all breeds/populations. The multitask Bayesian method estimates the SNP effects for each breed/population by also utilizing SNP effect information from other breeds/populations. The results showed that realized prediction accuracy of the single-task Bayesian method with the 50K SNP chip for RFI ranged from 0.22 ± 0.04 for Elora to 0.65 ± 0.03 for Angus. For ADG and DMI, the realized prediction accuracy of the single-task Bayesian ranged from 0.21 ± 0.03 for ADG in the Kinsella and TX populations to 0.57 ± 0.03 for DMI in Angus. The genomic prediction accuracies were improved by 0.02 to 0.17 with the simple data pooling Bayesian method except for RFI, in which the prediction accuracies were similar or slightly reduced by 0.02 to 0.07. The multitask Bayesian method yielded better prediction accuracy than the single-task Bayesian for most of the traits but did not perform better than the simple data pooling Bayesian method. Genomic prediction based on the imputed HD SNPs resulted in similar accuracies to that of the 50K SNP chip under all three methods. Further studies that include SNP functional information and/or intermediate phenotypes are underway to improve the genomic prediction accuracy for feed efficiency traits in Canadian beef cattle.

Key Words: beef cattle, genomic prediction, feed efficiency

0323 Use of multivariate statistical analyses to preselect SNP markers for GWAS on residual feed intake in dairy cattle. C. Dimauro[*1], E. Manca[2], A. Rossoni[3], E. Santus[3], M. Cellesi[4], and G. Gaspa[5], *[1]Università di , Italy, [2]Università di Sassari, Italy, [3]ANARB, Italian Brown Cattle Breeders' Association, Bussolengo (VR), Italy, [4]Università di Sassari, Italy, [5]Dipartimento di Agraria, University of Sassari, Italy.*

An index currently used to evaluate feed efficiency in cattle is the residual feed intake (RFI) whose heritability is around 0.20–0.40. Genome wide association studies (GWAS) can contribute to breeding programs aimed at improving RFI by detecting genomic regions and candidate genes that regulate it. However, the detection of significant SNP in GWAS with high density SNP platforms is often hampered by the severity of Bonferroni's *p*-value correction for multiple testing, due to huge number of tests. The pre-selection of markers could be an option to mitigate this problem. In the present research, a multivariate approach was used to select a pool of markers that could have any chances to be associated with RFI. Data consisted of 1092 Brown Swiss young bulls genotyped with the Illumina's 50K BeadChip. Animals were divided into two groups, according to RFI: high RFI (HRFI) for RFI > 0.5 standard deviations from the mean RFI; low RFI (LRFI) for animals with RFI < – 0.5 standard deviations from the mean. The two groups consisted of 266 and 280 animals, for LRFI and HRFI, respectively. Individuals that did not belong to the two groups were discarded.Three multivariate discriminant techniques were applied to data. The stepwise discriminant analysis was used to select 152 genome-wide most discriminant markers that were retained for the further analyses. The canonical discriminant analysis significantly separated the LRFI from the HRFI group, and the extracted canonical function was able to correctly assign 92% of animals to the correct group. Canonical coefficients associated to the 152 SNP in the canonical function were useful to rank markers according to their discriminant power. The ability of the selected SNP in depicting the RFI profile of calves was tested by developing a k-means cluster analysis that correctly classified 84% of individuals. For instance, a GWAS was also developed by regressing RFI phenotypes on SNP covariates. After *p*-values were corrected for multiple testing, no significant marker was obtained by using all original variables (41,183). When only the selected 152 SNP were used, 5 significant markers were obtained.

Key Words: SNP preselection, discriminant analysis, RFI

0324 Breed base representation in dairy animals of five breeds. H. D. Norman[*1], P. M. VanRaden[2], J. H. Megonigal[1], J. W. Dürr[1], and T. A. Cooper[2], *[1]Council on Dairy Cattle Breeding, Bowie, MD, [2]Animal Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD.*

Inheritance of DNA from different dairy breeds can be determined by genotyping, just as individual ancestors such as parents, grandparents, or even great grandparents can be identified correctly in a high percentage of cases by genotyping even if not reported or reported incorrectly in pedigrees. Numbers of crossbreds in the U.S. dairy herd have increased by about 400% in the last decade. A procedure developed to determine the extent that alleles of various breeds appear in these crossbreds and in apparent purebreds was used to document breed composition in animals genotyped. The procedure constructed purebred reference groups (PRG) containing registered AI bulls (with milking daughters) chosen to represent 5 different breeds: Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE). Any bull with an ancestor of another breed in his recorded 5-generation pedigree was excluded from the PRG. An exception was made for AY, for which other red breeds were permitted. The procedure was termed breed base representation (BBR) and estimated the similarity of alleles present in the 5 PRG to those of genotyped individuals. To measure BBR, the percentages of DNA contributed to a genotyped animal by each of the 5 breeds were calculated, summed, and then restricted to be between 0 and 100%. The more an animal's alleles resembled those in a PRG, the higher its BBR for that breed. The BBR help reveal the presence of either outcross bloodlines or crossbreeding, which are difficult to separate. Because animals vary even within breeds, the true source of the various breed alleles differs somewhat from BBR. Numbers of AI bulls in the reference populations in March 2016 were 442 AY, 5464

BS, 550 GU, 19,209 HO, and 3147 JE. Primary-breed BBR for those bulls were 97.2, 97.6, 97.8, 99.2, and 98.0%, respectively, which implies that they are purebreds; SD were 1.9, 1.2, 2.7, 1.2, and 1.0%, respectively. Mean primary-breed BBR were 94.8 for AY, 97.0 for BS, 97.8 for GU, 99.0 for HO, and 96.5% for JE for all genotyped males (201,283) and 95.0, 97.1, 96.9, 98.9, and 96.5%, respectively, for all genotyped females (994,949); SD ranged from 1.2 for males and 1.5% for females (HO) to 5.6 and 4.4% (AY), respectively. Genetic predictions for animals with crossbred genetics in their pedigrees could be obtained in the future by weighting marker effects from each breed by BBR.

Key Words: allele, genomics, purebred

0325 Estimation of the composition of four U.S. swine breeds using genomic data. S. A. Funkhouser[*1], R. O. Bates[2], C. W. Ernst[2], D. W. Newcom[3], and J. P. Steibel[2,4], *[1]Genetics Program, Michigan State University, East Lansing, [2]Department of Animal Science, Michigan State University, East Lansing, [3]National Swine Registry, West Lafayette, IN, [4]Department of Fisheries and Wildlife, Michigan State University, East Lansing.*

Lines of purebred pigs are essential for use in crossbreeding systems within the commercial industry. However, verification of breed purity can be challenging, and using color test matings to confirm white color in the Yorkshire or Landrace breeds is time-consuming and costly. Alternatively, advances in the availability and analysis of genomic data may enable rapid and precise determination of breed composition. Here, we have refined methods for determination of breed composition in U.S. populations of four swine breeds, and white color in Yorkshire or Landrace breeds using SNPs present on the GeneSeek Genomic Profiler for Porcine LD platform. These methods use a linear model in which unknown animal genotypes are regressed on a panel of allele frequencies, derived from reference Duroc, Landrace, Hampshire and Yorkshire purebred animals. Only SNPs that are not fixed across all reference animals and have a genotyping call rate of 90% or greater were used in the model. Model coefficients were constrained to be non-negative and to sum to 1.0, facilitating their interpretation as breed composition coefficients. By simulating 1000 admixed animals of known composition, a strong correlation was observed between the actual and estimated breed proportion of the simulated animals ($R^2 = 0.94$) so long as the actual breed of the simulated animals was reflected in the reference panel. Among a real dataset consisting of 920 Yorkshire sires, 95% of the animals were evaluated to have a Yorkshire breed proportion of 0.825 or greater. Determining that an animal may be highly purebred genome-wide does not preclude from failing a color test mating, in which alleles at particular genes such as *KIT* play a major role in color segregation patterns. Using seven SNPs flanking *KIT* (spanning chr8:43Mb– 44Mb), we have demonstrated that SNP haplotypes derived from the reference animals may be used to compute breed composition probabilities for a genomic segment flanking *KIT* of an unknown test animal. From the real Yorkshire sire dataset, 95% of the animals were estimated to have at least a 0.439 KIT-based breed composition probability of being a white breed. Dual use of genome-wide breed proportions and gene-based breed probabilities has great potential to inform swine breeders of the overall purity of an animal, as well as breed characteristics around particular key genes. Such knowledge may reduce the need to perform color test matings or other time-consuming and expensive procedures for breed verification.

Key Words: breed composition, swine, SNPs

0326 Genome-wide association study and accuracy of genomic prediction for teat number in Duroc pigs using genotyping by sequencing. C. Tan[*1,2], Y. Da[2], Z. Wu[3], D. Liu[3], X. He[2,3], N. Li[1], and X. Hu[1], *[1]State Key Laboratory for Agrobiotechnology, China Agricultural University, Beijing, [2]Department of Animal Science, University of Minnesota, St. Paul, [3]College of Animal Science, South China Agricultural University, Guangzhou.*

Swine teat number is related to a sow's ability to rear piglets to weaning age. The objective of this study was to identify genetic factors affecting swine teat number, evaluate the accuracy of genomic prediction, and evaluate the contribution of significant genes and genomic regions to the total genomic heritability and prediction accuracy using 84,151 autosome single nucleotide polymorphism (SNP) markers from genotyping-by-sequencing on 2936 Duroc boars. Heritability of teat number estimated using genomic restricted maximum likelihood estimation was 0.397 ± 0.033 for additive heritability and was 0.055 ± 0.027 for dominance heritability. Observed prediction accuracy calculated as the average correlation between the genomic best linear unbiased prediction and the phenotypic observations of validation individuals in a 10-fold validation study was 0.44 ± 0.04. Genome-wide association study (GWAS) and heritability estimates of individual SNPs identified a cluster of SNPs in or near the *PTGR2*, *FAM161B*, *VRTN* and *AREL1*genes in the 102.5–104.3 Mb region of chromosome 7 to have highly significant SNP effects on teat number. Fitting 10 SNPs in or near these four genes as fixed non-genetic effects in the model eliminated the significant effects in this region, reduced the additive heritability by 5.9% and reduced the prediction accuracy by 6.88%. Chromosomes 1, 2, 11, 12, 14, and 17 also had significant effects on teat number or substantial SNP heritabilities, and removal of those significant effects by fitting them as fixed non-genetic effects in the model reduced the prediction accuracy by 0.74–2.59% and reduced the total SNP additive heritability by 0–2.69%. The results indicated that swine teat number was