Microarray analysis applied to the study of milk protein *loci* in cattle

S. Chessa[1], B. Castiglioni[2], G. Ceriotti[1], G. De Bellis[3], G. Pagnacco[1]

[1] Dipartimento Scienze e Tecnologie Veterinarie per la Sicurezza Alimentare, Università di Milano, Italy

[2] Istituto Biologia e Biotecnologia Agraria, C.N.R., Milano, Italy

[3] Istituto Tecnologie Biomediche, C.N.R., Segrate, Italy

Corresponding author: Stefania Chessa. Dipartimento Scienze e Tecnologie Veterinarie per la Sicurezza Alimentare. Via Trentacoste 2, 20134 Milano, Italy – Tel: +39 02 50315757 – Fax: +39 02 50315746 – Email: stefania.chessa@unimi.it

RIASSUNTO – Applicazione della tecnologia microarray allo studio dei loci lattoproteici bovini. *Lo studio dei polimorfismi genetici lattoproteici riveste molta importanza a causa dei loro effetti sulle caratteristiche quali-quantitative del latte così come sugli aspetti tecnologici. I progressi della tecnica in campo della biologia molecolare hanno permesso negli ultimi anni di mettere in luce nuovi polimorfismi all'interno di regioni codificanti e non-codificanti dei geni lattoproteici. La tipizzazione simultanea di tutti i polimorfismi genomici identificati a livello dei loci lattoproteici, permetterebbe di approfondire le conoscenze sulle variazioni genetiche più importanti dal punto di vista delle implicazioni selettive. Con le metodiche attualmente disponibili questo approccio richiederebbe un grosso investimento in termini sia di tempo, sia di denaro. In questo studio si descrive la messa a punto di una metodica che prevede l'associazione della reazione di ligazione con un array universale per identificare i polimorfismi lattoproteici bovini, sfruttando la possibilità che la tecnologia microarray offre di tipizzare contemporaneamente numerosi polimorfismi.*

KEY WORDS: cattle, milk protein, polymorphisms, microarray.

INTRODUCTION – Milk proteins still arise much interest because of the proved effects of the most common milk genetic polymorphisms on quantitative and qualitative milk production, as well as on milk technological properties. The role of α_{s1}-casein (*CSN1S1*), β-casein (*CSN2*), k-casein (*CSN3*) and β-lactoglobulin (*LGB*) polymorphisms in the genetic improvement of milk production was already demonstrated in cattle, as reviewed by Di Stasio and Mariani (2000) and Martin *et al.* (2002). At least 8, 14, 12, and 12 alleles were identified respectively at *CSN1S1, CSN2, CSN3,* and *LGB loci,* and different PCR-based protocols were developed in order to identify the genetic variability of each *locus*. Moreover, polymorphisms were found in the non-coding regions of the milk protein genes (reviewed by Martin *et al.*, 2002), which can also affect milk quality and technological properties. Polymorphisms both at the lactoprotein coding and non-coding regions have to be considered in order to assess which ones are responsible for the variations in milk traits, and could be usefully included in breeding programs for dairy cattle genetic improvement. Nevertheless, this methodological approach would require much time and money investment for genotyping.

The application of the microarray technology to the identification of genetic polymorphisms is gaining wide attention thanks to the extreme degree of parallelisation (Ramsay, 1988). Here we describe an application of this technology which combines enzymatic processing and hybridisation in order to detect SNP (Single Nucleotide Polymorphism) at the bovine milk protein *loci*.

MATERIAL AND METHODS – The used procedure associated the discriminative properties of DNA Ligation Detection Reaction (LDR) with a universal array. The universal array was prepared using the CodeLink Activated slides (Amersham Pharmacia Biotech Inc, Uppsala, Sweden), and the MicroGrid II Compact arrayer (BioRobotics Ltd, Cambridge, UK), as described by Consolandi et al. (2003). A slide included eight identical subarray each composed by 208 spot organised in 16 columns and 13 rows. In this way it was possible to test simultaneously eight samples for each experiment. A subarray contained 47 Zip Code repeated four times, one repeated eight times along the outline of the subarray, one repeated six time and six blank spot. For each SNP three probes were used as described by Gerry et al. (1999). Two discriminating probes were designed on the sequence preceding and including the mutation. Being the mutation the last nucleotide of each probe and differing depending on the allele, each probe discriminated for one of the two alleles determined by the considered mutation and was labelled respectively with Cy3 and Cy5. The third probe was designed starting from the nucleotide following the mutation and was therefore the same for the two alleles. A complementary Zip Code (cZip Code) was coupled to this probe. Each probe was designed in order to have a melting temperature ranging from 67 to 69°C and was produced and purified by Thermo Hybaid GmbH (Ulm, Germany). In the LDR the discriminating probe and the common probe hybridised to the template DNA and a DNA ligase ligated the two probes only in case of perfect match. During the hybridisation phase on the universal array the cZip Code drove the LDR solution to the correspondent Zip Code on the array. Then the fluorescent signal produced by the labelled ligated probe was detected at the resolution of 5 mm using the ScanArray, Lite laser scanning system and the ScanArray 3.1 software (Perkin Elmer, Boston, MA). Target sequences included the most common alleles of milk protein genes in the Italian taurine breeds. SNP determining differences in the aminoacid sequence were chosen among the various polymorphisms discriminating each allele. A multiplex PCR was developed to amplify simultaneously a 462 bp fragment of the fourth exon of CSN3, a 333 bp fragment including part of the seventeenth exon of CSN1S1, a 545 bp fragment of the seventh exon of CSN2, a 275 bp fragment of the sixth exon of CSN2, a 404 bp fragment of the LGB third exon and a 494 bp fragment of the LGB forth exon. Primers were designed using Primer3 software, available on-line at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi. Referring sequences were GenBank Accession No. X59856 for CSN1S1, X14711 for CSN2, AY380228 for CSN3 and X14710 for LGB.

To check the deposition quality of the Zip Codes on the subarray, 1 pmol/mL of oligonucleotides poly(dT)10 Cy5 labelled in a solution 5X of SSC containing 0.1 mg/mL of salmon sperm DNA (Sigma-Aldrich) was hybridised on one of the slides produced in the same series.

To test the discriminating power of the probes and the validity of the applied method known homozygous and heterozygous individuals for each SNP were used. First each polymorphism was tested separately in a subarray using only the three specific probes in the LDR. Then all the polymorphisms of each locus were tested simultaneously in the same LDR.

RESULTS AND CONCLUSIONS – The following variants were chosen for the application of the proposed method: CSN1S1*B, CSN1S1*C, CSN2*A1, CSN2*A2, CSN2*A3, CSN2*B, CSN2*C, CSN2*I, CSN3*A, CSN3*B, CSN3*C, CSN3*E, CSN3*H, LGB*A, LGB*B. For each reference individual, the corrected fluorescent signal was detected in all the four repetitions of the Zip Codes associated to the SNPs considered to differentiate the alleles in the simultaneous genotyping. The SNPs analysed are shown in Table 1. As deducible from this table, choosing the SNPs was quite difficult when the variants in a locus were more than two, because some alleles shared the same nucleotide in different positions. To reconstruct the correct genotype the combination of the results at different SNP positions must be carefully analysed. Anyway, preliminary results show the great potential of the Ligase Detection Reaction + Universal Array method for a rapid and reliable identification of milk protein polymorphisms in cattle.

Other polymorphisms, possibly influencing quantitative and qualitative milk production, as well as milk technological properties will be analysed within the same slide. Fragments containing part of the promoter region of the CSN1S1 and of the LGB, and part of the CSN3 second intron will be included. The last one contains a

polymorphism of the Short Interspersed Nucleotide Element Bov-A2 (Damani et al., 2000) showing statistically significant effects on milk production traits (Damiani et al., 2001). Using GenBank Accession No. X14710 for LGB, AF549499-AF549500-AF549501-AF549502 for the four known variants of the CSN1S1 promoter, and X14908 for the Short Interspersed Nucleotide Element Bov-A2 as referring sequences, primers to amplify also these regions have already been designed and the same multiplex PCR cycle has been successfully used to amplify the fragments.

Table 1. SNP considered in each *locus* to distinguish the different variants by the proposed method.

Locus	SNP position	Nucleotide involved	*Alleles*	Reference sequence (GenBank accession number)
CSN1S1	17807	A	*B*	X59856
		G	*C*	
CSN2	6690	G	A^1-A^2-A^3-*B*-*I*	X14711
		A	*C*	
	8108	C	A^2-A^3-*I*	
		A	A^1-*B*-*C*	
	8178	A	A^1-A^2-A^3-*B*-*C*	
		C	*I*	
	8219	C	A^1-A^2-*B*-*C*-*I*	
		A	A^3	
	8266	C	A^1-A^2-A^3-*C*-*I*	
		G	*B*	
CSN3	12951	G	*A*-*B*-*E*-*H*	AY380228
		A	*C*	
	13065	C	*A*-*B*-*C*-*E*	
		T	*H*	
	13104	A	*A*-*C*-*E*-*H*	
		C	*B*	
	13124	A	*A*-*B*-*C*-*H*	
		G	*E*	
LGB	3984	A	*A*	X14710
		G	*B*	
	5263	T	*A*	
		C	*B*	

ACKNOWLEDGMENTS – We are grateful to MURST contract 2003078003_004 for financial support.

REFERENCES – **Consolandi**, C., Busti, E., Pera, C., Delfino, L., Ferrara, G.B., Bordoni, R., Castiglioni, B., Rossi Bernardi, L., Battaglia, C., De Bellis, G., 2003. Detection of HLA polymorphisms by ligase detection reaction and a universal array format: a pilot study for low resolution genotyping. Human Immunology. 64: 168-178. **Damiani**, G., Florio, S., Budelli, E., Bolla, P., Caroli, A., 2000. Single nucleotide polymorphisms (SNPs) within Bov-A2 SINE in the second intron of bovine and buffalo k-casein (CSN3) gene. Anim. Genet. 31:277-279. **Damiani**, G., Caroli, A., Leone, P., Budelli, E., Florio, S., 2001. Effect of Bov-A2 Sine elements on quantitative traits. Atti. XIV Congr. Naz. ASPA. 2:49-51. **Di Stasio**, L., Mariani, P., 2000. The role of protein polymorphism in the genetic improvement of milk production. Zoot. Nutr. Anim. 26: 69-90. **Gerry**, N.P., Witowski, N.E., Day, J., Hammer, R.P., Barany, G., Barany, F., 1999. Universal DNA Microarray method for multiplex detection of low abundance point mutations. J. Mol. Biol. 292: 251-262. **Martin**, P., Szymanowska, M., Zwierzchowski, L., Leroux, C., 2002. The impact of genetic polymorphisms on the protein composition of ruminants milks. Reprod. Nutr. Dev. 42:433-459. **Ramsay**, G., 1988. DNA chips: state of the art. Nat. Biotechnol. 16: 40.