SMARTER

SMAll RuminanTs breeding for Efficiency and Resilience

Research and Innovation action: H2020 – 772787
Call: H2020-SFS-2017-2
Type of action: Research and Innovation Action (RIA)
Work programme topic: SFS-15-2016-2017
Duration of the project: 01 November 2018 – 31 June 2023

Report on identified pleiotropic QTLs and potential undelying genes and pathways associated efficiency, resilience and re-production related traits

Rachel Rupp[1], Johanna Ramirez-Diaz[2], Christel Marie Etancelin[1], Brenda Vera[3], Gabriel Ciappesoni[3], Beatriz Gutierrez-Gil[1], Isabelle Palhiere[1], Dominique Hazard[1], Alexandra Stella[2], Karolina Kaseja[5], Joanne Connington[5], Gwenola Tosser-Klopp[1]

1-INRAE*

2-CNR

3-INIA-UY

4-UNILEON

5- SRUC

* Deliverable leader – Contact: rachel.rupp@inrae.fr

DELIVERABLE D3.4

Workpackage N°3
Due date: M54
Actual date: 25/08/2023
Dissemination level: Public

About the SMARTER research project

SMARTER will develop and deploy innovative strategies to improve Resilience and Efficiency (R&E) related traits in sheep and goats. SMARTER will find these strategies by: i) generating and validating novel R&E related traits at a phenotypic and genetic level ii) improving and developing new genome-based solutions and tools relevant for the data structure and size of small ruminant populations, iii) establishing new breeding and selection strategies for various breeds and environments that consider R&E traits.

SMARTER with help from stakeholders chose several key R&E traits including feed efficiency, health (resistance to disease, survival) and welfare. Experimental populations will be used to identify and dissect new predictors of these R&E traits and the trade-off between animal ability to overcome external challenges. SMARTER will estimate the underlying genetic and genomic variability governing these R&E related traits. This variability will be related to performance in different environments including genotype-by-environment interactions (conventional, agro-ecological and organic systems) in commercial populations. The outcome will be accurate genomic predictions for R&E traits in different environments across different breeds and populations. SMARTER will also create a new cooperative European and international initiative that will use genomic selection across countries. This initiative will make selection for R&E traits faster and more efficient. SMARTER will also characterize the phenotype and genome of traditional and underutilised breeds. Finally, SMARTER will propose new breeding strategies that utilise R&E traits and trade-offs and balance economic, social and environmental challenges.

The overall impact of the multi-actor SMARTER project will be ready-to-use effective and efficient tools to make small ruminant production resilient through improved profitability and efficiency.

Table of Content

1. Summary

The challenge for livestock breeding is to improve the resilience traits simultaneous with efficiency traits. Simultaneous breeding for multiple traits can fall foul of trade-offs between traits. The objective of the study reported here was to identify and characterize genomic regions that are associated with both efficiency- and resilience- related traits, which can explain such trade off, and disentangle potential underlying genes and pathways.

To do so, we used results from genome wide association studies (GWAS) in various sheep and goat datasets implemented in SMARTER or provided by partners from background data. In total we analysed 10 populations from 8 partners [INRAE, UNILEON, INIA-UY, SRUC, TexelS, CAPGENES] and 9 breeds. We found indication of genomic regions that control both R&E traits in four populations:

- In Romane sheep, for both growth and body condition score on chromosome 1
- In Uruguay Merino, for resilience (weight and wool traits) and resilience (parasite resistance and temperament) on chromosome 5, 6, 11 and 21
- In Lacaune sheep, for both mastitis (SCS & Staphylococci) and growth, and potentially milk production traits on chromosome 3
- In Saanen goat, for milk production traits, udder type traits, reproduction (semen traits) and longevity on chromosome 19

The "Close LInkage versus Pleiotropism test" developed by David et al. (2013), was used to further distinguish between pleiotropy (a single genetic variant affecting more than one trait) and/or close linkage, i.e. different variants that are physically close. In most cases, we could exclude the hypothesis of pleiotropy (same mutation associated with both R&E traits) suggesting that close linkage of distinct causal mutations is the general pattern in this common QTL regions. Investigation of underlying genes and pathways in the identified genomic regions that control both R&E traits highlighted the importance of lipid metabolism, as underlying mechanism.

Finally, we provided a list of variants in the regions of interest by systematic screening for polymorphisms in the latter regions, using public databases. This list provides candidate causal mutations for QTLs in the studied regions likely to include mutations responsible for trade-off between R&E traits.

2. Introduction

The resilience ability of animals in various farmed species is being eroded by selection for high levels of production (Rauw, 1998). **This erosion is supported by genetic antagonisms (trade-offs) between traits** related to efficiency such as growth, milk production with traits related to resilience such as resistance or tolerance to disease or longevity, or body condition score variation. A meta-analysis of such genetic antagonism was done previously within SMARTER (Deliverable D3.1) and published (Mucha et al., 2022; doi: 10.1016/j.animal.2022.100456). Indeed, pooled estimates of genetic correlations between resilience and efficiency traits in dairy goats, dairy sheep and meat sheep were estimated. Results of this meta-analysis provided little evidence of genetic antagonisms found between resilience and efficiency for dairy goats (SCS with fat content) and dairy sheep (SCS with protein content), and not for meat sheep.

Even though the pooled estimates were non-significant, antagonisms may exist but only in specific populations and environments.

The biological basis for those genetic trade-offs is mostly unknown. **One hypothesis includes that they are due to closely linked or pleiotropic genes that were simultaneously selected**. The difference between these two hypotheses (closely linked or pleiotropic genes) is fundamental, as it determines the possibility of jointly selecting unfavourably linked traits. In the case of closely related genes, strategies can be implemented to break the linkage (cross-breeding, gene editing) and select both traits. In the case of pleiotropic genes, the local trade-off cannot be resolved. Methods, such as the "Close LInkage versus Pleiotropism test" (CLip Test) developed by David et al. (2013; DOI: 10.1038/hdy.2012.70), can be used to distinguish between pleiotropy (a single genetic variant affecting more than one trait) and/or close linkage, i.e. different variants that are physically close. **Genome wide association studies (GWAS) based on both genotype and phenotype data provide an opportunity to identify and explore** such cases using the CLip Test. Interrogation of these data will enhance the understanding of the genetic basis for trade-offs between resilience and efficiency related traits and provide candidates genes which are likely to include putative causal mutations for such trade-offs.

In order to study this hypothesis in depth, we collated and analysed existing genomic and phenotypic datasets from Smarter partners to detect pleiotropic/linked QTL/genes between production and resilience traits. A comparison of results was made across populations and traits to retrieve QTL regions associated with both R&E traits. We then applied the Clip test method to statistically asses the possibility for pleiotropic underlying genes.

We further analysed the list of genes in the regions of interest to identify molecular functions and biological processes underlying the pleiotropic QTL regions. Genes were retrieved using Ensembl and analysed using Gprofiler and Ingenuity software.

Finally, systematic screening of polymorphisms for the latter genes in public sequence data for sheep (Ensembl) and goat (Vargoats ; https://www.goatgenome.org/vargoats.html) was done.

3. Genotype and gwas datasets

We collated datasets including 50K SNP chip information created in WP1, WP2 and WP3 and existing datasets provided by the project partners. The data had to include both genotyping and phenotyping for resilience and efficiency (R&E) traits. For a given population, genotype and phenotype data were analysed together in genome wide association analyses (GWAS). GWAS were either done on purpose for SMARTER (WP1, 2 or 3) or available from previous projects. Accordingly, the following dataset were gathered:

Meat Sheep data (N = 15,528)

1. INIA_UY provided new GWAS data (WP1,2,3) in **Corriedale** (N = 1548) and **Merino** (N = 2100) and **Texel** (N = 657) Urugay for health (parasite resistance), body condition score, fat traits, weight and wool traits
2. CNR & INRAE provided new GWAS data (WP2) in French **Romane** (N = 1030) for body condition score and body weight.
3. SRUC & TexelS provided new GWAS data (WP2) in UK **Texel** (N = 10,193) for health (foot-rot and mastitis) and production (birth weight, weaning weight, scan weight, and fat and muscle depth)

Dairy Sheep data (N = 5685)

4. UNILEON, produced new GWAS data for mastitis resistance (SCS) and milk production traits (milk yield, fat and protein percentage) in **Churra** sheep (n = 2958) using background data from the Churra Breeders' Association, ANCHE.
5. INRAE provided background GWAS data analysed outside SMARTER in **Lacaune** (N=1009 & N= 504) for mastitis and milk production (Rupp et al., 2015 & Oget et al., 2019)
6. INRAE provided new GWAS data for mastitis resistance (SCS) and fine milk composition traits (milk yield, fat and protein percentage, various caseins and fatty acids porcentage) using background genotype data from one experimental farm of **Lacaune** dairy sheep (N=1069) for mastitis and milk production and spectra.
7. INRAE provided background GWAS data analysed outside SMARTER in **Manech tete rousse** (N = 145) for mastitis and milk production (Oget et al., 2019)

Goat data (N = 3391)

8. INRAE provided background GWAS data in **Alpine and Saanen** female goats (N=1941) for milk production, mastitis and udder type traits (Martin et al., 2018)
9. INRAE provided background GWAS data in **Alpine and Saanen** artificial insemination males (N=1112) for functional longevity (Palhiere et al., 2018), semen quality traits and milk production (Talouarn et al., 2020)
10. INRAE (in collaboration with CAPGENES and IDELE) provided new GWAS (WP2) data in **Saanen and Alpine** female goats (N = 338) for BHB as a proxy for metabolic disorder

As a summary we collated GWAS info from 8 partners (INRAE, CNR, UNILEON, INIA-UY, SRUC, TexelS, CAPGENES, IDELE) with around 25,000 genotyped animals in 9 breeds: Churra, Lacaune, Manech Tête Rousse, Texel (Urugay & UK), Romane, Corriedale, Merino, Alpine and Saanen.

Phenotypes to characterise efficiency-related traits included milk production and quality, growth and weight traits, wool, as well as ultrasound fat measures. Phenotypes to characterise resilience-related traits included mastitis, footrot, gastro intestinal parasite (FEC, FAMACHA), functional longevity, Body condition Score (BCS), together with reproduction (semen traits, Pregnancy rate, lambing potential) and temperament in lambs

4. Identification of pleiotropic chromosomal regions from GWAS

Gwas analyses were done separately for each population and were described below. We agreed to qualify as possible pleiotropic QTL region, a chromosome area of less or equal to 5 Mb including QTL for both R&E trait(s). Results were summarised in the last paragraph of this section, and the identified regions of interest were subjected to further analyses (sections 5, 6 and 7).

Gwas in Corriedale, Merino and Texel Uruguay sheep

In Uruguay, different breeds and group of traits were analysed. We analysed a comprehensive set of traits in the Australian Merino: gastro intestinal parasite traits (FEC, FAMACHA), Temperament (isolation box), fat depth and Body condition Score (BCS) in lambs and reproduction traits as well as live weight and wool traits in ewes. Additionally, FEC was studied in Corriedale breed and Fat Depth and body depth were studied in Texel

Breed	Resilience & Efficiency lambs	Reference
Merino	Lambs: FEC1, FEC2, FAMACHA, Temperament, Fat Depth, BCS	Vera et al., PhD tesis. Romaniuk et al. MSc thesis
	Ewes: BCS, Fat depth, Live weight, wool (fibre diameter & greasy fleece weight) & reproduction	Ramos et al. 2023 and PhD thesis
Corriedale	FEC	Carracelas et al. 2022 and MSc thesis
Texel	Fat Depth, Body Weight	Peraza et al. unplished

Merino breed:

Lambs, Stud flocks and INIA´s Nucleus - FEC

Two samples of fecal material from naturally developed parasitic infections were collected, which correspond to two independent parasitic cycles separated from each other by an anthelmintic treatment. The first FEC (FEC1) was recorded at 7-9 months of age after having been applied an anthelmintic at weaning, while the second FEC (FEC2) was recorded at 10-14 months of age after a dose with an anthelmintic after the recording of FEC1. A total of 26,638 animals born in the period 2001-2020 and coming from 13 establishments had FEC1 records, 18,971 of these, had FEC2 records. A total of 1702 individuals were genotyped with the GeneSeek® Genomic Profiler™ Ovine 50K panel (GGP, 43705 SNP). Finally, 38,268 SNPs for 1,697 lambs were used in the analysis after QC. Chromosomal segments representing 0.2% or more of the additive genetic variation (gVar) were defined as QTL regions. SNPs within QTL regions were identified and mapped on the ovine Oar v3.1 genome ("Oar_v3.1 - oviAri3 - Genome - Assembly - NCBI" n.d.) using the Ensembl database. A range of 5kb upstream and downstream of the variant position was used to capture candidate genes.
Results showed highly significant QTLS (≥0.8% of explained variance) for FEC traits on OAR5, 6, 11 and 21 (**Figure 4.1**).

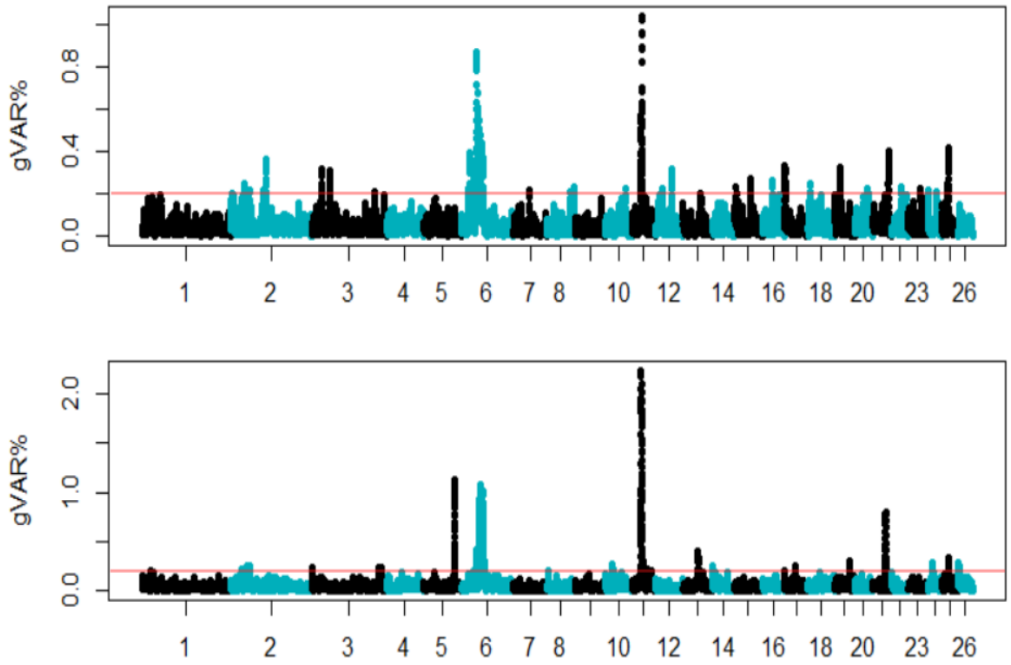


Figure 4.1: Results of ssGWAS for resistance to gastrointestinal nematodes, measured by a) FEC1 and b) FEC2 in Australian Merino sheep. The X axis shows the 26 autosomes and the y axis represents the proportion of additive genetic variance explained by 20 adjacent SNPs. Each point represents a window and the red line the threshold >0.2 gVAR%.

Lambs, Stud flocks and INIA´s Nucleus – Temperament (Isolation box)

A total of 4,317 lambs were tested for temperament between one- and three-months post-weaning, including four progeny: 2010, 2011, 2018 and 2019. Temperament was measured using the Isolation Box Test (IBT) (Blache and Ferguson, 2005). Each lamb was gently pushed inside the box, held there for 30 seconds, and agitation was objectively measured by an agitation meter. The agitation meter registered the vibrations of the box induced by the lamb's movements and high-pitched vocalizations. The higher the agitation score, the more nervous the sheep. There was no previous selection of animals and they had no previous experience with IBT.

Genotype data: A total of 1,697 lambs were genotyped for 43,705 SNP and the molecular information was obtained using the Geneseek Genotypong Profile panel (GeneSeek ® Genotyping Profile, GGP, Illumina, San Diego, CA). After QC, a total of 38,268 effective SNP were retained for subsequent genomic analyses. Pedigree data of 10,799 Merino sheep was used and the Off-Diagonal correlation between genomic information and pedigree data was 0.86.

The information on the most relevant SNP found in these regions about % var, shows that there are nine candidate genes for the temperament trait, which are: PYGM (21:42,295,599-42,307,126), SYVN1 (21:42,650,559-42,655,527), CAPN1 (21:42,712,976-42,740,799), LOC101110773 (10:29,275,771-29,457,586), LOC101110521 (10:28,986,741-29,188,660), GRID2 (6:30,768,380-31,534,647), FADS1 (21:39,652,537-39,665,108), SYT7 (21:39,390,965-39,426,334) and the GPRIN3 gene (6:35,511,293-35,513,635) (**Figure 4.2**).

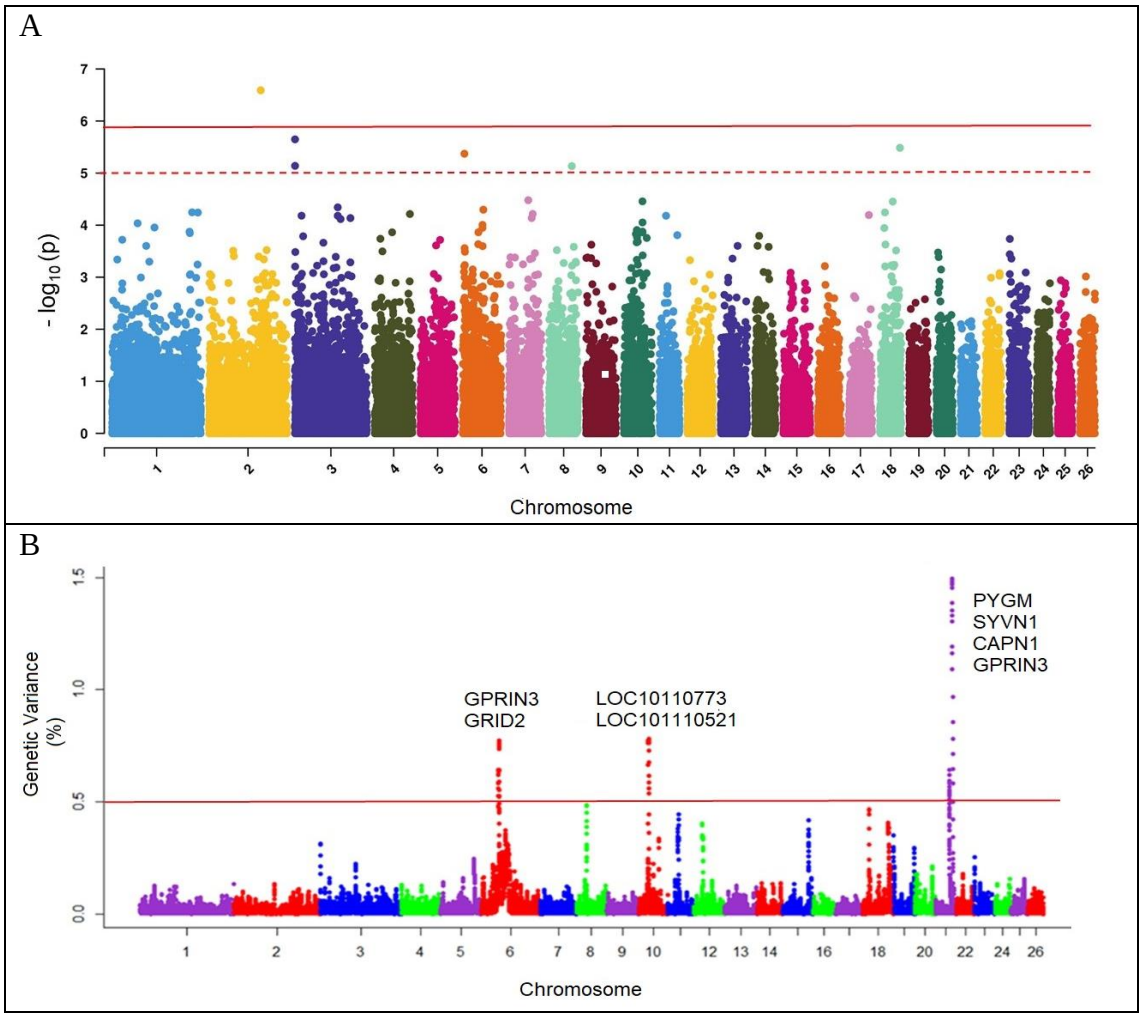


Figure 4.2 Manhattan plot of 38,268 effective SNP for 1,697 Australian Merino genotyped animals.

Ewes, INIA´s Nucleus, production, and reproduction

Data from approximately 5,700 mixed-sex yearling progeny and 2,000 mixed-age ewes born in a single Merino flock between 1999 and 2019 were analysed. The Merino flock was established in 1999, from 475 ewes provided by Uruguayan Merino stud breeders or commercial farmers. Each year, ewes were managed as a single flock and inseminated with either imported semen (from Australia and New Zealand) or flockborn rams. The selection of replacements was based on phenotypic and genetic criteria. During the first 10 years, the selection objective of this flock was to reduce Fiber Diameter (FD) (to produce 19.0 µm or finer wool) while allowing for a slight loss in CFW, whereas from 2011 to 2020, the breeding objective was focused on maintaining FD (less than 15.5 µm), while increasing both CFW and Live Weight (LW).

All genetic parameters and the Genome-Wide Association Studies (GWAS) were estimated utilizing the Julia for Whole-Genome Analyses Software (JWAS) software (Cheng et al., 2018). Estimates of (co) variance components were obtained utilizing a Bayesian method based on Markov chain Monte Carlo (MCMC) sampling (Blasco, 2017). After applying the quality control measures, 1,133 animals and 40,036 SNPs were retained and utilised in the analyses. For GWAS a single-step Bayesian approach with Bayes C linear mixed model that included the

genotyped and non-genotyped ewes was constructed. A total of 70,000 iterations were run after a burn-in of 5,000 cycles and a sampling interval each 10 interactions for both analyses. Results from GWAS highlighted a total of 13, 22, 42, 22, 24 and 29 genomic regions were significantly associated with the fibre diameter, clean fleece weight, live weight at mating, body condition score at mating, pregnancy rate and lambing potential, respectively. The GWAS results are presented as the proportion of additive genetic variation explained by the windows of 20 consecutive SNPs. Results for live weight and body condition score are presented in **Table 4.1.** Results for fibre diameter, clean fleece weight are presented in **Table 4.2.**

We detected several genes, some of which were novel, showing potential associations with the wool (IGF-1, TGFB2R, PRKCA), live weight (CAST, LAP3, MED28, HERC6), body condition score (CDH10, TMC2, SIRPA, CPXM1) or reproduction traits (ADCY1, LEPR, GHR, LPAR2) of mixed-age ewes. These results require validation using a larger dataset before their implementation in genomic selection among Uruguayan Merino sheep. Overall, our findings will be useful for further genomic studies and genetic improvement programs in Uruguay.

Complete results are presented in: Ramos, Z.; Garrick, D.J.; Blair, H.T.; Vera, B.; Ciappesoni, G.; Kenyon, P.R. Genomic Regions Associated with Wool, Growth and Reproduction Traits in Uruguayan Merino Sheep. Genes, 14(1), 167; https://doi.org/10.3390/genes14010167

Table 4.1. Chromosome, location, proportion of additive genetic variance (PVE, %) and candidate genes within the top 10 windows associated with the live weight (LWM) and body condition score (BCSM) at mating of Merino ewes.

Trait	Chr	Window Bounds	PVE (%)	Candidate Genes
LWM	5	93,416,569–93,461,942	1.54	*CAST*
	6	36,905,457–37,129,550	5.67	*LAP3, MED28*
	6	36,295,216–36,872,516	4.43	*BBS7, HERC6, CCNA2, LOC101120495*
	6	35,191,867–35,728,962	2.05	*GPRIN3, TIGD2*
	10	57,396,545–59,143,370	1.41	-
	11	13,795,276–16,306,848	8.68	*LOC101110777, AP2B1, CCT6B, ZNF830*
	13	85,702,745–58,406,417	0.74	*GPR158, MKX, SYNDIG1, PREX1*
	16	54,796,169–58,047,574	0.76	*MYO10, CPEB4*
	22	55,395,817–44,154,558	0.76	*A1CF, CTBP2, GRK5, XPNPEP1, CFAP46, DOCK1, INSYN2A, LIPA, MUOF, PLCE1*
	23	36,106,915–38,820,448	3.53	*MYOM1, DLGAP1, SMCHD1*
BCSM	1	173,862,929–190,851,636	0.88	*ATP6V1A, CD200, ATG3, CFAP44, CCDC191, NECTIN3, NEPRO, PLCXD2, SLC9C1*
	2	111,201,892–114,207,253	2.21	*HPGD, ECPAS, FBXO8, GLRA3*
	2	114,746,188–116,378,689	0.65	*GALNTL6*
	9	57,388,779–56,572,469	3.35	*STMN2, TPD52, ZBTB10*
	10	56,843,164–69,129,301	1.14	*LOC101115632, SPATA13*
	12	36,958,022–40,040,678	0.69	*FMO1, FMO2, FMO4, MTHFR, MFN2, PRRC2C, TNFRF1B*
	13	51,269,879–54,158,422	13.89	*TMC2, SIRPA, CPXM1, KCNQ2, RBBP8NL, DNAAF9*
	13	26,807,364–30,712,871	5.61	*ITGA8, FRMD4A, MINDY3, RSU1, ANKEF1, CUBN, FAM171A1, PTER, PRPF18, TRDMT1*
	13	42,377,205–45,741,876	4.68	*PGF2, LOC106990122, LOC101108592*
	20	52,019,528–6,887,963	4.33	*KHDRBS2, F13A1, GMDS, CDYL, HCRTR2, LRRC1, LOC101114063*

Chr, chromosome; PVE (%), proportion of additive genetic variance explained by each window; LWM and BCS, live weight and body condition score at mating, respectively.

Table 4.2. Chromosome, location, proportion of additive genetic variance (PVE, %) and candidate genes within the top 10 windows associated with the adult fibre diameter (FD microns) and greasy fleece weight (CFW in kg) of Merino ewes at pre-lambing shearing.

Trait	Chr	Window Bounds (bp)	PVE (%)	Candidate Genes
A_FD	1	221,133,506–241,462,240	0.66	-
	2	132,440,576–134,761,891	0.28	*HOXD10, OLA1, SP9, CHN1, CHRNA1, MTX2*
	3	171,178,810–174,258,094	0.29	*IGF-1, PAH, STAB2, NT5DC3, GLT8D2, SLC41A2, TDG*
	3	191,554,038–194,804,430	0.26	*PDE3A, C2CD5, ST8SIA1, HPCAL1, KCNJ8, PYROXD1, SLCO1C1*
	8	35,247,461–84,539,654	0.52	*ESR1, PLEKHG1, NT5E, NHSL1, ANKRD6, CGA, PLEKHG1*
	8	66,612,859–69,354,718	0.51	*ADGRG6, PHACTR2, UTRN, VTA1*
	9	48,994,813–16,528,100	0.26	*PRDM14, WWP1, EXT1, MATN2, PTDSS1, ZNF704*
	15	47,505,272–4,779,161	0.34	*DYNC2H1, LOC101106199, LOC101105437*
	16	52,688,043–60,285,005	0.68	*ARHGAP22, CDH18, TSNAX*
	25	36,631,465–40,822,020	0.30	*WAPL, GRID1*
A_CFW	1	111,294,404–122,825,051	0.58	*UHMK1, DDR2, NUF2, ATF6, INPP5B, RGS5*
	6	36,295,216–36,872,516	0.72	*BBS7, HERC6, CCNA2, LOC101120495,*
	6	36,066,911–36,286,475	0.56	*HERC3, HERC5, HERC6, PYURF, PIGY*
	6	36,905,457–37,129,550	0.51	*LAP3, MED28,*
	6	35,191,867–35,728,962	0.50	*GPRIN3, TIGD2*
	6	33,844,752–35,184,703	0.42	*MMRN1, CCSER1,*
	6	37,767,491–38,052,441	0.41	-
	9	77,283,695–85,378,072	0.61	*STK3, MTDH, MATN2, OSR2, VPS13B*
	11	66,432,553–10,722,809	0.50	*PRKCA, DHX40, LOC101102402, COIL, INTS2, PPM1E, SRSF1*
	19	4,811,675–54,605,752	0.45	*PBRM1, TGFBR2, BAC5, RBM6, CACNA2D3, DCP1A, MAP4*

Chr, chromosome; PVE (%), proportion of additive genetic variance explained by each window; A_FD = adult fibre diameter (microns) and A_CFW greasy fleece weight (in kg)

Corriedale Breed

In Uruguay, most prevalent parasites in sheep are *Haemonchus contortus* and *Trichostrongylus colubriformis.* One strategy to reduce the negative impact of GIN is by selection of genetically resistant animals, because GIN resistance is a moderately heritable trait. In Uruguay, since 1994, genetic resistance for GIN is included in the Corriedale National Genetic Evaluation using faecal egg counts (FEC) measured in lambs as selection criterion. Since 1999, the Merino breed start the genetic evaluation for FEC. Identification of quantitative trait loci (QTL) associated with GIN resistance or susceptibility could improve the selection process and the understanding of biological processes related to host immune response. Genome-wide association studies (GWAS) are used to identify genomic regions where candidate genes associated with a phenotypic trait are located.

The aim of this study was to identify genomic regions associated with GIN resistance in Corriedale sheep by single-step genome-wide association studies (ssGWAS) using 170, 507 and 50K single nucleotide polymorphisms (SNPs). Analysis included 19,547 lambs with faecal egg counts (FEC) records, a pedigree file of 40,056 animals and 454, 711 and 383 genotypes from 170, 507 and 50K SNPs, respectively. Genomic estimated breeding values (GEBV) were obtained with single-step genomic BLUP methodology (ssGBLUP), using a univariate animal model, which included contemporary group, type of birth and age of dam as class fixed effects and age at FEC recording as covariate. The SNP effects as wells as p-values were estimated with POSTGSF90 program. Significance level was defined by a chromosome-wise False Discovery Rate of 5%.

Significant genomic regions were identified in chromosomes 1, 3, 12 and 19 with the 170 SNP set, in chromosomes 7, 12 and 24 using the 507 SNP chip and only in chromosome 7 with the 50K SNP chip. Candidate genes located in these regions, using Oar_v4.0 as reference genome, were TIMP3, TLR5, LEPR and TLR9 (170 SNPs), SYNDIG1L and MGRN1 (507 SNP chip) and

INO80, TLN2, TSHR and EIF2AK4 (50K SNP chip). These results validate genomic regions associated with FEC previously identified in Corriedale and other breeds and report new candidate regions for further investigation.
Full results are presented in: Carracelas, B.; Navajas, E.A.; Vera, B.; Ciappesoni, G. Genome-Wide Association Study of Parasite Resistance to Gastrointestinal Nematodes in Corriedale Sheep. Genes 2022, 13, 1548. https://doi.org/10.3390/genes13091548

Texel breed

New data was recorded at INIA's Nucleus at Las Brujas and background data come from Central Progeny Testing (CPT) of the Texel breed. The CPT was continued at INIA from 2015 with the main aim of facilitating genetic linkage between studs-flocks and allows genetic evaluation of carcass and meat quality traits. In vivo data come from connected commercial flock as well. Two traits were analysed, *in vivo* fat depth (mm) by ultrasound and body weight at scanning (average age of 265±36 days), 9,011 phenotypic records, 12,675 animals in pedigree and 675 genotyped (50K Illumina and Affymetrix) animals were included. After QC 21,990 ere included. A total of 39 SNPs were found significantly ($-\log10 > 3$) associated with fat depth in chromosomes 1, 2, 3, 5, 6, 9, 10, 14, 17, 18, 19, 21 and 23 (Figure 4.4). For body weight only one SNP in Chr 3 shows significative association and 8 SNPs has a $-\log10 > 2.5$ in chromosomes 1, 4, 5, 15, 21 and 25.

Comparison of results in Uruguayan populations

The results of the different Uruguayan studies conducted independently are summarised in Annexe1. Only the nearby regions associated with different characteristics within each breed are included.
Noteworthy the following regions were associated with both resilience and efficiency traits in the same breed (Merino), with a percentage of variance explained over 1% for at least one trait:

- OAR5 (93.2-93.4 Mb): Live weight, fat depth, BCS, FAMACHA and FEC
- OAR6 (31.4-37.5 Mb): Live weight, fat depth, BCS, greasy fleece weight (wool), FAMACHA, FEC and temperament
- OAR11 (25.29-27.37 Mb): fat depth, BCS, FAMACHA and FEC
- OAR21 (39.13-42.71 Mb): fat depth, BCS, FAMACHA, FEC and temperament. Interestingly in the Texel breed we found a QTL in the same OAR21 region for fat depth and body weight (OAR21 – 35.7-42.4 Mb)

Accordingly, the following four regions were chosen for further insight (section 6 and 7).

Gwas in Romane meat sheep

Body Weight (BW- kg) and Body Condition Score (BCS) measurements were recorded from 1034 Romane ewes during three production cycles between 2006 and 2019 at the INRAE La Fage Experimental Farm in the south of France. BW and BCS was collected in eight physiological stage: 1) mating M, 15 days before mating; 2) early pregnancy - Pa, 39 ± 11 days after mating; 3) two-thirds of pregnancy - Pb, 101 ± 11 days after mating; 4) lambing -L, 5) early

suckling - Sa, 17 ± 10 days after lambing; 6) middle of the suckling period - Sb, 42±10 days after lambing; 7) weaning −W, 80 ± 10 days after lambing; and 8) post-weaning period -Wp, 149 ± 11 days after lambing. BCS was collected by scoring on a scale of 1 to 5 with increments of 0.1 from the adapted scale described by Russel et al., 1969. After the quality control 16,236 records were included on analysis.

Genotyping of individual SNPs was performed using the OvineSNP50 BeadChip (Illumina inc). Quality control excluded indels, non-autosome SNPs, SNPs with minor allele frequency (MAF < 0.05), individual and SNP missingness (0.05), sex discrepancy, and deviation from hardy-Weinberg equilibrium (HWE – 0.0005). Finally, the filtered reference data set, comprising 45,638 SNPs. Population structure analysis was performed in order to avoid false positive signal on GWAS analysis using PLINK.

SNPs effects were predicted by using the preGSf90 from BLUPf90 family software (Misztal et al., 2014), considering repeatability animal model and ssGBLUP, which combined pedigree and genomic information and included the parity order, age of first lambing, the litter size class and year of measurements as fixed effects. The additive genetic and permanent environmental effects were included as random components. The ssGBLUP has the same model as BLUP, except for the inverse of numerator relationship matrix A^{-1} , which was replaced by matrix H^{-1}. Gwas results are report as P-Value of individual SNP according to Aguilar et al., (2019). To control false positives, a Bonferroni correction was fitted to take into account multiple tests. One significant association reached the threshold (p-value < 10^{-6}) mapped on OAR1 and was associated with BCS. The related genes within 50-K base pairs upstream and downstream of the physical position of the significant SNP (pos: 40,821,987 bp) locus were examined. The LEPR gene occurred in a candidate region on OAR 1 that have a key role in the regulation of whole-body energy balance by acting on the central nervous system and influencing fat deposition in animals through the control of appetite and energy expenditure (Mace et al., 2022), reproductive seasonality traits (Lakhssasssi et al., 2020) and litter size (Mace et al., 2022) in sheep.

Although it was not significant, a two SNPs hit was observed in the same region for Body weight (**Figure 4.3**). The maximum was reached at position 32.66 Mb, however several SNPS were also close to significance closer to the BW QTL (40.8 Mb) around 39 Mb. Therefore, this region (39-41 Mb) was chosen for further insight (section 5).

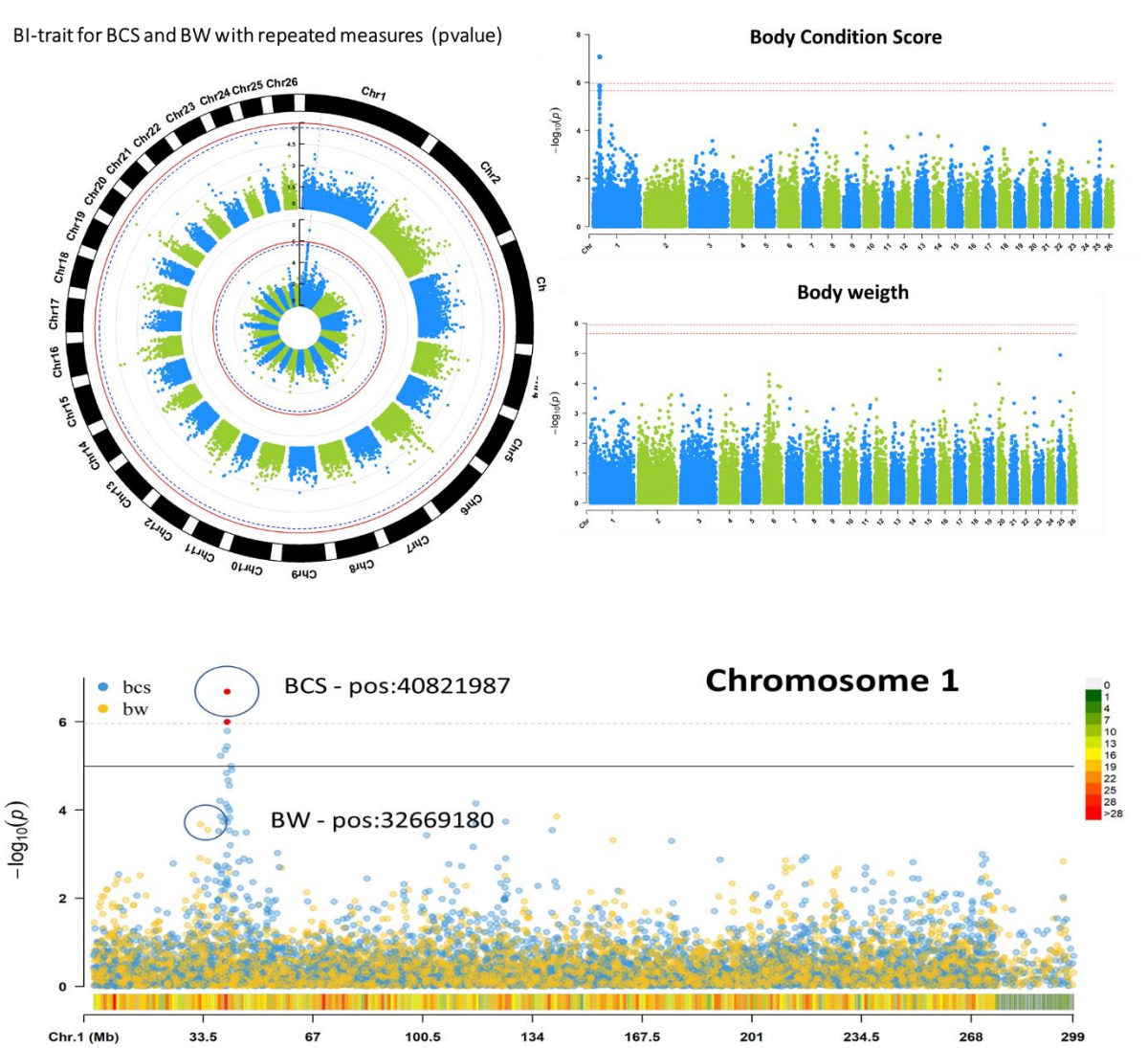


Figure 4.3 Results for bi-trait GWAS in 1034 Romane sheep both Body condition Score (BCS) and Body weight (BW).

Gwas in UK Texels

New GWAS in UK Texel (N= 10,193) for health (footrot and mastitis) and production (birth weight, weaning weight, scan weight, and fat and muscle depth) traits were performed (for WP2 and WP3). Results were drafted as a paper and are under review in Animal (Kaseja et al., "Genome-wide association study of health and production traits in meat sheep").

GWAS was performed to identify informative markers associated with health (footrot and mastitis) and production (birth weight, weaning weight, scan weight, and fat and muscle depth) traits using the available phenotypic and Single Nucleotide Polymorphism (SNP) data collected on the UK Texel sheep population. Initially, 10 193 genotypes were subject to quality control, leaving 9 505 genotypes for further analysis. Selected genotypes, recorded on four different Illumina chip types from low density (15k SNPs) to high quality (606 006 SNPs), were imputed to subset of 45 686 markers from 50k array, distributed on 27 chromosomes. Phenotypes collected on 32 farms across the UK for footrot and mastitis and extracted from the UK

National database (iTexel) for the production traits were used along with pre-estimated variance components to obtain de-regressed breeding values and used to perform GWAS.
The GWAS was performed using multi-locus mixed model algorithm (Segura et al. 2012) implemented in the R (R Development Core Team, 2011) package 'statgenGWAS' (Bart-Jan van Ross, 2022 ; https://cran.r-project.org/web/packages/statgenGWAS/vignettes/GWAS.html).
An additive genetic model was used. The input (de-regressed EBVs) was calculated, by adjusting for fixed effects using **BLUP** (Best Linear Unbiased Predictor) which already accounted for fixed and random effects (hence no covariate matrix in GWAS was used). Bonferroni correction was applied to obtain significance thresholds for each analysed trait. Markers were considered as being significant at the genome-wise level if the $-\log_{10}(\text{p-value}) > -\log_{10}(0.05/n)$, where n=45 686 (number of markers), giving the threshold of 5.96. Chromosome-wise significance of the marker was assessed by having $-\log_{10}(\text{p-value}) > -\log_{10}(0.05/n)$, where n is the number of markers on each chromosome.

Results showed three SNPs being significant on the genome-wise level ('OAR8_62240378.1' on chromosome 8 for birth weight, 's14444.1' on chromosome 19 for weaning weight and 's65197.1' on chromosome 23 for scan weight (Table 4.3). Fourteen subsequent SNPs were found to be significant at the chromosome-wise level (Table 4.4). Regrettably none of the significant SNPs identified in the current study were within a common genomic region less than 8 MB

Table 4.3 Summary of genome-wise significant SNPs in UK Texel sheep

Trait	SNP name*	Chromosome	bp	–log10(p-value)	Variance explained (%)
BWT	OAR8_62240378.1	8	57911837	6.59	0.21
WW	s14444.1	19	18951998	6.82	0.11
SWT	s65197.1	23	58809336	7.43	0.12

BWT - birth weight, WW - weaning weight, SWT - scan weight,

Table 4.4. Summary of chromosome-wise significant SNPs in UK Texel sheep

Chromosome	Chromosome-wise significance threshold	Trait	SNP name	–log10(p-value)
3	4.92	FD	OAR3_192372203.1	5.67
10	4.50	SWT	OAR10_9706403.1	4.65
11	4.30	MD	OAR11_32227171.1	4.89
		BW	OAR11_5688066.1	4.53
14	4.29	SWT	s22731.1	4.87
		CMT	s08817.1	5.16
15	4.45	WW	s60057.1	4.65
16	4.43	WW	OAR16_52790463.1	4.48
17	4.39	CMT	OAR17_32936496.1	4.51
19	4.31	FRT	OAR19_22759405.1	4.43
21	4.19	SWT	OAR21_28724590.1	5.50
			OAR21_31718415.1	4.21
22	4.26	WW	OAR22_13469217.1	4.38
23	4.28	FRT	s36409.1	5.50
26	4.21	FRT	s37597.1	5.23

FRT - footrot, CMT - california mastitis test, BWT - birth weight, WW - weaning weight, SWT - scan weight, MD - muscle depth, FD - fat depth

Gwas in Churra dairy sheep

In order to study potential genetic antagonisms (trade-offs) between production traits and resistance or tolerance to disease the UNILEON group performed a 4-multitrait genome-wide association analysis (GWAS) to assess the trade-offs between mastitis resistance (SCS) and milk production traits in the autochthonous Spanish Churra sheep breed. Data for the traits of interest (milk yield, fat and protein percentage and somatic cells score, SCS) were obtained from the Churra Breeders' Association, ANCHE. A total of 2958 animals were genotyped with the 50K OvineSNP50 Beadchip (Illumina, San Diego, CA). After debugging the phenotype database, estimating basic statistics, defining the analysis model, and quality control of the genotypes, the GWAS for SCS and Milk production Traits (n = 2958) was performed with the BLUPF90 family of programs.
GWAS for the four traits did not identify clear pleiotropic regions, due to the lack of genome-wise significant effects detected for the SCS trait. Hence, this analysis suggests a lack of pleiotropic segregating effects between milk production traits and resistance to mastitis in the studied population of Churra dairy sheep. These results agree with the low genetic correlations previously reported in Churra sheep of SCC with milk protein content (rg = 0.13) and milk fat content (rg = 0.04), although a moderate and negative correlation had been reported with milk yield (rg = 0.36) (Othmane et al., 2002).

Gwas in Lacaune sheep

Commercial farms

Gwas was previously carried out in commercial Flocks of French Lacaune and published (Rupp et al., 2015). Briefly QTLs associated with mastitis susceptibility were mapped by performing a genome scan (26 autosomes) in 1009 dairy sheep distributed in 33 half-sib families. The trait pertaining to mastitis susceptibility was the lactation average of somatic cell score (LSCS) as this trait is highly correlated with intra mammary infections. All animals were genotyped with the 50K OvineSNP50 Beadchip (Illumina, San Diego, CA). Haplotype-based linkage and association analyses were used to detect QTLs on at a 5% chromosome-wide threshold using the QTLMap software (Elsen et al., 1999; http://dga7.jouy.inra.fr/qtlmap/). For LA, interval mapping was performed by likelihood ratio test (LRT) using within-sire linear regression. The QTL effect (average substitution effect) was expressed in deviation units (SD) for the trait. GWAS was based on a regression analysis of the phenotypes on founder sires' haplotypes for every haplotype of 4 consecutive SNPs along the chromosome. Chromosome-wide significance levels were calculated with QTLMap, using the current family structure and phenotypes. For LA, the empirical 5% and 1% chromosome-wide significance levels of the test statistics were estimated from 1000 within-family permutations (40) for each chromosome. For GWAS, the empirical chromosome-wide significance level of the test statistics was estimated from 1000 simulations for each chromosome. The genome-wise thresholds were obtained by applying the Bonferroni correction $P_{\text{genome-wise}} = 1 - (1 - P_{\text{chromosome-wise}})n$, where n is the number of chromosomes, i.e., 26 in sheep. Five regions on chromosomes OAR3, 4, 11, 16 and 23 exceeded the 5% genome-wide threshold. The marker order and positions were based on the Ovine

Assembly v3.1 (http://www.livestockgenomics.csiro.au/sheep/oar3.1.php). One highly significant QTL on chromosome 3 (OAR3) was similarly located in the two association and linkage analyses. For this OAR3-QTL, the association study provided haplotypes of contrasting susceptibility based on four consecutive SNP in the following interval: 129,685,397bp - 130,103,393 bp. Fine mapping of the region, using full sequencing with 12X coverage in three animals, provided one strong candidate SNP that mapped to the coding sequence of a highly conserved gene, suppressor of cytokine signaling 2 (Socs2). Additionally, the size, weight and milk production in R96C homozygote sheep, were significantly increased by 24%, 18%, and 4.4%, respectively, when compared to wild type sheep, supporting the view that the point mutation causes a loss of SOCS2 functional activity. Altogether these results provide strong evidence for a causal mutation controlling mastitis in sheep and highlight the major role of SOCS2 as a tradeoff between the host's inflammatory response to mammary infections, and body growth and milk production, which are all mediated by the JAK/STAT signaling pathway.

Oget et al. (2019) further confirmed the QTL on LSCS, using a 960 custom-designed ovine single nucleotide polymorphism (SNP) chip in 504 Lacaune. Additional efficiency phenotypes were considered : milk yield, fat content, protein content and weight. The arithmetic averages of the first lactation test-days were then computed and corrected for year of sampling for fat content (FAT_L1), protein content. Milk yield was the 250 days cumulative milk yield. Weights were also available at birth, at 100 days and 250 days, after the and second lambing and at the age of 920 day. Weight phenotypes were corrected for year and feeding method (breastfeeding or artificial suckling). Regarding resilience traits *Staphylococcus spp.* abundance in milk was measured at three-time points during the first lactation by a qPCR- based technique. The three results were averaged for each ewe and corrected for the effects of month and year of sampling. Chronic mastitis was based on the presence of mammary abscesses, recorded by clinical examination. The 504 Lacaune sheep were genotyped with the 960 custom-designed chip which was designed and developed within the 3SR EU project, and included the SOCS2 SNP. GWAS were performed for each phenotype using the polygenic univariate mixed model approach implemented in the genome-wide efficient mixed-model association (GEMMA) software. The polygenic effect was fitted using a covariance structure according to the genomic relationship matrix. Corrections were applied to account for multiple testing.
Results from the GWAS highlighted the highly significant region on OAR3 in the Lacaune population. The most significant SNP (rs868996547, p value = 3.0E−07, position = 129,722,200 bp) was the causal mutation in the SOCS2 gene, previously reported by Rupp et al. (2015]. The lowest p values and highest estimates of effects for this SNP were observed for both mastitis traits (SCSC and Staphyloccocci) and four of the six weight traits. Corresponding effects varied from 0.33 SD for weight at 100 days to 0.50 SD for SCS.

La Fage Experimental farm:

In Lacaune, a complementary study was carried out specifically for this task in SMARTER using background data. From 2014 to 2022, dairy production traits of 1069 primiparous ewes reared

in the INRAE Experimental Unit of La Fage (UE 321 agreement A312031, Roquefort, France) were collected.

For each of the 5 milk controls at morning and evening milkings, milk yield (MY), milk somatic cell count (SCS), and milk fat and protein contents (FC and PC, respectively) were quantified at the Interprofessional Milk Analysis Laboratory (Agrolab's Aurillac, France). From these morning and evening milk records, the MIR spectra were retrieved to predict the fine profile of the milk proteins, i.e., the four caseins: alpha-S1-casein (α_{s1}-CN), alpha-S2-casein (α_{s2}-CN), beta-casein (β-CN) and kappa-casein (κ-CN), and two whey proteins: alpha-lactalbumin (α-lactalbumin) and beta-lactoglobulin (β-lactoglobulin), and of the milk fatty acids (FA), i.e., saturated FA, such as butyric acid (C4:0), caproic acid (C6:0), caprylic acid (C8:0), capric acid (C10:0), lauric acid (C12:0), myristic acid (C14:0) and palmitic acid (C16:0), and unsaturated FA, such as oleic acid (*cis-9* C18:1), rumenic acid (*cis-9 trans-11* C18:2) and alpha-linolenic acid (C18:3*n-3*) (Ferrand et al., 2012; Ferrand-Calmels et al., 2014), for the milk proteins and FA, respectively)

All the ewes were genotyped using a medium-density single nucleotide polymorphism (SNP) chip (Illumina OvineSNP50 BeadChip: 54,241 SNPs). After quality control, 42489 autosomal SNPs positioned on the 26 *Ovis aries* (OAR) autosomes and mapped to the *Ovis aries* genome assembly Oar_v3.1, plus one SNP corresponding to the Socs2 gene mutation (OAR3-129,722200 bp). The GWAS was performed using the single-trait model for each of the 5 milk recoding controls, taking into account the year of milk sampling and the stage of lactation of ewes. After solving the single step GBLUP model, all SNP effects were estimated back solving the breeding values estimates as proposed by Aguilar *et al*. (2019) using POSTGSF90 (Misztal *et al*., 2002). Multiple testing was corrected by using the false discovery rate (FDR), and a genome-wide SNP significance threshold of $P < 0.05$ was considered.

Depending on the milk recording control, between 2 to 9 QTL regions have been identified, involving between 4 to 16 traits. Overall, the number of QTLs increases with the stage of lactation. We have found QTL on OAR 6 for α_{s2}-CN, with a maximum around 84 Mbp, according to the milk recording control. This QTL was already identified on Lacaune (Boichard et al., 2014; Martinez Boggio et al., 2022) and Churra (Diez-Tascon et al., 2001). Regarding proteins, some punctual but significant signals were observed for α_{s1}-CN (OAR 9 at 46 Mbp), for α-lactalbumin (OAR 4 at 19.6 Mbp) and for β-lactoglobulin (OAR12 at 41 Mbp). We also identified a QTL on OAR 17 for short-chain saturated fatty acids (C4:0 and C6:0) that was significant in all dairy controls for the same SNP (OAR17_54879473.1, pos 50,418,326). This QTL is closed to the one we found (Martinez Boggio et al., 2022) for C6:0 and C8:0 in a subset of the data. García-Gámez et al. (2012) already reported a significant SNP (58.8 Mbp) on OAR17 that was associated with fat percentage. Significant QTL for C12:0 on OAR1 (at 260 Mbp), C14:0 on OAR10 (at 19 Mpb) and USFA C18:2 and C18:3 on OAR20 (at 3 Mbp) were observed. The most remarkable QTL region was on OAR3 at the 4th milk recording control. For 12 traits, significant QTL were observed, with the most significant SNP was either SOCS2 (pos: 129,722,200 bp) for MY, SCS, β-CN or 2 SNP before SOCS2 (OAR3_138290871.1, pos: 129,648,614) for FC, β-lactoglobulin, α_{s1}-CN, C6:0, C8:0, C10:0, C12:0 C14:0 and C16:0. This QTL region, which did not appear in the first 3 milk controls, remained significant in the 5th one, with mainly a maximum signal on SOCS2 SNP for MY and FC, but not for SCS. A paper in under construction by C Marie-Etancelin et al.

As a conclusion in the Lacaune breed, following the new and background analyses we

can suspect a pleiotropic gene (SOCS2) with an adverse effect on mastitis and growth (Rupp et al. 2015; Oget et al., 2019). The local association with fine milk composition traits (quantity, milk fat content, some saturated fatty acids and proteins) requires additional analyses to disentangle pleiotropy from closely linked genes. Accordingly, the following region was chosen for further insight (sections 5, 6 and 7).

Gwas in Manech tete rousse sheep

Oget et al. (2019) performed a GWAS in Manech tete rousse, using the 960 custom-designed ovine single nucleotide polymorphism (SNP) chip. Traits related to efficiency were milk yield, fat and protein contents and the resilience trait was mastitis (SCS). For GWAS analyses, authors used the daughter yield deviations from regular national genetic evaluations formilk production traits (MILK, PROTEIN, and FAT) and lactation average somatic cell scores (LSCS). DYD cor-respond to the average performance of the daughters of a ram, corrected for the environmental effects and the genetic value of the dams. GWAS were performed for each phenotype using the polygenic univariate mixed model approach implemented in the genome-wide efficient mixed-model association (GEMMA) software. The polygenic effect was fitted using a covariance structure according to the genomic relationship matrix. Corrections were applied to account for multiple testing.
Results from the GWAS highlighted a significant region on OAR16 associated with MILK, SCS and protein content at position 30,69 Mb, 30,69 Mb and 31,29 Mb, respectively. Another QTL region was identified on OAR2 but the regions was associated only with milk production trait and not with SCS.

Gwas in French Alpine and Saanen Goats

Milk production, mastitis and udder type

Milk production, mastitis (SCC) and udder type traitsMartin et al. (2018) published for the first-time gwas results in French Saanen and Alpine goats. The study was based on a population of 1,941 Alpine and Saanen females goats genotyped with the Illumina GoatSNP50 BeadChip (Illumina Inc., San Diego, CA). Gwas was performed with both linkage analyses and linkage disequilibrium using QTLmap software (http://dga7.jouy.inra.fr/qtlmap/). Interval mapping was performed with the likelihood ratio test using linear regressions. Traits analysed were mastitis (lactation average of monthly recorded SCS, LSCS), 11 type trait and milk production
Authors reported a large region on chromosome 19 in Saanen which was significantly linked to six traits (Figure 4.4): Milk yield, Fat Yield, protein Yield, two type traits (rear udder attachment and Udder flor position), and LSCS. The most likely position of LSCS (41.3 Mb) however suggested that this was a different QTL. Results from the single-trait approach using LD are presented in Figure 4.4a, and results from the multi-trait approaches using LA and LD are presented in Figure 4.4b. The PCA for the 5 traits provided a value on the first component (PC1) that explained 57% of the over- all variance.

The correlation of PC1 with MY, FY, PY, RUA, and UFP was 0.91, 0.91, 0.94, −0.29, and −0.48 respectively. The LD analysis for PC1 on CHI 19 showed a higher LRT estimate than for the traits analysed separately (Figure 3a, b). The LD analysis for PC1 also reduced the length of the confidence interval to 2.4 Mb (5% CI = 24.5–26.9 Mb) compared with the 4.2-Mb cumulative interval obtained from the 5 separate analyses (5% CI = 22.2–26.4 Mb).

Functional Longevity

A complementary gwas was carried out by Palhiere et al. (2018) for functional longevity. In this study, authors used 298 Saanen and 341 Alpine bucks with phenotypes for functional longevity which were genotyped with the Illumina GoatSNP50 BeadChip. After quality control 46,778 SNPs (Saanen) and 46,780 SNPs (Alpine) distributed on goat autosomes CHI 1 to CHI 29 were analysed. GWAS were performed using the univariate mixed model approach implemented in the Genome-wide Efficient Mixed Model Association (GEMMA) software.
GWAS identified no significant association with functional longevity in the Alpine breed. In Saanen, one region of chromosome 19 was highly significant with 5 SNPs above the 5% genome wise threshold spanning a region from 26.19 to 28.87 Mb (Figure 4.5A). The SNP snp10603-scaffold1377-32250 was the most significant with a p-value equal to 1.94e-09 (-log10(p) = 8.71).

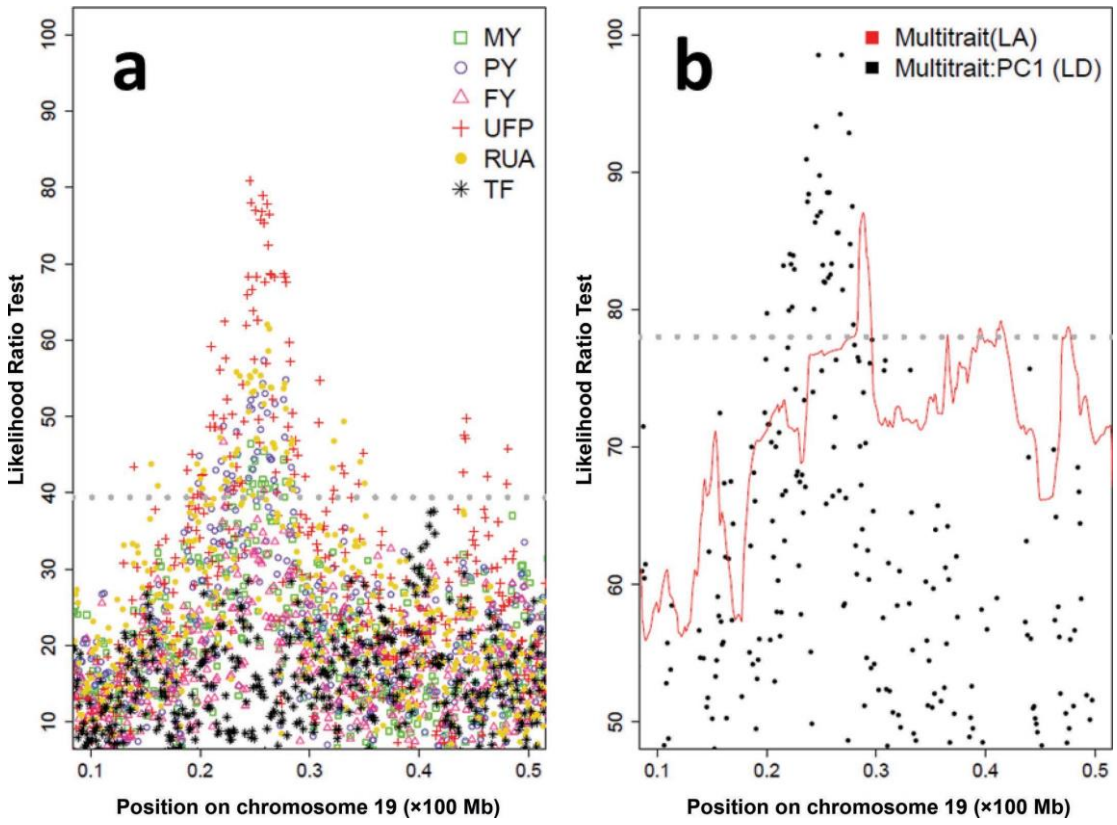


Figure 4.4 Global likelihood ratio test profiles chromosome 19 (CHI19) in the Saanen breed show significant effect of the same region on 5 traits (*Source*= Martin et al., 2018) Results were based on (a) association analyses for udder floor position (UFP), rear udder attachment (RUA), teat form (TF), milk production (MY), fat yield (FY), and protein yield (PY); and (b) multitrait linkage analysis(LA; curve) and association analysis (linkage disequilibrium, LD) for the principal component trait PC1 (■). The PC1 was from a principal component analysis carried out on the 5 MY, FY, PY, RUA, and UFP traits, using the SAS PRINCOMP procedure (SAS Institute, 2008). The horizontal dotted line represents the 5% genome-wide threshold.

Semen production

Talouarn et al. (2020) used the same French commercial population updated with new genotypes (483 Saanen and 629 Alpine) to carry out a gwas on three male semen traits in addition to milk production. Semen traits were recorded on artificial insemination (AI) bucks: semen volume in mL (SV), semen concentration in billions of spermatozoa per mL (SC) and number of spermatozoa in billions of spermatozoa (SN). For GWAS, Talouarn et al (2020) performed imputation with FImpute for 1129 French Alpine and Saanen males using within-breed and French panels on 23,338,436 filtered variants. The imputed sequences were subjected to single-trait association analysis semen production traits using mixed linear models with the mlma option of GCTA software.

Authors reported a large region on chromosome 19 in Saanen which was significantly linked to both semen volume and milk yield (**Figure 4.6**). For milk yield, 313 variants

reached the chromosome significance level ($p - value \leq 1.17 * 10-7$) for milk yield, all of which were situated between 23.55 and 27.68 Mb. For semen traits, a wide significant signal was found on chromosome 19 using within-breed imputation, spanning a region from 24.5 to 27 Mb. The signal was most significant for semen volume for which 209 variants reached the chromosome significance level.

Beta-hydroxybutyrate as proxy for metabolic disorder

Beta-hydroxybutyrate (BHB) was considered in SMARTER as a proxy for metabolic disorder. Indeed, pregnancy toxemia is a common metabolic disease in pregnant small ruminants and is characterized by an increased blood concentration of ketone bodies, especially beta-hydroxybutyrate (BHB), and is caused by inadequate hepatic metabolism of non-esterified fatty acids excessively mobilized from adipose tissue during negative energy balance.

For WP2 (Deliverable D2.5), a gwas was perfomed for newly phenotyped Alpine and Saanen goats. Briefly, the data set came from a total of 338 primiparous Alpine and Saanen goats from 5 different farms, sampled between 2020 and 2022 as part of the SMARTER project by INRAE and CAPGENES partners. The study is also reported in deliverable D2.5. All 338 goats and their 71 sires for BHB were genotyped using the Illumina Goat_IGGC_65K_v2 (59,727 SNPs). GWAS was carried out for each breed separately, using GEMMA software.

No significant SNP was found in the Alpine breed. Two QTL were significant at the chromosome-wise level in the Saanen breed. Noteworthy, one SNP mapped on CHI 19 (snp28387-scaffold303-1179769) at position 31.22 Mb (**Figure 4.5B**).

Summary of finding on CHI19 in Saanen goats

GWAS reported in four different study in French Saanen goats systematically highlighted a genomic region on chromosome 19 that was associated with many traits related to efficiency, resilience and reproduction:

- a region from 40.8 to 42.1 Mb for LSCS (Martin et al., 2018)
- a region from 24.5–26.9 Mb for a multi-trait Principal component combining Milk production, Fat and protein yield and 2 type traits (Martin et al., 2018)
- a region from 26.19 to 28.87 Mb for functional longevity (Palhiere et al., 2018)
- a region from 23.55 and 27.68 Mb for Milk yield (Talouarn et al., 2020)
- a region from 24.5 to 27 Mb for Semen traits (Talouarn et al., 2020)
- a region around 31 Mb for BHB (Smarter WP2)

Accordingly, the region spanning the interval from **23.5-21.5 Mb** (on ARS1 assembly) was used for further insight (section 5,6 and 7).

A – Functional longevity

B - Beta-hydroxybutyrate as proxy for metabolic disorder

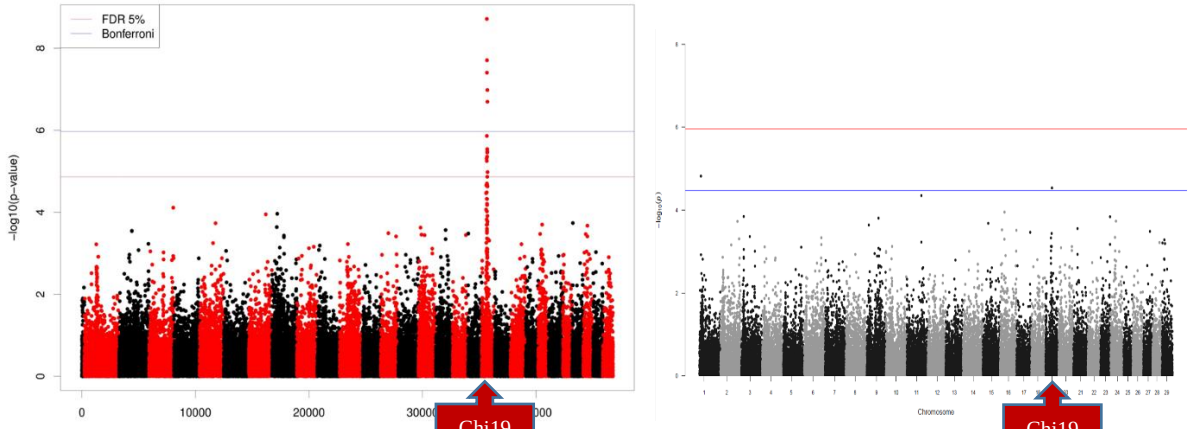


Figure 4.5 Manhattan plot from GWAS in French Alpine goats for (**A**) functional longevity (source: Palhiere et al., 2018) and (**B**) Beta-hydroxybutyrate as proxy for metabolic disorder (source: Palhiere et al., smarter, personal communication)

A - Semen volume

B - Milk production

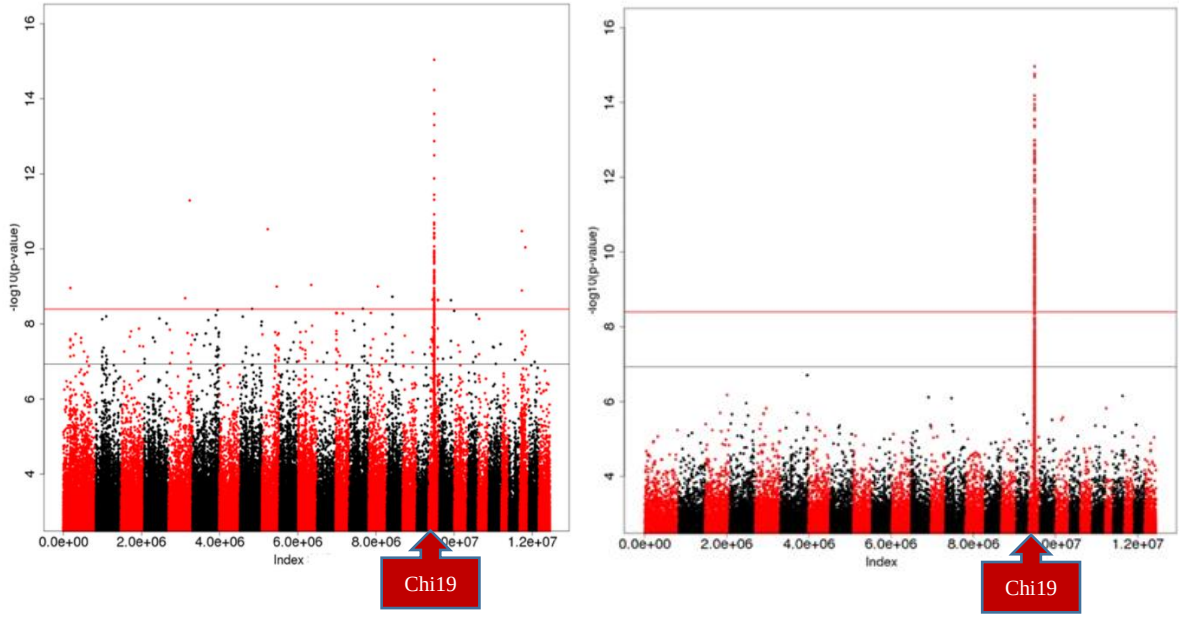


Figure 4.6 Manhattan plot from Gwas for semen volume (**A**) and milk production (**B**) in French Saanen goat (*source*: Talouarn et al., 2020)

Summary of pleiotropic QTL identified from GWAS

The regions we could qualify as possible pleiotropic QTL regions, i.e. a chromosome area of less or equal to 5 Mb including a QTL for both R&E trait, are summarised in the following **Table 4.5** and were subjected to further analyses (section 5, 6 and 7)

Table 4.5. Population used for GWAS and identification of chromosomal regions associated with both R&E traits, so called possible pleiotropic QTL

	Population Breed (Partner)		**Genotyped Animals** (assembly)	**Efficiency trait**	**Resilience trait**	**Possible Pleiotropic QTL** Chromosome – trait - QTL position (max) Interval to explore further (a)
1	Merino (INIA-UY)	Meat Sheep	2100	Live weight Fat depth Wool traits (greasy fleece weigh)	FEC, FAMACHA, temperamentBCS	OAR5 – fat depth -93.23 Mb OAR5 – BCS-93.23 Mb OAR5 – FAMACHA -93.23 Mb OAR5 – Live weight -93.42 Mb OAR5 – FEC2 -93.45 Mb OAR6 – FEC1 -35.49 Mb OAR6 – temperament -35.49 Mb OAR6 – FAMACHA -36.26 Mb OAR6 – greasy fleece weigh -36.29 Mb OAR6 – Live weight - 36.90 MB OAR6 – BCS- 37.27 Mb OAR6 – fat depth - 37.27 Mb OAR6 – FEC2 -37.53 Mb OAR11 – FAMACHA - 25.29 Mb OAR11 – fat depth - 25.29 Mb OAR11 – BCS - 25.29 Mb OAR11 – FEC2 - 25.29 Mb OAR11 – FEC1 - 27.26 Mb OAR21 – BCS- 39.13 Mb OAR21 – fat depth - 39.13 Mb OAR21 –FEC2- 39.65 MB OAR21 – FAMACHA - 42.26 Mb OAR21 – temperament - 42.29 Mb OAR21 – FEC1 - 42.71 Mb

						Interval = OAR5 93-93.5 Mb **Interval = OAR6 35.4-37.6 Mb** **Interval = OAR11 25-27.5 Mb** **Interval = OAR21 39-42.8 Mb**
1	Corriedale (INIA-UY)	Meat Sheep	1548		FEC	none
1	Texel (INIA-UY)	Meat Sheep	657	Weight, carcass and meat quality	FEC	OAR21 - fat depth 35.7 Mb OAR21 - body weight 42.4 Mb
2	Romane (INRAE& CNR)	Meat Sheep	1030	Weight	BCS	OAR1 - Weight 32.67 Mb OAR1 - BCS 40.82 Mb **Interval = OAR1 39-41 Mb**
3	UK Texel (SRUC & TexelS)	Meat Sheep	10193	Birth weight, weaning weight, scan weight, muscle depth, fat depth	footrot, mastitis (CMT)	none
4	Churra (UNiLEON)	Dairy Sheep	2958	Milk yield, fat and protein content	Mastitis (SCS)	none
5	Lacaune (INRAE) [**B**][2,3]	Dairy Sheep	1009 504	Milk yield, Fat and Protein % Weight (at birth, 100, 250 92 Days, 1rst and 2nd lambing)	Mastitis (SCS) Staphylococci Chronic mastitis (abces)	OAR3 – SCS 129.94 Mb OAR3 – SCS 129.7 Mb OAR3 – Staph 129.7 Mb OAR3 – W_100d 129.7 Mb OAR3 – W_250d 129.7 Mb OAR3 – W1rst lambing 129.7 Mb OAR3 – W_920d 129.72 Mb **Interval = OAR3 129-130 Mb**
6	Lacaune (INRAE) [**B**][1]	Dairy Sheep	1069	Milk yield, fat content, β-CN, α_{s1}-CN, β-lactoglobulin, predicted fatty acids : C6:0, C8:0, C10:0,C12:0 C14:0, C16:0	SCS	OAR3 – SCS 129.72 Mb OAR3 – MY 129.72 Mb OAR3 – β-CN 129.72 Mb OAR3 – FC 129.64 Mb OAR3 – α_{s1}-CN 129.64 Mb OAR3 – β-lacto 129.64 Mb OAR3 – C6:0 129.64 Mb OAR3 – C8:0 129.64 Mb

						OAR3 – C10:0 129.64 Mb OAR3 – C12:0 129.64 Mb OAR3 – C14:0 129.64 Mb OAR3 – C16:0 129.64 Mb **Interval=OAR3 128.9-130.1 Mb**
7	Manech Tete Rousse [**B**][3]	Dairy Sheep	145	Milk yield, fat and protein content	Mastitis (SCS)	OAR16 – SCS- 31,29 Mb OAR16 – Milk yield - 30,69 Mb OAR16 – Protein content- 30,69 Mb **Interval = OAR16 30-32 Mb**
8	Saanen & Alpine - goats (INRAE) [B]4	Dairy goat	1941	Milk yield, fat and protein yield	Mastitis (SCS) Udder type	*Saanen*: CHI19- LSCS 41.1 Mb CHI19- multi-trait Principal component combining Milk production, Fat and protein yield and 2 type traits 24.5–26.9 Mb
9	Saanen & Alpine -AI males (INRAE)		1112	Milk yield, fat and protein yield	Longevity[**B**][5] Semen traits [**B**][6]	Saanen: CHI19 – Longevity 25.26 Mb CHI19 – Milk Yield 23.55-27.68 Mb CHI19– Semen volume 24.5-27.0 Mb
10	Saanen & Alpine - goats (INRAE, IDELE, CAPGNES)		338		BHB	Saanen: CHI19 – BHB 31.22 Mb **Interval = CHI19 23.5-31.5 Mb**

a) The marker order and positions were based on the Ovine Assembly v3.1 (http://www.livestockgenomics.csiro.au/sheep/oar3.1.php) exept in Corriedale (Oar_v4.0)
[**B**][1] background genotype data
[**B**][2] background published GWAS analyses (Rupp et al., 2015)
[**B**][3] background published GWAS analyses (Oget et al., 2019)
[**B**][4] background published GWAS analyses (Martin et al., 2018)
[**B**][5] background published GWAS analyses (Palhiere et al., 2018)
[**B**][6] background published GWAS analyses (Talouarn et al., 2020)

5. Test for pleiotropy

The "Close LInkage versus Pleiotropism test" (CLIP) developed by David et al. (2013; DOI: 10.1038/hdy.2012.70) was implemented in order to try and distinguish between pleiotropy (a single QTL affecting more than one trait) and/or close linkage (different QTLs that are physically close). Briefly, the test compares two traits and rejects the hypothesis of a pleiotropic QTL if the square of the observed correlation between a combination of apparent effects at the marker level is below the minimal value it can take under the pleiotropic assumption.

OAR5 and 6 in Merino for parasite resistance, Body condition Score and Back fat Thickness

In Merino, the CLIP test was performed for Body Weight (BW) and Body Condition Score (BCS) in the OAR6 QTL region (95 SNPs ranging from 35.4 Mb to 37.6 Mb) and in OAR5 QTL region (77 SNPs ranging from 93 Mb to 93.5 Mb).

The results are presented in Table 5.1. The pleiotropic assumption was rejected in all situations, suggesting that the different traits related to parasite (FEC and Famacha) and efficiency (body condition score and Back fat Thickness) are controlled by different causal mutations and genes.

Table 5.1. Results of the CLIP test in 1695 Merino sheep from Uruguay.

chr	Trait1	Trait2	Correlation limit at 5%	Correlation observed	Significance (corr obs>corr limit)
OAR6	FEC1	FAMACHA	0.7264	0.2926	No, pleiotropism is rejected
OAR6	FEC2	FAMACHA	0.7468	0.338	No, pleiotropism is rejected
OAR6	FEC1	BFT	0.8689	0.7121	No, pleiotropism is rejected
OAR6	FEC1	BCS	0.9223	0.8023	No, pleiotropism is rejected
OAR5	FEC1	FAMACHA	0.6307	0.2603	No, pleiotropism is rejected
OAR5	FEC2	FAMACHA	0.7400	0.2130	No, pleiotropism is rejected
OAR5	FEC1	BFT	0.4123	1.29e-002	No, pleiotropism is rejected
OAR5	FEC1	BCS	0.4946	5.5e-002	No, pleiotropism is rejected

OAR1 in Romane for body weight and Body Condition Score,

In Romane, the CLIP test was performed for Body Weight (BW) and Body Condition Score (BCS) in the OAR1 QTL region (75 SNPs ranging from 39 Mb to 41 Mb).
The results are presented in Table 5.2. The pleiotropic assumption was rejected. The estimated correlation between effects of SNPs for BW and BCS was 0.39, which was clearly below

the threshold for significance (0.52). Therefore, we could conclude that the two BW and BCS QTLs are controlled by different causal mutations and genes.

Table 5.2. Results of the CLIP test in Romane for Body Weight and Body Condition Score in the OAR1 QTL region (75 SNPs ranging from 39 Mb to 41 Mb).

Trait 1	Trait 2	Correlation observed	Correlation limit at 5%	Significance (corr obs>corr limit)
Body Weight	Body Condition Score	0.392	0.521	No

OAR3 in Lacaune sheep for SCS, Staphylococci, weight and fine milk composition traits

First, the CLIP test was performed for SCS, Staphylococci, milk, and weight traits in the OAR3 QTL region (13 SNPs ranging from 129.6 Mb to 130.1 Mb) in the population of 504 Lacaune. The results are presented in Table 5.3. The pleiotropic assumption was never rejected in any of the two-trait analyses with SCS, Staphylococci and weight (250 d), thanks to strong observed correlations of SNP effects, especially between SCS and Staphylococci traits (r = 0.93). This suggests that all these traits may be controlled by the same causal mutation. In contrast, as soon as we included MILK trait, the observed correlations dropped down except for the analysis with W_DAY_250 and the pleiotropic assumption was even rejected for the analysis with STAPH_L1.

Table 5.3. Results of the CLIP test in Lacaune for SCS, Staphylococci, Milk and Weight traits

Trait 1	Trait 2	Observed correlation	Correlation limit at 5%	Significance (corr obs > corr limit)
SCS	Staphylococci	0.93	0.67	Yes
SCS	Weight_D250	0.65	0.58	Yes
SCS _L1	MILK	0.24	0.14	Yes
Staphylococci	Weight_D250	0.52	0.19	Yes
Staphylococci	MILK	6.31e-03	3.58e-02	No
Weight_D250	MILK	0.66	1.09e-02	Yes

On the experimental Lacaune population, the CLIP test was performed for SCS and 13 milk composition traits in the OAR3 QTL region (17 SNPs ranging from 128.9 Mb to 130.1 Mb – table 5.4). The pleiotropic assumption was never rejected when SCS is confronted with milk yield or protein content (Beta-casein, alphaS1-casein or beta-lactoglobulin). On the other hand, for the saturated fatty acids (from 8 to 16 C), the pleiotropic assumption with SCS was systematically rejected. For C6 and fat content, the pleiotropic assumption was not rejected but the observed correlations were very closed to the correlation limit at 5%. These results suggested that the SOCS2 mutation has an impact on protein fractions and milk quantity, but that a possible second gene, closed to SOCS2, has a more specific impact on fat and particularly saturated fatty acids.

Table 5.4. Results of the CLIP test in Lacaune for SCS, and fine milk composition traits

Trait 1	Trait 2	Observed correlation	Correlation limit at 5%	Significance (corr obs > corr limit)
SCS	Milk yield	0.793	0.597	Yes
SCS	FC	0.710	0.708	(Yes)
SCS	β-CN	0.739	0.581	Yes
SCS	α_{s1}-CN	0.677	0.585	Yes
SCS	β-lactoglobulin	0.815	0.559	Yes
SCS	C6:0	0.600	0.575	(Yes)
SCS	C8:0	0.587	0.607	No
SCS	C10:0	0.521	0.597	No
SCS	C12:0	0.494	0.605	No
SCS	C14:0	0.389	0.619	No
SCS	C16:0	0.427	0.578	No

Our results strongly suggest a direct pleiotropic effect of the *Socs2* gene mutation on the inflammatory response, but also on the control of the infection. This can be explained by changes in the regulation of signal transduction from several receptors like the well-described Janus Kinase (JAK)/signal transducers and activators of the transcription (STAT) pathway. The direct effect of the mutation on animal weight is also demonstrated in our study. The interaction between SOCS-2 protein and the growth hormone receptor and signalling may explain this effect. The effect of the *Socs2* gene mutation on milk production found by Rupp *et al.* (2015) is confirmed, but only in our experimental dataset. New impacts of *Socs2* gene mutation were revealed on some caseins (beta and alphaS1) and beta-lactoglobulin, but need to be confirmed in the Lacaune commercial population. Finally, we assumed the existence of a second QTL in the vicinity of *Socs2* gene mutation which has preferential impacts on fat, particularly the saturated fatty acids.

CHI19 in Saanen for SCS, longevity, and milk production traits

In Saanen, Martin et al. (2018) carried out the the CLIP test for Milk Yield, Fat Yield, Protein Yield and two type traits in the CHI19 QTL region (**Table 5.5**). The pleiotropic assumption (H0) was never rejected in any of the 2-trait analyses with RUA, UFP, MY, FY, and PY, except for RUA-FY. The observed correlation (0.52), however, was only slightly below the threshold (0.54). This suggests that all of these traits may be controlled by the same gene(s) or by genes too close together on the genome to be distinguished by our design.

Table 5.5. Results of the CLIP test in Saanen for milk production and type traits in the CHI19 QTL region (241 SNPs ranging from 20 Mb to 32 Mb).

Trait 1*	Trait 2*	Correlation observed	Correlation limit at 5%	Significance (corr obs>corr limit)
MY	FY	0,926	0,659	Yes
MY	PY	0,959	0,736	Yes
MY	UFP	0,791	0,704	Yes
MY	RUA	0,629	0,605	Yes
FY	PY	0,937	0,663	Yes
FY	UFP	0,697	0,641	Yes
FY	RUA	0,521	0,541	No
PY	UFP	0,751	0,71	Yes
PY	RUA	0,648	0,612	Yes

*Milk Yield (MY) , Fat Yield (FY), Protein Yield (PY) and type traits: Udder flor position (UFP) and Rear Udder Attachment (RUA).

Talouarn et al (2020) implemented the CLIP test on genotypes in chromosome 19 between 23 and 28 Mb for Milk yield and semen volume. As written in their discussion section and summarised in Table 5.6: "The CLIP test rejected the pleiotropy assumption. The observed correlation was estimated at 0.013 and the threshold not to reject pleiotropy was above 0.15. The two traits might therefore be controlled by two different mutations situated close to each other. Moreover, none of the top 10 variants is shared between the two traits. According to the estimated effects, the allele with the highest frequency in the QTL region decreases both SV (− 0.09 SD) and milk yield (− 0.51 SD)"

Table 5.6. Results of the CLIP test in Saanen for milk production and semen traits in the CHI19 QTL region (23 Mb to 28 Mb).

Trait 1*	Trait 2*	Correlation observed	Correlation limit at 5%	Significance (corr obs>corr limit)
Milk yield	Semen volume	0,013	0,15	No

As a summary the Saanen QTL located on CHI19 (around 29 Mb) regions is associated with Several R&E traits (milk production, udder type traits and semen traits, longevity, BHB). Milk yield and udder type seem to be controlled by a same underlying mutation whereas link with semen traits and mastitis (LSCS) might be controlled by different mutations located in close proximity. Such an association therefore shows a favourable condition for disrupting the local unfavourable trade-offs. The nature of the local link with functional longevity and BHB has yet to be established.

6. Investigation of underlying genes and pathways

Intervals listed in Table 4.5 were used for subsequent analysis. Using Biomart interface provided by Ensembl (https://www.ensembl.org/biomart/martview), Gene Stable IDs, Transcript Stable IDS and Gene Names were retrieved. Analysis were performed with 610/220 Gene Stable IDs/Gene Names for meat sheep, dairy sheep and goats, as broken down in Table 6.1 below. It can be noted that the current assemblies of sheep and goat are not very well annotated, since only 36% of the genes are annotated. "Telomere to telomere" assemblies are under construction in both species and should provide further information after they have been annotated.

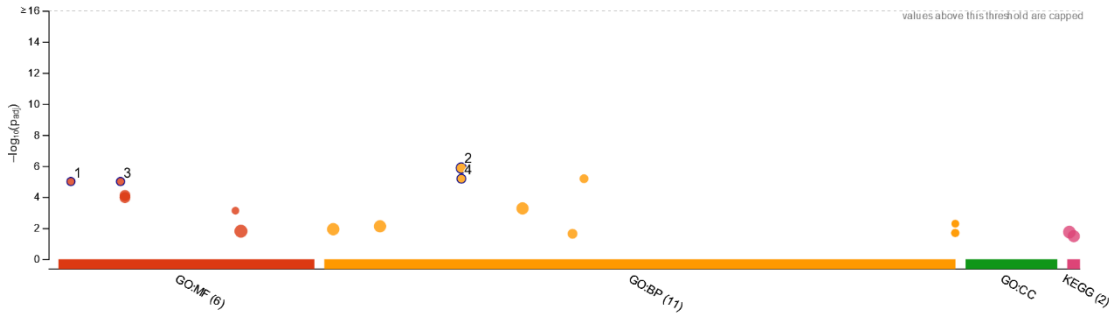
Table 6.1: Number of genes retrieved in possible pleiotropic QTL regions for R&E traits as found in gwas for 5 sheep and goat populations (Merino, Romane, Lacaune, Manech tete rousse, Saanen)

Breed	Gene Stable IDs	Transcript Stable IDs	Gene Names
meat sheep	349	375	33
dairy sheep	26	29	5
goat	235	327	182
Total	610	731	220

Gprofiler was used with either sheep or goat references and default parameters. No significant enrichment was found with the dairy sheep genomic intervals. Results are shown in Figure 6.1 for meat sheep and for goat. Results showed an over representation of functional terms linked to arachidonic and linoleic lipid metabolism as far as molecular function, biological process and KEGG pathways were concerned. Noteworthy, arachidonic can be synthetised from linoleic acid. These fatty acids are involved in numerous biological processes, such as the formation of cell membranes (within phospholipids), lipid signaling and the regulation of inflammation.

Another complementary analysis was performed using Ingenuity Pathway software (not shown) which confirmed the enrichment of the meat sheep genomic regions and the goat genomic regions in genes related to lipid metabolism. Results suggest that lipid metabolism is a key pathway involved in QTL regions with large spectra of action, controlling traits associated with diverse function such a efficiency (growth, milk production) and resilience traits (inflammation, regulation of body reserves).

A) Gprofiler analysis on 349 meat sheep EnsemblGene stable ID

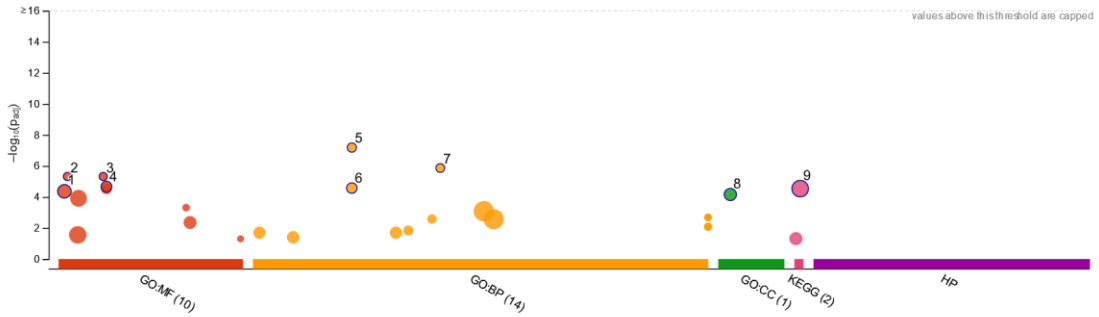


ID	Sour...	Term ...	Term Na...	p_{adj} (query_1)
1	GO:MF	GO:0004052	arachidonate 12(S)-lipoxygenase activity	9.959×10^{-6}
2	GO:BP	GO:0019369	arachidonic acid metabolic process	1.352×10^{-6}
3	GO:MF	GO:0016165	linoleate 13S-lipoxygenase activity	9.959×10^{-6}
4	GO:BP	GO:0019372	lipoxygenase pathway	6.561×10^{-6}

version	e109_eg56_p17_1d3191d
date	22/08/2023 17:22:55
organism	oaries

g:Profiler

B) Gprofiler analysis on 235 goat Ensembl Gene stable ID



ID	Sour...	Term ...	Term Na...	p_{adj} (query_1)
1	GO:MF	GO:0003774	cytoskeletal motor activity	4.325×10^{-5}
2	GO:MF	GO:0004052	arachidonate 12(S)-lipoxygenase activity	4.792×10^{-6}
3	GO:MF	GO:0016165	linoleate 13S-lipoxygenase activity	4.792×10^{-6}
4	GO:MF	GO:0016702	oxidoreductase activity, acting on single donors wi...	2.159×10^{-5}
5	GO:BP	GO:0019372	lipoxygenase pathway	6.559×10^{-8}
6	GO:BP	GO:0019369	arachidonic acid metabolic process	2.676×10^{-5}
7	GO:BP	GO:0043651	linoleic acid metabolic process	1.362×10^{-6}
8	GO:CC	GO:0016459	myosin complex	6.947×10^{-5}
9	KEGG	KEGG:04740	Olfactory transduction	2.919×10^{-5}

version	e109_eg56_p17_1d3191d
date	18/08/2023 09:54:16
organism	chircus

g:Profiler

Figure 6.1: results from G profiler analyses in meat sheep (A) and goat (B).
The x-axis represents functional terms that are grouped and colour-coded by data sources (e.g. Molecular Function from GO is red; biological Process from GO is orange; Cellular Component from GO is green; KEGG pathways in pink;). The y-axis shows the adjusted enrichment p-values in negative log10 scale. Plots are capped so that they can be compared to each other. The tables show the most significant results (p-values $<10^{-4}$).

7. Screening for polymorphisms in the list of potentially pleiotropic regions

Using the same intervals as in section 6, we extracted polymorphisms from available public variant datasets (60,248,438 SNPs from Ensembl for sheep and 28,645,747 SNPs from Var-Goats - https://www.goatgenome.org/vargoats.html- for goats). 204,844/58,587/197,534 SNPs were retrieved for meat sheep, dairy sheep and goats respectively.

For sheep, the SIFT annotation was retrieved. Caprine variants were annotated with the relevant GFF (General Feature Format from Ensembl) files and SnpEff prediction tool. We filtered out variants with a significant predicted impact (annotated as high/deleterious or moderate/tolerated). 2161/187/3026 SNPs with a significant predicted impact were identified for meat sheep, dairy sheep and goats, as broken down in Table 7.1 below. The list of SNPS is supplied in ANNEX2 (https://zenodo.org/record/8281277).

Table 7.1: Number of variants retrieved in possible pleiotropic QTL regions for R&E traits as found in gwas for 5 sheep and goat populations (Merino, Romane, Lacaune, Manech tete rousse, Saanen), according to impact predicted by SnpEff and SIFT prediction software

Breed	Total number of SNPs retrieved	SNP with high/deleterious impact	SNP with moderate/tolerated impact
meat sheep	204,844	764	1397
dairy sheep	58,587	72	115
goat	197,534	45	2981
Total	460,965	881	4493

This list provides candidate causal mutations for QTLs in the studied regions, some of which are likely to include mutations responsible for trade-off between R&E traits.

8. Conclusion

Comparison of Gwas carried out in smarter (Merino, Texel, Corriedale, Romane, Texel, Churra, Lacaune, Saanen and Alpine) and gwas published previously (Lacaune, Manech tete rousse, Saanen and Alpine) identified a limited number (N=9) of genomic regions that control both R&E traits. The results of possible pleiotropic genomic regions in Lacaune (OAR3) and Saanen (CHI19) were already reported in literature – and were refined. Results in Merino (OAR5, OAR6, ORA11 and OAR21) and Romane (OAR1) were new. Such regions can, to some extent, explain genetic correlations between R&E traits.

In most cases, we could exclude the hypothesis of pleiotropy (same mutation associated with both R&E traits) suggesting that close linkage of distinct causal mutations is the general pattern in this common QTL regions.
Further interrogation of these genomic regions that control both R&E traits, highlighted the importance of lipid metabolism, as underlying mechanism. Finally, we provided a list of variants in the regions of interest by systematic screening for polymorphisms in the latter regions, using both public and private databases. This list provides candidate causal mutations for QTLs in the studied regions likely to include mutations responsible for trade-off between R&E traits.
This study gathering many partners of smarter (CNR, INRAE, IDELE, INIA-UR, SRUC, UNILEON) and work done in WP1, 2 and 3, reported in this Deliverable produced four published papers, and 3 papers under construction.

Published/accepted

- Macé T, González-García E, Foulquié D, Carrière F, Pradel J, Durand C, Douls S, Allain C, Parisot S, Hazard D. 2022. Genome-wide analyses reveal a strong association between LEPR gene variants and body fat reserves in ewes. BMC Genomics. 1: 23(1):412. https://doi.org/10.1186/s12864-022-08636-z
- Carracelas, B.; Navajas, E.A.; Vera, B.; Ciappesoni, G., 2022. Genome-Wide Association Study of Parasite Resistance to Gastrointestinal Nematodes in Corriedale Sheep. Genes 2022, 13, 1548. https://doi.org/10.3390/genes13091548
- Ramos, Z.; Garrick, D.J.; Blair, H.T.; Vera, B.; Ciappesoni, G.; Kenyon, P.R., 2023. Genomic Regions Associated with Wool, Growth and Reproduction Traits in Uruguayan Merino Sheep. Genes, 14(1), 167; https://doi.org/10.3390/genes14010167
- Kaseja et al., Genome-wide association study of health and production traits in meat sheep, Animal in press

Under construction

- Brenda Vera et al., (PhD) Genomic regions associated with resistance to gastrointestinal parasites in Australian Merino sheep
- Christel Marie-Etancelin et al., Genome wide association for R&E traits in Lacaune
- Johanna Ramirez Diaz, Genome wide association in Romane

9. Deviations or delays

Deliverable expected M54 was submitted M58 with 4 months delay.
The total duration of the project, including the 8-month extension, has been used to finish the gwas analyses in WP1 and 2. Therefore some gwas data arrived late. We waited at the very last moment to re-run the gene pathway and variant screening analyses (section 6 and 7) to include these latest gwas results. Thus, the deliverable was finished and submitted only M58 with the final report.

10. References

- Aguilar, I., A. Legarra, F. Cardoso, Y. Masuda, D. Lourenco, and I. Misztal. 2019. Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. Genet. Sel. Evol. 51:28.
- Boichard D, Govignon-Gion A, Larroque H, Maroteau C, Palhière I, Tosser-Klopp G, et al. 2014. Déterminisme génétique de la composition en acides gras et protéines du lait des ruminants, et potentialités de sélection. INRA Prod Anim.27:283–98.
- David, I., J. M. Elsen, D. Concordet, 2013. CLIP Test: a new fast, simple and powerful method to distinguish between linked or pleiotropic quantitative trait loci in linkage disequilibria analysis. Heredity. 110(3): 232-238.
- Carracelas, B.; Navajas, E.A.; Vera, B.; Ciappesoni, G. 2022. Genome-Wide Association Study of Parasite Resistance to Gastrointestinal Nematodes in Corriedale Sheep. Genes 13, 1548.
- Diez-Tascón C, Bayón Y, Arranz JJ, De La Fuente F, San Primitivo F., 2021. Mapping quantitative trait loci for milk production traits on ovine chromosome 6. J Dairy Res.68:389–97.
- Ferrand M, Miranda G, Larroque H, Leray O, Guisnel S, Lahalle F, et al., 2012. Determination of protein composition in milk by mid-infrared spectrometry. In: Proceedings of the 38th ICAR Annual Meeting: 28 May 2012; Cork. 2012.
- Ferrand-Calmels M, Palhière I, Brochard M, Leray O, Astruc JM, Aurel MR, et al., 2014. Prediction of fatty acid profiles in cow, ewe, and goat milk by mid-infrared spectrometry. J Dairy Sci. 97:17–35.
- García-Gámez E, Gutiérrez-Gil B, Sahana G, Sánchez JP, Bayón Y, Arranz JJ., 2012. GWA analysis for milk production traits in dairy sheep and genetic support for a QTN influencing milk protein percentage in the *LALBA* gene. PLoS One 7:e47782.
- Lakhssassi K, Serrano M, Lahoz B, Sarto MP, Iguácel LP, Folch J, Alabart JL, Calvo JH. 2020. The LEPR Gene Is Associated with Reproductive Seasonality Traits in Rasa Aragonesa Sheep. *Animals*. 10(12):2448.
- Ma, S.; Ji, X.; Cang, M.; Wang, J.; Yu, H.; Liu, Y.; Zhang, W.; Wu, Y.; Zhao, S.; Cao, G.; et al. 2022. Association analysis between novel variants in LEPR gene and litter size in Mongolia and Ujimqin sheep breeds. Theriogenology 183:79–89.
- Macé T, González-García E, Foulquié D, Carrière F, Pradel J, Durand C, Douls S, Allain C, Parisot S, Hazard D. 2022. Genome-wide analyses reveal a strong association between LEPR gene variants and body fat reserves in ewes. BMC Genomics. 1: 23(1):412.
- Martin P, Palhière Is, Maroteau C, Clément V, David I, Klopp Tosser G, Rupp R, 2018. Genome-wide association mapping for type and mammary health traits in French dairy goats identifies a pleiotropic region on chromosome 19 in the Saanen breed. Journal of Dairy Science. 2018 101(6):5214–5226.
- Martinez Boggio, G., Meynadier, A., Buitenhuis, A.J. *et al.* 2022. Host genetic control on rumen microbiota and its impact on dairy traits in sheep. Genet Sel Evol 54, 77.
- Misztal, I., S. Tsuruta, D. A. L. Lourenco, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs. University of Georgia. http://nce.ads.uga.edu/wiki/lib/exe/fetch.php ?media=blupf90_all2.pdf

- Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee DH. 2002. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production: 19-23 August 2002; Montpellier. 2002.
- Mucha S, Tortereau F, Doeschl-Wilson A, Rupp R, Conington J. 2022. Animal Board Invited Review: Meta-analysis of genetic parameters for resilience and efficiency traits in goats and sheep. Animal 16(3):100456.
- Othmane MH, Carriedo JA, San Primitivo F, De La Fuente LF. Genetic parameters for lactation traits of milking ewes: protein content and composition, fat, somatic cells and individual laboratory cheese yield. Genet Sel Evol. 2002 Sep-Oct;34(5):581-96.
- Palhière P, C. Oget C and R. Rupp R, 2018. Functional longevity is heritable and controlled by a major gene in French dairy goats. WCGALP, Auckland, New-Zealand, 2018.
- Ramos, Z.; Garrick, D.J.; Blair, H.T.; Vera, B.; Ciappesoni, G.; Kenyon, P.R., 2023. Genomic Regions Associated with Wool, Growth and Reproduction Traits in Uruguayan Merino Sheep. Genes, 14(1), 167.
- Rauw W.M., E. Kanis, E.N. Noordhuizen-Stassen, F.J. Grommers, 1998. Undesirable side effects of selection for high production efficiency in farm animals: A review. Livestock Production Science 56, pp. 15-33
- Rupp R, Senin P, Sarry J, Allain C, Tasca C, Ligat L, et al. 2015. A point mutation in suppressor of cytokine signalling 2 (Socs2) increases the susceptibility to inflammation of the mammary gland while associated with higher body weight and size and higher milk production in a sheep model. PLoS Genet. 2015;11:e1005629.

11. ANNEX1 Significative associations of close regions for different traits in three Uruguayan breeds (Texel, Merino and Corriedale)

(chromosome, location, p-value or proportion of additive genetic variance (PVE, %)).

Breed	Trait	chromosome	Position	p-Value	% variance
Merino	FAMACHA	1	40,540,327		0.74
Merino	Pregnancy rate	1	40,646,909		0.38
Corriedale	FEC1	1	40,732,375	0.003	
Merino	BCS ewe	1	173,862,929		0.88
Merino	FAMACHA	1	185,760,112		0.27
Merino	BCS	2	110,373,614		0.25
Texel	fat depth	2	110,743,239	0.0003	
Merino	BCS ewe	2	111,201,892		2.21
Merino	BCS	2	112,388,825		0.28
Merino	FEC1	2	112,778,086		0.37
Texel	fat depth	2	112,815,979	5,00E-06	
Texel	fat depth	2	113,671,480	6,00E-05	
Merino	BCS ewe	2	114,746,188		0.65
Merino	FAMACHA	2	218,777,904		0.20
Merino	fat depth	2	219,781,182		0.20
Merino	FAMACHA	2	219,781,182		0.38
Merino	BCS	3	95,078,520		0.23
Merino	FAMACHA	3	95,078,520		0.26
Merino	Pregnancy rate	4	107,666,587		3.84
Merino	Lambing pot.	4	107,666,587		5.68
Merino	Pregnancy rate	4	113,244,910		0.61
Merino	FAMACHA	4	115,735,447		0.36
Merino	BCS	5	26,912,393		0.30
Texel	fat depth	5	28,490,054	0.001	
Merino	fat depth	5	93,233,297		1.94
Merino	BCS	5	93,233,297		0.37
Merino	FAMACHA	5	93,233,297		0.76
Merino	live weight mating	5	93,416,569		1.54
Merino	FEC2	5	93,453,349		1.13
Merino	temperament	6	31,453,177		0.77
Merino	FAMACHA	6	32,215,550		0.24
Merino	fat depth	6	32,215,550		0.36
Merino	temperament	6	35,184,703		0.62

Merino	live weight mat-ting	6	35,191,867		2.05
Merino	temperament	6	35,191,867		0.64
Merino	FAMACHA	6	35,254,368		0.59
Merino	fat depth	6	35,254,368		0.64
Merino	BCS	6	35,254,368		0.78
Merino	temperament	6	35,254,368		0.59
Merino	FEC1	6	35,491,698		0.87
Merino	temperament	6	35,491,698		0.77
Merino	temperament	6	35,512,106		0.55
Merino	FAMACHA	6	36,257,135		1.58
Merino	fat depth	6	36,257,135		1.10
Merino	BCS	6	36,257,135		6.57
Merino	live weight mat-ting	6	36,295,216		4.43
Merino	live weight mat-ting	6	36,905,457		5.67
Merino	FAMACHA	6	37,269,100		0.41
Merino	fat depth	6	37,269,100		2.65
Merino	BCS	6	37,269,100		8.57
Merino	FEC2	6	37,528,527		1.07
Merino	fat depth	6	38,270,152		1.10
Merino	BCS	6	38,270,152		1.68
Merino	FAMACHA	6	40,309,928		0.24
Merino	BCS	6	40,309,928		0.39
Corriedale	FEC1	7	33,565,208	0.000	
Merino	BCS	7	34,621,754		0.21
Merino	BCS ewe	9	57,388,779		3.35
Texel	fat depth	9	58,391,320	0.0004	
Merino	BCS ewe	10	56,843,164		1.14
Merino	live weight mat-ting	10	57,396,545		1.41
Merino	FAMACHA	11	25,298,833		0.23
Merino	fat depth	11	25,298,833		4.10
Merino	BCS	11	25,298,833		1.09
Merino	FEC2	11	25,971,167		2.24
Merino	fat depth	11	26,315,832		1.78
Merino	BCS	11	26,315,832		0.68
Merino	FEC1	11	27,258,896		1.04
Merino	fat depth	11	27,367,806		1.18
Merino	FAMACHA	12	24,130,322		0.31
Merino	BCS	12	24,130,322		0.71
Corriedale	FEC1	12	24,624,977	0.001	

Corriedale	FEC1	12	24,625,096	0.004	
Corriedale	FEC1	12	24,626,347	0.003	
Merino	fat depth	12	25,162,373		0.34
Merino	fat depth	12	49,792,166		0.31
Merino	FEC1	12	49,849,603		0.32
Merino	BCS ewe	13	42,377,205		4.68
Merino	FEC2	13	44,314,725		0.41
Merino	fat depth	14	12,525,704		0.32
Merino	BCS	14	13,562,123		0.29
Texel	fat depth	14	15,982,370	0.0007	
Merino	FAMACHA	14	54,194,578		0.27
Merino	fat depth	14	54,194,578		0.24
Merino	BCS	14	54,194,578		0.22
Merino	FAMACHA	15	46,552,672		0.38
Merino	fat depth	15	46,552,672		0.64
Merino	FEC1	15	47,489,709		0.27
Texel	BW	15	48,972,885	0.0016	2.79
Merino	Pregnancy rate	16	45,966,691		8.98
Merino	Lambing pot.	16	45,966,691		2.00
Merino	FAMACHA	18	63,854,026		0.60
Merino	fat depth	18	63,854,026		0.44
Merino	BCS	18	63,854,026		0.21
Merino	fat depth	19	47,919,516		0.33
Merino	BCS	19	47,919,516		4.70
Merino	FEC2	19	48,039,302		0.30
Texel	BW	21	33,814,204	0.0024	2.62
Texel	fat depth	21	35,746,173	0.001	
Merino	fat depth	21	39,131,113		1.63
Merino	BCS	21	39,131,113		0.86
Merino	temperament	21	39,432,569		0.54
Merino	temperament	21	39,646,260		0.59
Merino	FEC2	21	39,651,749		0.80
Merino	temperament	21	39,651,749		0.58
Merino	FAMACHA	21	42,263,495		0.34
Merino	BCS	21	42,263,495		0.25
Merino	temperament	21	42,295,749		1.47
Texel	BW	21	42,325,238	0.0025	2.61
Texel	BW	21	42,442,512	0.0032	2.50
Merino	temperament	21	42,601,748		1.45
Merino	temperament	21	42,654,067		1.33

Merino	temperament	21	42,714,381		1.16
Merino	FEC1	21	42,714,613		0.41
Merino	temperament	21	42,714,613		1.09
Merino	FEC2	25	21,582,203		0.34
Merino	FEC1	25	21,919,108		0.42
Merino	FAMACHA	25	33,366,256		0.37
Merino	BCS	25	33,366,256		0.21
Merino	fat depth	25	35,384,777		0.22
Merino	FAMACHA	25	36,389,257		0.24
Merino	fat depth	25	36,389,257		0.30

12. ANNEX2: lists of variants retrieved in possible pleiotropic QTL regions for R&E traits as found in gwas for sheep and goat populations, short lists of SNP variants with high/deleterious or moderate/tolerated impact,

(variant name, source dataset, chromosome, position, predicted impact).

Data is available here : https://zenodo.org/record/8281277