



Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA

Strillacci *et al.*

RESEARCH ARTICLE

Open Access

Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA

Maria G Strillacci^{1†}, Erika Frigo^{1†}, Fausta Schiavini¹, Antonia B Samoré², Fabiola Canavesi¹, Mario Vevey⁴, Maria C Cozzi¹, Morris Soller³, Ehud Lipkin³ and Alessandro Bagnato^{1,2*}

Abstract

Background: Mastitis is a major disease of dairy cattle occurring in response to environmental exposure to infective agents with a great economic impact on dairy industry. Somatic cell count (SCC) and its log transformation in somatic cell score (SCS) are traits that have been used as indirect measures of resistance to mastitis for decades in selective breeding. A selective DNA pooling (SDP) approach was applied to identify Quantitative Trait Loci (QTL) for SCS in Valdostana Red Pied cattle using the Illumina Bovine HD BeadChip.

Results: A total of 171 SNPs reached the genome-wide significance for association with SCS. Fifty-two SNPs were annotated within genes, some of those involved in the immune response to mastitis. On BTAs 1, 2, 3, 4, 9, 13, 15, 17, 21 and 22 the largest number of markers in association to the trait was found. These regions identified novel genomic regions related to mastitis (1-Mb SNP windows) and confirmed those already mapped. The largest number of significant SNPs exceeding the threshold for genome-wide significant signal was found on BTA 15, located at 50.43-51.63 Mb.

Conclusions: The genomic regions identified in this study contribute to a better understanding of the genetic control of the mastitis immune response in cattle and may allow the inclusion of more detailed QTL information in selection programs.

Keywords: Mastitis, GWAS, Valdostana Red Pied breed, SCS

Background

Mastitis is one of the most frequent inflammatory disease with a significant economic implication for the dairy herds and the resistance to this pathology may be improved by breeding.

The development of mastitis is the result of the interaction among three components: the individual genotype, the pathogens (ordinarily classified in contagious and environmental bacteria) and the environment (hygiene, housing, climate, milking machines, feeding) [1].

The resistance to an infection disease or the absence of susceptibility may be defined as the immune response

ability (immuno-competence capability) of an animal, to avoid the pathogens replication after the establishment of an infection. This implies that animals tend to vary in their genetic potential for immuno-competence [2]. The genetic resistance or the genetic susceptibility to mastitis involves interlinked biological mechanisms that activate and regulate the different levels of the immune response, as a consequence of the differences existing in the response to mastitis involving several pathogens [3]. A better understanding of the immune system and of the metabolic pathways involved in the response to various pathogens of resistant and susceptible animals may be used as complementary approach for the disease control.

The discovery of millions of SNP markers in animal genomes forming dense marker panels, and the concomitant decrease in genotyping costs have allowed the performing of genome-wide association studies (GWAS) [4]. The

* Correspondence: alessandro.bagnato@unimi.it

†Equal contributors

¹Department of Health, Animal Science and Food Safety (VESPA), University of Milan, Via Celoria 10, 20133 Milan, Italy

²Dipartimento di Scienze e Tecnologie Agro - Alimentari, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy

Full list of author information is available at the end of the article

availability of SNP dense genotypes have increased the power of the identification of QTL related to the traits of interest [5], allowing more accurate breeding values estimation with the use of genomic selection methodology and helping the understanding of the genetic control of the traits of interest [6]. Because of the established knowledge of the positive genetic correlation between clinical mastitis and SCS ranging from 0.6 to 0.8 [1], SCC is one of the traits used as an indirect measure of mastitis resistance/susceptibility in breeding programs in cattle and sheep. Many GWAS have detected QTL for SCC in cattle on BTAs 5, 6, 8, 11, 17, 18, 20 and 23 in cosmopolite improved dairy cattle breeds [1,7].

The high costs of screening large populations for marker allele frequencies can be decreased using the SDP approach, genotyping pooled DNA samples from selected individuals at each of the two phenotypic extremes of the trait distribution [8]. Equal amounts of DNA are pooled from individuals in the extreme tails, and pools are then genotyped to estimate allele frequency differences for each SNP among high and low tail pools. The significant identified candidate SNPs are then used for confirmatory association studies [9].

The aim of this study was to identify QTL associated with SCS as an indicator of mastitis. We performed a GWA study for SCS in the Valdostana Red Pied cattle, with a selective DNA pooling analysis, using the Illumina BovineHD Bead chip.

Results and discussion

Among the 2,417 bulls with DP-EBV values, 275 had semen samples available in the Valdostana Red Pied bio-bank that encompassed in total 373 sires samples spanning across generations.

The Valdostana Red Pied population counting at present about 11,000 milking cows, did not undergo focused selection for milk production only and no gene introgression from other populations have ever occurred. The breed is strongly adapted to harsh alpine environment because breed natural adaptation and because has been selected to maintain pasture capability (summer pasture is the common farming system), longevity, functionality and fertility. Thus the population is somehow a unique genetic resource to map mastitis resistance, a trait related to adaptation, functionality and longevity. The study used all the sire samples available in the Valdostana

Red Pied bio-bank thus highlighting the overall observable variability for productive and functional traits in this breed. The smaller number of sires available for the study respect to mapping, may limit the capacity to disclose QTL for mastitis resistance. Nevertheless the experimental design here used and the genetic makeup of the population allowed to identify several new QTL and confirm regions identified in the Italian and Swiss Brown population [10], another breed originating from alpine region, now strongly selected for milk production.

Descriptive statistics for the DP-EBVs and the size of the pools for each tail are reported in Table 1.

The initial dataset included 721,644 SNPs. After editing, the association analysis were performed with 655,665 SNPs for SCS DP-EBV.

Figure 1 shows the Q-Q plot of SNPs at marker level (p-values). Deviations from the identity line showed the amount of false positive tests resulted from the analysis of the data. Figure 2 showed the Manhattan plot of genome-wide associations for SCS trait.

A total of 171 significant SNPs in 24 chromosomes were identified above the Bonferroni genome-wide threshold of 0.05. The Additional file 1 showed the list of the 171 significant SNPs identified. The SNPs location and the gene annotation were reported for both the UMD3.1 and Btau4.6.1 assembly. Table in Additional file 1 included the indication of QTL, amongst the ones here disclosed, reported in the online AnimalQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/index>) for clinical mastitis, SCC and SCS.

Intragenic SNPs

Among the 171 significant markers, 52 SNPs were annotated within 36 genes (Table 2). In Table 2 the significant intragenic SNPs and their corresponding annotated genes in the Btau4.6.1 assembly are reported.

The *BovineHD0900019961* (*rs136413030*) SNP was associated to the *VNN1* (*vanin 1*) on BTA 9, the *BovineHD1500008135* (*rs134980659*) SNP was associated to the *THY1* (*Thy-1 cell surface antigen*) located on BTA 15 and the *BovineHD2100001405* (*rs133992914*) SNP was associated to the *IGF1R* (*insulin-like growth factor 1 receptor*), located on BTA 21.

Also the *BovineHD1500008366* (*rs41754552*) and the *BovineHD1500008367* (*rs110269361*) SNPs were located

Table 1 Details for DP-EBVs mean and their SD values and DP-EBVs reliability (REL) values for low and high tail pools

Pool	N° of samples	DP-EBV mean	Mean SD	DP-EBV REL mean	Pool	N° of samples	DP-EBV mean	Mean SD	DP-EBV REL mean
Low tail_1	20	-1.151	0.324	0.535	High tail_1	20	1.257	0.395	0.493
Low tail_2	19	-1.080	0.251	0.600	High tail_2	20	1.134	0.285	0.574

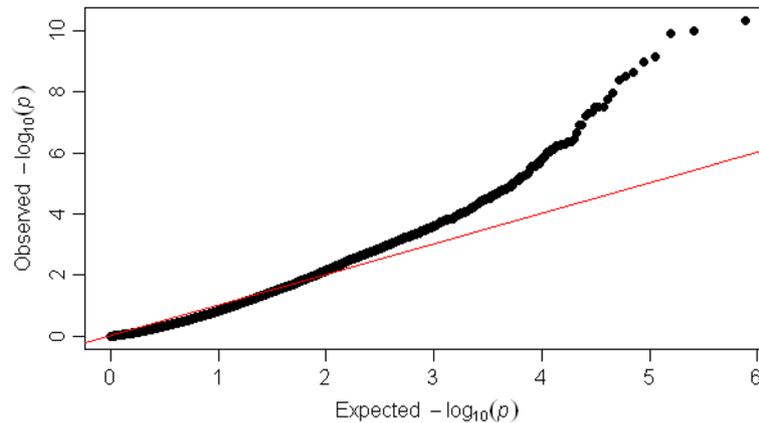


Figure 1 Q-Q plot of SNPs at marker level (p-values).

respectively at 594,104 bp and 601,630 bp from *THY1* on BTA 15.

THY1 is one of the genes differentially expressed between control quarters from cows infected with *E. coli* and *S. aureus* pathogens [11]. Also Moyes et al. 2010 [12] reported the *THY1* upregulation in *S. uberis* intramammary infections.

Sugimoto and Sugimoto [13] provided evidence that the *IGF1R* is involved in innate immunity through autophagy (general term for the degradation of cytoplasmic components within lysosomes, [14]) in bovine. In *Bos taurus*, in fact, a polymorphism in the 5'UTR region of *IGF1R* (BTA 21) was associated to mastitis incidence, determining the inhibition of autophagy in response to *S. Agalactiae* invasion.

Nearby genes SNPs

The *BovineHD0900019716* (*rs109049649*), the *BovineHD4100007550* (*rs41662465*) and the *Hapmap49339-BTA-84110* (*rs41662464*) SNPs were mapped near the

VNN1 (*vanin 1*) and the *VNN2* (*vanin 2*) located on BTA 9 respectively at 73.37 Mb and 73.39 Mb. On the same BTA 9, the *BovineHD0900019961* (*rs136413030*) SNP was close to *VNN2*. Jiang et al. [15] reported that *VNN1* and *VNN2* are related to resistance to bovine mastitis, being ranked among the 160 most mastitis relevant genes.

On BTA 19, at 55 Mb, *SOCS3* (*suppressor of cytokine signalling 3*) was found at 673,863 bp upstream the *BovineHD1900015066* (*rs132720248*) SNP. This gene, important for the mammary tissue homeostasis, encodes an intracellular inhibitor of cytokine signaling, thus playing an important role in the initial steps of the recognition of pathogen-associated molecular pattern (PAMP) of the innate immune cells. This leads to the activation and initiation of the innate and the adaptive immune responses. Heeg and Dalpke [16] and Brenaut et al. [17] found the *SOCS3* gene among the 39 differentially expressed genes in milk fat globules of goats in response to an experimental intramammary infection with *S. aureus*.

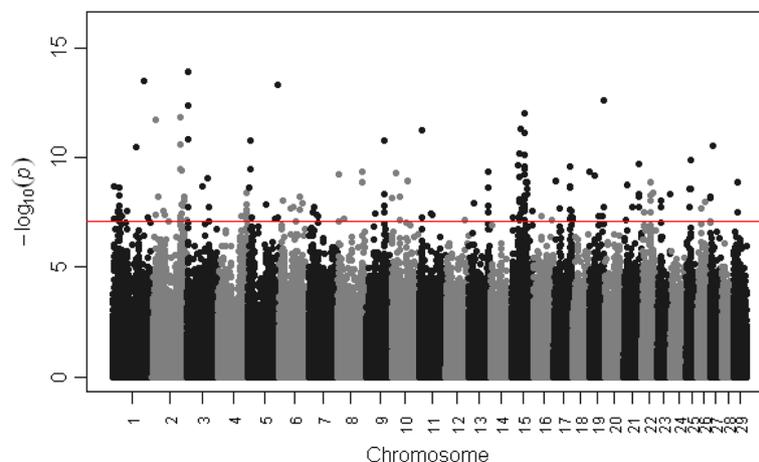


Figure 2 Manhattan plot of genome-wide associations for SCS trait. The red line represents the Bonferroni correction threshold.

Table 2 Significant intragenic SNPs above the Bonferroni genome-wide threshold of 0.05

llumina SNP name	Genbank SNP code	P-Value	BTA	SNP location	Gene symbol
<i>BovineHD0100007623</i>	<i>rs137585939</i>	4.43E-08	1	26254309	<i>ROBO1</i>
<i>BovineHD0100040084</i>	<i>rs43273786</i>	5.41E-08	1	141228619	<i>NEK11</i>
<i>BovineHD0200004154</i>	<i>rs110997154</i>	3.76E-08	2	15142189	<i>SSFA2</i>
<i>BovineHD0300000560</i>	<i>rs110459674</i>	inf	3	2612333	<i>TADA1</i>
<i>BovineHD0300002104</i>	<i>rs110093914</i>	1.87E-08	3	7276675	<i>DDR2</i>
<i>BovineHD0300018913</i>	<i>rs42371455</i>	1.87E-09	3	66866123	<i>LPHN2</i>
<i>BovineHD0300023699</i>	<i>rs135870054</i>	8.96E-10	3	87296944	<i>ALG6</i>
<i>BovineHD0400026934</i>	<i>rs109307332</i>	4.69E-08	4	98543003	<i>PLXNA4</i>
<i>BovineHD0500003126</i>	<i>rs134685896</i>	3.26E-10	5	12834395	<i>ACSS3</i>
<i>BovineHD0700010213</i>	<i>rs133885406</i>	4.73E-08	7	33526558	<i>HSD17B4</i>
<i>BovineHD0900019961</i>	<i>rs136413030</i>	3.70E-08	9	73355572	<i>VNN1</i>
<i>BovineHD1000004333</i>	<i>rs43612234</i>	5.32E-10	10	12722576	<i>MEGF11</i>
<i>BovineHD1000009424</i>	<i>rs43623003</i>	6.65E-08	10	28079552	<i>MIR2284Z-1</i>
<i>BovineHD1000009428</i>	<i>rs110034517</i>	5.56E-09	10	28102288	<i>MIR2284Z-1</i>
<i>BovineHD1000017503</i>	<i>rs42486408</i>	1.16E-09	10	60793897	<i>TRPM7</i>
<i>BovineHD1100003814</i>	<i>rs109489659</i>	5.53E-12	11	11771322	<i>CCT7</i>
<i>BovineHD1300006368</i>	<i>rs109943824</i>	1.12E-08	13	20845530	<i>PLXDC2</i>
<i>BovineHD1300022672</i>	<i>rs41710487</i>	4.53E-10	13	78416778	<i>KCNB1</i>
<i>BovineHD1500008135</i>	<i>rs134980659</i>	6.49E-08	15	28399876	<i>THY1</i>
<i>Hapmap40064-BTA-36665</i>	<i>rs41631137</i>	4.65E-12	15	33953859	<i>PIK3C2A</i>
<i>BovineHD1500015036</i>	<i>rs41769292</i>	5.15E-09	15	50730325	<i>NUP98</i>
<i>BovineHD1500015037</i>	<i>rs134338365</i>	2.50E-10	15	50733648	<i>NUP98</i>
<i>BTB-00604170</i>	<i>rs41769258</i>	5.84E-10	15	50753778	<i>NUP98</i>
<i>BovineHD1500015042</i>	<i>rs41769237</i>	3.41E-10	15	50765770	<i>NUP98</i>
<i>BovineHD1500015044</i>	<i>rs109649273</i>	6.55E-08	15	50769861	<i>NUP98</i>
<i>BovineHD1500015047</i>	<i>rs41768429</i>	7.66E-12	15	50774198	<i>NUP98</i>
<i>BovineHD1500015049</i>	<i>rs41768423</i>	1.42E-08	15	50780537	<i>NUP98</i>
<i>BovineHD1500015051</i>	<i>rs41768414</i>	6.94E-11	15	50784307	<i>NUP98</i>
<i>BovineHD1500015054</i>	<i>rs41768364</i>	7.12E-08	15	50792403	<i>NUP98</i>
<i>BovineHD1500015055</i>	<i>rs109966062</i>	1.99E-08	15	50795681	<i>NUP98</i>
<i>BovineHD1500015056</i>	<i>rs41768379</i>	9.17E-13	15	50799229	<i>NUP98</i>
<i>BovineHD4100012071</i>	<i>rs136525289</i>	7.93E-09	15	51638163	<i>PDE2A</i>
<i>BTA-18105-no-rs</i>	<i>rs109715014</i>	4.33E-09	15	62952170	<i>CCDC73</i>
<i>BovineHD1600009946</i>	<i>rs41798963</i>	4.62E-08	16	31290905	<i>CEP170</i>
<i>BovineHD1600021693</i>	<i>rs41819133</i>	7.33E-08	16	71743691	<i>CAMK1G</i>
<i>BovineHD1700002750</i>	<i>rs110828704</i>	1.11E-09	17	10472292	<i>NR3C2</i>
<i>BovineHD1700018352</i>	<i>rs135157738</i>	2.29E-09	17	64466089	<i>RPH3A</i>
<i>BovineHD1700019237</i>	<i>rs110644998</i>	2.43E-10	17	67344705	<i>CORO1C</i>
<i>BovineHD1700019238</i>	<i>rs134453171</i>	1.90E-09	17	67347843	<i>CORO1C</i>
<i>BovineHD1700020721</i>	<i>rs109085689</i>	4.61E-08	17	72364594	<i>MTMR3</i>
<i>BovineHD1700021131</i>	<i>rs135044766</i>	4.88E-08	17	73638738	<i>DEPDC5</i>
<i>BovineHD1700021132</i>	<i>rs135814317</i>	7.10E-08	17	73640453	<i>DEPDC5</i>
<i>BovineHD1900006167</i>	<i>rs134967563</i>	7.01E-10	19	20731168	<i>SSH2</i>
<i>BovineHD4100014346</i>	<i>rs29017164</i>	2.41E-13	19	57590345	<i>ATP5H</i>

Table 2 Significant intragenic SNPs above the Bonferroni genome-wide threshold of 0.05 (Continued)

Gene	SNP	P-value	Chr	Position	Gene
BovineHD2100001405	rs133992914	6.47E-08	21	6826694	IGF1R
ARS-BFGL-NGS-10830	rs109014211	1.74E-09	21	14303664	SLCO3A1
BovineHD2200009526	rs110064285	8.84E-09	22	33753508	FAM19A1
BovineHD2200009645	rs135018045	3.20E-08	22	34006051	FAM19A1
BovineHD2200009658	rs133223316	5.38E-09	22	34051778	FAM19A1
BovineHD2300014695	rs110724706	4.60E-09	23	50469508	TUBB2B
BovineHD2500003334	rs42064606	1.77E-08	25	13011549	SHISA9
BovineHD2500003336	rs109087355	2.45E-09	25	13017281	SHISA9

Genes and SNPs location as in the Btau4.6.1 assembly; gene symbol as in GenBank.

The gene encoding for the *serine dehydratase (SDS)* on BTA 17 was located 416,619 bp upstream of the *BovineHD1700018352 (rs135157738)* SNP. This gene is included in the glycine, serine and threonine metabolism, as reported by [18]. These authors demonstrated that the serine dehydratase is one of the enzymes that changed significantly in bovine affected to mastitis.

Four SNPs on BTA 9 (*BovineHD0900019961 (rs136413030)*, *BovineHD0900019716 (rs109049649)*, *BovineHD4100007550 (rs41662465)* and *Hapmap49339-BTA-84110 (rs41662464)*) mapped near *CTGF (connective tissue growth factor)*. The *ZNF1 (X1-type zinc finger-containing)* on BTA 13 was close to four SNPs (*BovineHD4100010442 (rs41634068)*, *BovineHD1300022626 (rs137320993)*, *BovineHD1300022630 (rs109123247)* and *BovineHD1300022672 (rs41710487)*). The *TRIM21 (tripartite motif containing 21)* was located 444,354 Mb upstream the strongest association chromosome region identified in BTA 15 (Table 3). The *CXCL2*

(*Chemokine (C-X-C motif) ligand 2*) and the *CXCL10 (Chemokine (C-X-C motif) ligand 10)* on BTA6 were significantly associated to the *BovineHD0600025253 (rs42615160)* SNP.

The genes above mentioned near to significant SNPs (*ZNF1, CTGF, TRIM21, CXCL2* and *CXCL10*) are significantly differentially expressed by the bovine mammary epithelial cells stimulated with *E. coli* crude lipopolysaccharide [19].

Jensen et al. 2013 [11] studied and compared the transcriptional responses of uninfected mammary gland quarters adjacent to quarters infected with *E. coli* and *S. aureus* in Holstein cows. The *CXCL2* resulted to be one of the genes differentially expressed between control quarters infected with both the pathogens, while the *CXCL10* resulted to be one of the genes differentially expressed in control quarters from animals infected with *S. aureus* for 24 and 72 hours.

Table 3 List of chromosome regions strongly associated to SCS

BTA	Start*	End*	Length (BP)	N. SNPs	Genbank SNP code
1	21625461	21632949	7488	3	rs110141424; rs42365792; rs42367069
1	27814460	28017039	202579	7	rs135454183; rs110174548; rs134436790; rs136371716; rs111001290; rs41586446; rs110002182
2	117668432	118739748	1071316	9	rs134103593; rs109545959; rs133621389; rs135143470; rs136343471; rs109908642; rs133815275; rs135205101; rs43320680
3	6388643	6396280	7637	3	rs110787209; rs42458782; rs132773940
4	117852857	118898784	1045927	4	rs133335423; rs43417362; rs133867064; rs136879377
9	72784616	72804256	19640	4	rs41662464; rs109049649; rs41662465; rs136413030
13	78273095	78416778	143683	4	rs41634068; rs137320993; rs109123247; rs41710487
15	28399876	28999494	599618	3	rs134980659; rs41754552; rs110269361
15	31285729	32027462	741733	5	rs135835073; rs29018094; rs110325464; rs43299708; rs43299703
15	50438721	51638163	1199442	14	rs137687321; rs108941833; rs41769292; rs134338365; rs41769258; rs41769237; rs109649273; rs41768429; rs41768423; rs41768414; rs41768364; rs109966062; rs41768379; rs136525289
17	67344705	67375670	30965	3	rs110644998; rs134453171; rs41850009
21 [#]	60154246	60175026	20780	4	rs29018575; rs42236250; rs42236274; rs109897238
22	33753508	34051778	298270	3	rs110064285; rs135018045; rs133223316

Start. End*: candidate region start and end (bp).

[#]Start and End position referred to Btau4.6.1 assembly.

The *BovineHD2200003506* (*rs110821186*) SNP on BTA 22 mapped close to the *MYD88* (*myeloid differentiation primary-response gene 88*) at 11.72 Mb, which plays a functional role in transducing pro-inflammatory molecule lipopolysaccharide (LPS) that are responsible for the majority of acute clinical cases of mastitis [20].

Chromosome regions associated to SCS and clinical mastitis

Table 3 reported a list of the chromosome regions defined by at least three SNPs that were strongly associated to SCS. The highest number of significant SNPs (14) exceeding the significant threshold for genome-wide significance signal was found on BTA 15 (located at 50.43-51.63 Mb). On the same BTA 15, also two smaller peaks consisting of three SNPs located at 28.39-28.99 and 5 SNPs located at 31.28-32.02 Mb were identified. These regions are located in QTL that were mapped, respectively, for clinical mastitis using a linkage analysis [21] and for SCS [22]. The region located at 50.43-51.63 Mb on BTA 15 has not been reported before in cattle breeds (<http://www.animalgenome.org/cgi-bin/QTLdb/index>), thus identifying a supposed candidate chromosome region associated to SCS. The chromosome region on BTA 9 (72.78-72.80 Mb) mapped in a QTL region previously identified for the general disease resistance (including clinical mastitis) and for SCS [23].

Lund et al. [21] found a QTL region associated to SCS located at 32.62-43.31 Mb on BTA 22. In our study, three significant SNPs were in this region.

Sahana et al. [24] in a study on the confirmation and fine-mapping of clinical mastitis and SCS QTL in Nordic Holstein cattle using BovineSNP50 BeadChip, found the highest number of significant associations on BTA 6 identifying a QTL region for clinical mastitis at 83.37-88.89 Mb (UMD3.1 assembly). This result was also confirmed in a recent study in German Holstein cattle [25]. In our study, two significant SNPs (*BovineHD0600023179* (*rs133319155*) and *BovineHD0600023185* (*rs136907262*)) were found respectively at 84.25 and 84.26 Mb on BTA 6 (UMD3.1 assembly; Btau4.6.1 assembly position was not available), being mapped within the QTL region described by the authors previously cited (see Additional file 1).

Annotation

Among the 36 genes listed in Table 2, the annotation data were available for 23 genes reported in the Additional file 2. This lists the biological processes (BP), the cellular components (CC), the molecular function (MF) and the metabolic pathways (KEGG) obtained with the annotation analyses performed with DAVID online Database.

The literature brings evidence that some of the genes reported in Table 2 map in QTL associated to traits of economic importance in bovine (<http://www.animalgenome.org/cgi-bin/QTLdb/BT/index>), as showed in Additional

file 3. Those mapping in QTL already associated to clinical mastitis and SCS reported in the QTLdb were only 4: the *PLXNA4* (*plexin A4*) on BTA 4, the *THY1* (*Thy-1 cell surface antigen*) on BTA 15 and the *SHISA9* (*known as CKAMP44, shisa homolog 9*) on BTA 25, the *FAM19A1* (*family with sequence similarity 19 (chemokine (C-C motif)-like), member A1*) on BTA 22 associated with SCS. This study thus highlighted possible QTL related to mastitis resistance in the other 19 genes annotated and considered in the GO analysis.

Conclusions

This is the first mapping for SCS in Valdostana Red Pied population, an autochthonous alpine dual purpose cattle breed whose selection is mainly focused on milk quality, meat production and functionality.

This study brings evidence of significant associations between SCS and SNP markers on several chromosomes in known and newly disclosed QTL regions. Some genes involved in mastitis resistance or variation of SCS content were in QTL on BTAs 9, 13, 15, 17, 19, 21, 22. In particular, the strongest associations were highlighted on BTA 15 with a total of 24 significant SNPs distributed in three regions.

The detection of genomic regions will help to understand which potential candidate genes may be responsible for the genetic variation in mastitis resistance/susceptibility, a trait of primary importance in dairy cattle breeding and farming.

Methods

Sampling

The Valdostana Red Pied cattle is the most common autochthonous dual purpose breed in the region Val d'Aosta (13,000 animals in 2013, almost all of them registered in the Herd Book), coming from the red pied cattle and dating back to the end of the fifth century. The National Association of Valdostana Breeders (A.N.A.Bo.Ra.Va.) provided commercial semen samples for 373 bulls and 725,337 test day records from milk routine recording from 45,410 cows. No animals were involved in the work.

The daily SCC were transformed into SCS [26]. Genetic parameters and estimated breeding values (EBVs) were calculated with a test day repeatability model on first parity cows. The model of analysis considered the fixed effects of days in milk (10 classes of 30 days each), herd-test day effect (32,870 levels), month of calving and age at calving (12 classes). Additive genetic and permanent environmental effects were considered as random. Three generations of ancestors were used for each individual extracting information from the National Herd Book for a total of 35,803 animals. Variance component estimations were calculated based on 258,680 test day records with the software VCE [27] and individual EBVs were obtained with the

package BLUPF90 [28]. Deregressed proofs (DP-EBV) were calculated for 2,417 bulls according to [29].

Pool constitution

The bull families structure was verified in terms of number of sons per bull, in order to avoid overrepresentation of a single sire. Only 1 bull had 6 sons, 4 bulls had 5 sons, 3 bulls had 4 sons and the rest of bulls had 3 or less sons. The sires were ranked according to DP-EBVs for SCS: the top 20% and bottom 20% sires were identified for the constitution of independent pools within tail of the DP-EBV distribution. In order to obtain two independent groups of different animals within tail with comparable phenotypic value, the selected samples for each tail were clustered (even and odds numbers) into 2 sub-pools.

A total of 79 samples were selected for the pools constitution as follows: 2 independent pools of 20 individuals each in the high tail and 2 independent pools of 20 and 19 individuals each in the low tail. Furthermore, for each pool, 2 DNA duplicate-pools were independently constructed from identical samples. Thus, a total of 4 pools per tail were produced.

DNA extraction and genotyping

Bulls DNA was extracted from semen samples using the ZR Genomic DNA™ Tissue MiniPrep (Zymo). The quality control was performed on each sample to verify the DNA integrity on Invitrogen E-Gel 1% Agarose Gel. The GloMax®-Multi Detection System instrument using the Quant-iT™ dsDNA Broad-Range (BR) Assay Kit (Life Technologies), determined the initial DNA concentrations. The DNA concentration for a single sample was evaluated three times and each read was verified twice (e.g. 2 instrument runs). Samples having concentration diverging ± 1 SD from the mean value were not included in the pools. Samples of DNA were normalized to a concentration of 10 ng/ul, which was reconfirmed with the same methods above described. DNA pools were constructed by taking equivalent amounts of DNA from each sample.

The final pools were concentrated to 50 ng/ul, as required for the Illumina array protocol. Each sub-pool was genotyped 3 times on different chips (array replicates). In all, 24 different chip positions on 3 microarrays were used for the pooled genotyping. Genotyping was performed using the Illumina BovineHD BeadChip (777,962 SNPs) according to the Infinium protocol. SNPs positions were accordingly to the UMB 3.1 bovine assembly.

Statistical analysis of pools

Pools were analysed according to the SDP approach. The B-allele frequencies being a good estimator of the allele frequency of the individuals in a pool for each array

replicate [30], were used in the analyses after obtaining them from the self-normalization algorithm of Illumina BeadStudio software®.

The multiple marker test

A pipeline in R software (<http://www.r-project.org/>) was adapted from [31] and [32] to perform a multiple marker test. The test statistic used for each SNP was:

$$Z_{test} = D_{test}/SD(D_{null})$$

where D_{test} is the difference of the B-allele frequencies means among tails; D_{null} is the difference of the B-allele frequencies means within tails. The test statistic was distributed as χ^2 with one degree of freedom under the null hypothesis of equal allele frequencies.

Quality control

We performed the analysis after excluding the 1% of SNPs that showed the highest variability as indicated by the size of the mean measures from the replicate array within tail [9]. In addition, the monomorphic SNPs were deleted from the dataset. Anderson-Darling, Shapiro-Wilk and Kolmogorov-Smirnov normality tests were performed on the D_{null} distribution [33-35].

The distribution of the p-values using the quantile-quantile (Q-Q) plot was examined to estimate the number and the magnitude of the observed associations between genotyped SNPs and DP-EBVs, compared to the statistics expected under the null hypothesis of no association.

Using the $-\log_{10}$ of the linkage test p-values for each SNP, a Manhattan plot was created. Manhattan plot is a SNP set out across the chromosomes for left to the right, and the heights correspond to the strength of the associations of the trait.

Bonferroni correction for multiple testing was applied in the analysis. The genome-wide significance threshold was set as a corrected p-value ≤ 0.05 , which equated to a nominal p-value of approximately 7.62×10^{-8} .

Annotation

The annotation analysis of significant SNPs was performed using UCSC, NCBI ENSEMBL and the Bovine SNP Annotation Tool (Snat), integrating the information from a variety of public bioinformatics databases (NCBI Entrez Gene, UniProt, Gene Ontology (GO), KEGG PATHWAY and AnimalQTLdb [36]). The Illumina BovineHD SNPs positions were converted from *Bos_taurus_UMD_3.1* to *Btau_4.6.1* assembly using the Batch Coordinate Conversion option in UCSC database as required by Snat tools. UCSC and NCBI databases were used to annotate those SNPs not included in Snat and to verify which of the significant SNPs were close (within 1 Mb [31,37]) to functional genes. GO and pathway analyses were performed

using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7.

Additional files

Additional file 1: List of the significant SNPs identified in the Valdostana Red Pied breed.

Additional file 2: List of the biological processes, cellular components, molecular function and metabolic pathways obtained with the annotation analyses performed with DAVID online Database.

Additional file 3: List of the genes mapping in QTL associated to traits of economic importance in bovine.

Competing interests

The authors declare they have no competing interests.

Authors' contributions

MGS and EF designed and performed the experiments, analysed the data and wrote the manuscript. FS and MCC produced and genotyped the pools, ABS calculated the EBVs and the DP-EBVs, FC contributed to the statistical analysis and wrote the R scripts, MV sampled and collected the semen samples and the phenotypes, MS and EL supervised the experimental design and the analysis, AB conceived the idea, designed the experiment and supervised the analysis and the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This study was funded by EC-FP7/2007-2013, agreement n°222664, "Quantomics". Authors gratefully acknowledge A.N.A.Bo.Ra.Va. for the availability of semen samples and phenotypes.

Author details

¹Department of Health, Animal Science and Food Safety (VESPA), University of Milan, Via Celoria 10, 20133 Milan, Italy. ²Dipartimento di Scienze e Tecnologie Agro - Alimentari, University of Bologna, Viale Fanin 46, 40127 Bologna, Italy. ³Department of Genetics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel. ⁴Associazione Nazionale Allevatori Bovini di Razza Valdostana (A.N.A.Bo.Ra.Va.), Fraz. Favret, 5, 11020 Gressan, AO, Italy.

Received: 1 July 2014 Accepted: 25 September 2014

Published online: 06 October 2014

References

1. Sender G, Korwin-Kossakowska A, Pawlik A, Galal Abdel Hameed K, Oprządek J: Genetic basis of mastitis resistance in dairy cattle – a review. *Ann Anim Sci* 2013, **13**:663–673.
2. Knap PW, Bishop SC: Relationships between genetic change and infectious disease in domestic livestock. *Br Soc Anim Sci* 2000, **27**:65–80.
3. Schukken YH, Günther J, Fitzpatrick J, Fontaine MC, Goetze L, Holst O, Leigh J, Petzl W, Schuberth HJ, Sipka A, Smith DG, Quesnell R, Watts J, Yancey R, Zerbe H, Gurjar A, Zadoks RN, Seyfert HM, Members of the Pfizer mastitis research consortium: Host-response patterns of intramammary infections in dairy cows. *Vet Immunol Immunopathol* 2011, **144**(3–4):270–289.
4. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME: Invited review: genomic selection in dairy cattle: progress and challenges. *J Dairy Sci* 2009, **92**(2):433–443.
5. Meuwissen TH, Solberg TR, Shepherd R, Woolliams J: A fast algorithm for BayesB type of prediction of genome-wide estimates of genetic value. *Genet Sel Evol* 2009, **5**:41–42.
6. Pryce JE, Bolormaa S, Chamberlain AJ, Bowman PJ, Savin K, Goddard ME, Hayes BJ: A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. *J Dairy Sci* 2010, **93**(7):3331–3345.
7. Meredith B, Lynn D, Berry D, Kearney F, Bradley D, Finlay E, Fahey A: A genome-wide association study for somatic cell score using the illumine high density bovine beadchip identifies several novel QTL potentially related to mastitis susceptibility. *Front Genet* 2013, **4**:229.
8. Darvasi A, Soller M: Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait locus. *Genetics* 1994, **138**:1365–1373.
9. Janicki PK, Vealey R, Liu J, Escajeda J, Postula M, Welker K: Genome-wide association study using pooled DNA to identify candidate markers mediating susceptibility to postoperative nausea and vomiting. *Anesthesiology* 2011, **115**(1):54–64.
10. Bagnato A, Soller M, Lipkin E, Samoré AB, Velayutham D, Schiavini F, Rossoni A, Dolezal MA: Genome Wide Association Analysis in Italian Brown Swiss for Somatic Cell Count [abstract]. In *Program and Book of Abstract 4th International Conference on Quantitative Genetics: Understanding Variation in Complex Trait*; 2012:234–235.
11. Jensen K, Günther J, Talbot R, Petzl W, Zerbe H, Schuberth HJ, Seyfert HM, Glass EJ: *Escherichia coli*- and *Staphylococcus aureus*-induced mastitis differentially modulate transcriptional responses in neighbouring uninfected bovine mammary gland quarters. *BMC Genomics* 2013, **16**:14–36.
12. Moyes KM, Drackley JK, Morin DE, Looor JJ: Greater expression of TLR2, TLR4, and IL6 due to negative energy balance is associated with lower expression of HLA-DRA and HLA-A in bovine blood neutrophils after intramammary mastitis challenge with *Streptococcus uberis*. *Funct Integr Genomics* 2010, **10**(1):53–61.
13. Sugimoto M, Sugimoto Y: Variant in the 5' untranslated region of insulin-like growth factor 1 receptor is associated with susceptibility to mastitis in cattle. *G3 (Bethesda)* 2012, **2**(9):1077–1084.
14. Mizushima N: Autophagy: process and function. *Genes Dev* 2007, **21**(22):2861–2873.
15. Jiang L, Sørensen P, Thomsen B, Edwards SM, Skarman A, Røntved CM, Lund MS, Workman CT: Gene prioritization for livestock diseases by data integration. *Physiol Genomics* 2012, **44**(5):305–317.
16. Heeg K, Dalpke A: TLR-induced negative regulatory circuits: role of suppressor of cytokine signaling (SOCS) proteins in innate immunity. *Vaccine* 2003, **21**(Suppl 2):S61–S67.
17. Brenaut P, Lefèvre L, Rau A, Laloë D, Pisoni G, Moroni P, Bevilacqua C, Martin P: Contribution of mammary epithelial cells to the immune response during early stages of a bacterial infection to *Staphylococcus aureus*. *Vet Res* 2014, **45**:16.
18. Wang C, Wang J, Ju Z, Zhai R, Zhou L, Li Q, Li J, Li R, Huang J, Zhong J: Reconstruction of metabolic network in the bovine mammary gland tissue. *Mol Biol Rep* 2012, **39**(7):7311–7318.
19. Gilbert FB, Cunha P, Jensen K, Glass EJ, Foucras G, Robert-Granié C, Rupp R, Rainard P: Differential response of bovine mammary epithelial cells to *Staphylococcus aureus* or *Escherichia coli* agonists of the innate immune system. *Vet Res* 2013, **44**:40.
20. Cates EA, Connor EE, Mosser DM, Bannerman DD: Functional characterization of bovine TIRAP and MyD88 in mediating bacterial lipopolysaccharide-induced endothelial NF-kappaB activation and apoptosis. *Comp Immunol Microbiol Infect Dis* 2009, **32**(6):477–490.
21. Lund MS, Gulbrandsen B, Buitenhuis AJ, Thomsen B, Bendixen C: Detection of quantitative trait loci in Danish Holstein cattle affecting clinical mastitis, somatic cell score, udder conformation traits, and assessment of associated effects on milk yield. *J Dairy Sci* 2008, **91**:4028–4036.
22. Rupp R, Boichard D: Genetics of resistance to mastitis in dairy cattle. *Vet Res* 2003, **34**:671–688.
23. Holmberg M, Andersson-Eklund L: Quantitative trait loci affecting health traits in Swedish dairy cattle. *J Dairy Sci* 2004, **87**:2653–2659.
24. Sahana G, Gulbrandsen B, Thomsen B, Lund MS: Confirmation and fine-mapping of clinical mastitis and somatic cell score QTL in Nordic Holstein cattle. *Anim Genet* 2013, **44**(6):620–626.
25. Abdel-Shafy H, Bortfeldt RH, Reissmann M, Brockmann GA: Short communication: validation of somatic cell score-associated loci identified in a genome-wide association study in German Holstein cattle. *J Dairy Sci* 2014, **97**(4):2481–2486.
26. Wiggans GR, Shook GE: A lactation measure of somatic cell count. *J Dairy Sci* 1987, **70**(12):2666–2672.
27. Gilmour AR, Gogel BJ, Cullis BR, Thompson R: *ASReml User Guide Release 3.0*. Hemel Hempstead, UK: VSN Int. Ltd; 2009.
28. Mistl I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee D: **BLUPF90 and Related Programs (BGF90)**. [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=28-07.pdf]
29. VanRaden PM, Wiggans GR: Derivation, calculation, and use of national animal model information. *J Dairy Sci* 1991, **74**(8):2737–2746.

30. Janicki PK, Liu J: Accuracy of allele frequency estimates in pool DNA analyzed by high-density Illumina Human 610-Quad microarray. *Intern J Genomics Proteomics* 2009, **5**:1.
31. Strillacci MG, Frigo E, Canavesi F, Ungar Y, Schiavini F, Zaniboni L, Reghenzani L, Cozzi MC, Samoré AB, Kashi Y, Shimoni E, Tal-Stein R, Soller M, Lipkin E, Bagnato A: QTL mapping for conjugated linoleic acid, vaccenic acid and Δ^9 -desaturase in Italian Brown Swiss dairy cattle using selective DNA pooling. *Anim Gen* 2014, **45**:485–499.
32. Bagnato A, Schiavini F, Rossoni A, Maltecca C, Dolezal M, Medugorac I, Sölkner J, Russo V, Fontanesi L, Friedmann A, Soller M, Lipkin E: Quantitative trait loci affecting milk yield and protein percent in a three-country Brown Swiss population. *J Dairy Sci* 2008, **91**:767–783.
33. Stephens MA: Tests based on EDF statistics. In *Goodness-of-Fit Techniques*. Edited by D'Agostino RB, Stephens MA. New York: Marcel D; 1986:97–193.
34. Royston P: A remark on Algorithm AS 181: The W test for normality. In *Applied Statistics*. 44th edition. Edited by Wiley for the Royal Statistical Society. 1995:547–551.
35. Marsaglia G, Tsang WW, Jingbo W: Evaluating Kolmogorov's distribution. *J Stat Softw* 2003, **8**:1–4.
36. Jiang J, Jiang L, Zhou B, Fu W, Liu JF, Zhang Q: Snat: a SNP annotation tool for bovine by integrating various sources of genomic information. *BMC Genet* 2011, **12**:85.
37. Pant SD, Schenkel FS, Verschoor CP, You Q, Kelton DF, Moore SS, Karrow NA: A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in holstein cattle. *Genomics* 2010, **95**(3):176–182.

doi:10.1186/s12863-014-0106-7

Cite this article as: Strillacci et al.: Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA. *BMC Genetics* 2014 **15**:106.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

