imputation tends to be high when minimum requirements are met. Nevertheless, a certain rate of errors is unavoidable. Such rate of errors tends to increase with the increase of the generational interval between reference and testing generations. Thus, it is reasonable to assume that accuracy of GEBVs will be affected by imputation errors. To evaluate the impact of multi-generational selection on the accuracy of SNP genotypes imputation on the reliability of resulting GEBVs, a simulation was carried out under varying updating of the reference population, distance between training and validation sets, and the approach used for the estimation of GEBVs. Using fixed reference populations, imputation accuracy decayed by around 0.5% per generations. In fact, after 25 generations, the accuracy was only 7% lower than the first generation. When the reference population was updated by either 1% or 5% of the top animals in the previous generations, decay of imputation accuracy was substantially reduced. These results indicate that low-density panels are useful, especially when the generational interval between reference and testing population is small. As the generational interval increases, the imputation accuracies decay, although not at an alarming rate. In absence of updating of the reference population, accuracy of GEBVs decays substantially in 1 or 2 generations with a decrease rate of around 20–25% per generation. When the reference population is updated by 1 or 5% every generation, the decay in accuracy was only 8 to 11% after 7 generations. These results indicate that imputed genotypes provide a viable alternative, as long the reference and training populations are appropriately updated.

Key Words: genotype imputation, genomic selection, accuracy

W83 Genome-wide association study for milk production traits in Russian dairy cattle. Alexander A. Sermyagin*[1], Elena A. Gladyr[1], Sergei N. Kharitonov[1], Alexander N. Ermilov[1,2], Ivan N. Yanchukov[2], Nikolai I. Strekozov[1], and Natalia A. Zinovieva[1], *[1]L.K.Ernst Institute of Animal Husbandry, Dubrovitsy, Moscow, Russia, [2]Regional Information Selection Center, Noginsk, Moscow, Russia.*

Genome-wide association study has been proven as a powerful tool for identifying genomic variants associated with economically important traits in domestic animal breeds. Our study is the first step toward the creating the reference population of Russian Holsteins to utilize the genomic information in the dairy cattle breeding programs in Russia. The objective of our study was to evaluate the association between SNPs and estimated breeding values (EBVs) for milk production traits. The genomic data were obtained by genotyping 195 progeny-tested and 61 young bulls using the Illumina Bovine SNP50 v2 BeadChip. The SNP's quality control was performed by using Plink (1.07) software. BLUP AM approach has been used to estimate the marker effects which were applied then to calculate the genomic EBVs for young bulls to increase the prediction reliability of the associations. Direct genomic and genomic EBVs (by GBLUP) were estimated for 305-d milk yield (MY), milk fat yield (FY), milk protein yield (PY), fat percent (FP) and protein percentage (PP). After the quality control, 41370 SNPs were selected for the association analysis. The average number of daughters per sire was about 240 and the reliability of EBVs amounted 87%. The linkage disequilibrium was $r^2 = 0.41$. The Bonferroni correction for detection significant associations was applied as $P < 1.2 \times 10^{-6}$. Two SNPs which had the most significant effect for MY were identified: ARS-BFGL-NGS-50172 on BTA17 ($P = 7.6 \times 10^{-8}$) and Hapmap54246-rs29017970 on BTA13 ($P = 1.7 \times 10^{-7}$). The association analysis for milk components revealed 3 SNPs significantly associated with FP: BTA-104917-no-rs on BTA9 ($P = 4.1 \times 10^{-8}$), ARS-BFGL-NGS-107379 on BTA14 ($P = 6.0 \times 10^{-7}$) and BTB-01604502 on BTA9 ($P = 1.1 \times 10^{-6}$). The effect on PP was shown for SNP Hapmap43278-BTA-50082 on BTA20 ($P = 8.0 \times 10^{-7}$). Few SNPs were found to have the effects on FY and PY traits. The significant effects of SNPs explained from 9.0 to 11.3% of additive genetic variances. Supported by the Russian Ministry of Education and Science (RFMEFI60414X0062).

Key Words: genome-wide association, breeding value, milk production

W84 Identification of copy number variable gene families in Holstein and Jersey cattle. Derek M. Bickhart*[1], Lingyang Xu[2,1], Jana L. Hutchison[1], Harris A. Lewin[3], and George E. Liu[1], *[1]United States Department of Agriculture, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD, [2]University of Maryland, Department of Animal and Avian Sciences, College Park, MD, [3]University of California, Department of Evolution and Ecology, Davis, CA.*

Copy number variants (CNV) represent a large proportion of genetic variation within the cattle genome that has yet to be accurately characterized by SNP genotyping arrays. While significant progress has been made in the identification of CNVs within individual animals using next generation sequence data, CNV frequencies within larger populations have not yet been estimated in cattle. In this study, we sequenced 28 individual bulls from 2 dairy breeds of cattle (22 Holstein bulls; 6 Jersey bulls) to identify dairy breed-specific copy number variation. Using a read depth method of CNV detection, we identified 1359 non-redundant CNV regions within all 28 animals. The number of variable bases contained within these CNV regions accounts for ~2% of the cattle genome, and the average CNV region frequency was 37.67%. This high average frequency suggests that a large proportion of CNVs were present in the ancestral population of both breeds of cattle rather than as a result of a large number of de novo events arising in subsequent generations after breed formation. We also assigned copy number values to each gene within each individual sequenced using the normalized sequencing read depth of non-overlapping genomic windows. Using a Vst approach on these gene copy number values, we identified 27 gene families with breed specific copy number expansions/contractions. We identified a Jersey-exclusive expansion of the CLEC5A gene, which is a regulator of osteoclastogenesis. Additionally, we identified a Holstein-exclusive duplication of the ASAP1 gene, which may be involved in cell membrane trafficking and the differentiation of fibroblasts into adipocytes. CNVs identified by this survey intersected gene families that may play a role in productive traits in dairy cattle and are therefore good candidates for novel genetic marker design.

Key Words: copy number variant, genomics, sequencing

W85 Single nucleotide polymorphisms in specific candidate genes are associated with phenotypic differences in days open for first lactation in Holstein cows. M. Sofia Ortega*[1], Anna C. Denicol[1], Daniel J. Null[2], John B. Cole[2], and Peter J. Hansen[1], *[1]Department of Animal Sciences, University of Florida, Gainesville, FL, [2]Animal Genomics and Improvement Laboratory, Agriculture Research Service, United States Department of Agriculture, Beltsville, MD.*

Previously, a candidate gene approach identified 51 single nucleotide polymorphisms (SNP) associated with genetic merit for reproductive traits and 26 associated with genetic merit for production in dairy bulls. We evaluated association of these 77 SNP with days open (DO) for first lactation in a population of Holstein cows grouped based on predicted

transmitting ability for daughter pregnancy rate (DPR): ≤ −1 (n = 1220) and ≥1.5 (n = 1053), and located on 11 farms in Florida and California. Cows were genotyped using a Sequenom MassARRAY assay. To evaluate phenotypes, farm records were retrieved from on-farm computers and combined with records from the national genetic evaluation system. The association of the genetic variants with DO was evaluated using the MIXED procedure of SAS V9.4 (SAS Institute, Inc., Cary, NC). The model included farm, number of copies of the minor allele, and the numerator relationship matrix to account for (co)variances among animals. For each SNP, the genotype was treated as a categorical variable to estimate additive and heterosis effects. Days open was lower ($P < 0.0001$) for cows in the high DPR group as compared with the low DPR group (97.8 ± 2.6 d vs 163.0 ± 2.9 d). There were 6 SNP with significant additive effects ($P < 0.05$) on DO (*COQ9, FCER1G, FST, GPLD1, MRGPRF* and *OCLN*) and an additional 6 SNPS with a tendency ($P < 0.10$) for an association (*ACAT2, CD14, PCCB, PMM2, RABEP2* and *SREBF1*). For example, DO for cows with 0, 1, or 2 copies of the minor allele for *COQ9* averaged 139.4 ± 3.5, 134.3 ± 2.8, and 123.6 ± 3.5 d, respectively. The DO for cows with 0, 1, or 2 copies of the minor allele for *FST* averaged 124.9 ± 3.3, 134.8 ± 2.6 and 135.8 ± 4.4 d, respectively. For 9 of 12 genes, the favorable allele for DO was also the favorable allele in the earlier report based on bulls. The SNP related to genetic and phenotypic estimates of fertility are likely to be informative markers for genetic selection. Moreover, the study of the role of these genes could provide new insights into the physiological regulation of fertility in dairy cattle (USDA AFRI 2013–68004–20365).

Key Words: single nucleotide polymorphism, days open, dairy cattle

W86 Animal selection for whole-genome sequencing by quantifying the unique contribution of homozygous haplotypes sequenced. Jana L. Hutchison*, John B. Cole, and Derek M. Bickhart, *United States Department of Agriculture, Agricultural Research Service, Animal Genomics and Improvement Laboratory, Beltsville, MD.*

Major whole-genome sequencing projects promise to identify rare and causal variants within livestock species; however, the efficient selection of animals for sequencing remains a major problem within these surveys. The goal of this project was to develop a library of high accuracy genetic variants found within diverse haplotypes that were in a homozygous state identified from animal genotypes in the national database. An inverted weight function that calculated the value of sequencing an animal based on the sum of the rarity of the haplotypes it had in its SNP-based genotype was used to calculate the estimate, as more common haplotypes would likely be represented within animals already sequenced in subsequent iterations. A weight value was assigned to each 75-SNP haplotype based on the inverse of its frequency within genotyped animals in the national database. Each individual's haplotype weights were summed, and the highest scoring animal was selected for sequencing. Haplotypes from selected animals were removed from future consideration, and the cumulative scores of all remaining animals were recalculated in the absence of those selected haplotypes. This iteration continued until all haplotypes above a frequency threshold of 4% had been selected for sequencing. There were a total of 3,680 75-SNP haplotypes above a frequency of 4% in the national database and 484,522 genotyped Holstein animals. We compared this method against the selection of animals for sequencing based on 3 additional algorithms: (1) an ascending relatedness weight function, (2) an unbiased predictor of imputation accuracy, and (3) a random selection of animals from the population. By calculating an iterative summed score based on the inverse value of an animal's unsequenced haplotypes, one can quickly determine the value of sequencing a new individual and avoid data redundancy that plagues projects that focus on sequencing highly related individuals in a population.

Key Words: sequencing, haplotype

W87 A GWAS on heat tolerance phenotypes for Italian Holstein bulls. Stefano Biffani[1], Umberto Bernabucci[2], Nicola Lacetera[2], Andrea Vitali[2], Paolo Ajmone Marsan[3], Nicolo PP Macciotta*[4], and Alessandro Nardone[2], *[1]IBBA-CNR, Lodi, Italy, [2]Dipartimento di Scienze e Tecnologie per l'Agricoltura, le Foreste, la Natura e l'Energia Università degli Studi della Tuscia, Viterbo, Italy, [3]Istituto di Zootecnica, Università CAttolica del Sacro Cuore, Piacenza, Italy, [4]Dipartimento di Agraria, Università di Sassari, Sassari, Italy.*

Heat stress is a key factor that negatively affects livestock productive and reproductive performance. A genome-wide scan was performed on a sample of 1,592 Italian Holstein bulls using 2 different measures of heat tolerance for milk yield and protein percentage. The first was a temperature-humidity index breeding value (THI EBV) recently proposed for the Italian Holstein breed. The latter was obtained by a principal component analysis carried out on milk test-day records corrected for environmental effects except for THI. Only the second principal component (PC2), which describes the individual patterns of corrected production data across different THI levels, was considered. Animals were genotyped with the Illumina BovineSNP 50 BeadChip. Monomorphic SNPs (7,140) and SNPs with a call-rate <95% (1,045) were discarded. In total, 45,546 SNPs were retained for the analysis. All bulls had a THI EBV, whereas the PC2 was available only for a sub-sample of bulls (423). Genome-wide scan was performed fitting the GRAMMAR approach through the GenABEL R package. Then, a Gene discovery analysis was carried out considering windows of 0.5 Mb surrounding the significant marker (0.25Mb up and down stream respectively). No significant associations were detected for milk THI EBV, apart from a weak signal on BTA 2 at about 32 Mb. In this region is located the solute carrier family 38, member 11 (SLC38A11) gene, reported to be involved in folliculogenesis in cattle. For PC2, 3 SNPs were detected on BTA 6, 16 and 26, respectively. The SNP on BTA26 is located in a region that hosts genes involved in the ovarian activity (FGF8). An interesting candidate for the SNP located on BTA16 at approximately 42.1 Mb is the dehydrogenase/reductase member 3 (DHRS3), involved in the embryonic development in humans. No significant associations were found for protein percentage THI EBV. However, 3 significant markers were detected for PC2 on BTAs 20, 14 and 8. Interestingly the BTA14 region hosts the junctophilin 1 (JPH1) gene, whose expression has been found to be upregulated in the hypothalamus of chickens subjected to heat stress. These preliminary findings suggest potential genomic regions linked to heat stress resistance in dairy cattle.

Key Words: heat stress, GWAS, principal component analysis

W88 A genome-wide association study of mastitis in US Holstein and the relationship to mammary microbiome profile identifies novel QTL. Heather Huson*[1] and Rodrigo Bicalho[2], *[1]College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, [2]College of Veterinary Medicine, Cornell University, Ithaca, NY.*

One of the most prevalent and costly obstacles facing dairy producers is the occurrence of mastitis. Mastitis is a worldwide endemic disease causing both short and long-term cow health and economic repercussions with production losses in terms of reduced milk yield, clinical treatment, culling of animals, and discarded milk. The objective of this study was