359 Comprehensive analyses of 723 transcriptomes enhance genetic and biological interpretations for complex traits in cattle. G. E. Liu*, *Animal Genomics and Improvement Laboratory, USDA-ARS*, *Beltsville, MD*.

To systemically establish connections at RNA level between tissue/cell types and complex traits in cattle, we uniformly analyzed 723 RNA-seq data from 91 tissues and cell types. We built a comprehensive gene atlas and studied tissue specificity of genes. We demonstrated that tissue-specific genes significantly reflected the tissue-relevant biology, showing distinct promoter methylation and evolution pattern (e.g., brain-specific genes evolve slowest, while testis-specific genes evolve fastest). Through integrative analyses of those tissue-specific genes with large-scale genome-wide association studies, we detected relevant tissues/cell types and candidate genes for 45 economically important traits in cattle, including blood/immune system (e.g., *CCDC88C*) for male fertility, brain (e.g., *TRIM46* and *RAB6A*) for milk production, and multiple growth-related tissues (e.g., *FGF6* and *CCND2*) for body conformation. We validated these findings by using epigenomic data across major somatic tissues and sperm. Collectively, our findings provided novel insights into the genetic and biological mechanisms underlying complex traits in cattle, and our transcriptome atlas can serve as a primary source for biological interpretation, functional validation, studies of adaptive evolution, and genomic improvement in livestock.

Key Words: cattle gene atlas, GWAS, RNA-seq

360 Imputation and investigation of sequence genotypes for 6,735,530 variants of 39,048 Holsteins. A. Al-Khudhair*[1], J. R. O'Connell[2], D. J. Null[1], and P. M. VanRaden[1], [1]*USDA/ Animal Genomics and Improvement Laboratory*, *Beltsville, MD*, [2]*The University of Maryland School of Medicine*, *Baltimore, MD*.

Previous US studies of Holstein genotypes from run5 of the 1000 Bull Genomes Project used sequence variants in exons and very close to genes, whereas current study of run7 genotypes also includes intronic and intergenic loci. After data cleaning/editing, sequence genotypes for 6,735,530 variants of 917 Holsteins were selected from run7 raw data, in addition to array genotypes from the Council on Dairy Cattle Breeding (Bowie, MD) database, which included either 79,294 SNP from routine predictions or 643,059 SNP from imputed high-density (HD) genotypes using Findhap, version 3. A total of 39,048 Holstein bulls had either sequence or imputed HD genotypes, and all were imputed to sequence. Editing and imputation tests combining sequence and HD array genotypes revealed higher genotype error rate with run7 genotypes than from previous run5 genotypes. Genome-wide association was performed with deregressed milk and fat phenotypes of Holstein bulls using a mixed model framework. That framework included an intercept and a polygenic random effect estimated with a genetic relationship matrix constructed from 79,294 markers from the imputed genotype file for December 2019 US genomic evaluations. Residual error was modeled using a diagonal matrix with deregressed animal-specific reliabilities. Milk and fat had 488 and 603 markers, respectively, with a P-value of < 1E-10. Known major loci, such as in DGAT1, ABCG2, and β-casein, had highest effects in official predictions, on the contrary, nearby linked loci had higher effects in imputed HD or imputed sequence data. This indicates that using more variants does not ensure localizing causal variants; however, official predictions included about 800,000 genotyped and phenotyped cows that were not included in the HD or sequence studies. Phenotypic effects were also estimated by multiple regression for 13 traits, but convergence was incomplete. Annotation of results and conditional analyses is underway to investigate if intronic and intergenic loci also directly affect phenotypes of interest and to identify additional candidate loci to be included in future genotyping chips.

Key Words: genome-wide association, genotype, sequence imputation

361 Genomic prediction with single-step genomic BLUP using a subset of genotypes in US Holstein. Y. Masuda*, S. Tsuruta, and I. Misztal, *University of Georgia*, *Athens, GA*.

As of January 2020, the US dairy database includes more than 3.8 million genotypes. Most of the genotypes are for heifers, and only a fraction of them have phenotypes. Although the use of all genotypes in genomic prediction is an ideal strategy, the same prediction-accuracy can be obtained using a subset of genotypes with a decrease of computing cost in single-step genomic BLUP (ssGBLUP). The objective of this study was to compare genomic predictions of young bulls between all genotypes and a subset of genotypes in US Holstein. We calculated the benchmarks using the full data set, provided by the Council on Dairy Cattle Breeding, including 61M phenotypes of 305-d protein yield, 36M pedigrees, and 2.3M genotypes. The benchmarks included daughter-yield-deviation (DYD) from pedigree BLUP, DYD from ssGBLUP, and GPTA from ssGBLUP. We cut off the last 4 years from the full data, and the truncated set included 841K genotypes up to 2014 (841K; ALL). Out of it, we created 2 sets of genotypes including bulls only (142K; BULL) and bulls and cows with records and with both parents known (256K; BULLCOW). For validation, we chose 3,250 bulls that had at least 50 phenotyped-daughters in 2018, but that had no daughters in 2014. The coefficient of determination (R^2) and the slope coefficient (b_1) were calculated from a linear regression of the benchmark on GPTA. For GEBV-ssGBLUP as the benchmark, R^2 was 0.82 for ALL, 0.76 for BULL, and 0.82 for BULLCOW. Whereas b_1 was around 0.90 in ALL and BULLCOW, BULL showed a lower value (0.83). We observed the same tendency in the other benchmarks. Using bull genotypes alone drop the accuracy probably because of limited information. The use of cow genotypes increases accuracy and reduces bias. The inclusion of genotyped heifers does not improve the accuracy of bull predictions. Statistics for validations based on GPTA-ssGBLUP indicate good stability of genomic predictions. Routine analyses by ssGBLUP can include only genotypes for bulls with daughters and cows, with the remaining animals predicted indirectly.

Key Words: genomic prediction, single-step GBLUP, selected genotypes

362 Accuracy of indirect predictions based on prediction error covariance from single-step genomic BLUP. D. Lourenco*[1], I. Aguilar[2], A. Legarra[3], A. Garcia[1], Y. Masuda[1], S. Tsuruta[1], and I. Misztal[1], [1]*University of Georgia*, *Athens, GA*, [2]*INIA*, *Las Brujas, Canelones, Uruguay*, [3]*INRA*, *Castanet Tolosan, France*.

One of the ways to deal with the ever-increasing number of genotyped animals in single-step genomic BLUP (ssGBLUP) evaluations may be to use only genotyped animals with complete information in the official evaluation and compute indirect predictions (IP) for the remaining young genotyped animals. However, if IP are going to be published, there is a need for a measure of accuracy that reflects the standard error of IP. This measure should be similar to the accuracy of GEBV to validate the usefulness of IP. The objective of this study was to implement formulas to compute accuracy of IP based on the prediction error covariance matrix from ssGBLUP. Using field data, complete ssGBLUP evaluations were run with up to 60k genotyped animals. Reduced ssGBLUP evaluations considered genotypes for up to 55k animals. BLUPF90 was used to compute both complete and reduced evaluations. Accuracy of GEBV in the complete evaluation was computed based on PEV, whereas in the reduced evaluation the left-hand side of the mixed model equations was stored. Using POSTGSF90, the submatrix of prediction error covariance (PEC) for GEBV of genotyped animals was extracted and converted to PEC of SNP effects. Using the same software, GEBV were converted to SNP ef-