variance component approaches. To remedy this situation, either some form of SNP filtering or external information is needed. Current methods for prioritizing SNP markers (i.e., BayesB, BayesC) are sensitive to the increased co-linearity in HD panels which could limit their performance. In this study, the usefulness of F_{st}, a measure of allele frequency variation among populations, as an external source of information in genomic selection was evaluated. A simulation was performed for a trait with heritability of 0.4. Data was divided into three subpopulations based on trait distribution (top 5%, bottom 5% and in between). Marker data was simulated to mimic 770K SNP marker panel. A ten chromosome genome with 200K SNPs was simulated. Several scenarios with varying number of QTLs and their associated effects were simulated. F_{st} empirical cutoff values of 0.004, 0.008, 0.01, and 0.02 were used to prioritize markers resulting in 4579, 2288, 1745, and 650 selected SNPs, respectively. Using all 200K markers and no filtering, the accuracy of genomic prediction (correlation between true and predicted breeding values) was 0.48. When SNPs were pre-selected based on F_{st}, accuracy was 0.41, 0.48, 0.49, and 0.53 for F_{st} cutoff values of 0.004, 0.008, 0.01, and 0.02, respectively. It is clear that the accuracy obtained using all SNPs can be easily achieved using only 0.5 to 1% of all markers. These results indicated that SNP filtering using already available external information could increase the accuracy of genomic selection. This is especially important as next generation sequencing technology is becoming more affordable and accessible to animal and plant applications.

Key Words: SNP prioritizing, genomic selection, high density

---

298 Selection of sequence variants to improve dairy cattle genomic predictions. M. E. Tooker*[1], P. M. VanRaden[1], D. M. Bickhart[1], and J. O'Connell[2], *[1]Animal Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD, [2]University of Maryland School of Medicine, Baltimore.*

Genomic prediction reliabilities improved when adding selected sequence variants from run 5 (July 2015) of the 1000 bull genomes project. High density (HD) imputed genotypes for 26,970 progeny tested Holstein bulls were combined with candidate sequence variants within or near genes for 444 Holstein animals. Variants with minor allele frequency (MAF) < 0.01, incorrect map locations, excess heterozygotes, or low correlations of sequence and HD genotypes for the same variant were removed. Individual genotype probabilities < 0.98 from Beagle and Mendelian conflicts between parents and progeny were set to missing. Test 1 included 481,904 candidate sequence SNP consisting of 107,471 exonic, 9422 splice, 35,242 untranslated regions at the beginning and end of genes, 329,769 SNP upstream or downstream of genes. Test 2 also included 249,966 insertions and deletions (indels). After merging sequence variants with 312,614 HD SNP and editing, Test 1 included 762,588 variants and Test 2 included 1003,453. Imputation quality was assessed by keeping 404 of the sequenced animals in the reference population and randomly choosing 40 animals as a test set. Their sequence genotypes were reduced to the subset in common with HD genotypes and then imputed back to sequence. Percentage of correctly imputed variants averaged 97.3% across all chromosomes in Test 1 and 97.2% in Test 2. Total time required to prepare, edit, and impute the sequence variants for 27,235 animals was about 5 d using < 20 processors. Computation of genomic predictions using deregressed evaluations from August 2011 for 33 traits and 19,575 bulls required about 3 d with 33 processors. Predictions were tested using 2015 data of 3983 U.S. bulls whose daughters were first phenotyped after August 2011. Many sequence variants had larger estimated effects than nearby HD markers, but prediction reliability improved only 0.6% points in Test 1 when sequence SNP were added to HD SNP, and only 0.4 higher than HD SNP in Test 2 when sequence SNP and indels were included. However, selecting the 17,000 candidate SNP with largest estimated effects and adding those to the 60,671 SNP used in routine evaluations improved reliabilities by 2.7% points (67.4% vs. 64.7%) on average across traits, compared with 35.2% parent average reliability. Accuracy of prediction can improve by adding selected sequence SNP to marker sets.

Key Words: causative variant, sequence data, genomic evaluation

---

0299 Genomic prediction of crossbred performance. B. Harlizius*[1], M. S. Lopes[1], J. Vandenplas[2], C. A. Sevillano[2], and J. W. M. Bastiaansen[3], *[1]Topigs Norsvin Research Center, Beuningen, Netherlands, [2]Wageningen University, Netherlands, [3]Animal Breeding and Genomics Centre, Wageningen University, Netherlands.*

The majority of the commercial slaughter pigs are crossbred animals. However, breeding efforts have been mainly focused on increasing genetic progress of purebred populations. The aim of this work is to evaluate different strategies to improve genomic prediction of crossbred performance taking into account the breed origin of alleles in crossbred populations (breed-specific effects). Previous work showed that marker effects estimated in one breed cannot predict performance in another breed (across-breed prediction). This might be due to breed-specific effects caused by differences in linkage disequilibrium between the marker and the QTL, as well as differences in allele frequencies and in genetic background of the breeds. For prediction of crossbred performance, marker effects estimated in single-breed data showed some predictive value but training on crossbred data achieved higher accuracies, although the breed origin of alleles was ignored. In this study, prediction accuracies of breeding values from a traditional genomic selection model (GS) were compared