Fine mapping of 7 QTL regions in dairy sheep confirms pleiotropic effect of the R96C mutation in the *Socs2* gene on SCC, bacterial infection, size and milk production

C. Oget[1], C. Allain[1], D. Portes[2], G. Foucras[3], G. Tosser-Klopp[1] & R. Rupp[1*]*

[1] GenPhySE, Université de Toulouse, INRA, INPT, ENVT, Castanet Tolosan, France
claire.oget@inra.frmailto:Riberc@univit.com (Corresponding Author)
[2] INRA, Domaine de La Fage, 12250 Saint-Jean-et-Saint-Paul, France
[3] IHAP, Université de Toulouse, INRA, INPT, ENVT, Castanet Tolosan, France
[*] These authors contributed equally to this work.

Summary

The aim of this study was to investigate the pleiotropic effects of a *Socs2* gene point mutation associated to lower resistance to mastitis, an inflammation of the mammary gland, mostly due to bacterial infection by Staphylococci in dairy sheep. GWAS were performed on SCC (Somatic Cell Count), bacterial infection, animal size and milk production using a 960 custom-designed ovine SNP chip and a data set of 504 Lacaune dairy ewes. The results confirmed the strong effect of this mutation on the inflammatory response (p-value of LSCS_L1 = 3.03e-07 and p-value of abscess = 2.83e-03), a slighter effect on animal size (p-value of W_ DAY_250 = 3.71e-04) and also showed for the first time an effect of this mutation on intramammary infections (p-value of STAPH_L1 = 6.80e-04). In contrast, no effect on milk production could be confirmed (p-value of MILK_L1 = 2.01e-01), suggesting an indirect role of SOCS-2 protein on milk production. Then pleiotropy tests (CLIP test) proved pleiotropic effects of the mutation on SCC, bacterial infection, size and more indirectly on milk production.

Keywords: dairy sheep, genetic resistance, mastitis, genome wide association study

Introduction

Mastitis is an inflammation of the mammary gland, which is mostly due to bacterial infections by Staphylococci in dairy sheep. Mastitis is a serious burden for the milk industry due to the altered quality of milk and increased cost for production and flock renewal. Beside hygienic measures, genetic selection for improved resistance to mastitis is now considered and implemented in breeding programs of several dairy breeds worldwide. However, resistance to mastitis is a highly complex trait which genetically-determined biological basis is mostly unknown.

In dairy sheep, several QTL regions controlling milk somatic cell count (SCC), an indirect indicator of udder infection, have been identified within the EU-funded 3SR project (Sustainable Solutions for Small Ruminants, FP7-KBBE-245140). These QTL regions were then confirmed by Banos *et al.* (2017) with four mastitis indicator traits (clinical mastitis occurrence, milk SCC, total viable bacterial count in milk and the California mastitis test). Amongst these QTLs, Rupp *et al.* (2015) identified a SNP (Single-Nucleotide Polymorphism) in the coding frame of *Socs2* associated with high SCC in the Lacaune breed. This point mutation was localized in the SH2 binding domain of the SOCS-2 protein leading to the loss

of ligand recognition. This strongly suggests that the identified R96C mutation was causal, although the effect of other SNPs in the QTL region could not be fully excluded. It was also demonstrated that size, weight and milk production were significantly increased in ewes carrying the mutated *Socs2* allele when compared to wild type sheep.

The objective of this work was (i) to confirm and fine map seven ovine QTLs controlling mastitis resistance using a 960 custom-designed ovine SNP chip, (ii) to assess the effect of confirmed QTLs on bacterial infection, and iii) to identify and prove possible pleiotropic effects of QTLs on animal size and milk production.

Material and methods

Animals and phenotypes

The data set included 504 Lacaune dairy ewes from a divergent selection based on extreme breeding values for the somatic cell score (SCS) at the INRA experimental facility of La Fage (UE 321, Roquefort, France) (Rupp *et al.*, 2009).

Monthly SCC were log-transformed for normality into SCS. Test-day SCS and milk yield were averaged over lactation and corrected for the year of control and feeding method. The considered variables were LSCS and MILK for the first and second lactation (L1 and L2, respectively).

The *Staphylococcus* spp. abundance in milk was quantified at three time points in first lactation by a qPCR-based technique developed at IHAP laboratory (ENV Toulouse). Briefly, the method combines an original extraction technique (milk centrifugation in 96-well plates, enzymatic lysis and semi-automated DNA extraction) and high throughput qPCR in 384-well format. Results were expressed as a bacterial titre (quantity of equivalent bacterial genomes in 10 µL of milk), expressed on a logarithmic scale (STAPH_L1). All results from each ewe were averaged and corrected for the effects of month and year of sampling.

Chronic mastitis was also assessed by the presence of mammary abscess, recorded by clinical examination. Animals were recorded as "1" (case) when examinations detected the presence of at least one abscess at least twice, while animals were recorded "0" (control) when they were examined healthy (without any abscess) at least three times.

Each ewe was weighed at birth (W_BIRTH), and at the age of 100 (W_DAY_100) and 250 days (W_DAY_250), and after 1st (on average 412 days, W_1ST_LAMB) and 2nd lambing (on average 744 days, W_2ND_LAMB), and at the age of 920 days (W_DAY_920). The variables of weight were corrected for year and feeding method. Table 1 provides the mean, the standard deviation and headcounts for each of these phenotypes.

Genotyping with a mastitis-specialized 960-SNP chip

All animals were genotyped with a 960 custom-designed ovine SNP chip (http://genoweb.toulouse.inra.fr/~tosser/3SR-WP3-960_snp_mastitis/). The chip was designed and developed within the "3SR" EU project (http://cordis.europa.eu/project/rcn/95054_en.html) based on several QTLs associated with SCC that were found in Spanish Churra, French Lacaune and Italian Sardinian-Lacaune Backcross populations. Seven regions of interest were selected based on commonalities among populations or on high significance on sheep chromosomes (OAR) 2, 3, 5, 16 and 18. SNPs were selected from the 54K or 800K Illumina ovine chips (Nicolazzi *et al.*, 2015) or were identified within the "3SR" project by novel genome sequencing. Details are provided

in Table 2. Genomic positions refer to *Ovis aries* assembly v3.1 (reference NCBI).

Genome wide association mapping

After quality control, 755 SNPs were kept for the study. Genome-wide association studies (GWAS) were performed for each phenotype using the univariate mixed model approach implemented in the Genome-wide Efficient Mixed Model Association (GEMMA) software (Zhou & Stephens, 2012). A Bonferroni correction of $\alpha = 5\%$ was applied (significance threshold = α/number of SNP). SNPs with a p-value < 6.63e-05 were considered to be significantly associated at the chip level. Chromosome-wide significance thresholds were also calculated with the same method.

Close Linkage versus Pleiotropism Test

Because QTL hits were found in the same genomic region for different traits (OAR3 and OAR16), the "Close LInkage versus Pleiotropism test" (CLIP) developed by David *et al.* (2013) 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.

Results and Discussion

Genome wide association mapping

The OAR3 region (around 129 Mb) was found significant for 9 out of 12 traits related to SCC, bacterial titre and weight. Table 1 provides for each analysis the p-values of the two most significant SNPs. The first SNP, ss1553223196, is in the coding frame of the *Socs2* gene at position 129,722,200 and the second, ss1553223200, is located in the 3'UTR of CRADD gene at position 129,927,538. The p-values of the *Socs2* SNP were strongly significant for the LSCS traits (p-values equal to 3.09e-07 for the 1st lactation), and, to a lesser extent, for weight phenotypes (p-values from 3.71e-04 to 1.54e-03). These results confirmed earlier findings (Rupp *et al.*, 2015) in a new independent population. The present study further showed a significant association of the OAR3 QTL region with bacterial titre and chronic mastitis. Finally, milk production traits were not found associated to any of the SNPs in this data set.

The SNP located in the CRADD gene was always less significant than the SNP in the *Socs2* gene, except for W_DAY_920. This is additional evidence that the *Socs2* gene is the causal mutation for mastitis resistance and weight, as strongly suggested by Rupp *et al.* (2015).

Some other SNPs from the 960-custom mastitis chip which are not on OAR 3, were found significant in our study: one SNP on OAR 16 for LSCS_L2 (p-value = 9.05e-07 ; MAF = 0.015), another SNP on OAR 19 for W_DAY_920 (p-value = 3.43e-05 ; MAF = 0.143), another SNP on OAR 16 for MILK_L1 (p-value = 1.87e-05 ; MAF = 0.062), and one SNP on OAR 19 for MILK_L2 (p-value = 1.58e-04 ; MAF = 0.219). These 4 SNPs, however, were isolated hits on a single trait and gave low confidence for true QTLs.

Pleiotropic effects of the mutation

The CLIP test was performed for LSCS_L1, STAPH_L1, W_DAY_250 and MILK_L1 traits in the OAR 3 QTL region (13 SNPs ranged from 129.68 Mb to 130.10 Mb) in order to determine whether those QTLs were a single pleiotropic QTL or different QTLs in close linkage. The results are presented in Table 3. The pleiotropic assumption was never rejected in any of the two-trait analyses with LSCS_L1, STAPH_L1 and W_DAY_250, thanks to strong observed correlations of SNP effects, especially between LSCS_L1 and STAPH_L1 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_L1 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.

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. In contrast, the small effect of the *Socs2* gene mutation on milk production found by Rupp *et al.* (2015) is not confirmed in any of the analyses with our data set. The differences observed in milk production between animals carrying or not the *Socs2* mutation could have been caused by an indirect effect on animal size rather than by a direct interaction between SOCS-2 protein and the prolactin hormone receptor.

Conclusion

Our study confirms one ovine QTL controlling mastitis resistance on chromosome 3 using a 960 custom-designed ovine SNP chip. GWAS and CLIP test were performed and confirmed that the QTL is caused by the R96C mutation in the *Socs2* gene with pleiotropic effects on SCC, bacterial infection, size and more indirectly on milk production.

List of References

Banos G., G. Bramis, S. J. Bush, E. L. Clark, M. E.B. McCulloch, J. Smith, G. Schulze, G. Arsenos, D. A. Hume, A. Psifidi, 2017. The genomic architecture of mastitis resistance in dairy sheep. BMC genomics. 18: 624.

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.

Nicolazzi, E. L., A. Caprera, N. Nazzicari, P. Cozzi, F. Strozzi, C. Lawley, A. Pirani, C. Soans, F. Brew, H. Jorjani, G. Evans, B. Simpson, G. Tosser-Klopp, R. Brauning, J. L. Williams, A. Stella, 2015. SNPchiMp v.3: integrating and standardizing single nucleotide polymorphism data for livestock species. BMC Genomics. 16: 283.

Rupp, R., D. Bergonier, S. Dion, M. C. Hygonenq, M. R. Aurel, C. Robert-Granié, G. Foucras, 2009. Response to somatic cell count-based selection for mastitis resistance in a divergent selection experiment sheep. J. Dairy Sci. 92(3): 1203-1219.

Rupp, R., P. Senin, J. Sarry, C. Allain, C. Tasca, L. Ligat, D. Portes, F. Woloszyn, O. Bouchez, G. Tabouret, M. Lebastard, C. Caubet, G. Foucras, G. Tosser-Klopp, 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 mild production in a sheep model. PLoS Genet. 11(12): e1005629.

Zhou, X. & M. Stephens, 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 42: 821-824.

Additionnal Figures and Tables

Table 1. Mean, Standard Deviation, Headcount and p-values of the two most significant SNPs located on chromosome 3 for each trait.

Trait	Mean (± SD)	Nb	P-value[1] ss1553223196	P-value[1] ss1553223200
LSCS_L1	3.08 ± 1.57	494	3.03e-07***	5.24e-07***
LSCS_L2	3.33 ± 1.92	391	2.61e-05***	6.41e-05***
ABSCESS	15 cases / 361 controls	376	2.83e-03*	2.24e-02
STAPH_L1	0.69 ± 0.83	404	6.80e-04*	1.61e-03*
W_BIRTH	3.9 ± 0.6	502	3.92e-02	4.62e-02
W_DAY_100	28.7 ± 3.0	492	5.97e-04*	1.17e-03
W_DAY_250	50.6 ± 4.1	472	3.71e-04*	2.27e-03
W_1ST_LAMB	64.7 ± 6.7	438	5.01e-04*	1.83e-03
W_2ND_LAMB	73.4 ± 7.7	393	1.54e-03*	2.60e-03
W_DAY_920	67.3 ± 7.2	378	9.48e-04*	9.34e-04
MILK_L1	999 ± 182	494	2.01e-01	2.44e-01
MILK_L2	1 114 ± 231	391	5.77e-01	7.65e-01

[1] ***Genome wide: SNP with a p-value < 6.63e-05; *Chromosome wide: 3.85e-03 for chromosome 3

Table 2. SNP content for the 3SR-mastitis-960-SNP chip

OAR	Discovery population(s)	QTL loc (Mb)	Covered interval min-max (Mb)		Origin of the SNPs 54K	800K	Seq	SNP number before-after qual. control	
2	Lacaune	125	124.6	125.5	19	10		29	24
2	Sardinian BC Churra	206.7 208.7	204.2	209.7	106	146		252	196
3	Lacaune	129.7	129.7	130.1	4	6	10	20	13
5	Lacaune	96.5	95.8	96.2	11	9		20	17
12	Churra Sardinian BC	18.3 18.4	17.9	20.4	54	76		130	101
16	Lacaune	36.2	28.3	36.8	144	110		254	202
19	Churra Lacaune	26.3 28.6	23.6	28.9	75	180		255	202

Table 3. Results of the CLIP test

Trait 1	Trait 2	Correlation observed	Correlation limit at 5%	Significance (corr obs > corr limit)
LSCS_L1	STAPH_L1	0.93	0.67	Yes
LSCS_L1	W_DAY_250	0.65	0.58	Yes
LSCS_L1	MILK_L1	0.24	0.14	Yes
STAPH_L1	W_DAY_250	0.52	0.19	Yes
STAPH_L1	MILK_L1	6.31e-03	3.58e-02	No
W_DAY_250	MILK_L1	0.66	1.09e-02	Yes

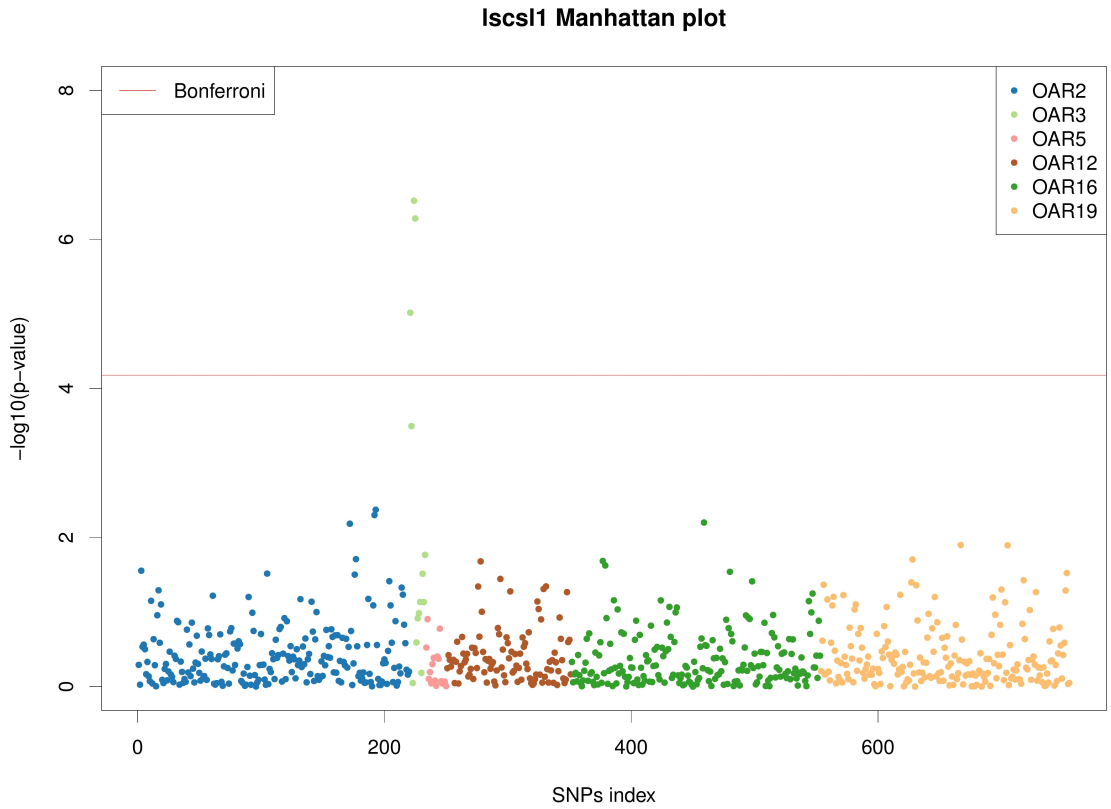

Figure 1. Manhattan plot of the Genome-Wide Association Study on LSCSL1. Each colour change corresponds to a chromosome.