University of Ljubljana,
Biotechnical Faculty

Database of cattle candidate loci for milk production and mastitis

Jernej OGOREVC (jernej.ogorevc@bfro.uni-lj.si), Tanja KUNEJ, and Peter DOVC
University of Ljubljana, Biotechnical Faculty, Chair of Genetics, Animal Biotechnology, and Immunology
SI-1230 Domzale, Slovenia

The ultimate goal of QTL analysis is identification of genes underlying complex traits, which is a very difficult task in large farm animals. In addition, the obtained results are less reliable due to the variation in genetic background and population specific interactions between loci. The recent developments in molecular biology opened the possibility to exploit heterologous animal models and allowed integration of information from different sources and complementation of different pieces of evidence to support involvement of candidate loci in investigated traits.

Integrative genomics approach was used to:
1. **collect** candidate loci involved in milk production and mastitis traits,
2. present the database in a form of **genetic marker map** (Figure 1),
3. ***in silico* analyze** the most promising candidate genes.

Materials and methods

Data mining:

- Published articles,
- Mouse Genome Informatics (MGI) (http://www.informatics.jax.org),
- Cattle QTL Database (http://www.animalgenome.org),
- cgQTL database (http://cowry.agri.huji.ac.il/).

***In silico* analysis:**

- The map locations – bovine human synteny map or NCBI's *Bos taurus* Build (4.0) (http://www.ncbi.nlm.nih.gov/),
- Expression level in lactating mammary gland - GNF BioGPS (http://biogps.gnf.org),
- Genetic variation data - Ensembl database (http://www.ensembl.org/),
- Pathway analysis - The Ingenuity Pathway Analysis programme (http://www.ingenuity.com),
- Putative target sites for mammary gland expressed miRNAs in candidate genes - Sanger's mir-Base Targets (http://microrna.sanger.ac.uk/), and polymorphic targets - Patrocles database (http://www.patrocles.org/).

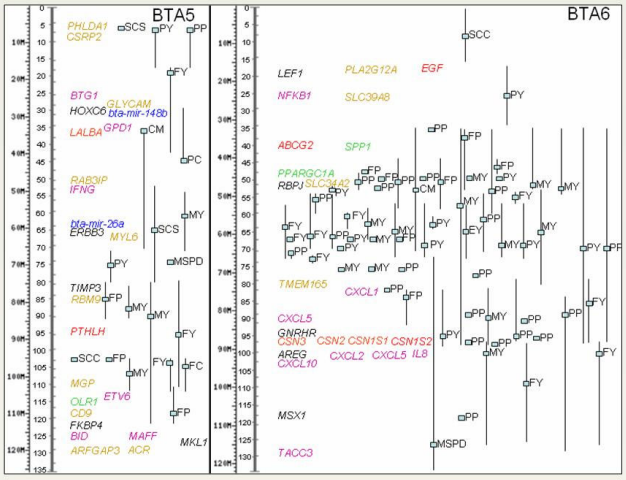

Figure 1: Example of BTA5 and BTA6 revealing several overlapping genomic regions.

Results and discussion

- Database consisting of 934 candidate loci available online: http://www.bfro.uni-lj.si/ Kat_genet/genetika/mammary_gland.xls
- Genetic marker map revealing positional candidates and regions with high density of loci (Figure 1),
- Pathway analysis of candidate genes revealing possible novel candidate genes (Figure 2),
- Identification and *in silico* analysis of 44 most promising candidate genes (Figure 3),
- Target search for cattle mammary gland expressed miRNAs identified 359 putative binding sites

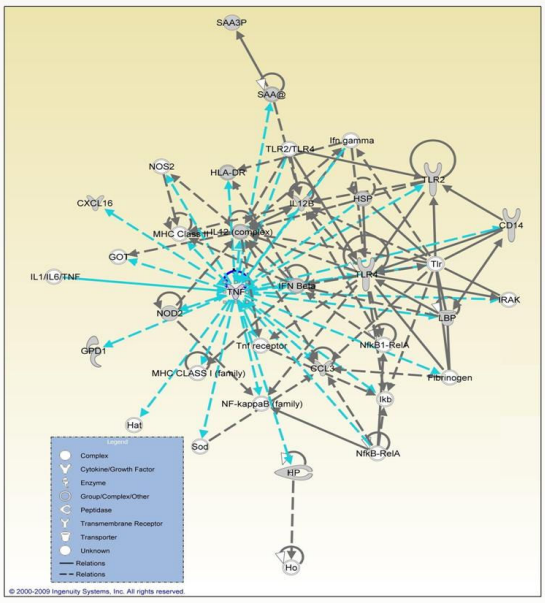

Figure 2: IPA proposed pathway for mastitis resistance, with *TNF* in central role, revealing potential novel infection response associated genes.

Cross-comparison of data obtained by different approaches was performed and the most promising candidates were *in silico* analyzed for expression levels in lactating mammary gland, genetic variability, and top biological functions in functional networks. miRNA target search for mammary gland expressed miRNAs revealed polymorphic target sites in *IL1B* and *CYP11B1*. Gene ontology can now be employed to find novel candidate genes in the regions underlying multiple QTL.

Gene	Gene name	Mouse KOs and transgenic experiments	Association studies		Expression studies		Milk protein genetic variants	Epigenetic studies	Expression in lactating mammary gland[2] (mouse)	Number of SNPs					Top functions[4]
			milk traits	mastitis traits	milk traits[3]	mastitis traits				Promoter (5 kb upstream)	5' UTR	Exons[3]	Introns (context 100 bp)	3' UTR	
Associated with mastitis resistance															
CD14	CD14 antigen					++++			***	0	1	1	0	0	A
ETS2	E26 avian leukemia oncogene 2, 3' domain	+				+			*	0	0	0	1	1	C
FEZF2	fez family zinc finger 2			+		+			**	0	0	1 (1)	1	0	E
IFNG	interferone gamma					+++			**	0	0	0	1	2	D
IL1B	interleukin 1 beta					++++			**	0	0	2	0	7	D
IL6	interleukin 6					++			**	0	0	1	1	1	B
IL8	interleukin 8					+++			NA	0	0	0	5	7	C
IL8RA	interleukin 8 receptor, alpha			++					**	0	0	0	0	0	C
LBP	lipopolysaccharide binding protein					++			***	2	0	5 (2)	0	0	A
SAA3	serum amyloid A3					++			**	0	0	0	0	0	A
TLR-2	toll-like receptor 2					++			**	0	0	3 (2)	0	0	A
TLR-4	toll-like receptor 4			++		++			**	0	1	28 (8)	4	2	A
TNF	tumor necrosis factor					++++			**	0	0	0	0	0	A

Figure 3: Detail from the table of the most promising candidates.

Conclusions: The created database represents a new research tool integrating different types of data offering genetic background for subsequent functional studies and opens as well the possibility to introduce complex decision making strategies which integrate multiple pieces of evidence supporting the candidate status of the selected region.