Understanding milk protein complexity to produce accurate phenotypes

G Miranda, Christelle Cebo, N Boumahrou, L Bianchi, Claudia Bevilacqua

Understanding milk protein complexity to produce accurate phenotypes

Miranda, G., Cebo, C., Boumahrou, N., Bianchi, L., Pinard, A., [1]Guillot, A., [1]Henry, C., Bevilacqua, C., Beauvallet, C., Bellier, S. and Martin, P.

Institut National de la Recherche Agronomique (INRA), UMR1313 Génétique animale et Biologie intégrative (GABI), Equipe «Lait, Génome & Santé » (LGS), 78 350 Jouy-en-Josas, France.

[1]Institut National de la Recherche Agronomique (INRA), Plateforme d'Analyse Protéomique de Paris Sud-Ouest (PAPPSO), 78 350 Jouy-en-Josas, France.

Email: patrice.martin@jouy.inra.fr

Milk is a complete and complex food containing a large number of biomolecules such as lipids, sugars and proteins. As far as proteins are concerned, they are found either at the colloidal state (micelles), soluble in the whey or associated with the fat globule membrane. For many years, milk has been considered as a raw material of which processing or "cracking" concentrated most of the added value. Therefore, the breeding target was over the past 30 years to produce larger amounts of milk with a high overall protein content, while controlling fat content, at the expense of quality and compromising health and fertility of high producing dairy cows. To reach such a goal, measurement and selection procedures have been developed and implemented based on milk yield, total fat and protein in milk. More recently, emphasis has been put on milk elementary components since many of them, in particular fatty acids and peptides, have putative or actual positive effects on human health. It is now of major interest to accurately measure elementary milk components, including proteins, to identify genes affecting fine milk protein composition. Genetic variants of a number of milk proteins have been shown to impact the protein composition in milk and explain, at least in part, the genetic variation in milk protein composition.

Progress made in the field of comparative and functional genomics, as well as in proteomics, has highlighted how such genetic polymorphisms are responsible for the extreme complexity and the large variability (qualitative and quantitative) of the milk protein fraction. At the quantitative level, general mechanisms controlling gene expression act both at the transcriptional and the post-transcriptional levels. Polymorphisms found in cis-regulatory elements, mainly within the 5'-flanking region of genes encoding several milk proteins (β-lactoglobulin, αs1- and αs2-caseins) have been shown to influence their transcription rate, in cattle. Polymorphisms found in the transcription unit, within intron as well as exon sequences, have been shown to be responsible for defects in processing of primary transcripts (exon skipping, usage of cryptic splice sites) impacting the amount and structure of messengers and consequently the primary structure of proteins. Such a situation, well-exemplified by the gene encoding αs1-casein in goats, may have dramatic biological consequences (protein and fat contents, casein micelle structure, secretion pathway, etc.).

Combining different proteomic approaches (liquid chromatography, electrophoresis and mass spectrometry) with exon-sequencing, we succeed in accurately characterizing, quantitatively and qualitatively, the protein fraction of milk from different cattle breeds, and identifying new genetic variants. Such a strategy was also effective to analyze milk from other species including mice for which a total of 34 SNP were identified in the coding and 3' untranslated regions of 3 milk protein genes, between mouse species.