Genomic breeding value estimation and QTL detection using univariate and bivariate models

Mario P.L. Calus[1*], Han. A. Mulder[1], Roel F. Veerkamp[1]

[1] Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, Lelystad, Netherlands
* Presenting author: Mario Calus, email: mario.calus@wur.nl

Background. The provided simulated dataset, including a quantitative and a binary trait, was analyzed with four univariate and bivariate linear models to predict breeding values for animals without phenotypes. Two models were used that estimated variance components with REML using a numerator relationship matrix (A), or a SNP based genomic relationship matrix (G), as well as two SNP based Bayesian models with one (BayesA) or two distributions (BayesC) for estimated SNP effects. The bivariate BayesC model sampled QTL probabilities for each SNP conditional on both traits. Genotypes were permutated 2000 times against the phenotype and pedigree data, to obtain significance thresholds for the posterior QTL probabilities.

Results. Estimated breeding values had correlations with phenotypes in the reference population ranging from 0.75 and 0.89 for the quantitative trait, and from 0.58 to 0.67 for the binary trait.
Correlations were calculated between all different estimated breeding values, across models and traits, for animals without phenotypes. Correlations between the different SNP based models were greater than 0.93 (0.87) for the quantitative (binary) trait. Correlations between both traits ranged from 0.48 to 0.77 for the SNP based models, and from 0.36 to 0.61 for model A. Correlations between both traits were on average 0.78 (0.55) for the bivariate (univariate) models. Estimated genetic correlations were 0.71 (0.66) for model G (A).
The bivariate BayesC model detected 17 significant SNPs at the genome-wide level and 24 significant SNPs chromosome-wide. Those SNPs clustered into 14 different windows of 2Mb, suggesting that 14 QTLs were detected.

Conclusions. Estimated breeding values of three different SNP based models, both in their univariate and bivariate forms, were in good agreement. Correlations between estimated breeding values of both traits indicated that bivariate models made better use of the data. Permutating the genotypes against phenotype and pedigree data in the BayesC model provided an effective way to derive significance thresholds for the posterior QTL probabilities.