Acta agriculturae Slovenica, suplement 2 (september 2008), 85–91.

http://aas.bf.uni-lj.si

Agris category codes: L73 COBISS Code 1.08

AN INTEGRATED MAP OF CATTLE CANDIDATE GENES FOR MASTITIS: A STEP FORWARD TO NEW GENETIC MARKERS

Jernej OGOREVC, Tanja KUNEJ and Peter DOVČ

Univ. of Ljubljana, Biotechnical Fac., Dept. of Animal Science, Groblje 3, SI-1230 Domžale, Slovenia, email: jernej.ogorevc@bfro.uni-lj.si

ABSTRACT

To facilitate the development of new genetic markers for mastitis resistance or susceptibility we used genome-wide comparative approach to review all known mastitis-associated loci. We assembled into a map 233 loci that were identified by six different study approaches (QTLs, association studies, expression experiments, AFLP-, miRNA- and epigenetic- studies). To integrate data from different sources and to identify overlapping regions we presented the results in the form of genetic map. The collected data represent genetic background for mastitis-related traits in cattle. Thirty most promising candidate genes (associated with mastitis in different study approaches or at least in two independent studies, or/and overlapping with QTL regions) were selected from database and *in silico* searched for genetic variability and putative miRNA target sites in 3'UTR. Thirty-one SNPs in the putative promoter/5' UTR (up to 2 kb upstream) of eight candidate genes were found. Bioinformatic analysis revealed that some promoter SNPs might potentially cause gain/loss of the putative transcription factors. Promoter SNPs were also present in CG dinucleotides and therefore possibly involved in gain/loss of CpG sites. In ten mastitis candidate genes we found 56 SNPs in exons of which 21 were non-synonymous substitutions. Additionally, 23 SNPs in intronic regions and 21 SNPs in 3'UTR were found. MiRNA target analysis revealed 89 putative target sites in 18 candidate genes; however, current SNPs were not identified in the miRNA target binding sites or miRNAs expressed in mammary gland. For SNPs with a putative regulatory role found in candidate genes, functional analyses and association studies are needed to facilitate identification of mastitis resistance or susceptibility alleles possibly involved in mastitis regulatory pathways.

Key words: cattle / mastitis / molecular genetics / candidate genes / quantitative trait loci / QTL

INTEGRIRANA KARTA KANDIDATNIH GENOV ZA MASTITIS PRI GOVEDU: KORAK NAPREJ PRI RAZVOJU NOVIH GENETSKIH OZNAČEVALCEV

IZVLEČEK

Pri iskanju kandidatnih genov za mastitis smo s pomočjo petih različnih študijskih pristopov (QTL-i, asociacijske študije, ekspresijski eksperimenti, miRNA, AFLP študije in epigenetski faktorji) pregledali 233 lokusov. Da bi sestavili podatke iz različnih virov v smiselno celoto in razkrili prekrivajoče se regije, smo rezultate predstavili v obliki genske karte. Zbrani podatki predstavljajo genetsko ozadje za lastnosti povezane z mastitisom pri govedu. Trideset najbolj obetavnih kandidatov (povezanih z mastitisom v različnih študijskih pristopih ali z vsaj dvema neodvisnima študijama ali/in ležeče v regijah prekrivajočih se s QTL-i) smo izbrali iz podatkovne zbirke in uporabili za *in silico* iskanje potencialnih, z mastitisom povezanih, molekularnih markerjev in domnevnih tarčnih mest za miRNA v 3' neprevedenih regijah. (3'UTR). V domnevni promotorski/5'UTR (2kb navzgor od začetka prevajanja) smo našli 31 SNP-jev v osmih kandidatnih genih. Analiza z Matinspectorjem je pokazala, da bi nekateri SNP-ji v promotorskih regijah lahko povzročili izgubo/pridobitev vezavnih mest za transkripcijske faktorje. Nekateri SNP-ji v domnevni promotorski regiji povzročajo tudi izgubo/pridobitev CpG mest. V desetih kandidatnih genih smo našli 56 SNP-jev v eksonih, med katerimi je bilo 21

zamenjav ne-sinonimnih. Poleg tega smo našli 23 SNP-jev v intronskih regijah in 21 SNP-jev v 3'UTR. Analiza tarč za miRNA je razkrila 89 domnevnih vezavnih mest za miRNA v 18 kandidatnih genih. V tarčnih zaporedjih za miRNA nismo našli SNP-jev. Za SNP-je z domnevno regulatorno vlogo v kandidatnih genih predlagamo funkcionalne analize in asociacijske študije, ki bi omogočile odkrivanje alel za odpornost oz. dovzetnost za mastitis in prispevale k razumevanju regulatornih poti pri mastitisu.

Ključne beside: govedo / mastitis / molekularna genetika / kandidatni geni / kvantitativni lokusi

INTRODUCTION

Mastitis is the most common and most costly disease in dairy cattle (Shook, 2006). As reported by Schutz (1994), estimated economic loses caused by mastitis range from $100 to $200 per cow per lactation. Milk production and manufacturing significantly supported genetic research related to milk production and udder health in the past.

During the last years many experiments have identified different QTL regions in cattle affecting functional traits such as mastitis (Schwerin *et al*., 2003). QTLs cover large chromosomal regions on average spreading from 10 to 40 cM (Stella and Boettcher, 2004), involving hundreds or thousands of genes. The ultimate goal of the QTL analysis is identification of causal gene itself, therefore fine mapping of the mastitis associated QTLs could make marker assisted selection (MAS) possible and eventually facilitate identification of resistance genes and alleles (Reinard and Riollet, 2005). Beside QTLs, a large number of genetic polymorphisms within the causal gene regions or genetic markers associated with mastitis traits have been identified in cattle. The high throughput technologies such as microarray analysis offer the possibility to study changes in expression profiles of thousands of genes simultaneously as a response to infection with the pathogen. The release of the cattle genome sequence enabled discovery of new markers and creation of synteny maps including data from other species. In addition, the newly discovered miRNA and epigenetic mechanisms have been associated with mastitis resistance or susceptibility (Silveri *et al.,* 2006; Vanselow *et al.,* 2006).

To facilitate development of new genetic markers for mastitis resistance or susceptibility we analyzed mastitis-associated genes identified by various approaches and integrated them into a single map. Gene map approach reveals positional overlaps of loci found with different approaches and exposes regions with high density of candidate loci. Best mastitis resistance or susceptibility candidates were selected and *in silico* searched for genetic variability and miRNA target sites in their 3' UTR. The aim of our work was to identify candidate regions for further functional studies for mastitis resistance or susceptibility.

MATERIALS AND METHODS

Database searches for candidate loci

Literature published up to December 2007 was reviewed by searching for the relevant publications through PubMed (http://www.ncbi.nlm.nih.gov) and Web of Science (http://isiknowledge.com) using key phrases: association, gene candidates, epigenetics, genetics, mammary gland, mastitis, methylation, milk, miRNA, QTL, SNP. QTLs were extracted from: Cattle QTL Database Release 5 (12/2007): http://www.animalgenome.org using ontology term "mastitis" (somatic cell score (SCS), clinical mastitis (CM)). Expression patterns associated with mastitis were collected from studies performed on cattle and mouse (Ogorevc *et al.,* unpublished).

Defining the map locations of the loci

The map location was retrieved from the NCBI database Bos taurus build (3.1). If the map location was not available, we identified the location of the locus using the bovine-human synteny map. The bovine – human synteny map was constructed through BLASTing 8 294 markers from MARC and RH maps (Everts-van der Wind *et al.,* 2004; Itoh *et al.,* 2005) with bovine contigs (Build 35.0) to obtain hits (defined as E < 10exp–19) with longer sequences. Hits were further BLASTed against the human genome, 6 231 putative human bovine orthologs were found. Positions on the human physical map were obtained using Map Viewer on NCBI. The synteny map was constructed using 6 023 orthologs sorted in 213 blocks of synteny. Each synteny block with at least 2 markers (singletons were excluded) is described by its position on the physical human map and on the bovine cytogenetic map.

In silico analysis of selected candidate genes for SNPs and miRNA target sites

Candidate genes associated with mastitis in at least two different approaches or reported by at least two independent studies and/or located in regions overlapping with QTLs were selected for further analysis. Selected candidate genes were *in silico* searched for SNPs in the putative promoter/5' UTR region (2 kb upstream), exon regions, intron regions, and 3' UTR. SNPs were retrieved from the Ensembl database (http://www.ensembl.org/). A search for transcriptional regulatory elements and SNPs involved in gain/loss of transcription factors binding sites was performed using MatInspector program (http://www.gsf.de/). Identification of miRNA target sites in 3' UTRs was performed using Sanger's miRBase (http://microrna.sanger.ac.uk/) and Patrocles database (http://www.patrocles.org/).

RESULTS

A mastitis resistance or susceptibility candidate gene map includes 233 loci (Table 1). To date there are 60 cattle QTLs associated with mastitis traits (clinical mastitis and somatic cell score). Six genes showed association between sequence variation and mastitis resistance or susceptibility. 107 genes with expression patterns associated with mastitis resistance or susceptibility were reported in 12 publications in cattle and mouse. Additionally, 27 AFLP markers associated with mastitis were found and the most promising marker, *CGIL4* was further characterized and mapped to BTA22 q24 (Sharma *et al.*, 2006). To date 32 miRNA genes were reported to be expressed in the bovine mammary gland (Gu *et al.,* 2007), but their involvement in mastitis is not known yet. Epigenetic factors, such as DNA-remethylation around the *STAT5*-binding enhancer in the *CSN1S1* promoter was shown to be associated with shutdown of α_{S1}-casein synthesis during acute mastitis (Vanselow *et al.,* 2006).

Table 1. Summary of the study approaches used for assembling the gene map of cattle gene candidates for resistance or susceptibility to mastitis

Study approach	Number of loci
QTLs	60
Association studies	6
Expression studies	107
AFLP markers	27
miRNAs expressed in mammary gland*	32
Epigenetic factors	1
Total	233

*not proved to be associated with mastitis

The results are presented in a form of genetic map (Fig. 1) with the highest number of candidate loci on chromosomes 7, 23, 19, 5 and 9 and the lowest on chromosome 28. The gene map shows mastitis candidate loci on all chromosomes except chromosome Y.

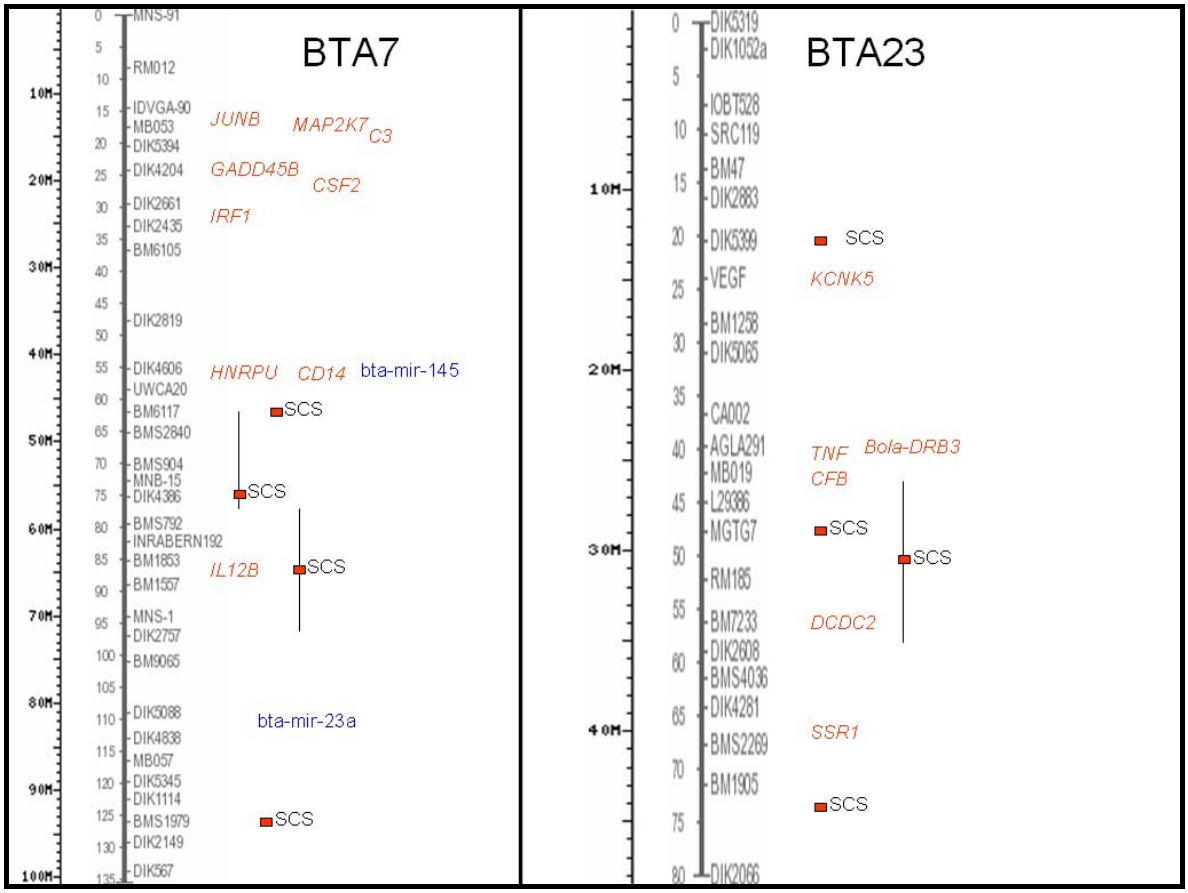

Figure 1. Gene map of cattle candidate genes for mastitis: examples are shown for BTA7 and BTA23.

Thirty most promising candidate genes (associated with mastitis in different study approaches or with at least two independent studies, or/and overlapping with QTL regions) were selected from database (Table 2) and *in silico* searched for genetic variability (in promoter/5' UTR (2 kb upstream), exon, intron, and 3' UTR) and putative miRNA target sites within the 3'UTR.

Ensembl genomic sequence and variation data was available for 22 of 30 selected candidate genes, but eight candidate genes (*CEBPB, C5AR1, FEZF2, IL8RA, KCNK1, PLCE1, PRKDC,* and *RELA*) have not been annotated yet. Thirty-one SNPs in the putative promoter/5' UTR region (2 kb upstream) of eight candidate genes were found (Table 3). Matinspector analysis revealed that gene promoter SNPs cause gain/loss of the putative transcription factor binding sites. Promoter SNPs also showed gain/loss of potential CpG sites. In ten candidate genes we found 56 SNPs in exons of which 21 were non-synonymous substitutions. Additionally, 23 SNPs in intronic regions and 21 SNPs in 3'UTR were found. Bioinformatics analysis revealed 89 putative miRNA target sites in 18 mastitis candidate genes. Bta-mir-142* with a putative target site in mastitis gene candidate *SAA3* was already experimentally confirmed to be expressed in bovine mammary gland (Gu *et al.*, 2007). To date, no SNPs were identified in the miRNA target binding sites or miRNA genes expressed in mammary gland.

Table 2. Most promising candidate genes associated with mastitis phenotype found in independent studies using the same or different study approaches

Gene	Gene name	Association studies	Expression studies	Inside QTL region
ACTB	actin, beta, cytoplasmic		+	+
BoLA-DRB3	major histocompatibility complex, class II, DRB3	+ + +		
C5AR1	complement component 5a receptor 1		+ +	
CD14	CD14 antigen		+ + + +	
CEBPB	CCAAT/enhancer binding protein (C/EBP), beta		+	+
DCDC2	doublecortin domain containing 2		+	+
FEZF2	fez family zinc finger 2	+	+	
HGF	hepatocyte growth factor		+	+
HP	haptoglobin		+	+
IFNG	interferon gamma		+ + +	
IL1B	interleukin 1 beta		+ + + +	
IL6	interleukin 6		+ +	+
IL8	interleukin 8		+ + +	
IL12B	interleukin 12b		+	+
JUN	Jun oncogene		+	+
KCNK1	potassium channel, subfamily K, member 1		+	+
LBP	lipopolysaccharide binding protein		+ +	+
LTF	lactoferrin	+	+	
MAIL	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta		+	+
MCL1	myeloid cell leukemia sequence 1		+	+
OSTF1	osteoclast stimulating factor 1		+	+
PLCE1	phospholipase C, epsilon 1		+	+
PRKDC	protein kinase, DNA activated, catalytic polypeptide		+	+
RELA	v-rel reticuloendotheliosis viral oncogene homolog A (avian)		+	+
S100A8	S100 calcium binding protein A8 (calgranulin A)		+	+
SAA3	serum amyloid A3		+ +	
TLR-2	toll-like receptor 2		+ +	
TLR-4	toll-like receptor 4	+ +	+ +	
TNFAIP3	tumor necrosis factor, alpha-induced protein 3		+	+
TNF	tumor necrosis factor		+ + + +	

+ = independent study.

Table 3. Number of SNPs in different regions of annotated candidate genes and miRNA target sites in 3'UTR

Mastitis candidate	SNPs					miRNAs with putative target site in mastitis gene candidates (bta-)
	Promoter (2 kb upstream)	5'UTR	Exon [1]	Introns	3'UTR	
ACTB	0	0	0	0	1	/
BoLA-DRB3	0	0	5 (5)	1	0	/
CD14	0	1	1 (0)	0	0	miR-193a
DCDC2	0	0	0	0	0	miR: 200a, 369-3p, 151, 200c, 200b, 27a, 21
HGF	0	0	0	0	0	miR: 26b, 200a, 26a, 199a-5p, 200c, 26b, 365, 200b, 455, 25, 92
HP	1	0	5 (1)	1	0	miR: 25, 204, 92,
IL8RB	7	0	0	5	0	/
IFNG	0	0	0	1	2	miR: 181a, 181b, 125b, 369-3p, 26b, 26a, 181c, 181b, 125a, 99a, 99b, 425-5p
IL1B	0	0	2 (0)	0	8	miR: 221, 320, 331, 31
IL6	1	0	1 (0)	1	1	miR: 181c, 22-5p, 455, 532, 23b, 23a, 221, 132 let: 7g, 7a, 7f
IL8	3	0	0	6	7	/
IL12B	1	0	3 (1)	2	0	miR: 369-3p, 380-5p, 425-5p
JUN	0	0	0	0	0	miR: 200b, 200c,
LBP	11	0	5 (2)	0	0	miR: 545*, 93, 20b, 20a, 545, 361, 142, 124a, 101, 145
LTF	0	2	1 (0)	1	0	/
OSTF1	0	0	0	1	0	miR: 200c, 200b, 148a, 128a, 126*
S100A8	2	0	0	0	0	miR: 126*, 98
SAA3	0	0	0	0	0	miR-142*
TLR-2	0	0	5 (4)	0	0	miR: 26b, 186
TLR-4	1	1	28 (8)	4	2	miR-151
TNFAIP3	0	0	0	0	0	miR: 16, 195, 205, 15a, 124a, 455
TNF	0	0	0	0	0	miR: 125b, 19a, 23a, 18b, 18a, 450, 19b, 125a

[1] Number in brackets represents number of non-synonymous coding SNPs.

DISCUSSION

Among the candidate genes identified in this study four genes (*BoLA-DRB3, FEZF2, LTF, TLR-4*) show association between the sequence variation and mastitis resistance or susceptibility. Eleven genes (*IL6, IL8, CD14, TLR-4, IL1B, LBP, TLR-2, C5AR1, TNF*, *IFNG, SAA*) were differentially expressed during mastitis in more than one (two to four) expression experiments. Moreover, six genes (*IL6, CD14, TLR-4, IL1B, TLR-2, SAA3*) were found to be differentially expressed in two species (cattle and mouse). Eighteen genes (*ACTB*, *CEBPB*, *DCDC2*, *HGF*, *HP*, *IL6*, *IL12B*, *JUN*, *KCNK1*, *LBP*, *MAIL*, *MCL1*, *OSTF1*, *PLCE1*, *PRKDC*, *RELA*, *S100A8*, and *TNFAIP3*) reported by association studies or expression experiments are located in regions overlapping with QTLs.

SNPs and putative miRNA target sites (in 3'UTR) were extracted from databases. SNPs in promoter regions were *in silico* analyzed for gain/loss of transcription factor binding sites. Gain/loss of transcription binding sites predicted *in silico* should be further confirmed with experimental methods (e.g. EMSA – electrophoretic mobility shift assay).

Putative miRNA target sites were found in 3'UTRs of candidate genes but to date no SNPs were identified inside target sites. The cross-section between the mammary-gland expressed miRNAs and identified putative miRNA targets revealed one miRNA (bta-miR-142*) having a target site within a mastitis gene candidate *SAA3* and therefore could potentially be involved in its regulation. To date, no genetic variability of the miRNA genes, their targets and silencing machinery in cattle is available in the Patrocles database. A fair amount of information is available for human miRNAs and it is expected that cattle data will be added soon enabling searches for polymorphisms in putative miRNA target sites.

Epigenetic factors were also proved to be involved in clinical mastitis, namely CpG remethylation around *STAT5*-binding enhancer in *CSN1S1*gene was involved in infection induced shutdown of αS1-casein synthesis. Therefore, promoter regions of candidate genes should be searched for STAT5 binding sites in the future. Promoter/5'UTR (2 kb upstream) SNPs in selected candidate genes also showed gain/loss of CpG sites. However, for SNPs with a putative regulatory role found in mastitis candidate genes, functional analyses and association studies are needed to facilitate identification of mastitis resistance or susceptibility alleles and understanding of mastitis regulatory pathways.

REFERENCES

Everts-van der Wind, A./ Kata, S.R./ Band, M.R./ Rebeiz, M./ Larkin, D.M./ Everts, R.E./ Green, C.A./ Liu, L./ Natarajan, S./ Goldammer, T./ Lee, J.H./ McKay, S./ Womack, J.E./ Lewin, H.A. A 1463 gene cattle-human comparative map with anchor points defined by human genome sequence coordinates. Genome Research, 14(2004), 1424–37.

Gu, Z./ Eleswarapu, S./ Jiang, H. Identification and characterization of microRNAs from the bovine adipose tissue and mammary gland. FEBS Letters, 581(2007), 981–8.

Itoh, T./ Watanabe, T./ Ihara, N./ Mariani, P./ Beattie, C.W./ Sugimoto, Y./ Takasuga, A. A comprehensive radiation hybrid map of the bovine genome comprising 5593 loci. Genomics, 85(2005), 413–24.

Ogorevc, J./ Kunej, T./ Razpet, A./ Dovc, P. Gene map of cattle candidate genes for milk traits. Domžale, Biotechnical faculty, Department of Animal Science (unpublished).

Rainard, P./ Riollet, C. Innate immunity of the bovine mammary gland. Veterinary Research, 37(2005), 369–400.

Schutz, M.M. Genetic evaluation of soamtic-cell scores for United-states dairy cattle. Journal of Dairy Science, 77(1994), 2113–29.

Schwerin, M./ Czernek-Schafer, D./ Goldammer, T./ Kata, S.R./ Womack, J.E./ Pareek, R./ Pareek, C./ Walawski, K./ Brunner, R.M. Application of disease-associated differentially expressed genes – Mining for functional candidate genes for mastitis resistance in cattle. Genetics Selection Evolution, 35(2003), S19–S34.

Sharma, B.S./ Jansen, G.B./ Karrow, N.A./ Kelton, D./ Jiang, Z. Detection and characterization of amplified fragment length polymorphism markers for clinical mastitis in Canadian Holsteins. Journal of Dairy Science, 89(2006), 3653–63.

Shook, G.E. Major advances in determining appropriate selection goals. Journal of Dairy Science, 89(2006), 1349–61.

Silveri, L./ Tilly, G./ Vilotte, J.L./ Le Provost, F. MicroRNA involvement in mammary gland development and breast cancer. Reproduction Nutrition Development, 46 (2006), 549–56.

Stella, A./ Boettcher, P.J. Optimal designs for linkage disequilibrium mapping and candidate gene association tests in livestock populations. Genetics, 166(2006), 341–50.

Vanselow, J./ Yang, W./ Herrmann, J./ Zerbe, H./ Schuberth, H.J./ Petzl, W./ Tomek, W./ Seyfert, H.M. DNA-remethylation around a STAT5-binding enhancer in the alphaS1-casein promoter is associated with abrupt shutdown of alphaS1-casein synthesis during acute mastitis. Journal of Molecular Endocrinology, 37(2006), 463–77.