C-135

A statistical learning approach to detect carriers of the HH1 haplotype in Italian Holstein Friesian cattle

Stefano Biffani[1], Stefania Chessa[1], Alessandra Stella[1]
Filippo Biscarini[2]

[1]*Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, Lodi, Italy*
[2]*Dipartimento di Bioinformatica, Parco Tecnologico Padano, Lodi, Italy*
Corresponding author: ste.bif68@gmail.com

The increasing availability of SNP (single nucleotide polymorphisms) genotype data in livestock is stimulating the development of new data analysis strategies, which can be applied in animal breeding. One possible application is the prediction of carriers of specific haplotypes, especially if they impact animal health. It is therefore convenient to have a practical and easy-to-implement statistical method for the accurate classification of individuals into carriers and non-carriers. In this paper, we present a procedure for the identification of carriers of the haplotype HH1 on BTA5 (Bos Taurus autosome 5), which is known to be associated with reduced cow fertility in Holstein-Friesian cattle. A population of 1104 Holstein bulls genotyped with the 54K SNP-chip was available for the analysis. There were 45 carriers (5.3%) and 1045 non-carriers (94.7%). Two complementary multivariate statistical techniques were used for the identification of haplotype carriers: Backward Stepwise Selection (BSS) to select the SNP that best fit the model, and Linear Discriminant Analysis (LDA) to classify observations, based on the selected SNP, into carriers and non-carriers. In order to explore the minimum-sized set of SNP that correctly identifies haplotype carriers, different proportions of SNP were tested: 2.5; 10; 15; 30; 50 and 100%. For each proportion of SNP, BSS and LDA were applied, and the classification error rate was estimated in a 10-fold cross-validation scheme. Data were split in 10 subsets. The first subset was treated as validation set, while the model was fit on the remaining nine subsets (the training set). The overall error rate for the prediction of haplotype carriers was on average very low (~1%) both in the training and in the validation datasets. The error rate was found to depend on the number of SNPs in the model and their density around the region of the haplotype on BTA5. The minimum set of SNPs to achieve accurate predictions was 8, with a total test error rate of 1.27. This work describes a procedure to accurately identify haplotype carriers from SNP genotypes in cattle populations. Very few misclassifications were observed, which indicates that this is a very reliable approach for potential applications in cattle breeding.

Acknowledgments
This research was financially supported by the national research project "GenHome" and the project "ProZoo", financed by "Regione Lombardia", "Fondazione Cariplo" and "Fondazione Banca Popolare di Lodi".

C-136

Fine mapping of loci on BTA8 associated to antibody response to *Mycobacterium avium* paratuberculosis in cattle

Giulietta Minozzi[1], Marcello Del Corvo[4], Isabella Della Noce[4], Mario Luini[2], Alessandra Stella[3], Giulio Pagnacco, John Luis Williams[4]

[1]*Dipartimento di Scienze Veterinarie e Sanità Pubblica, Università degli Studi di Milano, Italy*
[2]*Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Lodi, Italy*
[3]*Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, Lodi, Italy*
[4]*Parco Tecnologico Padano, Lodi, Italy*
Corresponding author: giulietta.minozzi@unimi.it

Paratuberculosis (ParaTB) or Johne's disease, caused by *Mycobacterium avium* subspecies paratuberculosis commonly known as MAP in cattle, is a chronic gastroenteritis characterized by diarrhoea, decreased milk production and ultimately death. MAP is responsible for huge economic losses, particularly in dairy cattle herds. Susceptibility to MAP infection has been found to be heritable with heritability estimates ranging from 0.06 to 0.102. The definition of an infected animal can be based either on the presence of anti-MAP antibodies in the serum, or by direct demonstration of MAP in tissue or faeces by culture or PCR. Several studies have addressed the identification of genetic loci associated with MAP susceptibility. The objective of this study was to refine a locus associated with antibody response to *Mycobacterium avium* paratuberculosi (MAP). Using a genome-wide association analysis, a single nucleotide polymorphism on Bos taurus autosome BTA8 namely the SNP rs43161947 at position 35398490 with a p-value of 7.02 e-05, has previously been identified by the authors as associated with MAP infection. Fine mapping of the region was conducted with 100 single nucleotide polymorphisms spanning a region between BTA8: 34422912 and BTA8: 364553881 covering 2 Mega bases (Mb) designed in to cover 1 Mb ahead and after the SNP identified on BTA8. The 2 Mb region on BTA8 was evaluated within a group of 966 Holstein cows collected from routine ParaTB screening in the province of Lodi in Italy, in an area with a high prevalence of ParaTB. Animals were defined as ParaTB positive based on the detection of serum antibodies produced in response to MAP infection using the ID-screen® ELISA test (ID VET Montpellier, France). Of the 966 samples, 483 were MAP antibody positive (cases) and 483 MAP antibody negative (MAP negative controls). All animals were female, and cases and MAP negative controls were from the same farm tested on the same day.Using a single marker association analysis, conducted within the R statistical environment, we identified 3 different QTLs within the 2 Mega base region, under the main QTL on BTA8 associated with antibody response to MAP, in position 34.700.000, 35.800.000 and 36.400.000 bp. This reveals the complexity of the genetic architecture of the

trait and confirms the need to further explore the genome with fine mapping approaches, or by the use of whole genome sequencing to investigate complex traits, such as disease resistance.

C-137

Heritability of milk b-hydroxybutyrate and its genetic association with milk yield and fat-to-protein ratio in Italian Holstein cows

Mauro Penasa[1], Denis Pretto[2], Alice Varotto[1,3], Massimo De Marchi[1]

[1]*Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente, Università degli Studi di Padova, Legnaro (PD), Italy*
[2]*Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, Tartu, Estonia*
[3]*Associazione Nazionale Allevatori Frisona Italiana, Cremona, Italy*
Corresponding author: mauro.penasa@unipd.it

Ketosis is a metabolic disease occurring when the cow experiences negative energy balance, typically in early lactation. Cows are more likely to suffer from subclinical ketosis, which is characterized by increase concentrations of ketone bodies such as β-hydroxyburyrate (BHB) in blood and milk. Subclinical ketosis has a negative economic impact on the dairy herd, mainly related to the loss of milk production. Despite ketosis is typically influenced by environmental factors, recent studies have reported that a genetic basis for this disease exists. The aim of this work was to estimate heritability of milk BHB (mmol/L) and to assess its genetic correlations with milk yield (MY, kg/d) and fat-to-protein ratio (F:P) in Italian Holstein dairy cattle. Test-day milk samples analyzed by mid-infrared spectroscopy were retrieved from the laboratory of the Breeders Association of Veneto region and included 86,908 records from 19,980 cows of parity 1 to 3 and between 5 and 305 days in milk. Samples were collected between April 2013 and May 2014 in 299 herds. The average number of records per animal within each lactation was 3.5 (range: 1 to 10). To achieve normality and homogeneity of variances, values of BHB were log-transformed to lBHB according to the formula lBHB = ln(BHB+1). The pedigree file (58,687 animals) included individuals with phenotypic records and all their ancestors up to 4 generations back. Genetic parameters for lBHB, MY and F:P were estimated within each lactation using a random regression animal model. Average heritabilities of lBHB were 0.14, 0.13 and 0.08 for parities 1, 2 and 3, respectively. Heritability was quite stable during lactation, with a slightly different pathway between first and second lactation. Mean genetic relationships between lBHB and MY were -0.21, -0.09 and -0.13 for parities 1, 2 and 3, respectively, and the highest negative values (-0.59, -0.36 and -0.20, respectively) were found in early lactation. Finally, mean genetic correlations between lBHB and F:P were 0.33, 0.28 and 0.31, respectively, and the highest positive values (0.63, 0.41 and 0.67, respectively) were assessed at the beginning of lactation. Results suggest that milk BHB routinely analyzed during test-days in Italian Holstein cows is a heritable trait and thus breeding strategies to produce progeny less susceptible to ketosis can be adopted.

Acknowledgments
The authors thank "CheeseBull" project (PSR Veneto - Misura 124) and ERMOS Grant (ERMOS104).

C-138

Genetic and phenotypic relationships of milk coagulation properties and curd firmness modeling with the cheese yield and curd nutrients recovery in bovine milk

Alessio Cecchinato, Claudio Cipolat-Gotet, Giorgia Stocco, Christos Dadousis, Giovanni Bittante

Dipartimento di Agronomia Animali Alimenti Risorse Naturali e Ambiente, Università degli Studi di Padova, Legnaro (PD), Italy
Corresponding author: alessio.cecchinato@unipd.it

The aim of this study was to elucidate the relationships of milk coagulation properties (MCP) and curd firmness (CF) modeling with the cheese yield (CY) and milk nutrient and energy recoveries (REC) in the curd, measured at the individual cow level. Data were collected from 1,167 Brown Swiss cows reared in 85 herds. A 2.0 L milk sample per cow was used for determining 10 phenotypes related to MCP, CF and syneresis as well as 7 cheese-making traits. The first set of traits comprised 4 traditional single point lacto-dynamographic properties [MCP expressed as RCT: rennet coagulation time, min; k20: time to a CF of 20 mm, min; a30 (a45): CF 30 (45) min after rennet addition], 4 parameters derived from modeling 360 CF data recorded over time (1 every 15 sec) for each milk sample (CFP: potential asymptotic CF at infinite time, mm; kCF: curd firming instant rate constant, %×min-1; kSR: syneresis instant rate constant, %×min-1; RCTeq: RCT obtained from modeling all CF measures) and 2 traits calculated from individual equations (CFmax: maximum CF, mm; tmax: time at CFmax, min). The second group of traits accommodated 3 CY traits expressing the weight (wt) of fresh curd (%CYCURD), curd dry matter (%CYSOLIDS), and curd moisture (%CYWATER) as % of wt of milk processed, and 4 REC traits (RECFAT, RECPROTEIN, RECSOLIDS, and RECENERGY calculated as the % ratio between the nutrient in curd and the corresponding nutrient in processed milk). Several multitrait analysis were performed, in a Bayesian framework, to estimate the phenotypic, additive genetic (rg), herd and residual relationships between the aforementioned traits. Results showed that a45, CFP and CFmax were genetically associated with all %CY (rg varying from 0.75 to 0.85 for a45; 0.49 to 0.58 for CFP; and 0.75 to 0.80 for CFmax) and REC traits (rg varying from 0.29 to 0.90 for