “Mixed Linear Model Association” (MLMA- LOCO; del inglés “Leaving – One – Chromosome – Out”)

N

NC: Número de Corderos Nacidos

NL: Número de Lactación

O

OAR: Cromosoma ovino

P

PG: Porcentaje de Grasa

PL: Producción de Leche

PP: Porcentaje de Proteína

PRL: Prolactina

Q

QTL: Loci relacionados con caracteres cuantitativos; del inglés “Quantitative Trait Loci”

R

RAM: Rebaño – Año – Mes

RCS: Recuento de Células Somáticas

S

SAM: Selección Asistida por Marcadores

SAS: Software estadístico desarrollado por el Instituto SAS; del inglés “Statistical Analysis System”

SNP: Polimorfismos de un solo nucleótido; del inglés “Single Nucleotide Polymorphism”

V

VG: Valor genético

RESUMEN

El programa de mejora genética de la raza Assaf utiliza como criterio de selección además de la producción de leche (PL), el porcentaje de proteína (PP) y porcentaje de grasa (PG), ambos relacionados con la calidad de la leche y el rendimiento quesero. Igualmente, desde una perspectiva de calidad higiénico-sanitaria, el Recuento de células somáticas (RCS) es un indicador de la salubridad de la ubre, y permite la eliminación precoz de animales con mamitis clínicas o subclínicas, pudiendo aplicar incentivos económicos en la venta. En el presente trabajo han sido estudiados factores genéticos relacionados con la producción y calidad de la leche, así como su calidad higiénico sanitaria en ovejas de raza Assaf, utilizando como estrategia análisis de asociación de genoma completo (GWAS) con plataformas de genotipado masivo de marcadores SNPs, y la posterior validación de resultados a través del genotipado en toda la población mediante KASP. En resumen, hemos identificado 2, 6, 3, 15, 25 y 5 SNPs asociados a los fenotipos PL, PP, PG, PL, (porcentaje de lactosa), ES (porcentaje de extracto seco), y RCS, respectivamente, mediante GWAS a nivel de significación cromosómica. De estos SNPs, 3 se han validado mediante estudios de asociación en toda la población de estudio, estando dos de ellos asociados al RCS (rs419096188 y rs424642424), en dos regiones del cromosoma OAR19 en la que están anotados genes relacionados con la respuesta inmune (*NUP210*, *HDAC11* y *ARPP21*). Además, el SNP rs419096188 se ha asociado al CL, presentando una correlación negativa con el RCS, de forma que la selección a favor del genotipo AA produciría una disminución del RCS e incremento del CL. Finalmente, el SNP rs419686662 se ha asociado al PP, y se encuentra localizado cerca del gen *FGFR2*, que codifica para un receptor de factores de crecimiento fibroblástico que se han asociado con en el desarrollo y funcionalidad de la glándula mamaria en producción.

ABSTRACT

The main focus of the genetic selection program of the Assaf sheep breed is to increase milk yield and quality (protein and fat content). Furthermore, traits related with milk health based on somatic cell count (SCC) as an indicator trait for mastitis resistance is being used and allows the elimination of animals with clinical or subclinical mastitis. In the present work, genetic factors related to milk production, quality and udder health in Assaf sheep have been studied, using as a strategy the genome-wide association analysis (GWAS), using a platform of thousands SNPs. Subsequent validation of results through genotyping by KASP was performed. In summary, we have detected 2, 6, 3, 15, 25 and 5 SNPs associated to milk production, protein percentage, fat percentage, lactose percentage traits, solids percentage and SCC, respectively, at chromosome-wise threshold. Validation studies confirmed the GWAS association data for three SNPs. The SNPs rs419096188 and rs424642424 were associated to SCC, being located in two different regions of the OAR19 chromosome. In this regions, three genes related to the immune response (*NUP210*, *HDAC11* y *ARPP21*) were annotated close to the SNPs. Furthermore, the SNP rs419096188 was also associated to lactose percentage trait, with a negative correlation with SCC. Finally, the SNP rs419686662 was associated to protein percentage, being this SNP located to the *FGFR2* gene that could be related to mammary development.

ÍNDICE

AGRADECIMIENTOS	i
ABREVIATURAS.....	iii
RESUMEN.....	v
ABSTRACT	vi
ÍNDICE	vii
ANEXOS.....	ix
I. Introducción	1
1.1 La importancia del ganado ovino lechero.....	1
1.1.1 El ganado ovino lechero en España.....	1
1.1.2 La explotación del ganado ovino lechero español.....	2
1.2 La raza Assaf como especie ovina española.....	2
1.3 Variables de estudio asociadas al ganado ovino lechero.....	5
1.3.1 Composición leche de oveja	6
1.3.2 Producción lechera raza Assaf	9
1.3.3 Factores asociados a composición y producción de la leche ovina.....	9
1.4 Programa de mejora genética en la raza Assaf.....	12
1.5 Uso de la información molecular en los esquemas de selección.....	13
1.5.1 Selección Genómica	14
1.5.2 Selección asistida por marcadores (SAM).....	15
1.5.3 Técnicas de detección de caracteres productivos relacionados con la producción láctea en la especie ovina.....	15
1.5.4 Genes asociados a rasgos productivos y de composición de leche en la especie ovina	16
II. Objetivos	19
2.1 Hipótesis.....	19
2.2 Objetivos: General y específicos.....	19
III. Metodología.....	21
3.1 Material animal y fenotipos analizados.....	21
3.1.1 Selección de animales genotipados con el chip ovino “Illumina AgResearch Sheep HD” de 680K.	22
3.1.2 Animales utilizados para la validación de los resultados GWAS	22
3.2 Análisis de asociación de genoma completo (GWAS).....	23
3.3 Genotipado mediante KASP	23

3.4	Análisis estadístico y bioinformático.....	25
3.4.1	Estimación del efecto animal para cada fenotipo	25
3.4.2	Control de calidad de los datos masivos procedentes del chip de SNPs de alta densidad 680K (Illumina AgResearch Sheep HD) y estudio de la estratificación de los datos.....	25
3.4.3	Análisis de asociación del genoma completo (GWAS).....	25
3.4.4	Análisis de asociación de SNPs candidatos	26
IV.	Resultados y discusión	29
4.1	Análisis de Asociación del Genoma Completo (GWAS).....	29
4.1.1	Fenotipos.....	29
4.1.2	Control de calidad y estudio de la estratificación de la población utilizada para llevar a cabo el GWAS	30
4.2	Validación de SNPs relacionados con los caracteres de leche mediante KASP.....	39
4.2.1	Resultados de los estudios de validación mediante KASP	39
V.	Conclusiones.....	43
VI.	Referencias bibliográficas.....	45

I. Introducción

1.1 La importancia del ganado ovino lechero

La mayor parte de la producción lechera ovina tiene lugar en Asia con 4.924.398 toneladas de leche en el año 2018, es decir, el 46% de la leche ovina mundial, África con 2.447.623 toneladas, el 23%, y Europa contribuye con 3.168.166 toneladas, siendo un 14,3% del censo mundial ovino lechero, que dadas sus elevadas producciones medias individuales aporta un 29% de la producción láctea a nivel mundial. Sus principales países productores son Grecia, Rumanía, España, Italia y Francia, en este orden (FAOSTAT, 2018).

1.1.1 El ganado ovino lechero en España

En 2018 el número de cabezas de ovino en España se situó alrededor de los 15,8 millones, siendo el segundo país de la UE tras Reino Unido con un 27% de la producción mundial. En España existen cerca de 7.000 explotaciones de producción de leche de oveja (6.742 en 2018), con un censo de unas 2.200.000 ovejas de ordeño (2.225.120 ovejas en 2018) (MAGRAMA, 2019).

España, con un 9,7%, es el cuarto país con mayor censo de la UE, por detrás de Rumanía, Grecia e Italia, cuya producción de leche de vaca, oveja y cabra es de más de 8.200.000 de toneladas durante el año 2019 según FEGA (Fondo Español de Garantía Agraria), correspondiendo el 6,22 y 5,8% a la producción de leche de oveja y cabra, respectivamente. Concretamente la leche de oveja contribuye en un 15% a la Unión Europea. A pesar de tener poco impacto como proveedor de leche a nivel mundial, su importancia es cualitativa por su repercusión social respecto a las poblaciones rurales, la conservación medioambiental, o el hecho de ser la base de elaboración de productos de alta calidad, como los quesos, que están circunscritos a ciertas regiones que contribuyen notablemente al desarrollo económico (Buxadé, 1996). De hecho, el aprovechamiento de recursos pastables de zonas marginales y la ocupación de mano de obra de carácter familiar alejada de los grandes centros urbanos, contribuyen al asentamiento de la población rural, evitando el despoblamiento de las zonas más desfavorecidas.

Por otra parte, su importancia varía entre continentes, países, según necesidades y condiciones particulares de cada lugar, siendo gran parte de la producción destinada al sector industrial para la transformación en productos derivados, no diferenciados, como el yogur, la cuajada y el queso de mezcla (Sánchez-Rodríguez, 2011). La utilización de la leche de oveja para la producción de queso artesanal puro ha aumentado discretamente en los últimos años con la idea de incrementar el valor

añadido del producto generado en la propia explotación, gracias a las diferentes políticas de fomento de la producción quesera bajo el amparo de las diferentes Denominaciones de Origen tales como Idiazábal, Roncal, Manchego, Zamorano, Queso de la Serena, Torta del Casar, etc. Estos quesos son muy apreciados en los mercados nacionales e internacionales y son elaborados fundamentalmente a partir de leche de razas locales, de producción mixta carne/ leche y explotadas mayoritariamente bajo sistemas de explotación tradicionales que van desde el extensivo hasta el semi-intensivo. El sector ovino español es la segunda potencia productora de leche de oveja en el ámbito comunitario, representando en términos económicos el 2,5% de la Producción Final Ganadera con un valor de 462 millones de euros en 2017 (INFOLAC, 2017). Además, en el año 2019 se logró recuperar la primera posición como principal país europeo en la elaboración de queso puro de oveja cuyo consumo mantiene una tendencia positiva en los hogares españoles (Ministerio de Agricultura, Pesca y Alimentación, 2019).

Según MAGRAMA (2019), si analizamos los datos por Comunidad Autónoma española del año 2018, la producción de leche de ganado ovino se concentra en dos regiones principalmente, Castilla y León y Castilla-La Mancha, donde se produjeron aproximadamente 300 millones (54,9%) y 175 millones (32%) de litros, respectivamente. En Aragón, la producción láctea de ovino supone un 0,8% del total producido en España, duplicándose casi su producción del 2015 al 2018.

1.1.2 La explotación del ganado ovino lechero español

La tradición española en el ganado ovino lechero ha sido la explotación de razas autóctonas adaptadas a sus zonas de origen, predominando la producción lechera ovina en Castilla y León con las razas Churra y Castellana, en Castilla-La Mancha con la raza Manchega y en el País Vasco y Navarra, con las razas Latxa y Carranzana. Sin embargo, en la actualidad más del 45% de la producción procede de animales de razas foráneas que aportan una mayor productividad como Assaf, Awassi y Lacaune, o del cruce de estas con razas autóctonas (Ugarte et al., 2001; Ugarte et al., 2002). Este cruce es un tipo de estrategia utilizada como mejora genética, junto con la selección de reproductores mejorantes mediante el control de rendimientos dentro de una explotación de animales de raza pura.

En la *Tabla 1*, se muestran de forma resumida censos y características productivas de las principales razas destinadas a la producción de leche ovina en nuestro país (Legaz, 2016).

1.2 La raza Assaf como especie ovina española

Tradicionalmente en España, la producción de ovino de leche se ha basado en razas autóctonas bien adaptadas a sus respectivas regiones de origen, en sistemas de explotación semi-extensivos (razas Churra y Castellana en Castilla y León, raza Manchega en Castilla-La Mancha, o razas Latxa y Carranzana en el País Vasco y Navarra).

Tabla 1. Características productivas de las principales razas destinadas a la producción de leche de oveja en nuestro país.

Raza y principales localizaciones	Producción	Sistema de Producción
Merina Extremadura	Leche: solo 40/50 días de ordeño. D.O.P Torta del Casar D.O.P. Queso de La Serena y queso de los Pedroches. Carne: Corderos de 20/25 kg/ pv en 60-90 días. Denominación específica Cordero de Extremadura "CORDEREX". Lana	Extensivo (Dehesa)
Manchega Castilla - La Mancha, Madrid, Valencia	Leche: 200 L/oveja/año. 165 L/120d. D.O.P Queso Manchego desde 1985. Carne: Corderos pesados, corderos Pascuales de 22-28 Kg/pv y corderos de 8-14kg. I.G.P de Cordero Manchego (1996).	Semi-extensivo hasta intensivo
Castellana Castilla y León	Leche: 120 L / 150 días. D.O.P queso Zamorano. Carne: Corderos de 10/12kg, 18/25 Kg y 30/35 kg/ pv. I.G.P Lechazo de Castilla y León.	Semi-extensivo/ semi-intensivo
Assaf Castilla y León, Madrid, Navarra, Castilla -La Mancha	Leche: 330 L/ 150 días. Carne: Corderos lechales 10 kg.	Intensivo/ semi-intensivo
Churra Castilla y León	Leche: 115 L/120 días. D.O.P queso Zamorano. Queso Castellano y Villalón, Burgos. Carne: Corderos lechales de 9-12 kg hasta 30 días. I.G.P. Lechazo de Castilla y León.	Semi-extensivo
Latxa País Vasco y Navarra	Leche: 126 L/120 días. D.O.P. Quesos Idiazabal y Roncal. Carne: Corderos lechales.	Extensivo / semi-extensivo

Lacaune Castilla-La Mancha, Castilla y León, Madrid, Andalucía, Extremadura	Leche: 350 L/ 150 días. Carne: Corderos lechales 10 kg.	Intensivo
---	--	-----------

Fuente: Legaz, 2016.

Actualmente el sector lácteo ovino se ha transformado, consecuencia de la incorporación de razas foráneas de alta producción, como Assaf y Lacaune, y Awassi en menor medida, además de una forma de explotación basada en sistemas intensivos de producción.

El presente trabajo se centra en la raza ovina Assaf, de origen Israel por el cruzamiento de Awassi y Milchscaf (ARO 1955, Organización para la Investigación Agrícola) obteniendo tanto las características de especialización lechera de Awassi, como las de prolificidad y precocidad de Milchscaf (<https://www.assafe.es/pagRaza/id/es/m/3>), convirtiéndose en una raza con alto potencial productivo que resulta interesante para la ganadería lechera. En España, fue reconocida como raza el 5 de agosto de 2003 (Orden APA/ 2420/2003), catalogada como raza de terceros países. La denominación de raza Assaf española, surge de un proceso de adaptación-selección, así como de sucesivos cruzamientos por absorción con las razas autóctonas españolas, que probablemente la llevan a diferenciarse genéticamente de la raza Assaf originaria de Israel.

El aspecto general es bastante característico, de color blanco, aunque pueden presentar coloraciones rojas o negras en partes distales como cara y extremidades, el perfil es subconvexo y tienen un tamaño grande (hembras 70 - 80 kg, machos 90 - 110 kg). En cuanto a la morfología de las mamas, son simétricas y desprovistas de lana, los pezones son de tamaño medio y están bien implantados en el borde inferior de la mama, con tendencia a situación lateral mejorando la aptitud lechera. Según ASSAF.E (Asociación de Criadores de Ganado Assaf de España) se caracterizan por su elevada producción, con lactaciones prolongadas medias que alcanzan los 210 días y pueden superarlos, con buena adaptación al ordeño mecánico. La producción media normalizada en controles oficiales a 150 días de lactación es de 350 litros (2,2 litros/ día), con composición media de la leche de 6,2% de grasa, 5,3% de proteína y 16,8% de extracto seco. Su aptitud carnicera se limita a la producción de corderos lechales (ASSAF.E, 2015; FEAGAS, 2015).

En la actualidad, en España están censados un total de 138.784 animales de la raza Assaf repartidos en 122 ganaderías distribuidas principalmente en Castilla y León. (MAGRAMA, 2019) (*Tabla 2*).

Tabla 2. Censo y distribución actual de la raza Assaf en España a 31/12/2019.

Comunidad autónoma	Total		Reproductores		Nº Explotaciones libro genealógico
	Hembras	Machos	Hembras	Machos	
Andalucía	654	24	553	24	1
Aragón	1,106	27	898	22	1
Castilla la mancha	6,163	342	5,364	336	6
Castilla León	125,232	6,226	104,766	5,981	107
Cataluña	462	18	431	18	2
Madrid	3,270	121	2,922	121	3
Navarra	1,897	58	1,927	59	2
Total	138,784	6,816	116,861	6,561	122

Fuente: MAGRAMA, 2019.

1.3 Variables de estudio asociadas al ganado ovino lechero

La raza ovina Assaf se caracteriza por su gran potencial productivo, sin embargo para mejorar los resultados de una explotación se utiliza el control de rendimientos (toma de datos fenotípicos) o control lechero regulado mediante el Real Decreto (CE) 368/2005, de 8 de abril, (última actualización en el Real Decreto (CE) 419/2015, del 29 de mayo). El control oficial del rendimiento lechero, es un conjunto de actuaciones cuyo objetivo es la evaluación genética de los reproductores de aptitud lechera para mejorar las producciones lácteas y que consistirá en la comprobación sistemática de la cantidad de leche producida y de sus componentes (composición de grasa, proteína, lactosa, extracto seco y células somáticas), así como la recogida de otra información para su incorporación a los esquemas de selección aprobados para las diferentes razas. El objetivo es conocer y evaluar la especie productora sobre su aptitud lechera para mejorar la productividad analizando tanto composición como cantidad, y además gracias a los datos generados, incorporar información de esquemas para la selección de las diferentes razas.

Por lo tanto, acorde a lo establecido se deberían considerar como caracteres de estudio tanto la producción como el contenido de los principales componentes de la leche, teniendo en cuenta aquellos factores que puedan influir en la calidad y productividad, tales como la época y el tipo de parto, la edad, el estado de nutrición, la lactación, salud y tamaño de la ubre y la duración del ordeño y periodo seco (Abecia y Palacios, 2018).

1.3.1 Composición leche de oveja

La bromatología de un alimento, leche ovina en este estudio, es imprescindible para lograr una mejor comprensión de sus propiedades y componentes (BEDCA, Base de Datos Española de Composición de Alimentos), y como se verá a continuación son modificados según determinados factores intrínsecos y extrínsecos.

1.3.1.1 Grasa

La grasa es un constituyente que se encuentra en una emulsión estabilizada gracias a pequeños glóbulos esféricos envueltos por una membrana lipoproteica. Aporta un elevado valor nutricional y efectos físico- químicos y sensoriales, además contribuye en la transformación de la leche a productos lácteos (Park et al., 2007). Según BEDCA el porcentaje de grasa en la leche de oveja es del 6,3%, sin embargo, según el Sistema Nacional de Información de Razas Ganaderas (ARCA), en leche de ovejas de raza Assaf es del 6,65%.

La leche de vaca varía en el contenido graso respecto a la leche de oveja, esta última contiene mayor proporción de ácidos grasos de cadena corta, esteárico y oleico, y una menor proporción de ácido palmítico (Gallardo, 2013). En la raza Assaf concretamente, según los análisis de un estudio realizado en el Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), el grupo principal de lípidos son los triglicéridos, siendo los ácidos grasos que predominan palmítico, oleico, esteárico, mirístico y cáprico, y en cuanto a los ácidos de cadena corta el ácido butírico supone el 2,64% de la leche (Estrada-Korta et al., 2013). Haciendo referencia a los porcentajes de ácidos grasos de cadena corta, media y larga, se obtuvieron unos resultados para la leche de raza ovina Assaf del 9%, 55%, 37%, respectivamente. Los ácidos de cadena corta y media resultan interesantes de evaluar, ya que son responsables de las características organolépticas de la leche y, por ende, de los productos lácteos obtenidos de su transformación (Park et al., 2007). En un estudio publicado recientemente se llevó a cabo la estimación de parámetros genéticos asociados con la lactancia y los rasgos productivos en ovejas lecheras cruzadas, dando como resultado una moderada heredabilidad del rendimiento en grasa y una elevada heredabilidad respecto al porcentaje de grasa láctea (Murphy et al., 2017).

1.3.1.2 Proteínas

Los compuestos nitrogenados pueden ser encontrados en forma de proteínas del lactosuero o también como caseínas, siendo estas últimas aproximadamente el 80% de las proteínas de la leche. La mayor parte de la leche ovina es destinada a la producción de quesos, yogures y cuajadas,

influyendo en su elaboración la cantidad de proteínas desnaturalizadas asociadas a las caseínas (Huppertz et al., 2004). Dada la importancia de este componente, ASSAF.E lleva un control oficial de la composición media de la leche obteniendo unos resultados del porcentaje de proteína del 5,3% (ASSAF.E, 2015, FEAGAS, 2015).

Gutiérrez et al. (2017) observó una característica que consiste en que las ovejas más productoras de la raza Assaf no necesariamente dan lugar a un porcentaje inferior de proteína, lo que resulta interesante ya que el contenido de grasa y proteína influye positivamente en el rendimiento quesero. Además, el rendimiento y porcentaje de proteína tienen una moderada y elevada heredabilidad. Esto sugiere que se puede llevar a cabo una selección genética, de forma que elegir futuros reproductores para la posterior transmisión a la descendencia podrían ser una buena estrategia (Murphy et al., 2017).

1.3.1.3 Extracto seco

El extracto seco (ES) también es medido por el control oficial lechero. Se trata del total de constituyentes sólidos, es decir, materia grasa, fracción proteica, hidratos de carbono y minerales sin contar con el porcentaje de agua que contiene la leche.

La composición de la leche ovina se caracteriza por un porcentaje de extracto seco del 12%. Según los datos publicados recientemente en ARCA en el año 2018, el porcentaje de extracto seco en la raza Assaf se sitúa en el 10,37%.

1.3.1.4 Lactosa

El contenido de lactosa en leche de oveja es del 4,6%, es el único componente que aumenta al principio de la lactación, después se mantiene constante, y finalmente disminuye en el último tercio del periodo de lactación (Nava-García et al., 2019). Respecto a su heredabilidad, es 0,43, siendo la selección genética del porcentaje de lactosa factible para la mejora de comercialización de la leche (Costa et al., 2019).

La cantidad de lactosa sintetizada es el principal mecanismo regulador del volumen de leche, por lo que afecta al rendimiento de la producción láctea, y está influenciada por diversos factores como el número de partos, la etapa de lactancia, el estado de salud de las ubres y la influencia individual del animal (Fox et al., 2015). De hecho, la síntesis y disponibilidad de lactosa en la ubre disminuye en presencia del alto recuento de células somáticas, debido no solamente a que los patógenos

utilizan lactosa como sustrato, sino también al aumento de la homeostasis mamaria durante la inflamación (Alessio et al., 2016; Costa et al., 2019).

Por otro lado, el contenido en lactosa influye en la calidad sensorial de los productos lácteos. Esto se debe a que la fermentación de la lactosa por los microorganismos y enzimas presentes en la leche da lugar a la formación de ácido láctico, ácido propiónico y otros componentes responsables del gusto y olor característicos del queso. Una reducción del contenido en lactosa afecta significativamente al rendimiento quesero, disminuye la vida útil, y reduce la aceptación del consumidor (Keefe y Elmoslemany, 2007).

1.3.1.5 Recuento de células somáticas

El recuento de células somáticas (RCS) es un parámetro cuya utilización es primordial para determinar la calidad higiénico-sanitaria de la leche de oveja, así como su precio (Raynal-Ljutovac et al., 2007). Generalmente se encuentra asociado a una inflamación mamaria e indica el número de leucocitos que afluyen de la sangre de la oveja a la leche como respuesta a una agresión local. Las CS están formadas por células epiteliales, sanguíneas y partículas citoplasmáticas, observándose que las proporciones de las diferentes categorías de células varían durante el curso de la lactancia y dependen también del estado sanitario animal. El número elevado de estas células afecta la calidad de la leche debido al aumento de las enzimas proteolíticas y lipolíticas, que provocan el incremento de las proteínas séricas y la disminución de caseína, así como el aumento de catalasas, fosfatasa alcalina y lipasas, que liberan ácidos grasos y dan lugar a sabores y olores extraños en la leche (Echeverri et al., 2010). Cabe destacar, que el RCS presenta gran variabilidad siendo particularmente elevado en el periodo de calostro y al final de la lactancia, pudiendo estar influenciado por varios factores tales como la edad del animal, nivel de producción, estrés y estado sanitario, entre otros.

Este parámetro es un indicador de la presencia de mastitis en la glándula mamaria, la enfermedad más común en ovejas lecheras, así como de microorganismos patógenos desencadenantes de reacciones inflamatorias, dando información además de las pérdidas de producción y de modificaciones composicionales de la leche, tales como la reducción del contenido en lactosa y de la proporción de caseínas con respecto a la proteína bruta (Martí, 2005). En la Unión Europea el límite legal del RCS es establecido por el Reglamento CE 853/ 2004, los operadores de la cadena alimentaria deben garantizar que la leche cruda cumpla con los siguientes criterios; respecto a la especie bovina, la leche cruda de vaca debe tener un contenido de células somáticas por mililitro ≤ 400.000 . Sin embargo, no se ha definido un límite legal de RCS en leche cruda de pequeños

rumiantes, lo que sí se establece es que la leche cruda de ovino y caprino debe contener $\leq 1.500.000$ colonias de gérmenes a 30°C por mililitro, y ≤ 500.000 si es leche cruda destinada a fabricación de productos lácteos sin ningún tipo de tratamiento térmico.

Las consecuencias negativas que ocasiona la mastitis en el sector lechero ha dado lugar a la estimación de parámetros genéticos de producción en relación al RCS en ovejas de raza Assaf española, obteniendo como resultado que la heredabilidad del carácter resistencia a mastitis basado en el fenotipo RCS, presenta valores moderados de 0,16. Por otro lado, se observaron correlaciones positivas entre los caracteres asociados a la producción de leche (0,39), indicativo de que los animales resistentes a mastitis (menor RCS), presentan una mayor producción de leche (Jurado y Jiménez, 2017). Sin embargo, en otro estudio, se observó baja heredabilidad y moderada repetibilidad para el carácter RCS (Murphy et al., 2017).

La búsqueda de animales con fenotipos mejorantes y la selección de estos para transmitir a su descendencia resistencia a la enfermedad, son alternativas interesantes, puesto que la susceptibilidad a padecer infección en las ubres es hereditaria (Tolone et al., 2016), sugiriéndose la selección genética como estrategia para la selección de animales resistentes a mastitis entre distintas razas de ovejas lecheras (Banos et al., 2017).

1.3.2 Producción lechera raza Assaf

Las ovejas de raza Assaf se caracterizan por su buena adaptación al ordeño mecánico y elevada producción lechera con lactaciones prolongadas medias que alcanzan los 210 días. En el año 2019 según ASSAF. E, se obtuvo un total de producción de 342.946.604 litros de leche, siendo la producción media normalizada en controles oficiales a 150 días de lactación, de 350 litros (2,2 litros/día). Según ARCA, 400 kg a 180 días de lactación.

Actualmente la leche de ovejas de raza Assaf en España, debido a su alto potencial productivo, ocupa el censo mayoritario ovino de aptitud lechera en la comunidad de Castilla y León aportando el 66,3% de la leche de oveja producida en España.

1.3.3 Factores asociados a composición y producción de la leche ovina

1.3.3.1 Raza

La raza es uno de los factores principales intrínsecos debido a las diferencias de potencial productivo, viéndose además una correlación negativa entre dicho potencial y la composición de

la leche. Los datos de ARCA muestran que la raza Assaf es una de las razas con mayor potencial productivo, con una producción de leche por lactación de 400 kg, sin embargo, los porcentajes de grasa y proteína son del 6,65% y 5,4%, respectivamente. Si comparamos con la raza Churra, la producción de leche media en una lactación es de 117 kg, y los porcentajes de grasa y proteína del 7,01% y 5,79%, respectivamente. Otra raza con menor producción de leche pero mayores porcentajes de grasa y proteína, es la oveja Manchega, con 180 kg por lactación media y unos porcentajes de 7% y 5,6% para grasa y proteína, respectivamente.

Vista la variabilidad entre razas es imprescindible tener en cuenta el factor raza a la hora de realizar un estudio genético sobre una población ovina para mejorar su potencial de productividad en tanto a rendimiento como en porcentajes.

1.3.3.2 Edad

La edad es un factor que afecta la productividad láctea ovina, generalmente conforme aumenta la edad de las ovejas de una explotación ganadera desciende significativamente la producción de leche (Hernández et al., 2011; Elvira, 2016). Esto también sucede en ovejas de raza Assaf, y se debe principalmente a que a mayor edad se da un menor pico de producción combinado con una menor persistencia de la lactación, desencadenando lactaciones más cortas (Pollot y Gootwine, 2004). La edad óptima del primer parto se sitúa entre los 390 y 450 días de edad, es decir, de 13 a 15 meses. Las corderas cuyo parto se encuentra fuera de este rango da lugar a una menor producción a lo largo de su vida productiva, por lo que aumenta la probabilidad de ser desechadas precozmente (Hernandez et al., 2011).

Este factor también influye sobre la composición de la leche ovina, afectando tanto a grasa como a proteína (Othmane et al., 2002). El perfil de los ácidos grasos sufre una modificación, observándose un aumento en los ácidos grasos de cadena corta y media conforme aumenta la edad del animal, y un descenso de los ácidos grasos monoinsaturados (MUFA) y ácidos grasos poliinsaturados (PUFA), todo ello en animales de 1 y 2 años, ya que a partir de los 2 años las diferencias solamente se observan en los PUFAs (De La Fuente et al., 2009).

1.3.3.3 Estado y número de lactación

El estado de lactación influye sobre la composición láctea, cuando la lactación alcanza su pico disminuye la concentración de grasa y proteína, sin embargo, estos aumentan progresivamente conforme disminuye la productividad. En general, la variabilidad de los componentes de la leche

de oveja durante la etapa de producción es muy amplia, principalmente el porcentaje de grasa, y en menor medida las proteínas y de lactosa (Kremer et al., 2015).

El número de lactación está asociado con la edad de la oveja, por lo general aumenta entre la primera y segunda lactación, y el máximo de productividad según la raza se alcanza en la tercera, cuarta o quinta lactación (Buxadé, 1996). La raza Assaf en condiciones intensivas, alcanza su máxima producción en la segunda lactación y disminuye a partir de la tercera, esto genera máximos rendimientos precoces y lactaciones más cortas (Pollott y Gootwine, 2004).

1.3.3.4 Tipo de parto

La prolificidad influye en la producción láctea ya que el número de crías nacidas aumenta la producción lechera (Kremer et al., 2015). Esto es debido a un mayor desarrollo placentario y a una mayor producción de esteroides ováricos, hormonas y lactógeno placentarios, provocando un desarrollo incrementado de la estructura mamaria que genera una mejor producción durante la lactación (Schoknecht et al., 1991). Sin embargo, los resultados de recientes y previas investigaciones mostraron una baja heredabilidad y repetibilidad del número de corderos nacidos por oveja. Además, se obtuvo una correlación genética insignificante entre este factor con los caracteres de rendimiento y porcentaje de grasa y proteína, y el carácter RCS (Hamann et al., 2004; Murphy et al., 2017).

1.3.3.5 Época del año

Este factor se asocia al año y mes de producción láctea. La estación y las características climáticas de cada año provocan variación en las temperaturas y disponibilidad del pasto, afectando la calidad y producción de la leche. La composición de la leche respecto a grasa y proteína varía según sea primavera u otoño. En primavera se caracteriza por un menor contenido en grasa y caseína, y por tanto el rendimiento quesero es mejor en otoño, donde, además, no afectan las altas temperaturas causantes del aumento del tiempo de cuajado y una menor firmeza de la cuajada (Sevi et al., 2004).

Recientemente, en ovejas de raza Lacaune ha sido detectada la influencia de la época de parto en la producción, observándose lactaciones menos productivas en ovejas nacidas en diciembre y febrero a diferencia de aquellas nacidas en septiembre. Esto indica asociación entre el retraso de la pubertad y una mayor producción lechera (Elvira, 2016).

En este apartado, se ha visto que la variabilidad ambiental, fisiológica y genética repercute en la capacidad productiva y la calidad de la leche de una explotación ganadera. En este último aspecto, el factor explotación que incluye manejo, alimentación y estado higiénico sanitario (entre otros

inherentes a la explotación), es también un factor que influye altamente en los fenotipos relacionados con la producción y calidad de la leche.

1.4 Programa de mejora genética en la raza Assaf

Desde el año 2010, existe un programa de Selección Genética propuesto para la raza Assaf basando la selección de la raza en pureza, a partir de los resultados de las valoraciones genéticas de los reproductores en cada rebaño del núcleo de selección. Este programa de mejora genética se desarrolla bajo la dirección y coordinación de ASSAF.E, encargada también de la gestión del libro Genealógico de la raza; junto con ella colaboran distintos organismos y su principal objetivo es aumentar la rentabilidad económica de las explotaciones que emplean esta raza. Para ello establecen como objetivos de selección, el mantenimiento del estándar racial, el incremento de la producción de leche por lactación y el incremento de la calidad de la leche (contenido de grasa y proteína).

El programa de mejora de la raza Assaf se basa en:

- Recolección de datos productivos y genealógicos. Esto se realiza mediante el control lechero. La producción lechera se toma en la misma explotación, mientras que la calidad de leche se mide en centros autorizados externos a la propia explotación.
- Inseminación artificial, para la conexión de rebaños y difusión de la mejora.
- Evaluación genética de los diferentes caracteres:
 - VG150: Valor genético en kg de leche tipificada a 150 días de lactación.
 - Índice de mérito genético combinado (ICO): incluye valor genético en Kg de grasa y proteína tipificadas a 150 días.
 - Valor genético para el recuento de células somáticas (RCS).
 - Valor genético de la morfología mamaria en 5 parámetros.

Se fijan como objetivos primarios la mejora de la cantidad y el rendimiento quesero de la leche. En ambos casos, la selección de los caracteres a través de los cuales se selecciona para estos objetivos se ha hecho en función de la heredabilidad de los mismos. La cantidad de leche se selecciona a través del carácter leche por lactación tipificada a 150 días de lactación, que presenta una estimación de la heredabilidad de 0,31 (Jurado y Jiménez, 2017). Para el rendimiento o extracto quesero se considera el porcentaje de proteína y de grasa en leche que presentan heredabilidades en la raza Assaf de 0,23 y 0,20, respectivamente (ASSAFE, 2011).

Por último, se ha de mencionar que la heredabilidad para el RCS también presenta valores medios de 0,16, mientras que los caracteres de morfología de la ubre muestran heredabilidades elevadas que van desde 0,236 (tamaño de los pezones) a 0,419 (profundidad de la ubre) (Jurado y Jiménez, 2017). En la raza Assaf no hay una estimación para la heredabilidad del porcentaje de lactosa o extracto seco, pero en otras razas italianas, como Comisana, Leccese, y Sarda, la heredabilidad para el carácter contenido en lactosa varía entre 0,14 y 0,18 (Selvaggi et al., 2017).

Dentro del esquema de selección de la raza, hay que destacar que desde el año 2003 y a consecuencia de la normativa europea de incluir criterios de selección relacionados con la resistencia a las encefalopatías espongiiformes transmisibles en ovino (EET), en el programa de selección de la raza Assaf también se incluyen planes de erradicación de genotipos sensibles al Scrapie. Estas decisiones afectan, principalmente, a los sementales ubicados en los centros de Inseminación Artificial (se eliminaron los animales del centro portadores del alelo VRQ y solo se admite la entrada de animales nuevos portadores de al menos un alelo ARR, es decir, machos pertenecientes a los grupos R1, R2 y R3).

Por último, es remarcable que la raza ovina Assaf se está incluyendo la información genómica en sus valoraciones genéticas, implantándose un esquema de selección genómica. En el apartado siguiente se explica con más detalle en qué consiste la selección genómica.

1.5 Uso de la información molecular en los esquemas de selección

La mayoría de los caracteres de interés en Producción Animal presentan una base genética compleja, al estar determinados por un número elevado de genes con un efecto individual pequeño y aditivo (poligenes o QTL), así como por el ambiente, e interacciones múltiples entre genes y ambiente. El avance de un programa de selección clásico, con información sobre méritos genéticos de los reproductores basados en el control de rendimientos de la progenie será más lento y con una menor fiabilidad en aquellos caracteres con heredabilidades bajas. Por otra parte, para los fenotipos mencionados se obtienen valoraciones genéticas fiables si los animales han tenido varios partos y por lo tanto han podido pasar varios años, además hay que tener en cuenta la variabilidad de la época del año de medición de estos caracteres.

El uso de herramientas moleculares permite superar algunas de las limitaciones de los métodos tradicionales de predicción del mérito genético. Las plataformas de genotipado masivo de SNPs (Single Nucleotide Polymorphism) y secuenciación masiva permiten explorar el genoma e identificar genes y polimorfismos responsables de caracteres de interés (por ejemplo, QTLs). La

posibilidad de utilizar genes/marcadores responsables de una parte significativa de la variabilidad fenotípica de caracteres de interés económico (ETLs) es, en el momento actual, una estrategia de alto interés para desarrollar con más eficiencia los programas de mejora en estos caracteres que en general presentan una baja heredabilidad y una difícil y costosa medida. La selección asistida por marcadores (SAM) puede incrementar la tasa de progreso genético para estos caracteres elevando la precisión de la selección y permitiendo la selección de los animales a edades tempranas (Rohrer, 2004) lo que conlleva una gran reducción del intervalo generacional.

1.5.1 Selección Genómica

El objetivo de la Selección Genómica (GS; por sus siglas en inglés “Genomic selection”) es predecir valores genómicos utilizando información disponible a lo largo del genoma en función de los miles de genotipos obtenidos a través de chips SNPs (Meuwissen et al., 2001). La selección genómica combina tres fuentes de información: control de producciones, libro genealógico y genoma de los animales. Una valoración genómica es, en el fondo, una valoración genética, pero con mucha más calidad, ya que, con esta tercera fuente de información, se aporta una mayor fiabilidad, algo que es muy importante para la selección, obteniendo el valor genómico de todos los reproductores, tanto machos como hembras. También se obtiene el valor genómico de todos los animales productores, además de los animales jóvenes, que son los corderos candidatos a futuros sementales, de los que únicamente se conocen los datos de padre y madre. La principal ventaja que tiene la selección genómica es que mejora la fiabilidad, aunque depende del tipo de animal que sea. En los sementales, es el tipo de animal en el que menos aumenta la fiabilidad, porque ya se tiene mucha información de ellos. En ovejas el aumento de la fiabilidad es bastante más importante, pero donde aumenta radicalmente es en los candidatos a futuros sementales, es decir, en los jóvenes corderos que entran en el centro de inseminación artificial. Se puede pasar de un 30% a un 50% de fiabilidad.

Para la selección genómica, se utilizan los chips de SNPs. En 2009, se desarrolló un chip de SNPs de densidad media (con alrededor de 50.000 SNPs) para el ganado ovino que fue denominado “Ovine SNP50 BeadChip” (ISGC, International Sheep Genomics Consortium) permitiendo estudios de asociación del genoma completo, así como la identificación de loci de carácter cuantitativo y estudios de genética comparativa. En la actualidad, se ha desarrollado un chip de alta densidad que dispone alrededor de 680000 SNPs (Illumina AgResearch Sheep HD) a lo largo de todo el genoma, y su utilización en programas de mejora genética ovina puede ofrecer la reducción del intervalo generacional (reducción de costes), el aumento de la ganancia genética y la posibilidad de inclusión de caracteres objeto de selección (Duchemin et al., 2012).

Hoy en día existen otros chips comerciales como el de Affimetrix que es de densidad media que también se están utilizando para este propósito.

1.5.2 Selección asistida por marcadores (SAM)

La incorporación de esta información molecular en los programas de mejora tradicionales se denomina de una forma genérica selección asistida por marcadores (SAM). La SAM se basa en la detección de QTL (loci relacionados con caracteres cuantitativos) o ETL (loci relacionados con caracteres de interés económico), términos que se refieren a marcadores de ADN que ejercen un efecto significativo sobre una o varias características fenotípicas. En este caso, la mutación genotipada no suele ser el QTL o mutación responsable, sino que está en desequilibrio de ligamiento (LD) con la misma. Esta propiedad de algunos genes consiste en no segregar de forma independiente, por lo que poseen una frecuencia de recombinación menor del 50%.

Las principales desventajas del uso de la SAM se deben a la limitada ganancia que se produce incorporando los marcadores y lo complejo de la puesta en marcha del sistema de genotipado de marcadores. Finalmente, si localizamos la mutación responsable del efecto fenotípico observado, y por lo tanto hacemos una selección basada en el marcador responsable del fenotipo directamente, la selección se denomina selección asistida por genes (GAS; del inglés “Gene Assisted Selection”); en principio, su aplicación es relativamente sencilla desde el momento en que se conoce exactamente la mutación funcional y el efecto que tiene sobre los caracteres. Un ejemplo de este proceso de GAS en Producción Animal es la utilización del genotipo del gen *PRNP* de resistencia al scrapie en los programas de selección del ganado ovino. En este sentido, lo que se suele hacer es aumentar la frecuencia del alelo favorable por selección o introgresión, mediante la preselección de reproductores con estos alelos favorables. Esta metodología se ha utilizado con caracteres de prolificidad en ovino (Calvo et al., 2020). Posteriormente en estos reproductores se llevará a cabo el control de producciones, y posterior estimación del VG para los caracteres en cuestión, eligiendo como reproductores aquellos animales con VG más elevado y con los alelos favorables. Un problema que puede presentar este tipo de selección es la disminución de la variabilidad de la población si seleccionamos sólo en función de unos pocos marcadores.

1.5.3 Técnicas de detección de caracteres productivos relacionados con la producción láctea en la especie ovina

En la búsqueda de variantes génicas, podemos distinguir dos grandes tipos de estudios, en función de si la aproximación del trabajo es mediante el estudio de todo el genoma (sin tener en cuenta

condiciones previas, es decir con un enfoque libre de hipótesis) o si es el estudio de un número de genes limitado (gen candidato), previamente seleccionados por estar potencialmente relacionados con el fenotipo (aquí tendríamos un enfoque con una hipótesis previa) como prolificidad, fertilidad, predisposición o resistencia a padecer una determinada enfermedad, etc. En el primer estudio destacan los análisis de asociación de genoma completo (GWAS; por sus siglas en inglés “Genome Wide Association Studies”) en el que se utilizan chips de SNPs que pueden contener desde miles de SNPs hasta 1 millón de SNPs en el caso de los chips de humana. Estos estudios nos permiten localizar posicionalmente en el genoma dónde se encuentran genes o regiones genómicas que afectan a un carácter en cuestión.

Con el enfoque de la hipótesis previa, contamos con la desventaja de que es necesario un conocimiento biológico previo para seleccionar esos genes candidatos, pero con la fortaleza de una gran eficacia estadística y de una comprensión biológica de las diferencias genéticas y fenotípicas observadas. Este segundo enfoque, además, permite la aplicación directa de los resultados en nuestra raza de estudio, lo cual es muy útil ya que en el caso de los estudios de todo el genoma se trata de una aproximación que generalmente obliga a buscar el polimorfismo responsable en genes situados en regiones del genoma que han sido significativamente asociados al fenotipo analizado, y por lo tanto no son de aplicación directa.

En este trabajo hemos llevado a cabo un análisis GWAS para los caracteres de producción y calidad de leche en la raza Assaf.

1.5.4 Genes asociados a rasgos productivos y de composición de leche en la especie ovina

El desarrollo de herramientas genómicas en el ganado ovino en los últimos años ha hecho posible la identificación de regiones asociadas a caracteres de interés en producción animal (García-Gómez et al., 2013; Zhao et al., 2011), y una mejor calidad de la secuencia genómica. En este sentido, actualmente se conoce la posición física de marcadores microsatélite y SNP a lo largo de todo el genoma, debido al desarrollo de las secuenciaciones masivas de nueva generación. En concreto, en ovino la versión más ampliamente utilizada es la v. 3.1 (Jiang et al., 2014) aunque actualmente existe una nueva versión con una calidad más alta de anotación de genes y polimorfismos que se denomina “Oar_rambouillet_v1.0” (2017; <https://www.sheepmap.org/>).

El uso de las metodologías descritas en el apartado anterior (estudio de todo el genoma y de gen candidato) ha permitido detectar un número alto de marcadores genéticos y/o regiones genómicas

asociados a la producción y calidad de la leche en ovino. En concreto, existe una base de datos de QTLs descritos en especies ganaderas y denominada “AnimalQTLdb” (<https://www.animalgenome.org/cgi-bin/QTLdb/index>), donde se recogen los principales QTLs descritos en las diferentes razas. En concreto en ovino, hay recogidos 3.411 QTLs de 178 publicaciones, afectando a 262 caracteres diferentes entre los que se encuentran los relacionados con la producción y calidad de la leche, incluyendo los relacionados con el RCS (<https://www.animalgenome.org/cgi-bin/QTLdb/OA/index>).

II. Objetivos

2.1 Hipótesis

El uso de herramientas de análisis masivos en poblaciones comerciales permitirá la identificación de variantes génicas (marcadores genéticos) asociadas a los fenotipos de producción de leche de ovino, contenido de proteína, grasa, lactosa, extracto seco y recuento de células somáticas (RCS), que una vez validadas podrá ser utilizada en selección asistida por marcadores (SAM) e incluso poder incluirlos en los chips utilizados para selección genómica.

2.2 Objetivos: General y específicos

El objetivo general del presente trabajo es el estudio de factores genéticos relacionados con la producción, y calidad de la leche mediante herramientas de análisis genéticos masivos. Este objetivo general se pretende conseguir a través de los siguientes objetivos:

- 1- Búsqueda de marcadores genéticos asociados a los fenotipos de producción de leche, contenido de proteína, grasa, lactosa, extracto seco y recuento de células somáticas (RCS) mediante la metodología de análisis de asociación de genoma completo (GWAS) con el microchip AgResearch de 680K (Illumina).
- 2- Validación de algunos SNPs asociados a los caracteres lecheros identificados mediante GWAS, mediante genotipado en toda la población de estudio.

III. Metodología

3.1 Material animal y fenotipos analizados

El material y diseño animal utilizado en este estudio, a través de muestras de ADN extraído de sangre en un proyecto anterior, cumple los principios éticos y de protección de los animales utilizados en experimentación y otros fines científicos, incluyendo la docencia.

El material animal utilizado en este trabajo proviene de dos proyectos FITE (Fondo de Inversiones de Teruel), ASCYTEQ y LACTOCYNARA, y de un convenio con el INIA y la Asociación Turolense de Productores de Leche y Queso. La Asociación Turolense de Productores de Leche y Queso lleva trabajando desde 2006 en la obtención de una figura de calidad diferenciada del Queso de Teruel (registrada actualmente como marca colectiva).

Tabla 3. Distribución de las explotaciones ovinas de estudio.

Explotaciones	Calidad y RCS	Producción de leche
A	574	574
B	893	--
C	427	427
Total de animales	1.894	1.001

La población sobre la que se realizó el presente estudio está conformada por animales de la raza Assaf pertenecientes a la citada asociación, cuya edad se encuentra entre dos y siete años, con al menos dos o más lactaciones y tres controles lecheros durante una lactación. Esta población está compuesta por tres explotaciones en las que se llevó a cabo el control lechero registrándose el porcentaje de proteína, grasa, lactosa, extracto seco, así como el RCS, y la producción de leche (en ml). Para los caracteres de calidad y RCS, se anotaron 6.118 registros de 1.894 animales en las tres explotaciones, mientras que para la producción de leche, fueron 2.697 registros de 1.001 animales, ya que una explotación no recogía datos de producción de leche. La distribución de las tres explotaciones se muestra en la *Tabla 3*.

Los diferentes análisis se realizaron con los datos fenotípicos brutos obtenidos del control lechero, excepto los del RCS. Los datos del carácter RCS se normalizaron mediante la siguiente ecuación: $SCS = \log_2 (SCC/100) + 3$ (Aka y Shook, 1980).

3.1.1 Selección de animales genotipados con el chip ovino “Illumina AgResearch Sheep HD” de 680K.

En la selección de animales para genotipar con el chip ovino “Illumina AgResearch Sheep HD” de 680K se estimó el valor genético animal en las 1.894 hembras una vez eliminados los efectos fijos y covariables (ver apartado 3.4.1). Esta valoración fue realizada para todos los caracteres de calidad, RCS y producción de leche (en este caso 1.001 hembras), y fue utilizado como fenotipo corregido en los estudios de GWAS. Finalmente, se seleccionaron 192 animales para los análisis de GWAS en función del fenotipo de RCS, ya que el objetivo principal de los proyectos en el cual está integrado este trabajo era la detección de marcadores relacionados con la resistencia a la mastitis. Así que se eligieron 96 animales con alto valor genético animal estimado para el RCS (H-RCS), y 96 con bajo valor (L-RCS). Las hembras de los grupos H-RCS y L-RCS pertenecían al 10% de las hembras con mayor valor genético estimado y al 14% de las hembras con menor valor estimado del total de la distribución de valores estimados para el RCS. Para intentar evitar el efecto de la estratificación en los análisis de GWAS se eligieron 64 hembras de cada explotación (n = 32 del grupo H-RCS; y n = 32 del grupo L-RCS) lo menos relacionadas posibles en función de los datos de pedigrí obtenidos en base a los registros de la explotación y test de paternidad con microsatélites. De esta manera se descartaron hermanas y medio-hermanas del análisis que tuviesen similar valor estimado en el RCS.

Para el resto de los caracteres se utilizaron los mismos animales elegidos para el RCS. Sin embargo, en estos casos (grasa, proteína, lactosa, extracto seco y producción de leche) los animales representaban todos los valores de la distribución de valores genéticos estimados para el carácter, y no sólo los extremos como en el caso del RCS.

3.1.2 Animales utilizados para la validación de los resultados GWAS

Para la validación de algunos de los SNPs que resultaron significativamente asociados con los caracteres estudiados, se utilizó la población total de 1.894 hembras, aunque 55 no se pudieron analizar por falta de ADN y sangre. Todas las muestras de ADN estaban ya extraídas, excepto 45 muestras que se extrajeron de nuevo de sangre mediante el kit comercial “Kit de Extracción de ADN Genómico en Sangre” (FavorPrep, Ibian, Zaragoza, España).

3.2 Análisis de asociación de genoma completo (GWAS)

Para el análisis GWAS se utilizó el chip ovino de SNPs de alta densidad 680K (Illumina AgResearch Sheep HD de 680K) diseñado por el Consorcio Internacional del Genoma ovino (Kijas et al., 2014), obteniendo los genotipos para 192 hembras de raza Assaf. Este chip tiene aproximadamente 680000 SNPs que están distribuidos de manera más o menos uniforme por todo el genoma. Todas las muestras de ADN se encontraban ya extraídas.

3.3 Genotipado mediante KASP

Para la validación de los resultados de GWAS se ha llevado a cabo el genotipado de 2 SNPs localizados en los cromosomas 19 (para el RCS), 13 (producción de leche), y 22 (porcentaje de proteína) (Tabla 4).

Tabla 4. Fenotipos y SNPs asociados.

Fenotipo ¹	dbSNPs ²	SNP en chip Agresearch de 680 k	Crs ³	Localización	Posición genoma Oar 3.1	Cambio nucleotídico
RCS	rs419096188	s36141.1	19	Intón 1 NUP210	Oar19: g.58334807	G → A
	rs424642424	oar3_OAR19_9410968	19	Espacio intergénico	Oar23: g.7432390	C → T
PL	rs411200126	oar3_OAR13_28464329	13	Espacio intergénico	Oar13: g.28464329	A → G
PP	rs419686662	oar3_OAR22_40605374	22	Espacio intergénico	Oar22: g.7375434	G → A

¹Fenotipo: RCS= Recuento de células somáticas; PL= Producción de Leche; PP= Porcentaje de Proteína. ²dbSNPs: Identificador en la base de datos de SNPs. Crs³: Cromosoma Ovino.

Ninguna de las mutaciones analizadas producía cambio aminoacídico.

El genotipado se ha llevado a cabo mediante discriminación alélica utilizando la técnica KASP (Kompetitive allele specific PCR), que es una PCR competitiva específica de alelo que se lleva a cabo en un termociclador a tiempo real, y permite la identificación de polimorfismos de un solo nucleótido (SNP) e inserciones y deleciones (Indels) en loci específicos. Para realizar esta técnica de genotipado se necesitan dos cebadores directos específicos (cada uno marcado con una secuencia marcadora, complementaria a la secuencia marcadora específica de dos casetes FRET, en función del alelo a detectar y diferenciados sólo en un nucleótido que corresponde a cada uno

de los alelos del polimorfismo a estudiar), y uno reverso común para los dos alelos de la secuencia, para amplificar la región que contiene el polimorfismo. Estos casetes FRET disponen de un fluorocromo alelo específico en el extremo 5', y de una molécula inhibidora de fluorescencia o "quencher". Cuando el cebador hibrida con el ADN durante la PCR, la secuencia marcadora específica se habrá unido con su casete FRET correspondiente (hibridan por complementariedad) que tendrá el fluorocromo correspondiente, incorporándose en el producto amplificado, separándose irreversiblemente del "quencher" en el casete FRET, y emitiendo la fluorescencia correspondiente. Los fluorocromos más comúnmente utilizados son HEX (emite a una longitud de onda de 556 nm) y FAM (emite a una longitud de onda de 510 nm).

En resumen, la técnica se basa en la extensión de cebador específico de alelo, y la liberación de una fluorescencia al hibridar los cebadores, que será posteriormente detectada por el sistema óptico del termociclador a tiempo real. En este trabajo, la discriminación alélica mediante la tecnología KASP se llevó a cabo en el termociclador a tiempo real Bio-Rad CFX96 (BioRad, Madrid, España).

El diseño de cebadores fue llevado a cabo por LGC Genomics, proveedor de los ensayos KASP, mediante el envío de las secuencias flanqueantes al SNP (LCG Genomics, Beverly, MA, USA). La empresa proveedora de los cebadores envía los ensayos premezclados para cada SNP, sin detallar la secuencia de los cebadores. La amplificación se realizó en un volumen final de 10 µl conteniendo 1 µl de DNA (50-100 ng), 5 µl de la mezcla de reacción KASP V4.0 2x con ROX estándar (LCG Genomics, Beverly, MA, USA), 0,14 µl de la mezcla de cebadores optimizada para la discriminación alélica (KASP-by-Design assay mix; LGC Genomics, Beverly, MA, USA), y el volumen de H₂O ultrapura necesario para alcanzar los 10 µl. La mezcla de reacción contiene los nucleótidos, los casetes FRET con la secuencia marcadora alelo específica complementaria a la del cebador directo para cada alelo (marcados con los fluorocromos FAM y HEX dependiendo del alelo), así como la enzima taq polimerasa, y el tampón de la reacción. Las condiciones de amplificación fueron: una fase de activación a 94°C durante 15 minutos; 10 ciclos de desnaturalización a 94°C durante 20 segundos de hibridación y elongación a 57 y 61°C para el exón 2 y 3, respectivamente (con una bajada de 0,6 °C por ciclo para el exón 3); 20 ciclos de desnaturalización durante 20 segundos a 94°C e hibridación y elongación a 55°C durante 1 minuto; y finalmente un ciclo a 30°C durante 1 minuto. Tras este ciclo se hace una lectura de la fluorescencia a tiempo final, para poder discriminar el genotipo. El software del PCR a tiempo real es capaz de discriminar los genotipos en función de las unidades de fluorescencia relativas para cada uno de los dos fluorocromos. En todos los ensayos KASP se incluyeron controles negativos

(sin ADN), y controles positivos de animales que conocíamos su genotipo previamente mediante secuenciación SANGER.

3.4 Análisis estadístico y bioinformático

3.4.1 Estimación del efecto animal para cada fenotipo

La estimación del efecto animal (o valor genético animal) fue realizada con el programa informático BLUPf90 (Miszta et al., 2018), que permite la obtención de una estima del valor genético animal para un carácter determinado, considerando los efectos fijos y covariables que pueden incidir sobre el carácter. Este programa utiliza el modelo mixto en los análisis de valoración.

En este estudio, han sido considerados como efectos fijos el número de corderos nacidos (tipo de parto), número de parto, y el efecto rebaño – mes – año de cuando se producen los partos; como efectos covariables la edad de la oveja y los días en lactación, y como efecto aleatorio el animal.

3.4.2 Control de calidad de los datos masivos procedentes del chip de SNPs de alta densidad 680K (Illumina AgResearch Sheep HD) y estudio de la estratificación de los datos

Los datos brutos de genotipado fueron depurados mediante un control de calidad a través del programa PLINK 1.9 (Chang et al., 2015) eliminando aquellos SNPs que no cumplieran las siguientes condiciones: SNPs con una frecuencia del alelo menos frecuente (MAF) $< 0,01$; SNPs que no estuviesen en equilibrio de Hardy – Weinberg ($P < 0.0001$); SNPs con una frecuencia de individuos genotipados $< 0,90$; SNPs sin localización cromosómica en el genoma ovino; y SNPs duplicados.

Igualmente, con el software PLINK 1.9 se realizó un estudio para conocer la posible estructuración de la población mediante escalado multidimensional (MDS), considerando LD entre SNPs menor que 0,2 ($r^2 < 0,2$) y ventanas de 50 SNPs. A partir de aquí se obtuvieron los coeficientes de parentesco IBS (Idéntico por estado) entre pares de individuos, y se representó gráficamente el componente 1 frente al componente 2 en función de estos coeficientes.

3.4.3 Análisis de asociación del genoma completo (GWAS)

El GWAS se realizó con el software GCTA (Genome – wide Complex Trait Analysis) (Yang et al., 2011) mediante un modelo lineal mixto (MLMA) y con el MLMA - LOCO (“leaving - one - chromosome - out”; en el cuál dejan un cromosoma fuera del análisis al azar). Este programa

permite estimar una matriz de correlaciones genómicas que permite tener en cuenta posibles problemas de estructuración y/o de relación entre muestras. Se aplicó el siguiente modelo:

“ $y_j = u + \text{SNP}_i + g_j + e_j$ ”; donde y_j es el valor estimado para cada fenotipo y cada animal genotipado; u es la media; SNP es el efecto de cada SNP (asumiendo valores de 0, 1 o 2, referido a los genotipos aa, Aa y AA); g_j es el efecto aditivo aleatorio y e_j es el error residual. Finalmente, los SNPs significativos fueron elegidos en base al valor P obtenido, utilizando la corrección de FDR (False Discovery Rate) para comparaciones múltiples tanto a nivel genómico, como a nivel cromosómico al 10%. La visualización gráfica de los estudios GWAS se llevó a cabo mediante los gráficos Manhattan y Q-Q usando el programa SNPEVG (Wang et al., 2012).

La anotación de genes en las regiones significativas (SNPs significativos para el GWAS) se realizó utilizando la versión 3.1 del genoma ovino, considerando un intervalo de 0,50 Mb, y situando el SNP significativo en la posición media del intervalo. Existen varias herramientas para la anotación de genes, sin embargo, en este estudio al tratarse de genoma versión 3.1, se ha seleccionado el servidor “ensembl” indicando en el motor de búsqueda el SNP significativo.

3.4.4 Análisis de asociación de SNPs candidatos

Como se ha descrito en el apartado 3.3, la validación de los resultados de GWAS se llevó a cabo mediante el genotipado de 2 SNPs relacionados con el RCS (rs419096188 y rs424642424), 1 con la producción de leche (rs411200126), y otro con el porcentaje de proteína (rs419686662) en 1.839 animales. El equilibrio de Hardy–Weinberg, heterocigosidad observada y frecuencia del alelo menor (MAF) para cada SNP se calculó usando el programa PLINK 1.9. Los estudios de correlación entre fenotipos se llevaron a cabo mediante el cálculo del coeficiente de correlación de Pearson con el software SAS (Versión 9.3; SAS Inst. Inc., Cary, NC). Los estudios de asociación entre los SNPs (rs419096188, rs424642424, rs411200126 y rs419686662) y los fenotipos estudiados (RCS, porcentaje de proteína, grasa, lactosa, extracto seco, y producción de leche) se llevaron a cabo mediante un modelo lineal mixto (MIXED) para medidas repetidas con el software SAS. El modelo incluyó el efecto rebaño-año-mes de parto (RAM; donde los niveles existentes son iguales a la combinación del número de rebaños en control, número de años que se consideran y los meses donde se producen los partos), número de corderos nacidos (CN), número de lactación (NL), y el genotipo del SNP (G; asumiendo valores de 0, 1 o 2, referido a los genotipos aa, Aa y AA) como efectos fijos; mientras que los días en lactación (DEL) y la edad (E) fueron considerados como covariables; y como efectos aleatorios el animal (A) del cual se toman los datos

y el residuo (e). La variable dependiente (y) fueron los datos del fenotipo bruto para cada una de las tomas para cada animal. La ecuación del modelo fue la siguiente:

$$y = NC + RAM + NL + G + b (DEL) + b (E) + A + e$$

Finalmente, se estimaron las medias mínimo cuadráticas (LSMeans) y error estándar de cada genotipo y SNP para cada fenotipo estudiado. Se utilizó el test de Bonferroni de comparaciones múltiples para comparar entre las medias mínimo cuadráticas. Todos los polimorfismos se analizaron independientemente con el mismo modelo estadístico.

IV. Resultados y discusión

4.1 Análisis de Asociación del Genoma Completo (GWAS)

4.1.1 Fenotipos

En las siguientes tablas se pueden observar la media y la desviación estándar, así como el número de datos utilizados para cada uno de los fenotipos analizados, así como la correlación obtenida entre los mismos (*Tabla 5* y *Tabla 6*).

Tabla 5. Número de datos utilizados para los análisis, media y desviación estándar de los mismos.

Fenotipo	N	Media	Desviación estándar
RCS ($\log_2(\text{SCC}/100,000)+3$)	6118	2,85	1,35
Grasa (%)	6118	5,69	1,53
Proteína (%)	6118	5,36	0,68
ES (%)	6118	11,11	0,62
Lactosa (%)	6118	4,83	0,30
Leche (ml)	2697	2267,00	949,73

Tabla 6. Coeficientes de correlación Pearson obtenidos entre los diferentes caracteres en la población estudiada.

	Grasa	Proteína	ES	Lactosa	Producción de leche
RCS	0,20	0,19	-0,08	-0,46	-0,19
Grasa	-	0,66	0,45	-0,36	-0,49
Proteína	-	-	0,83	-0,29	-0,37
ES	-	-	-	0,28	-0,17
Lactosa	-	-	-	-	0,39

Las correlaciones obtenidas son similares a las encontradas en otras poblaciones, por ejemplo, la correlación negativa entre la producción de la leche y el contenido de grasa y proteína. Es destacable, la alta correlación negativa encontrada entre el RCS y la lactosa, ya que la mayor parte de la leche de estas explotaciones se transforma en queso. La mayoría de los quesos de la asociación de Productores de quesos y leche de Teruel presentan una coagulación mixta, es decir enzimática y acidoláctica. La coagulación acidoláctica consiste en la acidificación natural por las bacterias lácticas, es decir, microorganismos que actúan sobre la lactosa y la transforman en ácido láctico. De esta forma una selección contra el RCS (mayor calidad higiénico sanitaria) podría producir un incremento de la lactosa en la leche, y por lo tanto incrementar la coagulación acidoláctica.

A partir de estos datos fenotípicos se estimó el valor genético animal para cada carácter, que fue el utilizado en el análisis GWAS (ver apartado 3.4.1).

4.1.2 Control de calidad y estudio de la estratificación de la población utilizada para llevar a cabo el GWAS

Inicialmente, tras el genotipado con el chip ovino “Illumina AgResearch Sheep HD” de 680K, se obtuvieron un total de 606.006 marcadores SNPs válidos, que tras el control de calidad se redujeron a 560.337 SNPs distribuidos por los cromosomas autosómicos y el cromosoma X, y estos SNPs se utilizaron en los estudios de asociación.

Con los SNPs que presentaron un LD menor que 0,2, se obtuvieron los coeficientes de parentesco IBS (Idéntico por estado) entre pares de individuos, y se realizó un análisis de escalado multidimensional (MDS) para conocer el grado de estructuración de la población. En total se utilizaron 86.146 SNPs, que se consideraron informativos para este análisis. En la *Figura 1* se puede observar la distribución de los animales en función de las tres explotaciones utilizadas, presentando cierta estructuración dependiente de la explotación de donde proviene la muestra, estando la población del rebaño B algo más diferenciada de las otras dos. El análisis GWAS utilizado con el programa GCTA considera una matriz de parentesco molecular entre los individuos de manera que se corrige la estructuración de la población en los análisis.

4.1.2.1 GWAS producción de leche (PL)

Los análisis mediante MLMA y MLMA - LOCO dieron los mismos resultados en cuanto a significación de SNPs. No hubo SNPs significativos a nivel genómico, sin embargo, 2 SNPs fueron significativos a nivel cromosómico tras corrección con FDR (en OAR13): oar3_OAR13_28464329 y oar3_OAR13_28473530. Estos SNPs estuvieron completamente ligados, es decir, presentaron los mismos genotipos para cada animal, con un MAF de 0,11, y un efecto negativo del alelo con menor frecuencia respecto a la producción láctea ($b = -1538,83$) (*Tabla 7*).

Tabla 7. SNPs significativos a nivel cromosómico tras corrección FDR al 10% para el carácter recuento de células somáticas (RCS) tras el análisis GWAS con el programa GCTA (procedimiento MLMA y MLMA-LOCO).

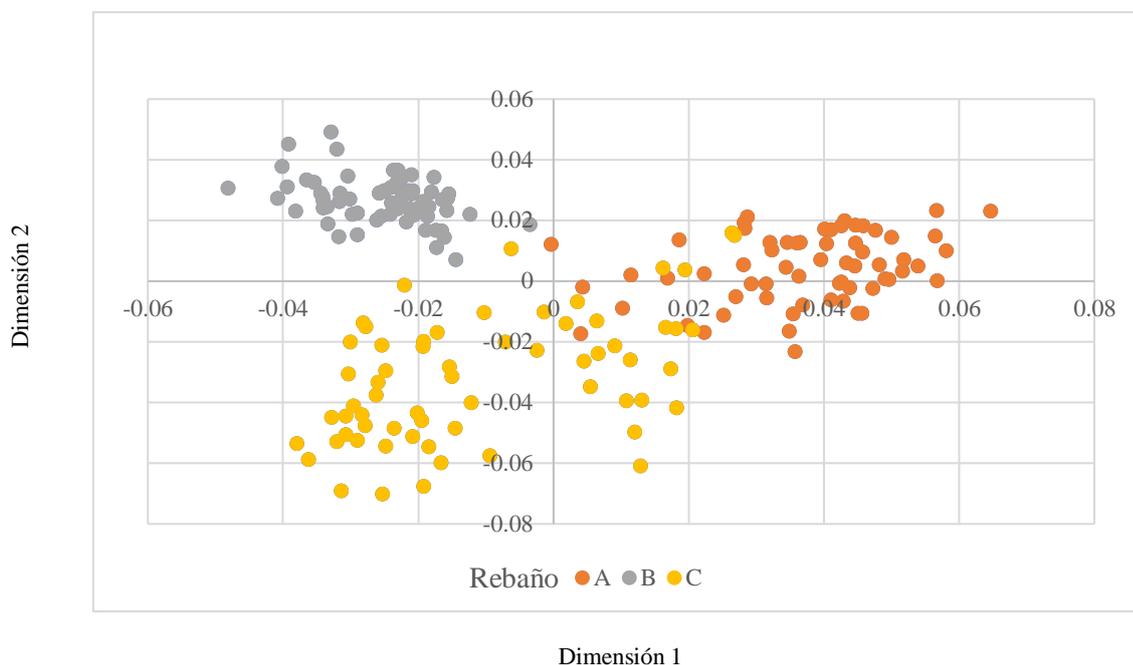
Fenotipo ¹	Crs	SNP en el chip Agresearch de 680 k	Posición en Oar 3.1	A1	A2	MAF (A2)	b	ES	P	FDR 10% ²	Modelo
PL	13	oar3_OAR13_28464329	Oar13: g.28464329	A	G	0,12	-1538,83	334,3	4,16 x 10 ⁻⁶	1,10 x 10 ⁻⁵	MLMA/LOCO
	13	oar3_OAR13_28473530	Oar13: g.28473530	G	A	0,12	-1538,83	334,3	4,16 x 10 ⁻⁶	1,10 x 10 ⁻⁵	MLMA/LOCO
PP	6	oar3_OAR6_5285000	Oar6: g.5285000	G	A	0,07	0,31	0,06	1,73 x 10 ⁻⁶	4 x 10 ⁻⁶	MLMA/LOCO
	17	oar3_OAR17_25027692	Oar17: g.25027692	A	C	0,40	0,14	0,03	6,43 x 10 ⁻⁶	4 x 10 ⁻⁶	LOCO
	17	oar3_OAR17_12384078	Oar17: g.12384078	C	A	0,38	0,15	0,03	5,50 x 10 ⁻⁶	6,5 x 10 ⁻⁶	MLMA/LOCO
	22	oar3_OAR22_40605374	Oar22: g.40605374	G	A	0,36	-0,14	0,03	1,06 x 10 ⁻⁵	2,7 x 10 ⁻⁵	MLMA/LOCO
	22	s37672.1	Oar22: g.40605508	A	G	0,36	-0,14	0,03	1,06 x 10 ⁻⁵	2,7 x 10 ⁻⁵	MLMA/LOCO
	22	oar3_OAR22_40606216	Oar22: g.40606216	A	G	0,36	-0,14	0,03	1,06 x 10 ⁻⁵	2,7 x 10 ⁻⁵	MLMA/LOCO
PG	14	oar3_OAR14_39364677	Oar14: g.39364677	A	G	0,21	-0,37	0,09	1,33 x 10 ⁻⁵	2,8 x 10 ⁻⁵	LOCO
	14	s32350.1	Oar14: g.39664630	A	G	0,28	0,32	0,08	2,55 x 10 ⁻⁵	2,8 x 10 ⁻⁵	LOCO
	14	oar3_OAR14_39913998	Oar14: g.39913998	G	A	0,40	-0,28	0,07	2,77 x 10 ⁻⁵	2,8 x 10 ⁻⁵	LOCO
CL	2	oar3_OAR2_49255325	Oar2: g.49255325	G	A	0,01	-0,52	0,11	1,82 x 10 ⁻⁶	3,7 x 10 ⁻⁶	MLMA/LOCO
	2	oar3_OAR2_48135063	Oar2: g.48135063	G	A	0,02	-0,44	0,09	2,07 x 10 ⁻⁶	3,7 x 10 ⁻⁶	MLMA/LOCO
	3	oar3_OAR3_64750845	Oar3: g.64750845	A	G	0,15	-0,17	0,04	4,58 x 10 ⁻⁶	6,1 x 10 ⁻⁶	MLMA/LOCO
	3	oar3_OAR3_64741679	Oar3: g.64741679	A	G	0,15	-0,17	0,04	4,64 x 10 ⁻⁶	6,1 x 10 ⁻⁶	MLMA/LOCO
	3	oar3_OAR3_73955311	Oar3: g.73955311	A	G	0,02	-0,45	0,10	5,96 x 10 ⁻⁶	6,1 x 10 ⁻⁶	MLMA/LOCO
	5	oar3_OAR5_49233977	Oar5: g.49233977	G	A	0,01	-0,54	0,12	9,57 x 10 ⁻⁶	1,1 x 10 ⁻⁵	MLMA/LOCO
	5	oar3_OAR5_49200052	Oar5: g.49200052	C	A	0,01	-0,53	0,12	1,04 x 10 ⁻⁵	1,1 x 10 ⁻⁵	MLMA/LOCO
	8	oar3_OAR8_84295105	Oar8: g.84295105	G	A	0,02	-0,42	0,09	6,19 x 10 ⁻⁶	1 x 10 ⁻⁵	MLMA/LOCO
	8	oar3_OAR8_77590661	Oar8: g.77590661	G	A	0,03	-0,30	0,07	6,72 x 10 ⁻⁶	1 x 10 ⁻⁵	MLMA/LOCO
	13	oar3_OAR13_67179806	Oar13: g.67179806	G	A	0,01	-0,63	0,12	2,11 x 10 ⁻⁷	5,5 x 10 ⁻⁶	MLMA/LOCO
	16	oar3_OAR16_24178409	Oar16: g.24178409	G	A	0,02	-0,37	0,07	3,78 x 10 ⁻⁷	1,9 x 10 ⁻⁵	MLMA/LOCO

	16	oar3_OAR16_14179399	Oar16: g.14179399	A	G	0,01	-0,52	0,11	$1,34 \times 10^{-6}$	$1,9 \times 10^{-5}$	MLMA/LOCO
	16	oar3_OAR16_14186759	Oar16: g.14186759	A	G	0,01	-0,52	0,11	$1,34 \times 10^{-6}$	$1,9 \times 10^{-5}$	MLMA/LOCO
	20	oar3_OAR20_13106461	Oar20: g.13106461	G	A	0,02	-0,43	0,10	$1,39 \times 10^{-5}$	$1,8 \times 10^{-5}$	MLMA/LOCO
	20	oar3_OAR20_11231797	Oar20: g.11231797	C	A	0,02	-0,40	0,09	$1,64 \times 10^{-5}$	$1,8 \times 10^{-5}$	MLMA/LOCO
ES	11	oar3_OAR11_44970125	Oar11: g.44970125	G	A	0,16	-0,21	0,05	$1,74 \times 10^{-5}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_22502645	Oar11: g.22502645	G	C	0,14	-0,23	0,06	$2,88 \times 10^{-5}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31971548	Oar11: g.31971548	A	G	0,13	-0,22	0,05	$2,94 \times 10^{-5}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_42884861	Oar11: g.42884861	A	G	0,22	-0,18	0,04	$5,54 \times 10^{-5}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31716346	Oar11: g.31716346	G	A	0,20	-0,18	0,05	$9,14 \times 10^{-5}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31732260	Oar11: g.31732260	G	A	0,29	-0,15	0,04	$1,01 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_34077455	Oar11: g.34077455	A	G	0,12	-0,20	0,05	$1,11 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_34085781	Oar11: g.34085781	A	G	0,12	-0,20	0,05	$1,11 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_34102262	Oar11: g.34102262	A	C	0,12	-0,20	0,05	$1,11 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31956799	Oar11: g.31956799	A	G	0,48	-0,15	0,04	$1,15 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_32608494	Oar11: g.32608494	G	A	0,17	-0,18	0,05	$1,27 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_37800259	Oar11: g.37800259	C	A	0,12	-0,22	0,06	$1,38 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31708633	Oar11: g.31708633	G	A	0,16	-0,19	0,05	$1,53 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31748081	Oar11: g.31748081	A	G	0,29	-0,16	0,04	$1,67 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31725295	Oar11: g.31725295	A	G	0,20	-0,17	0,05	$1,79 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31730623	Oar11: g.31730623	A	G	0,20	-0,17	0,05	$1,81 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31954402	Oar11: g.31954402	G	A	0,48	-0,15	0,04	$1,93 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
	11	oar3_OAR11_31970980	Oar11: g.31970980	G	A	0,48	-0,15	0,04	$1,93 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO
11	oar3_OAR11_31744746	Oar11: g.31744746	A	G	0,36	-0,14	0,04	$1,98 \times 10^{-4}$	$2,1 \times 10^{-4}$	LOCO	

	11	oar3_OAR11_31744803	Oar11: g.31744803	G	A	0,36	-0,14	0,04	1,98 x 10 ⁻⁴	2,1 x 10 ⁻⁴	LOCO
	11	oar3_OAR11_34067161	Oar11: g.34067161	A	C	0,12	-0,19	0,05	2,00 x 10 ⁻⁴	2,1 x 10 ⁻⁴	LOCO
	16	oar3_OAR16_14179399	Oar16: g.14179399	A	G	0,01	-0,73	0,16	4,43 x 10 ⁻⁶	1,3 x 10 ⁻⁵	MLMA/LOCO
	16	oar3_OAR16_14186759	Oar16: g.14186759	A	G	0,01	-0,73	0,16	4,43 x 10 ⁻⁶	1,3 x 10 ⁻⁵	MLMA/LOCO
	20	oar3_OAR20_32651642	Oar20: g.32651642	A	G	0,07	-0,32	0,08	1,72 x 10 ⁻⁵	2,5 x 10 ⁻⁵	MLMA/LOCO
	20	oar3_OAR20_46839037	Oar20: g.46839037	A	G	0,04	-0,34	0,08	2,47 x 10 ⁻⁵	2,5 x 10 ⁻⁵	MLMA/LOCO
RCS	19	s36141.1	Oar19: g.58334807	A	G	0,38	-0,61	0,15	2,91 x 10 ⁻⁵	4 x 10 ⁻⁵	MLMA/LOCO
	19	oar3_OAR19_9401922	Oar19: g.9401922	C	A	0,26	0,71	0,17	3,98 x 10 ⁻⁵	4 x 10 ⁻⁵	MLMA/LOCO
	19	oar3_OAR19_9407458	Oar19: g.9407458	G	A	0,26	0,71	0,17	3,98 x 10 ⁻⁵	4 x 10 ⁻⁵	MLMA/LOCO
	19	oar3_OAR19_9410968	Oar19: g.9410968	A	G	0,26	0,71	0,17	3,98 x 10 ⁻⁵	4 x 10 ⁻⁵	MLMA/LOCO

¹ Fenotipo: PL= producción de leche, PP= porcentaje de proteína, PG=porcentaje de grasa, P= porcentaje de lactosa, ES= porcentaje extracto seco, RCS= Recuento de células somáticas; Crs=cromosoma; A1= alelo1; A2= alelo 2; b= efecto de sustitución del alelo A2; ES= error estándar; FDR 10%= Límite de significación estadística para un FDR del 10% para el cromosoma y el carácter analizado; modelo: MLMA= SNP significativo con el modelo MLMA, y LOCO= SNP significativo con el modelo LOCO-MLMA.

Figura 1. Análisis de escalado multidimensional (Multidimensional scaling plot, MDS) de los 192 animales utilizados para el GWAS.



Como se explica en material y métodos, la anotación de genes se realizó utilizando la versión 3.1 del genoma ovino, y considerando un intervalo de 0,5 Mb. En este sentido, los genes anotados en esta región del cromosoma 13 fueron *FAM107B*, *CDNF*, *HSPA14*, *SNORD22*, *SUV39H2*, *DCLRE1C*, *MEIG1*, *OLAH*, *ACBD7*, *RPP38* y *NMT2* (Tabla suplementaria 1). De todos ellos, según la bibliografía consultada, los genes *CDNF*, *HSPA14* y *SUV39H2* podrían estar relacionados funcionalmente con el carácter PL en el ganado lechero.

Respecto al gen *CDNF*, es un factor neutrófico de dopamina cerebral. El control negativo de la dopamina desencadena la producción de prolactina (PRL), que contribuye en la diferenciación de la glándula mamaria durante la gestación. Además, se ha demostrado que en ganado bovino tiene efecto galactopoyético, es decir, produce una estimulación de la producción y eyección de la leche. Por otra parte, en un estudio se administró domperidona, antagonista de la PRL, resultando en un aumento de la concentración de PRL y el posterior incremento de la producción láctea (Lacasse y Ollier, 2015). Por otro lado, el gen *HSPA14* es miembro de la familia de proteínas de estrés térmico A (Hsp70). Entre sus vías metabólicas se encuentra la respuesta celular al estrés por calor, y se ha visto que en vacas lecheras este estrés produce efectos negativos como reducción de la ingesta de alimento, fertilidad comprometida y menor producción de leche (Nguyen et al., 2016). Finalmente, en cuanto al gen *SUV39H2*, entre sus vías metabólicas se encuentra la degradación de la lisina, un

aminoácido que junto con la metionina son limitantes en el proceso de síntesis lechera (Robinson et al., 2011).

En este cromosoma se han detectado SNPs/ genes asociados a la producción de leche (Gerlando et al., 2019, Usai et al., 2019), sin embargo, ninguno se localiza cerca de los SNPs significativos encontrados en el presente trabajo, de manera que los efectos detectados son en zonas del genoma diferentes a las descritas con anterioridad.

4.1.2.2 GWAS porcentaje de proteína (PP)

Los análisis mediante MLMA y MLMA - LOCO no mostraron SNPs significativos a nivel genómico. A nivel cromosómico, se encontraron 5 SNPs significativos para el análisis MLMA (1, 1 y 3 SNPs para los cromosomas OAR6, 17 y 22, respectivamente) (*Tabla 7*). Además de estos 5 SNPs, el análisis con la metodología MLMA – LOCO identificó otro SNP localizado en OAR17 (oar3_OAR17_12384078) (*Tabla 7*). Los 3 SNPs localizados en OAR22 estaban completamente ligados. La frecuencia del alelo menos frecuente fue alta para todos los SNPs excepto para el localizado en el cromosoma OAR6, que además presentó un efecto de sustitución mayor.

En el intervalo del SNP significativo localizado en OAR6, no aparecieron genes anotados, sin embargo, se encontró a 326kb del SNP el gen *MAD2L1* (*Tabla suplementaria 1*). Este gen participa en la síntesis proteica, y se ha relacionado con el rendimiento proteico de la leche (Suchocki et al., 2016). OAR6 ha sido uno de los cromosomas más estudiados debido a la presencia del cluster de las *caseínas* en el mismo (*CSNIS1*, *CSNIS2*, *CSN2* y *CSN3*). Sin embargo, este cluster se encuentra muy alejado en el genoma ovino del SNP significativo y, por lo tanto, no es esperable que el efecto encontrado sea debido a la existencia de LD con este cluster.

En la región cromosómica del OAR17 donde se encuentra el SNP significativo, se anotaron 6 genes (*Tabla suplementaria 1*). Destaca el gen *PCDH10* que presenta una relación indirecta con los caracteres de leche y no con el PP. Se ha visto que las mutaciones en este gen se asocian con las encefalopatías en el ganado bovino, dando lugar a síntomas entre los que destaca la disminución de la producción láctea (Biacabe et al., 2007). Finalmente, en cuanto a OAR22, se anotaron 5 genes (*Tabla suplementaria 1*). De estos genes, se debe destacar *FGFR2*, que codifica para el receptor de factores de crecimiento fibroblásticos (FGFs), y diversos autores han mostrado su influencia en el desarrollo y posterior funcionalidad de la glándula mamaria (Pond et al., 2013; Zhang et al., 2014; Sumbal et al., 2019). Aunque en este cromosoma se han detectado SNPs/ genes asociados al carácter PP (García- Gámez et al., 2013; Usai et al., 2019), ninguno se localiza cerca de los

SNPs significativos, de manera que los efectos detectados son en zonas del genoma diferentes a las descritas con anterioridad.

4.1.2.3 GWAS porcentaje de grasa (PG)

Los análisis mediante MLMA y MLMA - LOCO no mostraron SNPs significativos a nivel genómico. A nivel cromosómico, se encontraron 3 SNPs significativos pero sólo para el modelo MLMA-LOCO en el cromosoma OAR14 (oar3_OAR14_39364677, s32350.1 y oar3_OAR14_39913998) (Tabla 7). Los genes anotados en esta zona del cromosoma 14 fueron 8 (Tabla suplementaria 1), sin embargo, ninguno presentó funciones, o se encontró en la bibliografía información que indicase su posible asociación con el carácter PG.

4.1.2.4 GWAS porcentaje/contenido de lactosa (CL)

Los análisis mediante MLMA y MLMA - LOCO dieron los mismos resultados en cuanto a la significación de los diferentes SNPs. No hubo SNPs significativos a nivel genómico, sin embargo 15 SNPs fueron significativos a nivel cromosómico tras corrección con FDR (Tabla 7). Es destacable que todos los SNPs presentaron un MAF muy bajo, con excepción de los localizados en el cromosoma OAR3, que estuvieron completamente ligados, al igual que los del OAR16. Estos SNPs con un MAF tan bajo podrían ser asociaciones espurias, ya que con estas frecuencias podrían indicar que, de los 192 animales, 4 animales serían heterocigotos y el resto homocigoto para uno de los alelos, y solamente por azar ser significativos. Este es el caso del SNP oar3_OAR2_49255325.

En OAR13 se anotaron 7 genes (Tabla suplementaria 1). El gen *RALGAPB* codifica para una proteína que participa en la translocación de GLUT4 a la membrana plasmática, y se ha descrito su asociación con la cantidad de lactosa en leche de cabras, de forma que cuando se sobre expresa, se produce un aumento de la síntesis de lactosa (Zhu et al., 2014). Otro gen interesante es el *SLC32A1*, que codifica para proteína de membrana involucrada en la captación de ácido gamma-aminobutírico (GABA) que se ha relacionado con la PRL en ratas (Racagni et al., 1984). Además, otros autores, indican que el GABA puede estimular la liberación de PRL en dosis elevadas, mientras que en dosis bajas la inhiben (McCann et al., 1984). Este gen podría estar relacionado funcionalmente con el PL debido a su interacción con la PRL, ya que esta tiene efecto estimulante sobre la producción y eyección láctea, así como con la síntesis de lactosa mediante la estimulación del complejo enzimático lactosa sintetasa (McNeilly, 1987). Finalmente, el gen *PPP1R16B* en el

ganado bovino de raza Holstein se ha relacionado con la producción de leche, y con el rendimiento y porcentaje de grasa y proteína láctea (Jiang et al, 2019).

Los genes detectados en el intervalo del cromosoma OAR16 son 13 (*Tabla suplementaria 1*), y entre ellos cabe mencionar el *ESM1*, que codifica para una proteína que participa en la unión al receptor del factor de crecimiento de los hepatocitos, y que podría estar relacionado con la síntesis de glucosa, necesaria para la síntesis lactosa en la glándula mamaria (Payne et al., 1970). Otro gen anotado en este intervalo es el *CENPK*, que, aunque no se ha asociado al PL, es un componente del complejo CENPA-CAD que interactúa con el complejo CENPA-NAC, ambos asociados con el porcentaje proteico de la leche, y con el rendimiento de grasa láctea en el ganado bovino (según CattleQTLdb), además de estar asociado con la producción de leche (Jiang et al, 2019).

En la región del cromosoma 2 se encuentran 10 genes anotados (*Tabla suplementaria 1*), entre ellos el *FOXE1*, codifica para una proteína que actúa como un factor de transcripción, jugando un papel primordial en la morfogénesis tiroidea. Mutaciones en este gen producen hipotiroidismo, y se ha demostrado que las hormonas tiroideas intervienen en la producción láctea del ganado lechero. Concretamente se ha visto que la administración de tiroxina (hormona T4 tiroidea) en vacas lecheras, aumenta la producción de lactosa en un 25% (Davis et al., 1988). En cuanto al cromosoma OAR3, solamente se ha anotado el gen *NRXN1A*; en OAR8 los genes *PRKN*, *OPRM1* y *CNKSR3*; y en OAR5 16 genes (*Tabla suplementaria 1*), aunque de todos ellos no se ha encontrado una relación clara con el PL en la leche. Sin embargo, sería relevante citar el gen *APBB3* que es miembro de la familia de proteínas APBB, de las cuáles la *APBB2* se ha asociado con el contenido en grasa y proteína láctea (Cole et al., 2011); y el gen *CD14*, que ha sido estudiado en ganado lechero bovino Holstein y se ha visto su asociación con la producción de leche, así como con el porcentaje y rendimiento de grasa láctea (Beecher et al., 2010).

Por último, en el intervalo del cromosoma OAR20, se han anotado 14 genes (*Tabla suplementaria 1*). De todos ellos resulta interesante *GLP1R*, que codifica una proteína transmembrana que funciona como receptor de la hormona del péptido 1 similar al glucagón (GLP-1), estimulante de la secreción de insulina inducida por glucosa. Las enfermedades asociadas con *GLP1R* incluyen insulinoma, caracterizada por la producción de una cantidad elevada de insulina, favoreciendo la síntesis de grasas y proteínas en detrimento de la síntesis de lactosa (Harris y Bachman, 2003). Se debe destacar también el gen *TBC1D22B* (miembro 22B del dominio familiar TCB1), habiéndose encontrado el *TBC1D22A* (miembro 22A), que aunque en la bibliografía no esté asociado con el

carácter lactosa, sí que se ha asociado con el porcentaje de proteína láctea del ganado lechero bovino (Pausch et al., 2017).

4.1.2.5 GWAS porcentaje de extracto seco (ES)

No hubo SNPs significativos a nivel genómico, sin embargo 25 SNPs si fueron significativos a nivel cromosómico tras corrección con FDR (*Tabla 7*). Los análisis mediante MLMA y MLMA - LOCO detectaron consistentemente 4 SNPs localizados en los cromosomas OAR16 y OAR20. Además, mediante MLMA – LOCO se detectaron 21 SNPs significativos más, todos ellos en el cromosoma OAR11, y cercanos entre sí.

Es destacable que los dos SNPs del cromosoma 16 también fueron significativos para CL (oar3_OAR16_14179399 y oar3_OAR16_14186759). Los genes anotados en el intervalo del cromosoma OAR16 ya se discutieron en el apartado anterior. Respecto a OAR11, ninguno de los genes anotados se relacionó con el ES por función o por localización en otros estudios de asociaciones en esta región del genoma (*Tabla suplementaria 1*). Finalmente, respecto al intervalo en el cromosoma OAR20 se anotaron 8 genes (*Tabla suplementaria 1*), los cuáles tampoco se han asociado ni tienen una función relacionada con el ES.

4.1.2.6 GWAS recuento de células somáticas (RCS)

Los análisis mediante MLMA y MLMA – LOCO dieron los mismos resultados en cuanto a la significación de SNPs. No hubo SNPs significativos a nivel genómico, sin embargo 4 SNPs fueron significativos a nivel cromosómico tras corrección con FDR (en OAR19) (*Tabla 7*). Los 4 SNPs estaban en dos zonas diferentes del cromosoma Oar19. Los tres SNPs localizados alrededor de los 9,5 Mb del cromosoma Oar19, estuvieron completamente ligados (oar3_OAR19_9401922, oar3_OAR19_9407458, y oar3_OAR19_9410968). En estas dos regiones se anotaron los genes *NUP210* y *HDAC11* (SNP s36141.1), y *ARPP21*, *miRNA128-2* y *LRPFIP2* (SNPs oar3_OAR19_9401922, oar3_OAR19_9407458, y oar3_OAR19_9410968) (*Tabla suplementaria 1*). El SNP s36141.1 se sitúa en el intrón 1 del gen *NUP210*. Este gen codifica para una proteína que es un componente de los complejos núcleo-poro celulares, estando relacionados con en el ratio de linfocitos T CD4/CD8, y por lo tanto pudiendo influir en la respuesta inmune adaptativa (Borlido et al., 2018; Nieuwenhuijze et al., 2018). El gen *HDAC11*, que produce modificaciones epigenómicas, ha sido descrito como un modulador del sistema inmune actuando como factores pro- y anti- inflamatorios, y regulando células inmunes como neutrófilos, macrófagos y linfocitos T (Yanginlar y Logie, 2018, Liu et al., 2020). Finalmente, el gen *ARPP21* se ha encontrado sobre-

expresado en infecciones por *Salmonella Typhi* en intestino de humanos (Nickerson et al., 2018). Además, uno de sus intrones codifica para el microRNA *miRNA128-2*, que se asoció a un incremento de la expresión del mismo tras infección con *Streptococcus uberis* en células epiteliales mamarias bovinas en cultivo (Lawless et al., 2013). Aunque en estas regiones del cromosoma 19 no se han encontrado estudios en otras razas asociadas al RCS, los genes contenidos en los citados intervalos presentan funciones que nos hacen pensar que podrían estar asociados con el RCS.

4.2 Validación de SNPs relacionados con los caracteres de leche mediante KASP

Todos los SNPs significativos para los diferentes caracteres se localizaron en espacio intergénico o en intrones dentro de los genes, de forma que no producían un cambio aminoacídico. Además, tampoco se situaron en zonas reguladoras de los genes (promotores o regiones UTR de los genes). Por lo tanto, los SNPs detectados para los diferentes fenotipos, deben de estar en LD con la mutación responsable del efecto observado sobre la variabilidad del fenotipo.

Para la posible utilización de estos SNPs en MAS o en selección genómica, o para intentar el posible aislamiento del gen/ mutación responsable del carácter, es necesario validar previamente los resultados encontrados con GWAS (con 192 animales) en una población mayor y/o diferente, o llevar a cabo estudios funcionales. En nuestro caso la validación la hemos realizado en toda la población estudiada de 1.839 animales, de manera que podamos descartar asociaciones espurias en el análisis de GWAS con un número pequeño de animales. Debido al gran número de SNPs significativos, y la imposibilidad de validar todos ellos, se seleccionaron algunos SNPs que tuviesen frecuencias intermedias, y que por la función que presentaban los genes cercanos al SNP, pudiesen estar relacionados con el carácter lechero para el cuál era significativo. En este sentido se eligió 1, 1 y 2 SNPs para la producción de leche, porcentaje de proteína y RCS, respectivamente. No se eligió, ninguno para porcentaje de grasa ni extracto seco porque ninguno de los SNPs anotados en los distintos intervalos presentaba una función que nos hiciese pensar que estuviesen implicados en dichos fenotipos. Para el porcentaje de lactosa tampoco se eligió ninguno, ya que la mayoría presentaban un MAF inferior al 0,05 y aquellos que lo tenían superior no tenían una función clara relacionada con el fenotipo.

4.2.1 Resultados de los estudios de validación mediante KASP

La validación a través del genotipado mediante tecnología KASP, se realizó sobre 1.839 animales con 6.111 registros para el carácter PP y el RCS, a diferencia del carácter de producción de leche con 2.868 registros. Se debe tener en cuenta que para este último carácter no se tienen datos de tres explotaciones sino de dos.

Los 4 SNPs estuvieron en equilibrio de Hardy-Weinberg. En la Tabla 8, se pueden observar los resultados del análisis de asociación mediante el modelo lineal mixto (MIXED) para medidas repetidas con el software SAS. Los efectos rebaño-año-mes (RAM) y días en lactación (DEL) resultaron significativos para todos los fenotipos y SNPs. Igualmente, el efecto número de corderos (NC) fue significativo para todos los SNPs y los fenotipos PL, PG y PL. En cuanto al efecto del genotipo, el SNP rs419096188 fue significativo para los fenotipos RCS ($P = 0,003$) y CL ($P = 0,015$). El genotipo AA mostró una disminución del valor de RCS de $0,21 \pm 0,07$ y $0,26 \pm 0,08$ en relación a los genotipos AG ($P < 0,05$) y GG ($P < 0,01$), respectivamente. En cuanto al CL, el genotipo AA presentó un incremento del $0,046 \pm 0,016$ % de lactosa con respecto al genotipo GG ($P < 0,05$). Como se ha mencionado en el apartado 1.3.1.4, existe una relación entre el RCS y la concentración de lactosa, conforme el primero aumenta, el segundo disminuye ya que los patógenos podrían estar utilizando la lactosa como sustrato (Alessio et al., 2016; Costa et al., 2019). Además, en el presente estudio presentaron una correlación negativa (-0,46).

Respecto al SNP rs424642424, los resultados indicaron que existe asociación significativa entre este SNP y el carácter RCS ($P = 0,009$), presentando el genotipo GG una disminución de $0,163 \pm 0,053$ de este parámetro con respecto al AG ($P < 0,01$). Estos dos primeros SNPs, fueron seleccionados por su posible implicación en el fenotipo RCS y los estudios de validación han confirmado su implicación en el RCS, y lactosa (para el SNP rs419096188). Para el SNP rs411200126, el efecto del genotipo no fue significativo para ningún fenotipo. Este SNP fue seleccionado por su implicación en la PL, pero no hemos podido validar su efecto en toda la población. Esto puede ser debido, a que el estudio para este carácter sólo fue en 2 rebaños y casi la mitad de los animales. También podría deberse a que la asociación del análisis GWAS fue espuria. Finalmente, en cuanto al rs419686662, seleccionado por su posible implicación en el PP, el efecto del genotipo fue significativo ($P = 0,0001$) para el carácter PP. Los animales con genotipo AA presentaron un porcentaje de proteína mayor de $0,058 \pm 0,021$ y $0,134 \pm 0,034$ % que los genotipos GA ($P < 0,05$) y GG ($P < 0,001$), respectivamente. De esta manera, como en el caso de los SNPs elegidos para el RCS, se han validado los efectos de este SNP sobre el PP.

En resumen, hemos identificado 2, 6, 3, 15, 25 y 5 SNPs asociados a los fenotipos PL, PP, PG, PL, ES, y RCS, respectivamente, mediante herramientas de genotipado masivo (GWAS). De estos SNPs, 3 se han validado mediante estudios de asociación en toda la población de estudio, estando dos de ellos asociados al RCS (rs419096188 y rs424642424), y el SNP rs419686662 al PP. Los SNPs asociados al RCS se sitúan en una región genómica en la que están anotados genes relacionados con la respuesta inmune (*NUP210*, *HDAC11* y *ARPP21*). Además, el SNP rs419096188 se ha asociado al CL, presentando una correlación negativa con el RCS, de forma que la selección a favor del genotipo AA produciría una disminución del RCS e incremento del CL. Como se comentó anteriormente, los quesos de la Asociación de Productores de Quesos y Leche de Teruel presentan una coagulación mixta, es decir enzimática y acidoláctica. La coagulación acidoláctica se produce por bacterias láctica que actúan sobre la lactosa y la transforman en ácido láctico. De esta forma una selección a favor de un menor RCS, podría producir un incremento de la lactosa en la leche, y por lo tanto incrementar la coagulación acidoláctica, y el rendimiento quesero. Y por último, el SNP asociado al PP se encuentra localizado cerca del gen *FGFR2*, que codifica para un receptor de factores de crecimiento fibroblásticos. Estos factores de crecimiento fibroblásticos influyen en el desarrollo y funcionalidad de la glándula mamaria en producción (Pond et al., 2013; Zhang et al., 2014; Sumbal et al., 2019).

Los tres SNPs validados podrían ser utilizados en SAM, y en concreto se han introducido en un panel de 192 SNPs que se utilizan para llevar a cabo asignamientos de paternidad y genotipar algunos SNPs funcionales (Calvo et al., 2020).

Se debe destacar, que estos SNPs no son los responsables del efecto detectado sobre los fenotipos, y por lo tanto deben de estar en LD con la mutación responsable del efecto observado sobre la variabilidad del fenotipo. Estas mutaciones podrían estar en los genes anotados cercanos a los SNPs significativos, y de esta manera serían genes candidatos para detectar las mutaciones funcionales responsables del efecto observado sobre los genotipos. Una vez detectadas las mutaciones funcionales, y caracterizado su efecto de forma precisa sobre los diferentes fenotipos podrían ser utilizadas en GAS, y ser incluidas en los diferentes paneles de SNPs, y chips utilizados para selección genómica.

Tabla 8. Test tipo III para los distintos efectos considerados en los estudios de asociación y LSmeans para el efecto del genotipo (indicando la media y desviación estándar para cada genotipo).

SNP ¹	Fenotipo ²	P - VALORES						LSMEANS PARA EL EFECTO GENOTIPO ⁹		
		RAM ³	DEL ⁴	CN ⁵	E ⁶	NL ⁷	G ⁸	AA	AG	GG
rs419096188 s36141.1 (0,4)	RCS	<0,001	<0,001	0,549	0,252	0,071	0,003	2,71 ± 0,10a,c	2,92 ± 0,08b	2,97 ± 0,09d
	PL	<0,001	<0,001	<0,001	0,884	<0,001	0,337	2309,87 ± 74,02	2361,59 ± 58,98	2284,92 ± 63,53
	PG	<0,001	<0,001	0,007	<0,001	<0,001	0,204	5,43 ± 0,09	5,53 ± 0,07	5,47 ± 0,08
	PP	<0,001	<0,001	0,632	0,730	<0,001	0,755	5,27 ± 0,04	5,28 ± 0,03	5,26 ± 0,03
	ES	<0,001	<0,001	0,528	0,317	0,546	0,075	11,03 ± 0,04	11,01 ± 0,03	10,97 ± 0,04
	CL	<0,001	<0,001	0,035	0,138	<0,001	0,015	4,83 ± 0,02a	4,81 ± 0,02ab	4,78 ± 0,02b
rs424642424 oar3_OAR19_9410968 (0,25)	RCS	<0,001	<0,001	0,515	0,289	0,088	0,009	2,94 ± 0,13a,b	2,99 ± 0,09a	2,83 ± 0,08b
	PL	<0,001	<0,001	<0,001	0,787	<0,001	0,144	2081,7 ± 135,31	2331,85 ± 61,34	2336,26 ± 57,33
	PG	<0,001	<0,001	0,007	<0,001	<0,001	0,158	5,59 ± 0,12	5,54 ± 0,08	5,46 ± 0,07
	PP	<0,001	<0,001	0,643	0,720	<0,001	0,505	5,29 ± 0,05	5,28 ± 0,03	5,26 ± 0,03
	ES	<0,001	<0,001	0,502	0,359	0,625	0,719	11,00 ± 0,05	11,01 ± 0,04	10,99 ± 0,03
	CL	<0,001	<0,001	0,033	0,183	<0,001	0,929	4,80 ± 0,03	4,80 ± 0,02	4,80 ± 0,02
rs411200126 oar3_OAR13_28464329 (0,12)	RCS	<0,001	0,016	0,400	0,971	0,267	0,278	3,22 ± 0,41	3,21 ± 0,30	3,05 ± 0,28
	PL	<0,001	<0,001	<0,001	0,703	<0,001	0,127	1992,65 ± 250,5	1938,71 ± 180,08	2072,34 ± 175,43
	PG	<0,001	<0,001	0,002	0,040	<0,001	0,365	5,30 ± 0,31	5,19 ± 0,21	5,09 ± 0,20
	PP	<0,001	<0,001	0,382	0,323	0,066	0,248	5,30 ± 0,14	5,31 ± 0,10	5,25 ± 0,09
	ES	<0,001	<0,001	0,453	0,114	0,975	0,732	11,00 ± 0,15	11,02 ± 0,10	10,99 ± 0,10
	CL	<0,001	<0,001	0,006	0,264	0,001	0,462	4,78 ± 0,08	4,78 ± 0,06	4,81 ± 0,05
rs419686662 oar3_OAR22_40605374 (0,36)	RCS	<0,001	<0,001	0,564	0,474	0,121	0,032	3,01 ± 0,24	2,93 ± 0,24	2,79 ± 0,26
	PL	<0,001	<0,001	<0,001	0,650	<0,001	0,731	2008,8 ± 176,89	2031,67 ± 175,78	2069,74 ± 188,04
	PG	<0,001	<0,001	0,014	<0,001	<0,001	0,048	5,20 ± 0,16	5,10 ± 0,16	5,05 ± 0,17
	PP	<0,001	<0,001	0,754	0,912	<0,001	<0,001	5,32 ± 0,06a,c	5,26 ± 0,07b	5,18 ± 0,07d
	ES	<0,001	<0,001	0,504	0,461	0,760	0,106	11,05 ± 0,07	11,02 ± 0,07	10,98 ± 0,08
	CL	<0,001	<0,001	0,043	0,182	<0,001	0,081	4,81 ± 0,03	4,83 ± 0,03	4,85 ± 0,04

¹ SNP: Identificador en la base de datos de SNPs-dbSNPs (rxXXXXXX). Identificación del SNP en el chip Agresearch de 680 k (OAR3_XXX_XXXX). Entre paréntesis se indica el MAF. ²Fenotipo: PL= producción de leche; PP= porcentaje de proteína; PG= porcentaje de grasa; CL= contenido en lactosa; ES= porcentaje extracto seco; RCS= recuento de células somáticas; ³RAM: Rebaño- Año - Mes; ⁴DEL: días en lactación; ⁵CN: número corderos nacidos; ⁶E: edad oveja; ⁷NL: número de parto; ⁸G: genotipo; ⁹Diferentes letras indican diferencias significativas entre genotipos tras corrección de bonferroni para p(a,b)< 0,05 y p (c,d)< 0,01.

V. Conclusiones

Las conclusiones de este trabajo fueron las siguientes:

1. Los estudios de asociación genómica GWAS no mostraron ningún SNP significativo a nivel genómico para ninguno de los caracteres.
2. A nivel cromosómico se han encontrado 2, 6, 3, 15, 25 y 4 SNPs para los caracteres de producción de leche (PL), porcentaje de proteína (PP), porcentaje de grasa (PG), contenido en lactosa (CL), porcentaje de extracto seco (ES) y recuento de células somáticas (RCS), respectivamente.
3. Los estudios de validación mediante genotipado de toda la población mediante KASP nos permitieron asegurar la implicación de tres SNPs en los caracteres analizados, no pudiendo validar el SNP relacionado con la PL.
4. El SNP rs419096188, se sitúa en el intrón 1 del gen *NUP210* y cercano al gen *HDAC11*, ambos relacionados con la respuesta inmune, y se asoció a los fenotipos RCS y CL, favoreciendo el genotipo AA la disminución del RCS e incremento del CL.
5. El SNP rs424642424 se asoció al RCS, estando el gen *ARPP21* anotado en la región donde se localiza el SNP. Dicho SNP mostró que el genotipo GG presentaba una disminución del RCS con respecto al AG .
6. Finalmente, el SNP rs419686662, seleccionado por su posible implicación en el PP, se asoció con el carácter PP, presentando los animales con genotipo AA un mayor PP que los otros dos genotipos.

VI. Referencias bibliográficas

- ASSAF.E Asociación de Criadores de Ganado Ovino de la raza Assaf. (2015).
- ASSAF.E Asociación de Criadores de Ganado Ovino de la raza Assaf. (2011).
- Banos, G., Bramis, G., Bush, S.J., Clark, E.L., [et al.] (2017). "The genomic architecture of mastitis resistance in dairy sheep". *BMC genomics*, 18 (1), pp.
- Beecher, C., Daly, M., Childs, S., Berry, D.P., [et al.] (2010). "Polymorphisms in bovine immune genes and their associations with somatic cell count and milk production in dairy cattle". *BMC genetics*, 11(1), pp. 99.
- Biacabe, A., Jacobs, J.G., Bencsik, A., Langeveld, J.P.M., [et al.] (2007). "H-Type Bovine Spongiform Encephalopathy". *Prion*, 1(1), pp. 61-68.
- Borlido, J., Sakuma, S., Raices, M., Carrette, F., [et al.] (2018). Nuclear pore complex-mediated modulation of TCR signaling is required for naïve CD4+ T cell homeostasis. *Nature immunology*, 19(6), 594–605.
- Buxadé, C., 1996, Zootecnia. Bases de producción animal. Producción ovina, Vol VIII, 1ª edición Edition. Mundi-Prensa Libros, Madrid.
- Calvo, J.H., Dervishi, E., Sarto, P., González-Calvo, L., [et al.] (2013). "Structural and functional characterisation of the α S1-casein (CSN1S1) gene and association studies with milk traits in Assaf sheep breed". *Livestock Science*, 157(1), pp. 1-8.
- Calvo, J., Serrano, M., Tortereau, F., Sarto, P., [et al.] (2020). Development of a SNP parentage assignment panel in some North-Eastern Spanish meat sheep breeds. *Spanish Journal of Agricultural Research*, 18(4), e0406.
- Cole, J.B., Wiggans, G.R., Ma, L., Sonstegard, T.S., [et al.] (2011). "Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows". *BMC genomics*, 12(1), pp. 408.
- Costa, A., Lopez-Villalobos, N., Visentin, G., De Marchi, M., [et al.] (2019). "Heritability and repeatability of milk lactose and its relationships with traditional milk traits, somatic cell score and freezing point in Holstein cows". *Animal (Cambridge)*, 13(5), pp. 909-916.
- Davis, S. R., Collier, R. J., McNamara, J. P., Head, H. H., [et al.] (1988). Effects of thyroxine and growth hormone treatment of dairy cows on mammary uptake of glucose, oxygen and other milk fat precursors. *Journal of animal science*, 66(1), 80–89.
- Duchemin, S.I., Colombani, C., Legarra, A., Baloche, G., Larroque, H., Astruc, J.-., Barillet, F., Robert-Granié, C. y Manfredi, E. (2012). "Genomic selection in the French Lacaune dairy sheep breed". *Journal of dairy science*, 95(5), pp. 2723-2733.
- Echeverri Zuluaga, J, Jaramillo. MG y Restrepo Betancur. LF, (2010). "Evaluación comparativa de dos metodologías de diagnóstico de mastitis en un ható lechero del Departamento de Antioquia". *Revista Lasallista de Investigación*, 7 (1), pp. 116.

- Estrada Korta, O., Molino Gahete, F., Joy Torrens, M., Ariño Moneva, A.A. y Juan Esteban, T. (2013). Composición en ácidos grasos de la leche de oveja Assaf y modificación en el perfil de ácidos grasos libres del queso de Teruel durante la maduración. Zaragoza.
- FEAGAS, 2015 Razas de ganado ovino. En Catálogo Oficial de Razas de Ganado de España, Federación Española de Asociaciones de Ganado Selecto.
- Fox PF, Uniacke-Lowe T, McSweeney PLH and O'Mahoni JA (2015). Dairy chemistry and biochemistry. Springer International Publishing, Cham, Switzerland.
- Gallardo, B. 2013. Utilización de diferentes fuentes lipídicas en la ración de ovejas churras en lactación: efecto sobre los rendimientos y calidad de los productos. Tesis Doctoral Universidad Valladolid.
- García-Gómez, E., Gutiérrez-Gil, B., Suárez-Vega, A., de la Fuente, L. F., & Arranz, J. J. (2013). Identification of quantitative trait loci underlying milk traits in Spanish dairy sheep using linkage plus combined linkage disequilibrium and linkage analysis approaches. *Journal of dairy science*, 96(9), 6059–6069.
- Harris B, Bachman KC. (2003). Nutritional and management factors affecting solids-not-fat, acidity and freezing point of milk. University of Florida, Institute of Food and Agricultural Sciences, *Florida Cooperative Extension Service*, p. 5.
- Hernandez, F., Elvira, L., Gonzalez-Martin, J. V., Gonzalez-Bulnes, A., & Astiz, S. (2011). Influence of age at first lambing on reproductive and productive performance of Lacaune dairy sheep under an intensive management system. *The Journal of dairy research*, 78(2), 160–167.
- Huppertz, T., Fox, P.F. y Kelly, A.L. (2004). "High pressure-induced denaturation of α -lactalbumin and β -lactoglobulin in bovine milk and whey: a possible mechanism". *Journal of Dairy Research*, 71(4), pp. 489-495.
- Jiang, Y., Xie, M., Chen, W., Talbot, R., [et al.]. 2014. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, 344 (6188) pp. 1168-1173.
- Jiang, J., Ma, L., Prakapenka, D., VanRaden, P.M., [et al.]. (2019). "A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle". *Frontiers in genetics*, 10, pp. 412.
- Jurado, J.J. y Jiménez, MA (2017). "Estimación conjunta de parámetros genéticos de producción, conformación de la ubre y recuento de células somáticas en la raza Assaf". Departamento de Mejora Genética Animal. INIA: XVII Jornadas sobre Producción Animal, pp. 519-521.
- Keefe, G. P., & Elmoslemany, A. (2007). A consumer acceptance of fluid milk after raw milk selection using bulk tank bacteriologic and somatic cell count criteria. in Proceedings of the National Mastitis Council. San Antonio, Texas.
- Kijas, J. W., Porto-Neto, L., Dominik, S., Reverter, A., [et al.] (2014). Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip. *Animal genetics*, 45(5), 754–757.

- Kremer, R., Giordano, J.P., Rosés, L. y Rista, L. (2015). "Producción de ovejas Milchschaf en un sistema lechero en pastoreo". *Veterinaria (Montevideo)*, 51(199), 12- 23.
- Lawless, N., Foroushani, A. B., McCabe, M. S., O'Farrelly, C., & Lynn, D. J. (2013). Next generation sequencing reveals the expression of a unique miRNA profile in response to a gram-positive bacterial infection. *PloS one*, 8(3), e57543.
- Legaz, E.A. 2016. Caracterización de la raza Assaf en España. Tesis doctoral UCM.
- Liu SS, Wu F, Jin YM, Chang WQ, Xu TM (2020). "HDAC11: a rising star in epigenetics". *Biomedicine & pharmacotherapy*. .
- MAGRAMA. 2019. El sector ovino y caprino de leche en cifras. Principales indicadores económicos.
- McNeilly, A. S., & Fraser, H. M. (1987). Effect of gonadotrophin-releasing hormone agonist-induced suppression of LH and FSH on follicle growth and corpus luteum function in the ewe. *The Journal of endocrinology*, 115 (2), 273–282.
- Murphy, T. W., Berger, Y. M., Holman, P. W., Baldin, M., [et al.] (2017). Estimates of genetic parameters, genetic trends, and inbreeding in a crossbred dairy sheep research flock in the United States. *Journal of animal science*, 95 (10), 4300–4309.
- Nava-García, A., Martínez-Rojero, R.D., Mastache-Lagunas, ÁA. y Ulloa-Arvizu, R. (2019). "Curva de rendimiento y composición de leche en ovejas criollas de la Montaña de Guerrero, México". *Ecosistemas y recursos agropecuarios*, 6 (17), pp. 391-398.
- Nickerson, K. P., Senger, S., Zhang, Y., Lima, R., [et al.] (2018). Salmonella Typhi Colonization Provokes Extensive Transcriptional Changes Aimed at Evading Host Mucosal Immune Defense During Early Infection of Human Intestinal Tissue. *EBioMedicine*, 31, 92–109.
- Pausch, H., Emmerling, R., Gredler-Grandl, B., Fries, R., [et al.] (2017). Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. *BMC genomics*, 18(1), 853.
- Payne, J. M., Dew, S. M., Manston, R., & Faulks, M. (1970). The use of a metabolic profile test in dairy herds. *The Veterinary record*, 87(6), 150–158.
- Pond, A. C., Bin, X., Batts, T., Roarty, K., [et al.] (2013). Fibroblast growth factor receptor signaling is essential for normal mammary gland development and stem cell function. *Stem cells* (Dayton, Ohio), 31(1), 178–189.
- Racagni, G., Apud, J. A., Cocchi, D., Locatelli, V., [et al.] (1984). Regulation of prolactin secretion during suckling: involvement of the hypothalamo-pituitary GABAergic system. *Journal of endocrinological investigation*, 7 (5), 481–487.
- Raynal-Ljutovac, K; Pirisi, A; De Crémoux, R.; Gonzalo, C. (2007). Somatic cells of goat and sheep milk: Analytical, sanitary, productive and technological aspects. *Small Rum Res*, 68:126- 144.

- Schoknecht, P.A., Nobrega, S.N., Petterson, J.A., Ehrhardt, R.A., [et al.] (1991), Relations between maternal and fetal plasma concentrations of placental lactogen and placental and fetal weights in well-fed ewes. *J Anim Sci* 69, 1059-1063.
- Selvaggi, M., D'Alessandro, A. G. and Dario, C. (2017) "Environmental and genetic factors affecting milk yield and quality in three Italian sheep breeds," *Journal of Dairy Research*, Cambridge University Press, 84 (1), pp. 27–31.
- Suchocki, T., Wojdak-Maksymiec, K. y Szyda, J. (2016). "Using gene networks to identify genes and pathways involved in milk production traits in Polish Holstein dairy cattle". *Czech Journal of Animal Science*, 61(No. 11), pp. 526-538.
- Sumbal, J., & Koledova, Z. (2019). FGF signaling in mammary gland fibroblasts regulates multiple fibroblast functions and mammary epithelial morphogenesis. *Development* (Cambridge, England), 146(23), dev185306.
- Tolone, M., Larrondo, C., Yáñez, J. M., Newman, S., [et al.] (2016). Assessment of genetic variation for pathogen-specific mastitis resistance in Valle del Belice dairy sheep. *BMC veterinary research*, 12 (1), 158.
- Usai, MG, Casu, S., Sechi, T., Salaris, SL, [et al.] (2019). Mapeo de las regiones genómicas que afectan los rasgos de la leche en ovejas Sarda mediante el uso de OvineSNP50 Beadchip y componentes principales para realizar un análisis combinado de ligamiento y desequilibrio de ligamiento. *Genética, selección, evolución: GSE*, 51 (1), 65.
- van Nieuwenhuijze, A., Burton, O., Lemaitre, P., Denton, A. E., [et al.] (2018). Mice Deficient in Nucleoporin Nup210 Develop Peripheral T Cell Alterations. *Frontiers in immunology*, 9, 2234.
- Yanginlar, C., & Logie, C. (2018). HDAC11 is a regulator of diverse immune functions. *Biochimica et biophysica acta. Gene regulatory mechanisms*, 1861(1), 54–59.
- Zhang, X., Martinez, D., Koledova, Z., Qiao, G., [et al.] (2014). FGF ligands of the postnatal mammary stroma regulate distinct aspects of epithelial morphogenesis. *Development* (Cambridge, England), 141(17), 3352–3362.
- Zhu, L., Yu, Q., Lin, J., Zhang, Q y Yang, Q. (2014). "Cloning and functional analysis of goat glucose transporter 4". *Molecular biology reports*, 41 (2), pp. 757-768.

ANEXOS

Tabla suplementaria 1. Genes anotados dentro del intervalo 0,5 Mb de los SNPs localizados en posición central del intervalo y asociados a caracteres lecheros.

Fenotipo¹	Crs²	Snp³	Genes
PL	13	oar3_OAR13_28464329	<i>ENSOARG00000006835 (FAM107B)-ensoarg00000004517 (FAM107B)-*-CDNF-HSPA14-snord22- SUV39H2-DCLRE1C-MEIGI-ensoarg00000007897-ACBD7-RPP38</i>
	13	oar3_OAR13_28473530	<i>ENSOARG00000006835 (FAM107B)-ensoarg00000004517 (FAM107B)-*-CDNF-HSPA14-snord22- SUV39H2-DCLRE1C-MEIGI-ensoarg00000007897-ACBD7-RPP38-NMT2</i>
PP	6	oar3_OAR6_5285000	<i>*-MAD2L1 (situado a 326 Kb)</i>
	17	oar3_OAR17_12384078	<i>ZNF827-ensoarg00000008671-MMAA *-SMAD1</i>
	17	oar3_OAR17_25027692	<i>PCDH10- *-ensoarg000000014661 (RNF144A)</i>
	22	oar3_OAR22_40605374	<i>FGFR2- *-ATE1-ensoarg00000006090-NSMCE4A-TACC2</i>
	22	s37672.1	<i>FGFR2- *-ATE1-ensoarg00000006090-NSMCE4A-TACC2</i>
	22	oar3_OAR22_40606216	<i>FGFR2- *-ATE1-ensoarg00000006090-NSMCE4A-TACC2</i>
PG	14	oar3_OAR14_39364677	<i>CMTR2-HYDIN*-POP4-PLEKHF1-ensoarg00000004057-CCNE1-UR11- *-ZNF536</i>
	14	s32350.1	<i>HYDIN-POP4-PLEKHF1-ensoarg00000004057- *-CCNE1-UR11</i>
	14	oar3_OAR14_39913998	<i>PLEKHF1-ensoarg00000004057-CCNE1-UR11- *-ZNF536</i>
CL	13	oar3_OAR13_67179806	<i>CCNO-MCIDAS-CDC20B-GPX8-GZMA-ENSOARG00000007978-GZMK- *-ESM1-ENSOARG00000008055</i>
	16	oar3_OAR16_24178409	<i>TRAPPC13-TRIM23-PPWD1-CENPK-ADAMTS6*</i>

	16	oar3_OAR16_14179399	<i>TRIM23-PPWD1-CENPK-ADAMTS6-*</i>
	16	oar3_OAR16_14186759	<i>TBC1D2-*-ENSOARG00000011118-ENSOARG00000011131-NANS-ENSOARG00000011154-HEMGN-TRMO-FOXE1</i>
	2	oar3_OAR2_49255325	<i>ENSOARG00000008196-*.ALG2</i>
	2	oar3_OAR2_48135063	<i>(ausencia de genes)</i>
	3	oar3_OAR3_64750845	<i>(ausencia de genes)</i>
	3	oar3_OAR3_64741679	<i>*-ENSOARG00000004322 (NRXN1A)</i>
	3	oar3_OAR3_73955311	<i>ENSOARG00000004788 (PRKN)*</i>
	8	oar3_OAR8_84295105	<i>OPRM1-ENSOARG00000004106*</i>
	8	oar3_OAR8_77590661	<i>ENSOARG00000017846-SRA1-APBB3-SLC35A4-CD14-TMCO6-IK-NDUFA2-DND1-HARS1-HARS2-ZMAT2-*.PCDHA2-PCDHA5-PCDHA11-PCDHAC2</i>
	5	oar3_OAR5_49233977	<i>SLC4A9- ENSOARG00000017846-SRA1-APBB3-SLC35A4-CD14-TMCO6-IK-NDUFA2-DND1*-HARS1-HARS2-ZMAT2-PCDHA2-PCDHA5-PCDHA11-PCDHAC2</i>
	5	oar3_OAR5_49200052	<i>DNAH8-GLP1R-SAYSD1-KCNK5*-KCNK17-KCNK16-KIF6</i>
	20	oar3_OAR20_13106461	<i>FGD2-PIM1-TMEM217-TBC1D22B*-RNF8-CMTR1-CCDC167</i>
	20	oar3_OAR20_11231797	<i>FGD2-PIM1-TMEM217-TBC1D22B*-RNF8-CMTR1-CCDC167</i>
ES	16	oar3_OAR16_14179399	<i>TRAPPC13-TRIM23-PPWD1-CENPK-ADAMTS6*</i>
	16	oar3_OAR16_14186759	<i>TRIM23-PPWD1-CENPK-ADAMTS6-*</i>
	20	oar3_OAR20_32651642	<i>DCDC2-NRSN1-*</i>
	11	oar3_OAR11_44970125	<i>LYZL6-ENSOARG00000010725-WNT9B-WNT3-NSF*-ENSOARG00000011103</i>
	20	oar3_OAR20_46839037	<i>SNRNP48-DSP-*.RIOK1-CAGE1-SSR1-RREB1</i>

11	oar3_OAR11_22502645	INPP5K-ENSOARG00000012974-SLC43A2-SCARF1-PRPF8-TLCD2-WDR81-SERPINF2-SERPINF-SMYD4*-RPA1-RTN4RL1-DPH1-OVCA2-SMG6
11	oar3_OAR11_31971548	<i>COX10-ENSOARG00000016450-*</i>
11	oar3_OAR11_42884861	<i>NBR1-TMEM106A-ARL4D-DHX8-ETV4*-MEOX1-SOST-ENSOARG00000005929-CFAP97D1-MPP3</i>
11	oar3_OAR11_31716346	<i>COX10-ENSOARG00000016450-*</i>
11	oar3_OAR11_31732260	<i>COX10-ENSOARG00000016450-*</i>
11	oar3_OAR11_34077455	<i>PRPSAP2-SHMT1-SMCR8-TOP3A-MIEF2-FLII-LLGL1-ALKBH5-MYO15A-DRG2-GID4-ATPAF2*-DRC3-SREBF1-RAI1</i>
11	oar3_OAR11_34085781	<i>PRPSAP2-SHMT1-SMCR8-TOP3A-MIEF2-FLII-LLGL1-ALKBH5-MYO15A-DRG2-GID4-ATPAF2*-DRC3-SREBF1-RAI1</i>
11	oar3_OAR11_34102262	<i>PRPSAP2-SHMT1-SMCR8-TOP3A-MIEF2-FLII-LLGL1-ALKBH5-MYO15A-DRG2-GID4-ATPAF2*-DRC3-SREBF1-RAI1</i>
11	oar3_OAR11_31956799	<i>COX10-ENSOARG00000016450-*</i>
11	oar3_OAR11_32608494	<i>TEKT3-CDRT4-TVP23C-FBXW10-ENSOARG00000016719*-ENSOARG00000016792-ZNF624-ZNF287-LRRC75A-TRPV2-UBB-CENPV-PIGL</i>
11	oar3_OAR11_37800259	<i>SKAP1-SNX11-ENSOARG00000007305-NFE2L1-COPZ2-CDK5RAP3</i>
11	oar3_OAR11_31708633	<i>ENSOARG00000016403-COX10*-ENSOARG00000016450</i>
11	oar3_OAR11_31748081	<i>COX10*-ENSOARG00000016450</i>
11	oar3_OAR11_31725295	<i>COX10*-ENSOARG00000016450</i>
11	oar3_OAR11_31730623	<i>COX10*-ENSOARG00000016450</i>
11	oar3_OAR11_31954402	<i>COX10-ENSOARG00000016450-*</i>

	11	oar3_OAR11_31970980	<i>COX10-ENSOARG00000016450-*</i>
	11	oar3_OAR11_31744746	<i>COX10*-ENSOARG00000016450</i>
	11	oar3_OAR11_31744803	<i>COX10*-ENSOARG00000016450</i>
	11	oar3_OAR11_34067161	<i>PRPSAP2-SHMT1-SMCR8-TOP3A-MIEF2-FLII-LLGL1-ALKBH5-MYO15A-DRG2-GID4-ATPAF2*-DRC3-SREBF1-RAI1</i>
RCS	19	s36141.1	<i>NUP210*-HDAC11</i>
	19	oar3_OAR19_9401922	<i>ARPP21 (mir128-2)-*-LRPFIP2</i>
	19	oar3_OAR19_9407458	<i>ARPP21 (mir128-2)-*-LRPFIP2</i>
	19	oar3_OAR19_9410968	<i>ARPP21 (mir128-2)-*-LRPFIP2</i>

* Indica la localización del SNP, dentro del gen (GEN*) o en espacio intergénico (-*-). ¹Fenotipo: PL= Producción de Leche; PP= Porcentaje de proteína; PG= porcentaje de grasa; CL= contenido en lactosa; ES= porcentaje extracto seco; RCS= recuento de células somáticas. ²Crs: Cromosoma. ³SNP: Polimorfismo de un solo nucleótido.