

## Evidence of genetic links between rumen bacteria and milk somatic cell score in Lacaune dairy sheep

G. Martinez Boggio<sup>1</sup>, A. Meynadier<sup>1</sup>, R. Rupp<sup>1</sup>, C. Allain<sup>2</sup>, A.J. Buitenhuis<sup>3</sup> and C. Marie-Etancelin<sup>1\*</sup>

<sup>1</sup> GenPhySE - INRAE – ENVY – Université de Toulouse, 24 Chemin de Borde Rouge, 31326, Castanet-Tolosan, France; <sup>2</sup> INRAE – Unité Expérimentale de La Fage, 12250, Saint-Jean et Saint-Paul, France; <sup>3</sup> Center for Quantitative Genetics and Genomics, Aarhus University, Blichers Allée 20, DK-8830 Tjele, Denmark; \*christel.marie-etancelin@inrae.fr

### Abstract

Mastitis is a high prevalence infectious disease in dairy sheep. Even if selection by milk somatic cell score (SCS) has been implemented to increase mastitis resistance, we wonder if rumen bacteria could influence udder inflammatory response. The objectives are to estimate genetic correlations between rumen bacteria and average SCS over lactation (LSCS), and to identify underlying mechanism between traits. 795 Lacaune ewes were studied using metabarcoding (16S rRNA gene) for rumen bacteria characterization and annual SCS records. Two-trait genomic models were used for the parameter estimation and GWAS. Results showed significant genetic correlations, and colocalized QTLs for LSCS and nine bacterial operational taxonomic units (OTUs) on chromosomes 3, 17, and 20. We found a genetic association between resistance/susceptibility to mastitis, using LSCS, and the ruminal bacteria abundance. These preliminary results pave the way for better understanding the biological basis and causalities of an intra-mammary inflammations in sheep.

### Introduction

As in cattle, mastitis is one of the most important infectious diseases in small dairy ruminants, due to its high frequency and the price reduction paid to the farmer of up to 10% (95€ per 1,000 litres of ewes' milk). The selection was focused on the milk SCS, as a tool to decrease subclinical intra-mammary infections, which is the main problem in dairy sheep. The existence of a genetic basis for mastitis resistance has been demonstrated, and the underlying mechanisms involved such as the *suppressor of cytokine signalling 2* (Socs2) gene (Rupp *et al.*, 2015). Recently, the influence of the rumen microbiome on subclinical ketosis has been demonstrated in cows (Gebreyesus *et al.*, 2020), but no work is available in sheep. Therefore, could rumen bacteria impact on intra-mammary inflammations in sheep? To address this, and after quantifying host control over rumen bacteria, we 1) estimated genetic correlations between rumen bacteria and LSCS, and 2) identified possible underlying mechanisms with genome-wide association analysis (GWAS).

### Materials & Methods

**Animals phenotyping.** A total of 795 Lacaune dairy ewes were reared at INRAE Experimental Unit of La Fage (UE 321 agreement A312031, Roquefort, France), between 2015 to 2019. All ewes (weighing 77 kg on average) raised indoors, and fed 93% meadow hay and silage plus 7% of concentrates (on dry matter basis). The daily somatic cell count (SCC) was quantified with a Fossomatic cell counter (Nanterre, France), as part of the official milk recordings of the flock. Then, the average LSCS was calculated as log of somatic cell count, including at least four milk test-days. Rumen sampling of adult ewes between 28 to 133 days in milk was done once per ewe using a vacuum pump and a medical gastric tube. The protocol received approval from the French Ministry of Higher Education, Research and Innovation – Animal Ethics Committee

(approval number: APAFIS#6292-2016080214271984 v8). The RNA sequences of the 795 samples were processed with the FROGS 3.0 pipeline, and OTUs with abundances below 0.005% were removed (Bokulich *et al.*, 2013). The resulting abundance table included 2,059 OTUs, but we focused on the 306 OTUs with heritability significantly different from zero (empirical significant threshold of 0.10 obtained after running 10,000 analyses where OTUs abundancies were shuffling randomly across individuals) using an additive relationship matrix.

**Animals genotyping.** DNA extraction from blood samples and genotyping were performed for 795 Lacaune ewes. From those ewes, 743 were genotyped using a medium-density SNP chip (Illumina Ovine SNP50 BeadChip, 54,241 SNPs). The remaining 52 ewes were genotyped with a low-density SNP chip (Illumina Ovine SNP15, 16,681 SNPs) and imputed to a medium-density SNP chip genotypes in the framework of the Lacaune dairy sheep genomic selection program (Larroque *et al.*, 2017). SNPs were removed using a call rate <0.99 for SNPs, and <0.95 for individuals, and minor allele frequency <0.05. In total, for the 26 *ovis aries* autosomal chromosomes (OAR), 35,496 SNPs remained for the analyses, plus one SNP corresponding to the Socs2 gene mutation (OAR3).

**Parameter estimation.** To estimate the heritability of the 306 rumen OTUs and LSCS, and the genetic correlations between each OTU and LSCS, a two-trait animal model was used:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \quad (1)$$

where  $\mathbf{y}_1$  and  $\mathbf{y}_2$  are the vectors of observations for LSCS and one OTU at a time, respectively;  $\mathbf{b}_1$  and  $\mathbf{b}_2$  are the vectors of fixed effects,  $\mathbf{b}_1$  including for LSCS: litter size, number of milking test-day, and  $\mathbf{b}_2$  for the OTUs: sampling year, total number of DNA sequences per sample, and three effects nested on year sampling: lactation number, sequencing run, and order-time of sampling;  $\mathbf{a}_1$  and  $\mathbf{a}_2$  are the vectors of random additive effects; and  $\mathbf{e}_1$  and  $\mathbf{e}_2$  are the vectors of residual effects.  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are incidence matrices relating fixed effects to vector  $\mathbf{y}_1$  and  $\mathbf{y}_2$ , respectively;  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are incidences matrices relating the additive effects to vector  $\mathbf{y}_1$  and  $\mathbf{y}_2$ , respectively. The distributional assumptions are  $\mathbf{a} \sim N(0, \mathbf{G} \otimes \mathbf{W})$ , and  $\mathbf{e} \sim N(0, \mathbf{I} \otimes \mathbf{R})$ ;  $\mathbf{G}$  is a genomic relationships matrix as  $\mathbf{G} = \mathbf{Z}\mathbf{Z}'/2\sum p_i(1 - p_i)$  (VanRaden, 2008);  $\mathbf{I}$  is an identity matrix;  $\mathbf{W}$  and  $\mathbf{R}$  are the variance-covariance matrices for the random additive and residual effects, respectively. The heritability was estimated as  $h^2 = \sigma_a^2/\sigma_y^2$ , and the genetic correlations as  $r_g = \sigma_{a12}/\sqrt{(\sigma_{a1}^2\sigma_{a2}^2)}$ . The variance components were estimated using AIREMLF90 (Misztal *et al.*, 2002). The significance was considered as genetic correlation value greater than the standard error.

**Genome-Wide Association.** The GWAS was performed using the two-trait model presented in Equation 1. 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). To correct for multiple testing the false discovery rate (FDR) was used as implemented in the p.adjust package in R software. A SNP was considered significant at an FDR of  $P < 0.10$ .

## Results

The average heritability estimated for LSCS was  $0.39 \pm 0.07$ . Genetic parameters of the 14 rumen bacteria having significant genetic correlations with LSCS are presented in Table 1.

**Table 1. Heritability ( $h^2$ ) and genetic correlations ( $r_g$ ) between lactation somatic cell score and ruminal bacteria abundance.**

Genus	OTU name	$h^2$ (SE)	$r_g$ (SE)
<i>Lachnospiraceae_NK3A20_group</i>	OTU506	0.10 (0.06)	0.68 (0.63)
<i>Prevotella</i>	OTU871	0.17 (0.07)	0.58 (0.43)
<i>Monoglobus</i>	OTU1066	0.13 (0.06)	0.51 (0.40)
<i>Leifsonia</i>	OTU1479	0.14 (0.07)	0.48 (0.43)
<i>Lachnospiraceae/unknown genus</i>	OTU309	0.15 (0.07)	0.47 (0.41)
<i>Mailhella</i>	OTU2331	0.15 (0.06)	0.39 (0.34)
<i>F082/unknown genus</i>	OTU1009	0.14 (0.06)	0.35 (0.29)
<i>Acetitomaculum</i>	OTU612	0.23 (0.07)	0.31 (0.21)
<i>Lachnospiraceae_NK3A20_group</i>	OTU278	0.20 (0.07)	0.30 (0.24)
<i>Lachnospiraceae/unknown genus</i>	OTU1666	0.19 (0.07)	-0.21 (0.20)
<i>Prevotella</i>	OTU1496	0.22 (0.07)	-0.32 (0.24)
<i>Prevotella</i>	OTU196	0.16 (0.06)	-0.35 (0.28)
<i>Prevotella</i>	OTU136	0.13 (0.06)	-0.36 (0.34)
<i>F082/unknown genus</i>	OTU2431	0.20 (0.07)	-0.37 (0.22)

SE= standard error.

The GWAS for LSCS detected on OAR3 four significant QTLs: in decreasing order of significance, the SNP for Socs gene (129,722,200 bp), SNP rs412514556 (129,423,378 bp), SNP rs425386363 (23,336,791 bp) and SNP rs416770789 (133,977,235 bp). The microbiome GWAS detected 387 significant SNPs for 67 OTUs among the 306 OTUs with heritability significantly different from zero. SNPs were scattered on all chromosomes, with the highest p-values and highest number of signals on OAR1, OAR2, OAR3, and OAR5. But, in this study we focus on the OTUs with significant signals located close to the LSCS QTLs on OAR3 (Table 2).

**Table 2. Summary of the significant results obtained from the microbiome GWAS.**

Genus	OTU name	SNP	Position (bp)	$-\log_{10}$ (P-value)
<i>Prevotella</i>	OTU1419	rs425386363	23,336,791	5.59
<i>Lachnospiraceae/unknown genus</i>	OTU1666 <sup>1</sup>	rs425386363	23,336,791	6.90
<i>Ruminococcus</i>	OTU602	Socs	129,722,200	5.70

<sup>1</sup> OTUs with significant genetic correlations with LSCS.

## Discussion

The LSCS heritability was greater than the 0.15 reported by Rupp *et al.* (2009), using the whole Lacaune population. Even if the genetic correlations have a high standard error which can lead to spurious ones, most genetic correlations between LSCS and OTUs were positive, meaning that higher bacterial abundance is genetically associated with susceptibility to mastitis. There is also a higher variety of genera within the positive correlations than for the negative correlations where the most frequently associated is *Prevotella*, a bacteria belonging to the *Bacteroidetes* phylum. Zhong *et al.* (2018) obtained *Bacteroidetes* as the most abundant in the group with low SCC, even if they showed that the rumen microbiome in cows was globally stable at different SCC levels. The link between gut bacteria and mastitis is probably not direct but would be through bacterial metabolites absorbed by the host, and which would have pro- or anti-inflammatory properties (Hu *et al.*, 2019). A QTL on OAR3 was detected for

*Ruminococcus* OTU602 at the Socs gene position, and the Socs2 genotype known to have more LSCS in milk is associated with highest abundance of OTU602 (results not shown). In contrast, Chuang *et al.* (2021) found that *Ruminococcus* genus was phenotypically reduced in the rumen of cows with clinical mastitis. As well on OAR3, two QTLs were detected, one for *Prevotella* OTU1419 and one for *Lachnospiraceae* OTU1666. The SNP genotype associated with more LSCS in milk is linked with more *Prevotella* OTU1419 abundance, which is contrary with the main genetic correlations for this genus, but on others OTUs. The same SNP genotype associated with mastitis susceptibility is also linked with less *Lachnospiraceae* OTU1666 abundance, which is consistent with the negative but weak genetic correlation of -0.21 with LSCS. This result on OTU1666 is in agreement with Wang *et al.* (2021) finding that this family is phenotypically associated with anti-inflammatory molecules such as 2-Phenylbutiric acid. But, for the three others OTUs belonging to the *Lachnospiraceae* family, an increase of their abundances is genetically linked with more LSCS. The identification in this study of QTLs on OAR3 common for LSCS and OTUs in addition to the results obtained for genetic correlations, contribute to the hypothesis of a genetic link between rumen bacteria and mastitis. We can conclude that host genetic may influence ruminal bacteria abundances and udder inflammation in sheep via mainly positive genetic correlations, but the underlying metabolic pathways are still difficult to identify with only the identification of bacterial genera. The genetic study of these bacteria's functions is underway in order to unravel the various results obtained at the OTU level.

## References

- Aguilar I., Legarra A., Cardoso F., Masuda Y., Lourenco D., *et al.* (2019). *Genet Sel Evol* 51:28. <https://doi.org/10.1186/s12711-019-0469-3>
- Bokulich N.A., Subramanian S., Faith J.J., Gevers D., Gordon J.I., *et al.* (2013). *Nat Methods* 10:57–59. <https://doi.org/10.1038/nmeth.2276>
- Chuang S.-T., Li K.-Y., Tu P.-W., Ho S.-T., Hsu C.-C., *et al.* (2021). *Animals* 11:3108. <https://doi.org/10.3390/ani11113108>
- Gebreyesus G., Difford G.F., Buitenhuis B., Lassen J., Noel S.J., *et al.* (2020). *J. Dairy Sci.* 103:4557–4569. <https://doi.org/10.3168/jds.2019-17824>
- Hu X., Li S., Fu Y., Zhang N., (2019). *Eur. J. Clin. Microbiol. Infect. Dis.* 38:1409–1423. <https://doi.org/10.1007/s10096-019-03549-4>
- Larroque H., Chassier M., Saintilan R., Astruc J.M. (2017). *Proc. of EAAP, Tallinn, Estonia.*
- Misztal I., Tsuruta S., Strabel T., Auvray B., Druet T., *et al.* (2002). *Proc. of 7th WCGALP, Montpellier, France.*
- Rupp R., Bergonier D., Dion S., Hygonenq M.C., Aurel M.R., *et al.* (2009). *J. Dairy Sci.* 92:1203–1219. <https://doi.org/10.3168/jds.2008-1435>
- Rupp R., Senin P., Sarry J., Allain C., Tasca C., *et al.* (2015). *PLoS Genet* 11:e1005629. <https://doi.org/10.1371/journal.pgen.1005629>
- VanRaden P.M., (2008). *J. Dairy Sci.* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Wang Y., Nan X., Zhao Y., Jiang L., Wang M., *et al.* (2021). *J Animal Sci Biotechnol* 12:36. <https://doi.org/10.1186/s40104-020-00543-1>
- Zhong Y., Xue M., Liu J. (2018). *Front. Microbiol.* 9:3217. <https://doi.org/10.3389/fmicb.2018.03217>