Breeding and Genetics Symposium: Systems Biology in Animal Breeding: Identifying relationships among markers, genes, and phenotypes

56 Building SNP-derived regulatory networks. A. Reverter,* *CSIRO Livestock Industries, Brisbane, Queensland, Australia.*

The advent of cheaply available high-throughput genetic and genomic techniques has equipped animal geneticists with an unprecedented ability to generate massive amounts of molecular data. As a result, large lists of genes differentially expressed in many experimental conditions of interests have been reported and, likewise, the association of an ever-growing number of DNA variants with phenotypes of importance is now a routine endeavor. While these studies have greatly improved our understanding of the genetic basis of complex phenotypes, they have also revealed the difficulty in explaining more than a fraction of the genetic variance. Inspired by this data rich - knowledge poor dichotomy, systems biology aims at the formal integration of seemingly disparate data sets allowing for a holistic view of the system and where the key properties emerge in a natural fashion. Herein, I present 2 examples of rigorous ways of integrating molecular data anchored in the power of gene network inference. The first example in concerned with the onset of puberty in cows bred in tropical regions of Australia. Using the results from genome-wide association studies across a range of phenotypes, we developed what we termed an association weight matrix to generate a gene network underlying cattle puberty. The network was mined for the minimal set of transcription factor genes whose predicted target spanned the majority of the topology of the entire network. The second example deals with piebald, a pigmentation phenotype in Merino sheep. Two networks were developed: a regulatory network and an epistatic one. The former is inferred based on promoter sequence analysis of differentially expressed genes. The epistatic network is build from 2-locus models among all pair wise associated polymorphisms. At the intersection between these 2 networks, we revealed a set of genes and gene-gene interactions of validated and de novo predicted relevance to the piebald phenotype. These new approaches render attractive a search for genetic mechanisms underlying phenotypes of importance in livestock species.

Key Words: systems biology, gene network

57 Networks and pathways to guide genomic selection. W. M. Snelling*[1], R. A. Cushman[1], J. W. Keele[1], C. Maltecca[2], M. G. Thomas[3], M. R. S. Fortes[4,5], and A. Reverter[4], *[1]USDA, ARS, US Meat Animal Research Center, Clay Center, NE, [2]Animal Science, North Carolina State University, Raleigh, [3]Animal Sciences, Colorado State University, Fort Collins, [4]Cooperative Research Center for Beef Genetic Technologies, CSIRO Livestock Industries, Brisbane, QLD, Australia, [5]The University of Queensland, School of Veterinary Medicine, Gatton, QLD, Australia.*

Many traits affecting profitability and sustainability of meat, milk and fiber production are polygenic, with no single gene having an overwhelming influence on observed variation. No knowledge of the specific genes controlling these traits has been needed to make dramatic improvement through selection. Gains have been made through phenotypic selection, enhanced by pedigree relationships and continually improved statistical methodology. Genomic selection, recently enabled by assays for dense SNP located throughout the genome, promises to increase selection accuracy and accelerate genetic improvement by emphasizing the SNP most strongly correlated to phenotype, although the genes and sequence variants affecting phenotype remain largely unknown. These genomic predictions theoretically rely on linkage disequilibrium (LD) between genotyped SNP and unknown functional variants, but familial linkage may increase effectiveness for predicting individuals related to those in the training data. Genomic selection with biologically relevant SNP genotypes should be less reliant on LD patterns shared by training and target populations, possibly allowing robust prediction across unrelated populations. While the specific variants causing polygenic variation may never be known with certainty, several tools and resources can be employed to identify those most likely to affect phenotype. Dense SNP associated with phenotype provide a one-dimensional approach to identify genes affecting specific traits, while associations with multiple traits allow defining networks of genes interacting to affect correlated traits. Such networks are especially compelling when corroborated by existing functional annotation and established molecular pathways. The SNP occurring within network genes, mined from public databases or derived from genome and transcriptome sequences, may be classified according to expected effects on gene products. Coupling evidence from livestock genotypes, phenotypes, gene expression and genomic variants with existing knowledge of gene functions and interactions may provide greater insight into the genes and genomic mechanisms affecting polygenic traits, and enable functional genomic selection for economically important traits. USDA is an equal opportunity provider and employer.

Key Words: genomic prediction, functional genomics, gene network

58 Causal graphical models in quantitative genetics and genomics settings. G. J. M. Rosa* and B. D. Valente, *University of Wisconsin, Madison.*

Phenotypic traits may relate to each other through complex causal relationship networks, which may transcend across transcriptome, proteome, metabolome, and endpoint phenotypes, such as economically important traits in livestock. Moreover, such systems of phenotypic interrelationships are often modulated by the joint action of genetic and environmental effects. For example, transcriptional levels may be affected by reproductive status or developmental stage, which in turn may be mediated by other physiological variables, with genetic polymorphisms and environmental components contributing to animal-to-animal variation. A probabilistic representation of such biological systems can be accomplished through causal graphical models, which allow expressing complex relationships between variables. Applying such models, however, involves a central task of inferring the causal structure underlying phenotypic networks, i.e., determining the subset of traits that directly affects each trait of interest. This challenge may be tackled by using algorithms that perform a data driven search for plausible causal network structures, based on some specific assumptions such as the causal sufficiency. In this context, quantitative genetics and genomics have been utilized either to relax the causal sufficiency assumption from removing polygenic pleiotropic effects or by using QTL information to aid in causal structure search and edge orienting. In this presentation we will discuss such applications to phenotype network studies, focusing on the use of genomic information to aid inferences. Additionally, we will provide examples of application of such methods, and discuss their potential benefits in animal sciences, such as optimal management decision and improvement of multiple traits in livestock.

Key Words: graphical models, causal inference, quantitative genetics

59 A systems biology definition for semen quality. D. Froman*[1], D. Rhoads[2], and S. Burgess[3], *[1]Oregon State University, Corvallis, [2]University of Arkansas, Fayetteville, [3]University of Arizona, Tucson.*

We have shown that semen quality can be defined as the number of mobile sperm produced per male per day. Moreover, semen quality is subject to genetic selection when so defined. Sperm mobility is a quantitative trait in poultry, and male fertility is a function of sperm mobility phenotype. The term sperm mobility denotes sperm that move against resistance at body temperature. In other words, all mobile sperm are motile but not all motile sperm are mobile. Immobile sperm contain dysfunctional mitochondria. Such mitochondria are damaged before ejaculation. The ratio of mobile to immobile sperm within an ejaculate varies among roosters and is independent of age. We hypothesized that premature mitochondrial failure was due to sperm cell and reproductive tract attributes that interact to affect sperm in a stochastic manner. Our current work tested this hypothesis in terms of (1) reproductive tract throughput, as defined by the number of mobile sperm produced per gram testis per day, and (2) the testis transcriptome. Test subjects were roosters from 2 experimental lines of chickens developed at Oregon State University. Distinct sperm mobility phenotypes have been maintained between lines for more than a decade without inbreeding. Our experimental outcomes demonstrated that deferent duct transit time does indeed differ between lines. In this regard, it is noteworthy that this duct contains the bulk of a rooster's extragonadal sperm reserve. Whereas differential gene expression was first observed in terms of sperm cell glycolytic enzymes (downregulated in sperm from low line males), we have now observed differential gene expression between lines at the level of the testis transcriptome. Analyses based upon reads per kilobase of mRNA per million reads identified 3000 transcripts that were differentially expressed at 1.5-fold or more up or down ($P <$ 0.01). Those genes were mapped to the original GWAS data to identify SNPs in mRNAs differentially expressed between lines. Three genes are now under investigation: ATP5a1, CKMT2, and ARMC2. We have 2 long-term goals: (1) to predict adult reproductive potential using DNA obtained from chicks, and (2) to improve fertility within pedigree lines of chickens through SNPlotype-based selection.

Key Words: fertility, sperm, transcriptome

60 A systems-genetics analysis of bovine skeletal muscle iron content. J. E. Koltes*[1], R. G. Tait Jr.[1], E. R. Fritz[1], B. P. Mishra[1,2], A. L. Van Eenennaam[3], R. G. Mateescu[4], D. L. Van Overbeke[4], A. J. Garmyn[4], Q. Liu[1], G. Duan[1], D. Nettleton[5], D. Beitz[1], D. Garrick[1], and J. M. Reecy[1], *[1]Department of Animal Science, Iowa State University, Ames, [2]National Bureau of Animal Genetic Resources, Karnal, India, [3]Department of Animal Science, University of California, Davis, [4]Department of Animal Science, Oklahoma State University, Stillwater, [5]Department of Statistics, Iowa State University, Ames.*

Regulation of cellular iron is critical in immune, mitochondrial, and erythrocyte function. Improper iron homeostasis results in disease both in the case of excessive iron (hemochromatosis) and when iron is deficient (anemia). Therefore, genetic regulation of iron is important to both human and animal health. Since beef is an excellent source of dietary iron, our objective was to investigate the genetic mechanisms responsible for variation in skeletal muscle iron content. To determine the genetic architecture of skeletal muscle iron content, we analyzed Illumina 54k bovine SNP genotypes (n = 2259) using a Bayes C genomic selection model in GenSel software. To investigate the transcriptional control of skeletal muscle iron content, we used RNA-seq to contrast mRNA levels in 5 high and 5 low (n = 10) total iron content Longissimus dorsi samples. Sequence tags were aligned to the UMD3 genome build using Cufflinks software and analyzed in R assuming a Poisson distribution accounting for fixed effects of iron, sex, age and contemporary group. False discovery rate was controlled using qvalues. GWAS results indicated that markers with the largest effect were located on chromosomes1, 7, 15, and 17. Markers on chromosomes 7 and 17 were near GDF15 and SMAD1 genes with known roles in human hemochromatosis. RNA-seq identified 3010 differentially expressed genes ($q < 0.05$). Pathway studio software was used to identify pathways including known regulators of iron homeostasis: BMP6 ($P < 6.38E^{-8}$), and STAT1 ($P < 1.3\ E^{-27}$). The SMAD3 pathway was identified as a potential novel regulator of iron homeostasis ($P < 3.07E^{-33}$). Additional pathway analyses and genome-wide re-sequencing are underway to identify novel iron regulatory mechanisms. Interestingly, our initial results indicate that many of the same genes involved in human hemochromatosis may also regulate bovine skeletal muscle total iron content.

Key Words: bovine skeletal muscle, genome-wide association study (GWAS), iron homeostasis