All these changes can affect the robustness of genetic evaluation since the definition of the trait is not the same across time. To address this issue, and to guarantee a better comparison of bull and cow EBVs across time and between national and foreign bulls, a formula to derive final score from linear conformation traits EBVs has been developed using the selection index theory. At first a new phenotypic definition of Final Score was defined combining 15 different linear traits according to the last recommendations of the international harmonisation group. Subsequently genetic correlations were estimated between the new definition of Final Score and each of the 15 linear traits scored in Italy. Selection index theory was then applied to maximize genetic gain for Final Score making use of the correlated traits. A total of 15 linear traits were used to estimate Final Score resulting in 31% contribution linked to frame and dairy strength traits, 21% to feet & legs and 48% to udder traits. The new Final Score will be introduced officially in Italy during 2009.

Key Words: EBV, final score, linear conformation traits

37 Modelling technical parameters of individual extended lactation curves in Italian Holsteins. R. Steri[1], E. L Nicolazzi[2], G. Gaspa[1], F. Canavesi[2], C. Dimauro[1], and N. P. P. Macciotta*[1], [1]*Dipartimento di Scienze Zootecniche, Università di Sassari, Sassari, Italia*, [2]*Associazione Nazionale Allevatori Frisona Italiana, Cremona, Italia*.

In several countries, the average length of lactation is increasing due both to reproductive failure but also to management strategies. The investigation of basic traits of individual extended lactations may be of interest for selection programmes. In this study, 68,899 individual lactations of 45,132 Italian Holstein cows were fitted with two mathematical functions that have been previously tested on extended lactations: the Ali and Schaeffer multiple regression function (AS) and the non linear modification of the Dijkstra model (DJ). Lactations were of three parities (first, second and thirty) and were classified into four lactation length classes (1<350 d; 2= from 351 to 450d; 3= from 451 to 650d; 4>651 to 1000d). Both models were fitted to individual lactation curves for milk yield (MY), fat (FP) and protein (PP) percentage and somatic cell score (SCS). Peak yield (Ym), time of peak yield (Tm), time et inflection point (Tf) and final test day production (for AS) or asymptotic production (for DJ) were calculated. AS and DJ function show a good fit when modeling lactations curves of milk yield and protein content in 651-1000d class: the percentage of curves having an R-square higher than 0.7 was about 75% for MY, 60% for PP, 15% for FP and 15% for SCS. Estimates of technical parameters show a remarkable variability. As an example, for milk yield the median of asymptotic level of production estimated with DJ was kg -7.3, kg 11.5 and kg 8.51 for first, second and third parity respectively, whereas the production at last test estimated by AS function was kg 5.5 kg, kg 8.4 and kg 5.5 for the same parities. Time of inflection (Tf) also show a great variability between models: for first parity cows; for example, it was estimated as 137.7 d by DJ and 596 by AS. Finally, time at peak occurrence (Tm) was at days 82.5, 53.6 and 52.9 for DJ and 49.5, 33.7 and 41 days for AS. Both models were able to highlight some basic traits of the curve shape for individual extended lactations, although a great variability has been also detected.

Key Words: extended lactations, individual curves, mathematical modelling

Breeding and Genetics: Molecular Genetics I

38 Hybridization quality diagnostics using control probes on long-oligonucleotide microarrays: An application to the Pigoligoarray. J. P. Steibel*[1], M. Wysocki[2], V. D. Rilington[1], A. M. Ramos[1,3], J. K. Lunney[2], and C. W. Ernst[1], [1]*Michigan State University, East Lansing*, [2]*ANRI, BARC, ARS, USDA, Beltsville, MD*, [3]*Wageningen University, Wageningen, the Netherlands*.

Long oligonucleotide microarrays are composed of 40- to 70-mer oligonucleotides spotted on glass slides. This type of expression profiling platform has been recently made available to several livestock species. Such microarrays commonly include negative control oligonucleotides (scrambled sequences) and perfect match/mismatch oligonucleotide sets (PM/MM). Graphical diagnostics (e.g. MA plots) are commonly used in exploratory microarray analysis, but they do not explicitly consider control oligonucleotides. The objective of this work was to develop graphical diagnostics that use control oligos to evaluate hybridization specificity. We re-analyzed data from 8 previously conducted experiments using the Pigoligoarray. A total of 160 microarray slides were processed under a range of stringency and hybridization conditions that resulted in various degrees of non-specific hybridization. The arrays were classified as "good" or "poor" hybridization quality depending on the extent of cross-hybridization. Negative control oligos had median signal intensities similar to or lower than non-control oligos for the "good" arrays whereas for the "poor" arrays negatives exhibited an almost symmetric distribution of intensities around the median for non-control oligos. The analysis of PM/MM sets was even more informative, clearly indicating whole experiments or arrays with extensive non-specific binding. In experiments with specific hybridization, individual probe analyses of PM/MM sets revealed the expected decrease in signal intensity with increased MM number. Additionally, the dispersion around the median intensity decreased as the number of mismatches decreased. Such patterns were not present on arrays with nonspecific hybridization. While these diagnostics were developed using the Pigoligoarray, the proposed applications could be straightforwardly implemented for other long-oligonucleotide microarrays that include the same types of control probes, such as for the bovine and turkey. In summary, we successfully used PM/MM and negative oligos to elaborate graphical diagnostics of hybridization quality.

Key Words: microarrays, cross-hybridization, graphical diagnostics

39 Low density SNP chip for non-genotyped animals. H. Wang*[1] and R. Rekaya[1,2], [1]*Department of Animal and Dairy Science*, [2]*University of Georgia, Athens*.

Availability of reliable and inexpensive genotyping platforms for single nucleotide polymorphism markers (SNP) has made the possibility for genomic breeding value estimation in several livestock species a reality. Unfortunately, at above two hundred dollars per animal this technology is still too expensive for massive use at the commercial level. Currently, this technology is mainly used for genotyping top animals and then used in two step procedures for estimating genomic breeding values. For its use at the population level, the SNPs genotypes of non-typed animals have to be inferred somehow from the already genotyped animals and their relationships. In fact, several attempts have been proposed ranging from the calculation of gene content to the construction of a covariance matrix similar to the classical additive relationship matrix.

However, in all cases the only information used is the one available from the allele frequencies and the relationship between animals. This information could be limited especially in pedigrees that span several generations as it is the case in food producing animals. In this study, we propose using low-density chips with few hundred SNPs for large scale genotyping. This low cost chip will provide an additional source of information, linkage disequilibrium, in inferring genotypes of non-typed animals. For that purpose a simulation was conducted where 1 and 5% of the population was genotyped for 50 K SNPs and 10 and 25% of the remaining animals were genotyped for 500, 1,000 and 2,000 SNPs selected jointly using an ant colony algorithm. The results showed an increase in the probability of predicting the true SNP genotypes by 5 to 12% and an increase in the accuracy of breeding values estimation by 3 to 6% depending on the number of SNPs in the low cost chips and number of animals genotyped.

Key Words: low-cost chip, SNP, GBV

40 An approach to predict and manage Mendelian sampling variation based on dense SNP data. G. Abdel-Azim*, *Genex Cooperative Inc., Shawano, WI.*

An approach to predict the variance of Mendelian sampling (MS) in future progeny, based on high density SNP marker genotypes of pre-selected parents, is introduced. Simple formulas have been derived to calculate the variance of MS for each SNP marker according to their natural segregation from parents to offspring; pre-estimated SNP effects, and population allele frequencies were utilized. The total variance of MS was obtained as the sum of single MS variances predicted from each SNP. Formulas were validated for theoretical accuracy by simulation. In addition, variability across families was studied and found to be generally sufficient for selection. Comparing variances of MS in future progeny has many applications in genetics. One obvious application is selective genotyping to reduce cost. Another application is management of genetic variability in future progeny of pre-selected parents. A stochastic simulation of 200 sires and 200 dams with 10,000 SNP markers was run a 100 times. The simulation showed a correlation of up to 0.68 between marker-predicted MS variance and realized within-family absolute deviations. The correlation increased from 0.21 to 0.68 with increasing full-sib family size from 1 to 40. A 19% average reduction in deviations from parent average was possible by selecting the bottom 10% of candidate parents for marker-predicted MS variance. Extending formulas to non additive genetic variation and application to multiple traits add extra benefit to the current approach.

Key Words: genetic variation, high density SNP, Mendelian sampling

41 Selection of SNPs for an optimal low-density assay for genomic prediction of transmitting abilities. A. Vazquez*, G. de los Campos, K. A. Weigel, G. J. M. Rosa, and D. Gianola, *University of Wisconsin, Madison*.

A 50K single nucleotide polymorphism (SNP) bovine chip seems to predict transmitting abilities (PTAs) of young dairy sires effectively. However, the cost of the assay may limit its use in commercial herds. A low-cost chip with a subset of informative SNPs may be useful for evaluating young males and females. This study: (a) evaluated different strategies for selecting subsets of SNPs; and (b) quantified the predictive ability of chips of different sizes. Data were from the USDA and consisted of SNP genotypes (32,518 SNPs, after editing) and progeny-test PTAs for 4,783 sires. A training set (sires born before 1998) was used to fit the models, and predictive ability was evaluated using bulls born between 1999 and 2003. SNP effects were estimated using the Bayesian LASSO. First, "full models" (32,518 SNPs) were fitted to PTAs of net merit (NM), productive life (PL), fat (F), and protein (P). Three selection strategies were evaluated: (1) top SNPs (Top) selected based on the absolute value of their estimated effects on the full model for the corresponding trait; (2) top SNPs chosen based on the absolute value of their estimated effects on the full model for NM (TopNM); and (3) SNPs evenly spaced (ES) along the chromosomes. TopNM and ES yield the same chip for all traits, while other Top chips are trait-specific. For each selection strategy, different chip sizes were evaluated (300, 500, 750, 1000, 1250, 1500 and 2000 SNPs), and predictive ability was assessed via the correlation between genomic PTAs and progeny test PTAs in the testing set. Full models had the best predictive ability. Among models based on subsets of SNPs, predictive ability increased with assay size for all traits; however, benefits of going from 1,000 to 2,000 SNPs were relatively small. Pre-selecting SNPs either with Top or TopNM increased predictive ability relative to ES, however the difference diminished as the size of the chip increased. The TopNM chip with 1500 SNPs achieved correlations between progeny test and genomic PTAs of 0.56, 0.61, 0.60, 0.63, which represent about 90% of those of the full model for NM, PL, F and P, respectively.

Key Words: feature selection, genomic PTAs, low density chip

42 Transcriptional profiling during fetal skeletal muscle development of Piau and commercial pigs. B. P. Sollero*[1,2], V. D. Rilington[1], R. J. Tempelman[1], S. E. F. Guimarães[2], J. D. Guimarães[2], M. S. Lopes[2], N. E. Raney[1], J. P. Steibel[1], and C. W. Ernst[1], [1]*Michigan State University, East Lansing*, [2]*Federal University of Viçosa, Viçosa, MG, Brazil*.

Mammalian myogenesis and prenatal skeletal muscle fiber formation are under genetic control. However, little is known about the genes involved in this process or how differences in genotype affect gene expression. The aim of this study was to identify differentially expressed genes in *longissimus dorsi* (LD) of pigs at 40 and 70 d of gestation (stages encompassing the transition from primary to secondary fiber formation) in U.S. commercial crossbred pigs (Yorkshire × Landrace) and Brazilian native Piau pigs. Total RNA was isolated from fetuses obtained from gilts at each gestational age (n=3 commercial gilts; n=4 Piau gilts) and RNA from 3 fetuses per litter was pooled. Transcriptional profiling was performed using the Pigoligoarray (www.pigoligoarray.org) comprised of 20,400 70-mer oligos. Thirteen slides were screened using the Amino Allyl MessageAmp II aRNA Kit (Ambion), and Cy3 and Cy5 dyes. Fluorescence intensity data was LOESS normalized and analyzed with a mixed model using SAS. A total of 937 unique genes with Human Gene Nomenclature (HGNC) annotation were differentially expressed (false discovery rate, FDR < 0.05) between 40 and 70 d gestation in either commercial or Piau pigs. Interestingly, only 226 (24%) of these genes were common to the two breed types, whereas 422 (45%) were preferential to the Piau breed and 270 (29%) were preferential to the commercial pigs. The remaining 19 genes (2%) had multiple oligos on the array for which different oligos were differentially expressed in each breed. Among the 226 genes, most exhibited similar fold-changes between the two ages in each breed. However, a few genes exhibited much greater differences in one breed including DYSFIP1 and NRAP with higher abundance in 40 d Piau fetuses (5.2- vs. 2.6-fold and 4.7- vs. 2.0-fold, respectively), and CA3, FABP4 and GRN with higher abundance in 40 d commercial fetuses (12.2- vs. 5.5-fold, 9.0- vs. 4.0-fold and 3.8- vs.

1.8-fold, respectively). This study reveals transcriptional profiles in LD at 40 and 70 d gestation for commercial and Piau pigs, which helps elucidate phenotypic differences between these breed types.

Key Words: myogenesis, microarray, skeletal muscle

43 Extent of linkage disequilibrium in purebred and crossbred beef cattle. D. Lu*[1], M. Sargolzaei[1], M. Kelly[1], G. Vander Voort[1], Z. Wang[2], J. Mah[2], G. Plastow[2], S. Moore[2], and S. Miller[1], [1]*University of Guelph, Guelph, Ontario, Canada*, [2]*University of Alberta, Edmonton, Alberta, Canada*.

Knowledge of linkage disequilibrium (LD) is important in genomic selection, QTL mapping and single nucleotide polymorphism (SNP) chip design. Knowledge of LD in crossbred beef cattle is limited. The objective of this study was to understand the extent of LD and persistence of LD phase in beef cattle. 56,947 SNP were genotyped (38,745 used) across 29 chromosomes of 60 Angus (AN), 43 Piedmontese (PI) bulls, and 400 crossbred beef cattle. Average distance between adjacent SNP was 60kb. The squared correlation of the alleles at 2 loci (r^2) and r were used to measure LD and phase persistence, respectively. Pattern of LD across the genome was investigated using the sliding window approach. Average LD within a given sized window was calculated as it slid along the chromosome. Amount of LD decreased rapidly from 0.31 to 0.22 to 0.15 when the distance range between markers changed from 0-30kb to 30-60kb and then to 60-100kb, respectively. The effect of distance between markers, breeds, and chromosomes on LD decay was tested using least squares. Breeds and chromosomes had significant effects ($P<0.001$) on LD decay. There were significant interactions between distance and chromosomes ($P<0.001$) as well as between breeds and chromosomes ($P<0.05$). There existed 1Mb-regions of consistently high LD ($r^2>0.3$) on chromosomes 1, 4, 5, 7, 9, 14, 16, 19 in purebred and crossbred animals. Correlations of LD phase were high between crossbred animals and either AN or PI (0.82), and between AN and PI (0.78) in 60kb-regions. These dropped when the regions expanded. The marker density in this study should be sufficient for accurate genomic selection within breed when LD of 0.2 between marker and QTL is required. High LD chromosomal segments should be investigated further. Marker effect estimates across breed may not be similar due to observed change in LD phase.

Key Words: linkage disequilibrium, linkage phase, single nucleotide polymorphism

44 Construction of LD maps for SNPs linked to susceptibility loci. L. Gomez-Raya*, *University of Nevada*, *Reno*.

The extension of linkage disequilibrium (LD) in human and animal genomes allows linkage disequilibrium mapping without requiring a family structure. The objective of this study was the construction of a maximum likelihood model for LD mapping in which only estimates of a disequilibrium parameter and penetrances (Ψ_{TT} ; Ψ_{Tt} ; and Ψ_{tt} ; where Ψ_i =probability of developing the disease given the i-th genotype) corresponding to the three genotypes at a susceptibility locus (T/t) are estimated. A second objective was to carry out a MonteCarlo computer simulation to test if the novel maximum likelihood method can provide unbiased estimates of the linkage disequilibrium parameter and penetrances. It is shown that all disequilibria between a pair of alleles at two loci can be put in terms of allele frequencies and a single disequilibrium parameter, which allows the construction of maximum likelihood equations. Montecarlo computer results are presented for a variety of scenarios regarding number of individuals, linkage disequilibrium parameter and penetrances for the three genotypes. For example, a population was simulated with 200 individuals, an incidence of disease of 30% and frequencies of 0.5 for both the susceptibility allele and the SNP alleles. The simulated disequilibrium parameter was 0.125 (50% of the maximum possible disequilibrium). Penetrances were Ψ_{TT} =0.60; Ψ_{Tt}=0.20 ; and Ψ_{tt} =0.20. The average of the disequilibrium parameter over 1500 replicates was 0.132 (SD=0.048). The averages of estimates of penetrances over replicates were Ψ_{TT} =0.62 (SD=0.17); Ψ_{Tt}=0.20 (SD=0.13); and Ψ_{tt} =0.19 (SD=0.12). Empirical statistical power was 0.77 at a significance level of 0.05. The results suggest some confounding in the estimation of the disequilibrium parameter and penetrances which may lead to bias in the estimation of the linkage disequilibrium parameter. The proposed maximum likelihood method is suitable for fine LD mapping since it allows estimation of a single disequilibrium parameter for each of a number of SNPs in a chromosomal fragment where a susceptibility locus has been mapped.

Key Words: linkage disequilibrium, SNP, genetic mapping

45 Characterization of a whole-genome map of single nucleotide polymorphisms applied to two selection lines in British dairy cattle. G. Banos*[1] and M. P. Coffey[2], [1]*Faculty of Veterinary Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece*, [2]*Sustainable Livestock Systems, Scottish Agricultural College, Edinburgh, Scotland, UK*.

The objective was to study genotyping results from the bovine 50K SNP chip and assess the extent of Linkage Disequilibrium (LD) in two divergent selection lines. DNA extracted from 299 Holstein cows was used to determine genotypes in 54,001 SNP loci with the Illumina BovineSNP50 array. These animals were from two selection experiment lines (166 genetically selected vs. 133 control) raised on the Scottish Agricultural College research farm. Data edits removed loci with a major allele frequency greater than 0.95, call rate less than 65% and missing valid chromosomal assignment or position. The Hardy-Weinberg equilibrium (HWE) state was assessed in each locus with a chi-square test. Pairwise haplotypes were determined using parsimony. LD values were calculated both within and across selection lines for all possible syntenic SNP pairs within 1 Mb as the squared correlation between alleles at two loci. After edits, 41,859 loci (78% of the original total) were kept for analysis. There were, on average, 17.5 SNP per Mb with the highest concentrations on chromosomes 1, 2 and 4. The majority (96%) of the loci were in HWE ($P>0.05$). Observed heterozygosity levels were very close to theoretical expectation. The average LD calculated across all chromosomes was 0.0691, 0.0705 and 0.0746 for all, control and selected line cows, respectively. LD was not uniformly distributed across regions. Looking at adjacent SNP, overall average distance was 0.063 Mb and average LD 0.125 in the three datasets. Of all SNP pairs studied, 6.7-12.8% were in LD greater than 0.30, which is considered the minimum useful for mapping purposes and genomic selection. A few SNP loci pairs (844, 984 and 1070 in the three datasets) were found in nearly perfect (>0.99) LD, suggesting that genotyping only one of them would be sufficient. The overall product-moment correlation of LD values calculated within the control and selected lines was 0.79 ($P<0.001$), ranging from 0.71 to 0.84 for different chromosomes. Results suggest that genetic selection has influenced the LD level in the studied population. Results can be used to support gene detection and genomic prediction studies.