Abstracts of the 2017 American Dairy Science Association® Annual Meeting

June 25–28, 2017
Pittsburgh, PA

Journal of Dairy Science®
Volume 100, Supplement 2

and are reported since May 2016 for all 1.6 million genotyped dairy animals. Animals with > 94% of any breed were rounded to 100%, and contributions of other breeds were set to 0%. All-breed scale GPTAs were first computed for each pure breed for traits milk, fat, protein, productive life, somatic cell score, daughter pregnancy rate, cow conception rate, livability, and net merit. These estimates included foreign information from multi-trait across-country evaluation (MACE) and foreign dams converted from within-breed to the all-breed base. Then, marker effects for each breed were blended by BBR to compute evaluations for crossbreds (<94% purebred) for those same traits. Conformation traits do not have an all-breed scale, so only the Jersey marker effects were applied to the crossbreds, and results seemed reasonable. Calving traits are not predicted for crossbreds, and instead a common mean was used for all crossbreds as is the current practice for breeds other than Holstein and Brown Swiss. All-breed GPTAs were then converted to within-breed GPTAs. Correlations of GPTAs for purebreds computed on the all-breed vs. current within-breed scales were 0.97 to 0.99 for most traits and breeds. Crossbred GPTAs were then computed for 44,023 crossbreds, 20,367 of which had no previous GPTAs because of breed check edits. The new GPTAs were for 1,822 Jersey × Holstein crossbreds with >40% of both breeds (F1 crosses), 75 Brown Swiss × Holstein F1, 7,237 Holstein backcrosses with >67% and <94% Holstein, 7,820 Jersey backcrosses, 313 Brown Swiss backcrosses, 1,763 other crossbreds of various mixtures, and 1,337 purebreds that had previously failed breed checks. Additional automation and redesign of many downstream programs is required for the new all-breed system to be used in weekly, monthly, and full releases. The new system is expected to provide accurate predictions for crosses among the 5 dairy breeds evaluated.

Key Words: crossbreeding, genomic prediction, breed composition

462 Genetic trends from single-step GBLUP and traditional BLUP for production traits in US Holstein. Y. Masuda*[1], I. Misztal[1], P. M. VanRaden[2], and T. J. Lawlor[3], [1]*University of Georgia, Athens, GA,* [2]*USDA, AGIL, Beltsville MD,* [3]*Holstein Association USA Inc., Brattleboro, VT.*

The objective of this study was to compare genetic trends from a single-step genomic BLUP (ssGBLUP) and the traditional BLUP (tradBLUP) models for milk production traits in US Holstein. We used 764,029 genotyped animals in this study. Phenotypes were 305-d milk, fat, and protein yield from 21,527,040 cows recorded between January, 1990 and August, 2015. The pedigree file included 29,651,623 animals limited to 3 generations back from recorded or genotyped animals. We applied a 3-trait repeatability model with the same genetic parameters used in the US official genetic evaluation. Unknown parent groups were incorporated into the inverse of a relationship matrix (H^{-1} in ssGBLUP and A^{-1} in tradBLUP) with the QP-transformation. In ssGBLUP, 18,359 genotyped animals were randomly chosen as core animals to calculate the inverse of genomic relationship matrix with the APY algorithm. Computations with tradBLUP took 6.5 h and 1.4 GB of memory, and computations with ssGBLUP took 13 h and 115 GB of memory. Estimated breeding values were adjusted to a genetic base on recorded cows born in 2000 in each model and converted to GPTA in ssGBLUP and PTA in tradBLUP. For genotyped sires with at least 50 daughters with phenotype(s) born between 2000 and 2010, the genetic trend of GPTA was always greater than PTA in all traits. The difference in 2 genetic trends was almost constant for the sires born up to 2008 (on average, 11 kg in milk, 0.5 kg in fat, and 0.3 kg in protein yield) and the difference was greater in the last 2 years. The difference between the GPTA means for the bulls born in 2010 was 35 kg for milk, 2.2 kg for fat, and 1.2 kg for protein yield. For genotyped cows with phenotype(s), the GPTA trend was identical to or slightly greater than the PTA trend up to 2006. Two trends started to diverge obviously in 2007 and the GPTA trend kept rising while the PTA trend remained at the same level. The single-step method provides very similar genetic trends to the traditional evaluations except for the last few years. The recent lower PTA trend can be due to a downward bias caused with genomic pre-selection of young animals.

Key Words: genomic evaluation, genetic trend, PTA

463 A Genetic Diversity Index method to improve imputation accuracies of rare variants. A. M. Butty*[1], F. Miglior[1,2], P. Stothard[3], F. S. Schenkel[1], B. Gredler[4], M. Sargolzaei[1,5], and C. F. Baes[1], [1]*Centre for Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, ON, Canada,* [2]*Canadian Dairy Network, Guelph, ON, Canada,* [3]*Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada,* [4]*Qualitas AG, Zug, ZG, Switzerland,* [5]*Semex Alliance, Guelph, ON, Canada.*

Different methods to select animals for sequencing have been developed, which rely on pedigree-based relationship matrices, genomic relationships matrices, or on haplotype frequencies. Relationship-based methods select representative key animals of a population whereas haplotype frequency methods aim for better coverage of rare variants. Good average accuracies of imputation from SNP chip to whole-genome sequence (WGS) for common haplotypes were reached with the relationship-based methods. Imputation of rare variants, however, still needs to be improved, which can possibly be accomplished with a newly developed Genetic Diversity Index (GDI). This algorithm optimizes the count of unique haplotypes present in a group of animals composed of already sequenced individuals and a fixed number of sequencing candidates. Optimization is run iteratively, exchanging one candidate at a time and computing the GDI of the new group. Use of the simulated annealing algorithm defines whether the last individual added to the group should be kept. Simulated annealing has the advantage of searching for a global optimum in a situation where multiple local optima are present. The previously mentioned key ancestor and haplotype-based methods for selecting sequencing candidate were assessed and compared with the GDI algorithm using simulated cattle WGS data. Average squared correlation coefficients were used to assess imputation accuracy. A preliminary study showed that the accuracy was 1.5% higher when using GDI to enlarge the reference population than the second-best method. Application of the different methods of selection in North American Holstein data showed that the GDI algorithm selected animals carrying a higher percentage of rare haplotypes than other methods examined. Principal component analysis of the population showed that the animals selected with all tested methods were similarly distributed over the pool of candidates. When representative animals of a population are already sequenced and good overall imputation accuracies are reached, sequencing of genetically diverse animals improved the accuracy of the imputation of rare variants to the WGS density level.

Key Words: sequencing, simulation, imputation

464 Determination of quantitative trait variants by concordance via application of the a posteriori granddaughter design to the US Holstein population. J. I. Weller*[1,2], D. M. Bickhart[2], G. R. Wiggans[2,3], M. E. Tooker[2], J. R. O'Connell[4], J. Jiang[5], and P. M. VanRaden[2], [1]*Agricultural Research Organization, The Volcani Center, Rishon LeZion, Israel,* [2]*Agricultural Research Service, Beltsville, MD,* [3]*Council on Dairy Cattle Breeding, Bowie, MD,* [4]*University of*

Maryland Medical School, Baltimore, MD, [5]University of Maryland, College Park, MD.

Experimental designs that exploit family information can provide substantial predictive power in quantitative trait variant discovery projects. The a posteriori granddaughter design was applied to the US Holstein dairy cattle population. Twenty-nine trait-by-chromosomal segment effects were found with probabilities $< 10^{-20}$ that a segregating quantitative variant was detected by chance. Polymorphism genotypes for 79 grandsires and 16,236 sons were determined by imputation for 3,148,506 polymorphisms across the entire genome; 444 Holstein bulls had complete genome sequence, including 38 of the grandsires. Concordance between quantitative trait locus genotype and polymorphism was determined for all 29 effects. Complete concordance was obtained only for daughter pregnancy rate on chromosome 18 and protein percentage on chromosome 20. For each quantitative trait locus, effects of the 20 polymorphisms with the highest concordance scores for the analyzed trait were computed by stepwise regression. The effects for stature on chromosome 7, daughter pregnancy rate on chromosome 18, and protein percentage on chromosome 20 met the following 3 criteria: complete or nearly complete concordance, significance of the polymorphism effect after correction for all other polymorphisms, and a marker coefficient of determination that was > 50% of the total multiple-regression coefficients of determination for the 20 polymorphisms with highest concordance. An intronic variant SNP on chromosome 5 at position 93,945,738 explained 7% of the variance for fat percentage and 85% of the total variance explained by the multiple-marker regression. Variants identified in this study are likely to provide improved predictive power for genomic evaluation of dairy cattle.

Key Words: genomic selection, granddaughter design, quantitative trait variant

465 Impact of SNP selection on genomic prediction for different reference population sizes. D. A. L. Lourenco*[1], I. R. Menezes[2,1], B. O. Fragomeni[1], H. L. Bradford[1], S. Tsuruta[1], and I. Misztal[1], [1]*University of Georgia, Athens, GA,* [2]*University of Sao Paulo, Pirassununga, SP, Brazil.*

Methods for SNP selection can improve prediction accuracy over genomic BLUP, but in practice, the improvement is trait and population specific. This study investigates the importance of SNP selection in populations with 2000 to 25,000 genotyped animals. Populations were simulated with effective population sizes (Ne) of 20 or 100, and assuming that 10, 50, or 500 QTL were affecting a trait with heritability of 0.3. Pedigree information was available for 6 generations; phenotypes were recorded for the 4 middle generations. Animals from the last 3 generations were genotyped for 45,000 SNP. Single-step genomic BLUP (ssGBLUP) and weighted ssGBLUP (WssGBLUP) were used to estimate genomic EBV (GEBV). For WssGBLUP, 2 iterations of weights were calculated and were used to derive SNP variances and to construct a weighted genomic relationship matrix (G). Improved prediction accuracies are expected in WssGBLUP because more weight is placed on important SNP. Prediction accuracies were calculated for 1000 genotyped animals in the last generation. Reference populations included 2000, 5000 and 25,000 genotyped animals. The latter genotyped set was used to assess the dimensionality of genomic information (number of effective SNP or effective chromosome segments - Me). This was calculated as the number of the largest eigenvalues explaining 98% of the variation in the genomic relationship matrix with and without the weights. For the data sets with Ne = 20 and 10 QTL, the accuracy gain from WssGBLUP was 12, 9, and 4 points for 2000, 5000, and 25,000 genotyped animals, respectively. With Ne = 100, this gain was 8, 10, and 7 points, respectively. For both Ne of 20 and 100, the gain assuming 50 QTL was halved, and no gain was observed assuming 500 QTL. The number of effective SNP was about 4-fold less in weighted G (~1512) than in unweighted G (~5790), explaining the greater gain in accuracy with fewer genotyped animals. The impact of SNP selection decreases with increasing size of the reference population and number of QTL. In large populations, the detection of chromosome segments is more difficult, requiring more genotyped animals.

Key Words: accuracy, variable selection, weighted ssGBLUP

466 Optimum selection of core animals in the efficient inversion of the genomic relationship matrix. H. L. Bradford*, I. Pocrnic, B. O. Fragomeni, D. A. L. Lourenco, and I. Misztal, *University of Georgia, Athens, GA.*

The objective was to determine the effect of using core animals from different generations in single-step genomic BLUP with the Algorithm for Proven and Young (APY). Effective population size and number of independent chromosome segments (ICS) are limited in livestock populations indicating limited dimensionality of genomic information. The APY takes advantage of this dimensionality and assumes that breeding vales (BV) for noncore animals are functions of the BV for core animals. The core animals represent the same information as the ICS. Simulations comprised a moderately heritable trait for 95,010 animals and 50,000 genotypes for animals across 5 generations. Genotypes consisted of 25,500 SNP distributed across 15 chromosomes. Core animals were defined based on individual generations, equal representation across generations, and at random. For a sufficiently large core size, core definitions had the same accuracies ($r^2 = 0.90 \pm 0.01$) and biases ($\beta_1 = 1.02 \pm 0.01$) for young animals, even if the core animals had imperfect genotypes because of imputation. Using the youngest generations as core caused an increase in the number of rounds to convergence indicating some numerical instability with these core definitions. When 80% of genotyped animals had unknown parents, accuracy and bias were significantly better ($P \leq 0.05$) for random and across-generation core definitions ($r^2 = 0.71 \pm 0.01$; $\beta = 0.75 \pm 0.01$) than for single generation core definitions ($r^2 = 0.61 \pm 0.01$; $\beta = 0.53 \pm 0.01$). This difference could result from improved relationship estimates between animals in different generations, because all generations were represented in the core partition that was directly inverted in APY. Thus, any subset of genotyped animals can be used to approximate the ICS when pedigrees are complete, but core animals should represent all generations when pedigrees are incomplete.

Key Words: APY, genomic selection, single-step genomic BLUP

467 Including causative variants into single-step genomic BLUP. B. D. Fragomeni*[1], D. A. L. Lourenco[1], Y. Masuda[1], A. Legarra[2], and I. Misztal[1], [1]*University of Georgia, Athens, GA,* [2]*INRA, Castanet-Tolosan, France.*

The purpose of this study was determining, by simulation, whether (single-step) GBLUP is useful for genomic analyses when causative Quantitative Trait Nucleotides (QTNs) are known. Simulations included 180k animals in 11 generations. Simulated population mimicked a cattle population with weak selection intensity (Ne ~200). Phenotypes were available for animals in generations 6–10. Genotypes were available for 24k parents and 5k young animals in generation 11, and included 60k regular SNPs in 10 chromosomes, with genetic variance fully accounted for by 100 or 1,000 biallelic QTN, with effected sampled from a gamma distribution with shape parameter equal 0.4. LD (r^2) between SNPs