NCCC204: The Interface of Molecular and Quantitative Genetics in Plant and Animal Breeding

Date of Annual Report: 4/30/2009

REPORT INFORMATION

- Annual Meeting Date: 02/22/2009.
- Period the Report Covers: 01/2008 to 12/2008.
- Annual meeting held at Galveston, TX.
- Meeting preceeded 12th Gordon Conference on Quantitative Genetics and Genomics.
- NCCC204 website: http://nimss.umd.edu/homepages/home.cfm?trackID=8920

PARTICIPANTS

- David Cowley (New Mexico State University)
- Jack Dekkers (Iowa State University)
- Ignacy Misztal (University of Georgia)
- William Muir (Purdue University)
- Guilherme Rosa (University of Wisconsin)
- Frank Siewerdt (University of Maryland)
- Juan Pedro Steibel (Michigan State University)
- Ignacio Aguilar (University of Georgia) *Guest*
- John Bastiaansen (Wageningen University, The Netherlands) *Other*
- Marco Bink (Wageningen University, The Netherlands) *Guest*
- Hossein Jorjani (Interbull Centre, Sweden) *Guest*
- Marek Lukaszewicz (Polish Academy of Sciences, Poland) *Guest*
- Herman Mulder (Wageningen University, The Netherlands) *Guest*
- Fabiano Pita (Dow AgroSciences LLC) *Other*
- Jiuzhou Song (University of Maryland) *Guest*
- Addie Werejhen (Hendrix Genetics, The Netherlands) *Guest*
- Craig Beyrouty (Purdue University) *Administrative Advisor*

TECHNICAL PROGRAM

The meeting began at 8:30 a.m. on Sunday, February 22 with opening remarks by Frank Siewerdt, Chair. Current members and new participants were welcome and a tentative agenda was set for the meeting. The morning presentations were given by the following experiment stations: University of Maryland (Siewerdt and Song), Michigan State University (Steibel), and New Mexico State University (Cowley; also presenting for Miller). Discussions on specific station projects were done as the reports were presented. The floor was then given to Craig Beyrouty, Administrative Advisor to the group. The midterm report is due on December 31 of this year, which is the third year of the project after its latest renewal. Low attendance in even-numbered years has been an issue and the continuity of this group may be threatened, despite the group being successful in achieving the goals set forth in the project documentation. Beyrouty also updated the group on the online NIMSS resources and urged the leadership of the group to maintain that information updated to help higher level decision makers in the NIMSS structure properly evaluate the progress and impact of NCCC204 and the need for continuity. The next station report was done by University of Wisconsin (Rosa), closing the morning session.

Station reports resumed in the afternoon, after the business meeting was conducted. Presentations in the afternoon were by representatives of University of Georgia (Misztal), Purdue University (Muir), Wageningen University (Bastiaansen), and Iowa State University (Dekkers) with discussions done as the presentations proceeded. The meeting adjourned at 5:00 p.m.

BUSINESS MEETING

The business meeting was conducted in the afternoon after the group reconvened. The new officers of NCCC204 were elected unanimously: Ignacy Misztal (University of Georgia) for Chair, Juan Pedro Steibel (Michigan State University) for Secretary and Fabiano Pita (Dow Agro) as the local host for the 2010 meeting. The location for the next meeting was not determined but three possibilities were considered: meet with PAG (Tuesday and Wednesday afternoon), meet with the Poultry Breeders Roundtable in May (St. Louis) or meet at Dow Agro in Indianapolis, in late January or early February. Guilherme Rosa (University of Wisconsin) proposed to hold a course on Genomic Selection or a similar topic for an audience of faculty and senior researchers that wish to become knowledgeable in this topic as an additional stimulus for increasing attendance.

No participant volunteered to fulfill the duties of Secretary for the 2009 meeting so the current Chair (Siewerdt) accumulated the functions for the second year in a row and will be responsible for collecting station reports and preparing the annual report, to be submitted by April 30, 2009.

Regarding the mid-term progress report, William Muir (Purdue University) and Frank Siewerdt (University of Maryland) agreed to be the focal points for producing the document. Each participant will have a chance to review the document and to attest for completeness and accuracy of matching publications with the specified goals.

SUMMARY OF INDIVIDUAL EXPERIMENT STATIONS REPORTS

Nine stations submitted activity reports summaries for calendar year 2008. A list of specific accomplishments and impact statements is included with each individual experiment station report summary; these reports are reproduced below. Some participants that were not able to attend also sent station reports.

LIST OF REPORTING STATIONS	**REPORTING MEMBER(S)**
1. University of Delaware	Schmidt
2. University of Georgia	Misztal
3. Iowa State University	Dekkers
4. University of Maryland	Siewerdt, Song, Schierholt
5. Michigan State University	Steibel, Wang
6. New Mexico State University	Cowley, Thomas
7. Purdue University	Muir
8. Virginia Tech and State University	Lewis
9. University of Wisconsin	Rosa

Station 1: U Delaware (C. Schmidt)

We aim to address objective 1. The major goal of the University of Delaware participant is to provide integrated Web based tools for analysis and evaluation of genomic data. Both physical (sequence) and genetic data (QTL) have been integrated via the web based Gallus GBrowse. (http://birdbase.net/cgi-bin/gbrowse/gallus08/). Currently, the Gallus GBrowse database maps many different types of data including QTLs, Genes (including exons, start and stop codons) Gene Ontology terms for genes Links to the Gallus embryonic in situ hybridization project Genetic markers including SNPs and microsatellites Expressed sequence tags A mapping of all probes available on different Gallus microarray platforms Mapping of copy number variations in a variety of chicken lines including representatives from broilers, layers and red jungle fowl.Mapping of proteome results to validate gene expression.

Station 2: University of Georgia (I. Misztal)

Personnel

Ignacy Misztal, PI
Shogo Tsuruta (Associate Research Scientist)
Ignacio Aguilar (Ph.D. student)
Andres Legarra (Research Scientists, INRA, Toulouse)

Genetic evaluation including phenotypic, full pedigree and genomic information

Currently the genomic evaluations use multiple step procedures, which are complicated and prone to errors. For many traits, predictions involving estimation of SNP effects or BLUP using a genomic relationship matrix are equivalent. A single step procedure may be applicable by modifying the numerator relationship matrix A in a regular evaluation to $H= A+\Delta$, where Δ includes deviations from original relationships. However, the traditional mixed model equations require H^{-1}, which is difficult to obtain for large pedigrees. The computations with H are feasible when the mixed model equations are expressed in an alternate form given by Henderson that also applies for singular H, and when those equations are solved by the conjugate gradient techniques. Then the only computations involving H are in the form of Aq or Δq, where q is a vector; the product Ac can be calculated efficiently in linear time using Colleau's indirect algorithm. The alternative equations have a nonsymmetric left-hand side. Several H are possible. A naïf possibility is to substitute the relationships of genotyped animals with the genomic relationship matrix. However, this results in incoherencies because the genomic relationship matrix includes information on relationships among ancestors and descendants. Another possibility is to condition the genetic value of ungenotyped animals on the genetic value of genotyped animals via the selection index (e.g., pedigree information), and then use the genomic relationship matrix for the latter. This results in a joint distribution of genotyped and ungenotyped genetic values, with a pedigree-genomic relationship matrix H. In this matrix genomic information is transmitted to the covariances among all ungenotyped individuals. Both possibilities allow for an efficient computing in the form of Δq. The proposed methodology may allow upgrading of an existing evaluation to incorporate the genomic information.

Future work

The methodology will be implemented for analysis of large data sets. It will be compared with the traditional genomic selection (multiple-step approach) on dairy and possible beef data.

Station 3: Iowa State University (J. Dekkers)

Objective 1. Develop and compare statistical methodology to map genes

Predicting Allele Frequencies in DNA Pools Using High Density SNP Genotyping Data

B. L. Peiris[1], J. Ralph[2], S. J. Lamont[1] and J. C. M. Dekkers[1]
[1] Department of Animal Science, Iowa State University, Ames, IA 50011
[2] Aviagen Ltd. UK

Recent scientific and technological advances have made possible the genotyping of large numbers of individuals for large numbers of single nucleotide polymorphisms (SNPs). Although declining, costs per sample are still prohibitive for many applications. One way to reduce costs is to genotype DNA pools. To evaluate the ability to use DNA pools with the Illumina Infinium genotyping platform, two sets of gradient pools were created using two different pairs of Iowa State University highly inbred chicken lines. Pools containing 0, 10, 20, 40, 60, 80, 90, and 100% of line A vs. B or line C vs. D DNA were created. Two replicate pools were prepared for each gradient, except for the 0 and 100% pools, resulting in 28 pools. All pools were genotyped for 12,046 SNPs. Fixed SNPs were removed for the analysis. Three frequency estimation methods proposed in the literature (Standard, heterozygote-corrected, and normalized) were compared to three alternate methods proposed herein based on MSE, bias, and variance of estimated versus true allele frequencies and the fit of a simple linear regression of estimated on true frequencies. The three new methods had average square root MSE of 4.6, 4.6 and 4.7%, compared to 5.2, 5.5, and 11.2% for the three literature methods. Average absolute biases of three literature methods were 2.4, 2.7, and 8.2% compared to 2.4% for all three of the new methods. The standard deviations of estimates were also smaller for the new methods at 3.1, 3.2 and 3.2%, compared to 3.5, 4.0 and 5.0% for the literature methods. In conclusion, intensity data from the Illumina Infinium Assay can be efficiently used to estimate allele frequencies in pools.

Comparison of Chi-Square, ANOVA, and MANOVA for analysis of high-density SNP data from case-control pools

B. L. Peiris[1], S. J. Lamont[1], K. A. Watson[2], J.C.M. Dekkers[1]
[1] Iowa State University; [2] Aviagen Ltd

Genome-wide case-control association studies using high-density SNP panels are effective to detect SNPs associated with disease. To reduce genotyping costs, DNA pools can be used to estimate and compare SNP allele frequencies between cases and controls. Frequencies estimated from DNA pools are subject to more technical errors than individual genotyping and the power of tests for association depends in part on accurate estimation of technical error variance. The conventional test for association in case-control pool designs is the Chi-square test, with technical error variance estimated across, and assumed equal across, SNPs. Single-SNP analysis of variance (ANOVA), instead, allows for SNP-specific error variance estimates. Neither test, however, allows for simultaneous analysis of neighboring SNPs, of which several may be associated with the disease locus, or for which the associations may differ between lines that are analyzed together. To address these shortfalls, multivariate analysis of variance (MANOVA) can be used to analyze the joint effect of several SNPs. We compared Chi-square, ANOVA and MANOVA to identify SNPs or regions associated with disease using case-control pool data from 4 commercial broiler chicken lines. Thirty case and 30 control pools were constructed by combining equal amounts of anticoagulated blood from 50 case/control individuals from each line. Pools were genotyped for 12,046 SNPs using the Illumina Infinium assay, and 2083 SNPs on chromosome 1 were used for the current analyses. Allele frequencies were estimated and normalized for each pool. Chi-square tests for both main effects across lines and allowing for

line-specific effects were conducted, with p-values from standard Chi-square tables. The ANOVA and MANOVA analyses were based on a model with line, disease status, and their interaction. MANOVA used a moving window of 5 SNPs. Standard F-tests were used for ANOVA and F-tests based on Wilk's lambda for MANOVA. The three analyses resulted in substantial consistency in regions identified as significant. The p-values from Chi-square and ANOVA had a correlation of 0.925. Assuming equal type I errors, power increased from Chi-square to ANOVA to MANOVA. In conclusion, (1) ANOVA is an effective substitute for the conventional Chi-square test to identify significant SNPs, and (2) MANOVA can be used to identify the joint effect of a region in case-control studies across lines.

Comparison of the pattern of linkage disequilibrium in purebred, crossbred and admixed populations

Ali Toosi, Rohan L. Fernando and Jack C.M. Dekkers

The extent and magnitude of linkage disequilibrium (LD) determine the feasibility of association mapping studies that seek to localize genes underlying complex traits and diseases. Here we present a comparison of the patterns of LD across various populations of interest to animal breeders, based on a simulated genome of size 1 M with 40 segregating SNPs per cM. A base population of unrelated individuals was stochastically simulated and used as the ancestral population of four pure breeds that were used to create admixed and crossbred populations. The base population was randomly mated for 1000 generations, with an effective size (N_e) of 500. To simulate the pure breeds, at generation 1001 four independent random samples of 100 animals were drawn from the base population, and each was randomly mated for another 50 generations, with N_e of 100. These breeds were used to generate two-breed combined, F_1, F_2, three- and four-way crosses, and admixed populations of 1000 individuals. We compared the distribution of LD (as measured by r^2) in each population by comparing the average distances between flanking markers at various levels of r^2.

There were significant differences in the extent and the rate of decay of LD between the populations. The slowest and the sharpest rates of decline of average LD were in the purebred and the four-way cross populations, respectively. The average distance between pairs of markers with weak LD ($r^2 \geq 0.1$) was significantly larger in the crossbred, two-breed combined and admixed populations than in the purebred population. On the other hand, the distance between SNPs with strong LD ($r^2 \geq 0.70$) was significantly shorter in the admixed and the crossbred populations than that in the purebred population. The average distance between pairs of markers with strong LD decreased as the number of breeds contributing to the population increased.

Our results indicated that in crossbred and admixed populations marker intervals with strong LD are, on average, much shorter than in purebred populations. Thus, crossbred or admixed populations are more suitable for QTL fine mapping than purebred populations, provided marker density is sufficient. This study confirms that, when high-density markers are available, purebred populations provide more power to detect QTL with their higher level of and yet more extensive LD, but at the cost of a lower mapping resolution compared to multi-breed populations.

Objective 2. Examine the efficiency of incorporating molecular tools in breeding programs through theoretical modeling, computer simulations, and biological testing in actual breeding populations

Response and inbreeding from genomic selection

Jack C. M. Dekkers[1*], Hong-hua Zhao[1,2], Jennifer M. Young[1], David Habier[1,3] and Rohan L. Fernando[1]
[1] Department of Animal Science, Iowa State University, Ames, USA
[2] Pioneer Hi-Bred Int., Johnston, Iowa, USA
[3] Institute of Animal Breeding and Husbandry, Christian-Albrechts University of Kiel, Germany

Genomic Selection (GS) using breeding values (GS-EBV) estimated from dense marker data is promising for genetic improvement. Our objective was to evaluate responses from GS to selection on traditional BLUP-EBV over multiple generations. A trait with 100 or 200 QTL with heritability 0.3 and phenotyping prior to selection was simulated. GS-EBV were estimated using Bayes-B using 1000 individuals from the training generation only (GS-1) or with updating using data from all generations (GS-all). BLUP-EBV used data from all generations (BLUP-all). Response for GS-1 was similar to BLUP-all in initial generations but then fell behind because of declines in accuracy. GS-all had greater response than BLUP-all. Doubling the number of QTL increased response, in particular for GS-all. Rates of inbreeding increased from GS-1 to GS-all and BLUP-all. GS-all resulted in the fastest loss in variance, followed by BLUP-all and GS-1. Lost variance was to a greater degree from drift and loss of favorable QTL alleles for BLUP-all than GS. Despite lower inbreeding, GS had greater variance of response than BLUP-all because of variation in the accuracy of GS-EBV. Deterministic predictions of response were similar to those observed from simulation in the initial generations. Deterministic predictions of rates of inbreeding were similar to observed rates, but observed rates were higher in generations immediately following training because of the impact of relationships on GS-EBV. In conclusion, GS offers great opportunities to further advance genetic improvement programs. GS, as simulated here, primarily capitalizes on historic linkage disequilibrium.

Genomic selection in admixed and crossbred populations

Ali Toosi, Rohan L. Fernando, and Jack C.M. Dekkers

In livestock, genomic selection (GS) has primarily been investigated by simulation of purebred populations. Traits of interest are, however, often measured in crossbred or mixed populations with uncertain breed composition. If such data are used as the training data for GS without accounting for breed composition, estimates of marker effects may be biased due to population stratification and admixture. To investigate this, we simulated a genome of 100 cM with varying marker densities (5 to 40 segregating markers per cM). After 1000 generations of random mating in a population of effective size 500, four lines with effective size 100 were isolated and mated for another 50 generations to create 4 pure breeds. These breeds were used to generate combined, F_1, F_2, three- and four-way crosses, and admixed training data sets of 1000 individuals with phenotype for an additive trait controlled by 5 or 100 segregating QTL and heritability of 0.30. The validation data set was a sample of 1000 genotyped individuals from one pure breed. Method Bayes-B was used to simultaneously estimate the effects of all markers for breeding value estimation. With 40 markers per cM, the correlation of true with estimated breeding value of selection candidates (accuracy) was highest (0.85) when data from the same pure breed was used for training. When the training data set consisted of crossbreds, the accuracy ranged from 0.79 to 0.83. Admixed training data sets gave accuracies in between those for crossbred and purebred training data. However, accuracy was greatly reduced when genes from the target pure breed was

not included in the admixed or crossbred population. This implies that, with high-density markers, admixed and crossbred populations can be used to develop GS prediction equations for all pure breeds that contributed to the population, without a substantial loss of accuracy compared to training on purebred data, even if breed origin has not been implicitly accounted for. In addition, using GS based on high-density marker data, purebreds can be accurately selected for crossbred performance without the need for pedigree or breed information. Results also showed that haplotype segments with strong linkage disequilibrium are shorter in crossbred and admixed populations than in purebreds, providing opportunities for QTL fine mapping.

Station 4: University of Maryland (F. Siewerdt, J. Song, A. Schierholt)

Our goal is to contribute to objectives number 1 and 2 of the project. F. Siewerdt and J. Song are the official representatives of the University of Maryland Experiment Station.

Influence of associative genes on detection of major genes

A.S. Schierholt, R.F. Euclydes, F.Siewerdt

Associative effects are produced by the influence that the social environment has on gene expression in group companion animals. This social environment is defined in part by genes responsible for the behavior of the animals that composes the group. The objective of this work was to evaluate the influence of associative genes in identification analyses of major genes. The data were obtained by simulation. A genome and the allele frequencies of 200 QTLs were simulated. The population in study was made by selecting randomly 3740 simulate animals, offspring of 20 parents and 337 females, forming 11 groups composed by different numbers of animals with the dominant allele for the associative gene in each group. Three simulation of a major gene with direct effects were made in a way that this gene had a different level of effect in each simulation. The associative major gene was simulated as having effect both before and after a hierarchal order be made or having effect only after the hierarchy is formed. For the associative major gene were considered four levels of effects. The result showed that for the simulated effect levels of the genes the associative effect could not mask the identification of major genes with direct effect. The analysis testing whether the associative major gene would be identified as a gene with direct effect showed that, for direct gene effect levels that contributed with more than 3% of the phenotypic mean of the population, the associative major gene was identified as a major gene with direct effect.

Alteration of composition of broiler excreta through selection

F. Siewerdt, A.K. Sasikala-Appukuttan

The ultimate goal is to attempt to reduce the content of nitrogen and phosphorus in the excreta through selection. A proposal was sent to a major broiler breeding company, who is considering the possibility of providing access to their facilities for this work. Expected outcomes from this work are: (1) traditional estimation of genetic parameters of excreta composition; (2) estimate the genetic correlations between excreta composition traits and commercially important traits (e.g., growth, feed utilization efficiency); (3) search for QTL that influence excreta composition; (4) identify methods to incorporate this information into formal breeding programs. Preliminary work led to heritability estimates for percentage of nitrogen in the excreta of $h^2 = 0.14$ in a 197-bird sample and $h^2 = 0.16$ for a 279-bird sample) while heritability estimates in those same samples for percentage of phosphorus in the excreta were, respectively, 0.002, and 0.20. Genetic correlations with residual feed consumption, 3- to 6-wk growth rate, and breeding values for deboned white meat were favorable or neutral, indicating that selection for reduced nitrogen or phosphorus in the excreta will be in synergy with current breeding goals.

J. Song

My research deals with statistical genomics and epigenetics.

Station 5: Michigan State University (J.P. Steibel, D. Wang)

Optimizing design of two-stage experiments for transcriptional profiling (Steibel)
Gene expression microarrays are powerful tools for simultaneously screening the transcriptional profile for thousands of genes across different treatments. Despite their continually improving sensitivity and dynamic range, microarrays are commonly regarded as a first screening step, with a level of precision often deemed unacceptable to use as a standalone technology. This limitation has prompted genomics researchers to validate a statistically significant subset of their microarray results using a second technique, typically quantitative reverse transcription polymerase chain reaction (qRT-PCR). The problem of optimizing such two-stage transcriptional profiling experiments in order to maximize sensitivity, while controlling the false discovery rate (FDR), is addressed. This optimization is based on partitioning the set of available biological replicates into two groups, one for each of the microarray (Stage 1) and qRT-PCR (Stage 2) experiments. It is demonstrated how the significance level should be determined for Stage 2, after selecting a fixed percentage of the genes to validate from Stage 1, in order to maximize the sensitivity of detection of differentially expressed genes for a desired overall FDR. The results indicate that most of the available replicates (typically >60%) should be consumed in Stage 1. Even though the optimization scheme assumes independent genes and known variances, simulation results show that this approach is robust to moderate departures from those assumptions. A procedure to optimize a validation experiment, conditional upon an existing microarray assay that was not optimized for two-stage testing, is also introduced. The results indicate that generally liberal significance levels (i.e., alpha > 0.05) could be used for gene-specific Stage 2 tests in typical studies to properly control FDR.

QpowR: Interactive power calculator for two-stage genetic association studies of quantitative traits (Steibel)
An interactive power calculator for genetic association studies of quantitative traits is described. The calculator supports two-stage designs and can help identify designs with optimal power (for a given cost) or minimum cost (for a given power). In its current version, QpowR assumes that a sample of N unrelated individuals is genotyped following a two-stage design. In a two-stage design the sample is partitioned into two subsets. On one subset, comprising a proportion of the individuals, all (M) available markers are typed (stage one) and tested for association with a quantitative trait. A fraction of the markers with stronger evidence of association to the quantitative trait are typed in the rest of the samples (stage two). Using this calculator, different design scenarios may be explored interactively. Once one scenario (defined by sample partition between stages, proportion of markers to be re-genotypes) is selected, the cost may be decreased or the power may be maximized using convenient menu options. The calculator facilitates the implementation of joint analysis by computing appropriate test statistic thresholds.

A powerful and flexible linear mixed model framework for the analysis of relative quantification RT-PCR data (Steibel)
Quantitative reverse transcription polymerase chain reaction (qRT-PCR) is currently viewed as the most precise technique to quantify levels of messenger RNA. Relative quantification compares the expression of a target gene under two or more experimental conditions normalized to the measured expression of a control gene. The statistical methods and software currently available for the analysis of relative quantification of RT-PCR data lack the flexibility and

statistical properties to produce valid inferences in a wide range of experimental situations. In this paper we present a novel method for the analysis of relative quantification of qRT-PCR data, which consists of the analysis of cycles to threshold values (C_T) for a target and a control gene using a general linear mixed model methodology. Our method allows testing of a broader class of hypotheses than traditional analyses such as the classical comparative C_T. Moreover, a simulation study using plasmode datasets indicated that the estimated fold-change in pairwise comparisons was the same using either linear mixed models or a comparative C_T method, but the linear mixed model approach was more powerful. In summary, the method presented in this paper is more accurate, powerful and flexible than the traditional methods for analysis of qRT-PCR data. This new method is especially useful for studies involving multiple experimental factors and complex designs.

Station 6: New Mexico State University (D. Cowley, M. Thomas)

Objective: Incorporate molecular and quantitative genetics into a captive propagation program for an endangered species, a nontraditional application of animal breeding and genomics. (David Cowley, Douglas Tave, Alison Hutson)
Accomplishments: The New Mexico Interstate Stream Commission (NMISC) has completed construction of a flowing water habitat for propagation of the endangered Rio Grande silvery minnow. Fish produced by the facility will be stocked into the middle Rio Grande of New Mexico to supplement the native stock of the species. In order to monitor the genetics of the captive stock, work has begun in collaboration with the NMISC personnel (Douglas Tave, Alison Hutson) to develop a large number of genetic markers that will be used in intra-generational and inter-generational monitoring of the endangered species at the facility. The goal will be to detect when significant declines in genetic diversity occur and to screen for evidence of domestication selection occurring in the stock.
Broader Impacts: This work will represent one of the first applications of genomics to captive propagation of an endangered species. It will evaluate the utility of using genetic markers to avoid unintended domestication selection occurring during captive propagation.

Objective: Identification of Molecular Markers to Improve Fertility in Beef Cattle. (Milton Thomas, Jim Reecy, Rohan Fernando, Sunday Peters)
Accomplishments: Since whole-herd data collection systems for beef cattle are slowly developing in the U.S., applied genetic research is greatly needed to develop strategies that provide enhanced economic opportunities for beef producers. A multidisciplinary team of scientist with expertise in reproductive physiology and genetics established a long-term goal to understand genetic pathways regulating reproductive performance in beef cattle, with the intent of using the information to develop gene-assisted improvement programs for fertility. Preliminary data collected by this team from beef cattle involving important reproductive hormones as well as publications involving dairy-type cattle known for twinning formulated the *hypothesis* "we can discover and test functionality of markers in chromosomes that are associated with reproductive performance in beef cattle." The markers to be evaluated will be single nucleotide polymorphisms (SNP). Access to data and DNA resources from large beef organizations is available so the hypothesis can be tested by pursuing these objectives: 1) Conduct a SNP-based whole-genome scan to identify important chromosome regions associated with heifer pregnancy rate. 2) Develop data and DNA resources from large commercial beef operations for validation and technology transfer. These resource populations will be from varied commercial production systems and environments representing the U.S. beef industry. These resources allow expansion of the research efforts to include other economically relevant reproductive traits such as heifer rebreeding rate and stayability. Project also being expanded to evaluate hypothalamic-transcriptome sequence among pre- and post-pubertal Brangus heifers (collaboration with NCGR-

Santa Fe, NM). We have completed 12 months of the project and have obtained Illumina-bovineSNP50 data from 835 Brangus heifers. We have also collected DNA-blood-cards from >10,000 heifers representing 19 Agricultural Experiment Stations and varied beef cattle production systems across the U.S. We will be working in 2009 to analyze data and build a database from the blood cards and phenotypes.

Broader Impacts: Project will enhance understanding of the genetic pathways regulating reproductive performance in beef cattle, with the intent of using the information to develop gene-assisted improvement programs for fertility. Development of resources within this project allows expansion of the research efforts to include other economically relevant reproductive traits such as heifer rebreeding rate and stayability. Project has also expanded to include a transcriptome sequencing effort for the hypothalamus, which is a key tissue regulating the reproductive endocrine axis. This resource will expand the project into the realm of gene discovery.

Station 7: Purdue University (W. Muir)

1. **PRINCIPAL LEADER:** W.M. Muir

2. Collaborators within NCR-204: Shizhong Xu, Guilherme Rosa, IgnacyMiztal, Frank Siewerdt

3. **PROGRESS OF WORK AND PRINCIPAL ACCOMPLISHMENTS:**

Objective 2: Examine the efficiency of incorporating molecular tools in breeding programs through theoretical modeling, computer simulations, and biological testing in actual breeding populations.

A. A mixture model approach for the analysis of small exploratory microarray experiments s. In collaboration with Guilherme Rosa and Shizhong Xu

The microarray is an important and powerful tool for prescreening of genes for further research. However, alternative solutions are needed to increase power in small microarray experiments. Use of traditional parametric and even non-parametric tests for such small experiments lack power and have distributional problems. A mixture model was developed that is performed directly on expression differences assuming that genes in alternative treatments are expressed or not in all combinations (i) not expressed in either condition, (ii) expressed only under the first condition, (iii) expressed only under the second condition, and (iv) expressed under both conditions, giving rise to 4 possible clusters with two treatments. The approach is termed a Mean-Difference-Mixture-Model (MD-MM) method. Accuracy and power of the MD-MM was compared to other commonly used methods, using both simulations, microarray data, and quantitative real time PCR (qRT-PCR). The MD-MM was found to be generally superior to other methods in most situations. The advantage was greatest in situations where there were few replicates, poor signal to noise ratios, or non-homogenous variances.

B. Use of Functional Genomics to Understand Physiological Mechanisms Associated with Aggression and Stress in Poultry. In collaboration with Heng-Wei Cheng and Anthony Giannani

.

Animal well-being is important in all aspects of agriculture. Some practices used in the poultry industry, such as high densities, low light levels, and beak trimming, may result in stress for the animal. The KGB and DXL bird lines differ greatly in levels of stress response and aggression in social environments. In this study, we examined the gene profiles of these two bird

lines to uncover genes that are differentially expressed. The objective of this study was to identify genes responsible for innate behavioral differences between the lines when a stressor, transportation, was placed on chickens. We assessed gene expression profiles from the hypothalamus using the Affymetrix GeneChip chicken genome array to identify genes involved in behavior. About 3,500 genes were found to be of significant difference between the two lines. Using Gene Ontology (GO) analysis, the 3,000 most significant genes were grouped together into categories based on biological processes. A highly significant node classified as sensory perception to smell was identified and deemed relevant for confirmation. Quantitative real-time polymerase chain reaction was used to verify the given genes. Significant differentially expressed genes in this node were DNAJC17, Cor4, OR5AS1 (262), and OR5AS1 (141). All genes except one, OR5AS1 (262) were verified to be significant. The ontology for each of the significantly differentially expressed genes has not as yet been proven but inferred by association. GO molecular function terms of these genes include rhodopsin-like receptor activity, signal transducer activity, and olfactory receptor activity.

C. Genomic Selection. In collaboration with Hans Cheng (ARS) Martien Groenen and co-workers (The Netherlands), Theo Meuwissen (Norway), Mario Calus and John Bastiaansen (The Netherlands), Gane Wong (U. of Alberta), Cobb Vantress (AR) - Albert Paszek and Ron Okimoto, Hendrix Genetics (The Netherlands) - Gerard Albers, Addie Vereijken, Annemieke Jungerius, DNA Landmarks (Canada) - Charles Pick and Tun-Ping Yu

Genome-wide marker-assisted selection (GWMAS) utilizes markers spanning the entire genome to increase accuracy and efficiency of estimating breeding values (EBV). We will evaluate the power of GWMAS in a multi-generational selection experiment. We will utilize 2 broiler and 3 layer lines with different selection objectives and traits with low to high heritability and varying extent of LD. The first step was to develop a 60k SNP chip which has now been completed. The next step is to collect genotypes and actually implement GWMAS.

4. Work planned for next year.

A. In collaboration with Hans Cheng and Martien Groenen, we will utilize allele specific expression to determine cis- and trans-acting elements affects MD resistance.

B. Mapping Quantitative Trait Loci for Growth in Chickens Paul Siegel at Virginia Polytech Institute have over the past 50 years conducted a bi-directional selection experiment for body weight in chicken. As a first step towards identifying the genetic mechanisms underlying the observed selection response, chickens from generation 43 of the high- and low- body weight lines were used as founders for an F_2 intercross for QTL mapping (Jacobsson *et al.*, 2006). In this intercross, a radial network of 4 interacting QTL was identified and predicted to have caused nearly half of the nearly nine-fold difference in body weight between the selected lines (Carlborg *et al.*, 2006). This is the first empirical evidence for how interactions in gene networks cause response to selection, release of standing genetic variation and bias estimates marginal effects estimated for individual QTL. The Virginia lines are thus a fantastic resource for further basic research to understand the complex genetics underlying phenotypic evolution and genetic evolvability. As the high- and low- body weight lines also display large phenotypic differences in a wide range of metabolic as well as immunological traits, they are also a valuable resource as a model for studying the genetics underlying important traits in agriculture as well as human and veterinary medicine. With the funding that ÖC received from one of the 2006 EURYI Awards, an effort will be made to develop a novel resource population – a 4 QTL introgression line - based

on the original Virginia selection lines to facilitate in-depth studies of the gene interaction network.

Station 8: Virginia Tech and State University (R. Lewis)

SOLVING AN EDUCATIONAL DILEMMA THROUGH COLLABORATION: A GRADUATE DISTANCE-LEARNING CURRICULUM IN ANIMAL BREEDING AND GENETICS
Ron Lewis

Diminishing numbers of faculty in the area of animal breeding and genetics have reduced opportunities for specialized coursework that would prepare future scholars in the field. Existing faculty must often focus on addressing basic content, limiting their ability to offer specialized instruction that could serve to develop higher-level skills and knowledge in their students. As a result, graduate students in this field are often not able to access necessary curricula that would build their own expertise and abilities in animal breeding and genetics.

A consortium of universities, led by Virginia Tech, is combining efforts to address this challenge by developing a series of online courses designed to supplement graduate-level instruction for existing degree programs in animal breeding and genetics. Faculty from Virginia Tech, Colorado State, Michigan State, and Cornell University are working together to create a curriculum that may be taken by students at universities and colleges across the United States. This project is funded by the Higher Education Challenge Grant Program in the U.S. Department of Agriculture Cooperative State Research, Education, and Extension Service, and began in August 2007.

Three of the seven planned courses have been developed and implemented, with students from 18 land grant universities participating. Each course enjoyed an enrollment of approximately 25 students. An eighth course with the aim of providing foundational knowledge in the field of quantitative genetics is also being added to the curriculum. Student feedback from the courses has been quite positive. One student, referring to the first course on matrix algebra, stated, "I thought the class was very organized and structured. Expectations were clearly outlined. The subject material was comprehensive, yet intuitive and understandable. Very useful class overall".

Station 9: University of Wisconsin (G. Rosa)

1. PRINCIPAL LEADER: Guilherme J. M. Rosa

WI Collaborating Faculty: Daniel Gianola, Kent Weigel, Hasan Khatic and Brian Kirkpatrick

WI Students and Staff: Ana Vazques, Gustavo de los Campos, Nick Wu, Oscar Gonzalez-Recio, Nanye Long

Collaborators in NCCC204: Bill Muir, Juan Pedro Steibel

2. PROGRESS OF WORK AND PRINCIPAL ACCOMPLISHMENTS

Objective 1: Develop and compare statistical and computational methodology for analysis of molecular genetic and genomic data associated with quantitative traits.

A. Mixture Model Approach for the Analysis of Microarray Data

Use of traditional parametric and even non-parametric tests for small microarray gene expression experiments lack power and have distributional problems. We describe a mixture model that is performed directly on expression differences assuming that genes in alternative treatments are

expressed or not in all combinations (i) not expressed in either condition, (ii) expressed only under the first condition, (iii) expressed only under the second condition, and (iv) expressed under both conditions, giving rise to 4 possible clusters with two treatments. We term this approach a Mean-Difference-Mixture-Model (MD-MM) method. Accuracy and power of the MD-MM was compared to other commonly used methods, using both simulations, microarray data, and quantitative real time PCR (qRT-PCR). The MD-MM was found to be generally superior to other methods in most situations. The advantage was greatest in situations where there were few replicates, poor signal to noise ratios, or non-homogenous variances.

B. Optimal Designs for Genetical Genomics Studies with Outbred Populations

Genetic analysis of transcriptional profiling experiments is emerging as a promising approach for unraveling genes and pathways that underlie variation of complex biological traits. However, these genetical genomics approaches are currently limited by the high cost of microarrays. We studied five different strategies to optimally select subsets of individuals for transcriptional profiling, including 1) maximizing genetic dissimilarity between selected individuals, 2) maximizing the number of recombination events in selected individuals, 3) selecting phenotypic extremes within inferred genotypes of a previously identified quantitative trait locus (QTL), 4) purely random selection and 5) profiling animals with the highest and lowest phenotypic values within each family-gender subclass. A simulation study was conducted based on a linkage map and marker genotypes derived from data on Chromosome 6 for 510 F2 animals from an existing pig resource population and on a simulated biallelic QTL with pleiotropic effects on performance and gene expression traits. Bivariate analyses were conducted for selected subset sample sizes of 80, 160 and 240 individuals under three different correlation scenarios between the two traits. The genetic dissimilarity and phenotypic extremes within genotype methods had the smallest mean square error on QTL effects and maximum sensitivity on QTL detection, thereby outperforming all other selection strategies, particularly at the smallest proportion of samples selected for gene expression profiling (80/510).

C. Two-Stage Designs for Gene Expression Assays

Gene expression microarrays are powerful tools for simultaneously screening the transcriptional profile for thousand of genes across different treatments. Despite their continually improving sensitivity and dynamic range, microarrays are commonly regarded as a first screening step, with precision often deemed unacceptable to use as a standalone technology. This limitation has prompted genomics researchers to validate a statistically significant subset of their microarray results using a second technique, typically quantitative reverse transcription polymerase chain reaction (qRT-PCR). The problem of optimizing such two stage transcriptional profiling experiments in order to maximize sensitivity while controlling the false discovery rate (FDR) is addressed. This optimization is based on partitioning the set of available biological replicates into two groups, one for each of the microarray (Stage 1) and qRT-PCR (Stage 2) experiments. It is demonstrated how the significance level can be determined for Stage 2, after selecting a fixed percentage of the genes to validate from Stage 1, in order to maximize the sensitivity of detection of differentially expressed genes for a desired overall FDR. The results indicate that most of the available replicates (typically > 60%) should be consumed in Stage 1. Even though the optimization scheme assumes independent genes and known variances, it is demonstrated with simulation studies that this approach is robust to moderate departures from those assumptions. It is also demonstrated the problem of optimizing a validation experiment, conditional upon an existing microarray assay that was not optimized for two-stage testing. The results demonstrate that generally liberal significance levels (i.e., α>0.05) for gene-specific Stage 2 tests could be used to properly control FDR in typical studies.

D. Non-Parametric Aproach for Incorporating Genomic Information into Genetic Evaluations

Four approaches using SNP information (F-infinity-metric model, kernel regression, reproducing kernel Hilbert spaces (RKHS) regression, and a Bayesian regression) were compared with a standard procedure of genetic evaluation (E-BLUP) of sires using mortality rates in broilers as a response variable, working in a Bayesian framework. Late mortality (14-42 days of age) records on 12,167 progeny of 200 sires were precorrected for fixed and random (nongenctic) effects used in the model for genetic evaluation and for the mate effect. The average of the corrected records was computed for each sire. Twenly-four SNPs seemingly associated with late mortality were included in three methods used for genomic assisted evaluations. One thousand SNPs were included in the Bayesian regression, to account for markers along the whole genome. The posterior mean of heritability of mortality was 0.02 in the E-BLUP approach, Suggesting that genetic evaluation could be improved if suitable molecular markers were available. Estimates of posterior means and standard deviations of the residual variance were 24.38 (3.88), 29.97 (3.22), 17.07 (3.02), and 20.74 (2.87) for E-BLUP, the linear model on SNPs, RKHS regression, and the Bayesian regression, respectively, suggesting that RKHS accounted for more variance in the data. The two nonparametric methods (kernel and RKHS regression) fitted the data better, having a lower residual sum of squares. Predictive ability, assessed by cross-validation, indicated advantages of the RKHS approach, where accuracy was increased from 25 to 150%, relative to other methods.

E. Marker-Assisted Assessment of Genotype by Environment Interaction

Interplay between genetic and environmental factors, genotype x environment interactions (G x E), affect phenotypes of complex traits. A methodology for assessing G x E was investigated by detecting hygiene (low and high) environment-specific SNP subsets associated with broiler chicken mortality, followed by an examination of consistency between SNP subsets selected from the 2 environments. The trait was mean progeny mortality rate in 253 sire families, after adjusting records for nuisance effects affecting mortality at the individual bird level. Over 5,000 whole-genome SNP were narrowed down via a machine-learning (filter-wrapper) feature selection procedure applied to mortality rates in each of the 2 environments. For both early and late mortality, it was found that the selected SNP subsets differed across hygiene environments, in terms of either across-environment predictive ability or extent of linkage disequilibrium between the subsets. Reduction in predictive ability due to G x E was assessed by the ratio of 2 predicted residual sum of squares statistics, one associated with SNP selected from the same hygiene environment and the other associated with the SNP subset from a different environment. Reduction was 30 and 20% for early and late mortality, respectively. An extremely low level of linkage disequilibrium between SNP subsets selected under low and high hygiene also indicated G x E. Findings suggest that there may not be a universally optimal SNP subset for predicting mortality and that interactions between genome and environmental factors need to be considered in association analysis of complex traits.

3. WORK PLANNED FOR NEXT YEAR

Continue work on research related to items D and E above, as well as on another project aiming the integration of gene sequence information into the statistical analysis of microarray data.

LIST OF REFEREED AND INVITED PUBLICATIONS FOR NCCC204

Abasht, B., Beach, E., Arango, J.A., Settar, P., Fulton, J.E., O'Sullivan, N.P., Fernando, R.L., Dekkers, J.C.M., and Lamont, S.J. 2008. SNP Markers for production and egg quality QTLs in layer lines. Proc. Conf Avian Genomics, Mississippi State Univ, Starkville, MS

Aguilar, I., and I. Misztal. 2008. Recursive algorithm for inbreeding coefficients assuming non-zero inbreeding of unknown parents. J. Dairy Sci. 91:1669-1672.

Beach, E., Andreescu, C., Abasht, B., Lamont, S.J., Fulton, J., O'Sullivan, N. and Dekkers, J.C.M. 2008. Level and consistency of linkage disequilibrium in commercial layer chicken populations. Proc. Plant & Animal Genome XVI, San Diego, CA.

Bohmanova,J., F. Miglior, J. Jamrozik, I. Misztal and P.G. Sullivan. 2008. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. J. Dairy Sci. 91:3627-3638.

Cai, W., D.S. Casey, and J.C.M. Dekkers. 2008. Selection response and genetic parameters for residual feed intake in Yorkshire swine. J. Anim. Sci. 86:287-298.

Cardoso, F. F., Rosa, G. J. M., Steibel, J. P., Ernst, C. W., Bates, R. O. and Tempelman, R. J. (2008) Selective transcriptional profiling and data analysis strategies for eQTL mapping in outbred F2 populations. Genetics 180(3): 1679-1690.

Fontoura Júnior, J.A., F.Siewerdt, N.J.L.Dionello, M.N.Corrêa, 2009. Theorerical model of reproductive performance of beef cows based on body condition score. *Revista Brasileira de Zootecnia*. (in press)

Gonzalez-Recio, O., Gianola, D., Rosa, G. J. M., Weigel, K. A. and Kranis, A. (2009) Genome-assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. Genetics Selection Evolution (accepted)

Gonzalez-Recio, O., Gianola, D., Long, N., Weigel, K. A., Rosa, G. J. M. and Avendano, S. (2008) Nonparametric methods for incorporating genomic information into genetic evaluations: an application to mortality in broilers. Genetics 178(4): 2305-2313.

Groenen, MAM, Per Wahlberg, HH Cheng, Hendrik-Jan M, R Crooijmans, I Gut, M Latrop, Bill Muir, GKS Wong and L Andersson. 2008. A second generation consensus linkage map of the chicken genome. Genome Research (In Press)

Guo, X., S.G. Gordon, D. Wang, E. Helliwell, T. Smith, S.A. Berry, S.K. St. Martin, and A.E. Dorrance. 2008. Genetic mapping of QTLs underlying partial Sclerotinia stem rot resistance in soybean PI 391589A and PI 391589B. Crop Sci. 48:1129-1139

Habier, D., J.C.M. Dekkers, and R.L. Fernando. 2008. Genomic selection using low-density SNPs. Abstract 521. ASAS/ADSA annual meeting, Indianapolis.

Habier, D., J.C.M. Dekkers, and R.L. Fernando. 2008. Genomic selection using low-density SNPs. QTL-MAS workshop. Uppsala, Sweden.

Habier, D., R. L. Fernando and J. C. M. Dekkers. 2008. A Mixture Genetic Model for whole-genome analyses. 59th Annual meeting of the European Association of Animal Production, Vilnius, Lithuania. 2008.

Hasenstein, J.R., Hassen, A., Dekkers, J.C.M., and Lamont, S.J. 2008. Whole-genome QTL analysis of response to *Salmonella* in advanced intercross line chicks. Proc. Plant & Animal Genome XVI, San Diego, CA.

Hassen, A., S. Avendano, W.G. Hill, R.L. Fernando, S.J. Lamont, and J.C.M. Dekkers. 2008. The effect of heritability estimates on high-density SNP analyses with related animals. J. Anim. Sci. 21: (on line).

Ibáñez-Escriche N., R. L. Fernando, J. C. M. Dekkers. 2008. Genomic selection of purebreds for crossbred performance using a model with breed-specific SNP effects. 58th Annual meeting of the European association for animal production. Vilnius, Lituania.

Li, Y., H.N. Kadarmideen, and J.C.M. Dekkers. 2008. Selection on Multiple QTL with Control of Gene Diversity and Inbreeding for Long-term Benefit. J. Anim. Breed. Genet. 125: 320-329.

Long, N., Gianola, D., Rosa, G. J. M., Weigel, K. A. and Avendano, S. (2008) Marker-assisted assessment of genotype by environment interaction: A case study of single nucleotide polymorphism-mortality association in broilers in two hygiene environments. Journal of Animal Science 86: 3358-3366.

Misztal, I. 2008. Reliable Computing in Estimation of Variance Components. J. Anim. Breed. Genet. 125:363-370.

Muir, W.M., G.K. Wong, Y. Zhang, J. Wang, M.A.M. Groenen, R.M.A. Crooijmans, H.-J. Megens, H.M. Zhang, J.C. Mckay, S. Mcleod, R Okimoto, J.E. Fulton, P. Settar, N.P. O'sullivan, A. Vereijken, A. Jungerius-Rattlnk, G.A.A. Albers, C. Taylor-Lawley, M.E. Delany and H.H. Cheng. 2008. Review of the initial validation and characterization of a 3K chicken SNP array. World's Poultry Science Journal 64: 219-226.

Muir, W.M. , G.J.M. Rosa, B.R. Pittendrigh, S. Xu, S.D. Rider, M. Fountain and J. Ogas 2008. A mixture model approach for the analysis of small exploratory microarray experiments. J. Computational Statistics and Data Analysis. In Press doi:10.1016/j.csda.2008.06.011

Muir, WM., GKS Wong, Y Zhang, J Wang, MAM Groenen, RPMA Crooijmans, HJ Megens, H Zhang, R Okimoto, A Vereijken, A Jungerius, GAA Albers, C Taylor-Lawley, ME Delany, S MacEachern, and HH Cheng. 2008. Genome-Wide Assessment of World-Wide Chicken SNP Genetic Diversity Indicates Significant Absence of Rare Alleles in Commercial Breeds. Proceeds of the National Academy of Sciences 105:17312-17317

Olmstead, J.W., A.M. Sebolt, A. Cabrera, S.S. Sooriyapathirana, S. Hammar, G. Iriarte, D. Wang, C.Y. Chen, E. van der Knaap, and A.F. Iezzoni. 2008. Construction of an intra-specific sweet cherry (Prunus avium L.) genetic linkage map and synteny analysis with the Prunus reference map. Tree Genet. Genomes 4:897-910.

Povh, JA, RP Ribeiro, RN Sirol, DP Streit Jr., HLM Moreira, F Siewerdt, NM Lopera-Barrero, CA Mangolin, L Vargas, 2009. Microsatellite analysis of the parental contribution of *Piaractus mesopotamicus* to the production of offspring in the semi-natural system of reproduction**.** Brazilian Archives of Biology and Technology **(in** press).

Pribyl, J., H. Krejcova, J. Pribylova, I. Misztal, S. Tsuruta, and N. Mielenz. 2008. Models for evaluation of growth of performance tested bulls. Czech J. Anim. Sci. 53:45-54.

Pyiasatian, N., R.L. Fernando, and J.C.M. Dekkers. 2008. Introgressing multiple QTL in breeding programs of limited size. J. Anim. Breed. Genet. 125: 50-56.

Qu L., Nettleton D., Dekkers J.C.M., 2008. Improved Estimation of False Discovery Rates Based on Subsampling with Applications to Microarray Data. Spring Meeting of Eastern Eastern North American Region/International Biometric Society. Arlington, Virginia, USA

Sánchez, J. P., I. Misztal, and J. K. Bertrand. 2008. Evaluation of methods for computing approximate accuracies in maternal random regression models for growth trait in beef. J. Anim. Sci. 86:1057-1066.

Schmidt, C.J., Michael Romanov, Oliver Ryder, Vincent Magrini, Matthew Hickenbotham, Jarret Glasscock, Sean McGrath, Elaine Mardis, and Lincoln D. Stein (2008) Gallus GBrowse: a unified genomic database for the chicken. Nucleic Acids Res. 2008 January; 36(Database issue): D719–D723.

Siewerdt, F., A.K. Sasikala-Appukuttan, 2008. Genetic variation in fecal nitrogen and phosphorus levels in broilers: opportunity for improving both birds and the environment. Proceedings of the 57th National Breeders Roundtable, St. Louis, MO. CD Proceedings, p.377-391. (Invited)

Steibel, J. P., Rosa, G. J. M. and Tempelman, R. J. (2008) Optimizing design of two-stage experiments for transcriptional profiling. Computational Statistics and Data Analysis, 2008 (in press).

Tudor, C.O., Carl J Schmidt, and K Vijay-Shanker, Mining for Gene-Related Key Terms: Where Do We Find Them?, Third International Symposium on Semantic Mining in Biomedicine SMBM (2008), Turku Finland.

Tudor, C.O., K Vijay-Shanker, and Carl J Schmidt, Mining the Biomedical Literature for Genic Information,(2008) BioNLP Columbus Ohio,

Tsuruta, S. and I. Misztal. 2008. Computing options for genetic evaluation with a large number of genetic markers. J. Anim. Sci. 86:1514-1518.

Uthe, J.J., Y. Wang, S.M.D. Bearson, A.M. O'Connor, J. McKean, D. Nettleton, J.C. Dekkers, and C.K. Tuggle. 2008. Identifying genes associated with bacterial shedding through transcriptional profiling. Abstract 41, Midwest Animal Science Meeting, Des Moines.

Van Eenennaam, A. and W.M. Muir 2008. Animal Biotechnologies And Agricultural Sustainability. In The Role of Biotechnology in a Sustainable Food Supply. Eds (in press)

Wang, D., and D. Grant. 2009. Molecular mapping of quantitative trait loci (QTLs) in soybean. p. (in press). In M. Ratnaparkhe and K. Bilyeu, eds. Soybean Genomics.

Weebadde, C.K., D. Wang, C.E. Finn, K.S. Lewers, J.J. Luby, J. Bushakra, T.M. Sjulin, and J.F. Hancock. 2008. Using a linkage mapping approach to identify QTL for day-neutrality in the octoplid strawberry (Fragaria ×ananassa Duch ex Rozier). J. Plant Breed. 127:94-101.

Wensheng Zhang , J. Song, 2008 Term-Tissue Specific Models for Prediction Of Gene Ontology Biological Processes Using Transcriptional Profiles of Aging in D. melanogaster. BMC Bioinformatics 2008, 9:129doi:10.1186/1471-2105-9-129

Wensheng Zhang , Hong-Bin Fang, J. Song, 2008 Principal component tests: applied to temporal gene expression data. BMC Bioinformatics 2008, 10 (Suppl 1):S26. doi: 10.1186/1471-2105-10-S1-S26

Wiggans, G. R., S. Tsuruta, and I. Misztal. 2008. Adaptation of an Animal-Model Method for Approximation of Reliabilities to a Sire-Maternal Grandsire Model. J. Dairy Sci. 91:4058-4061.

Young, J.M., H. Zhao, R.L. Fernando, D. Habier, and J.C.M. Dekkers. 2008. Responses to genomic selection in livestock. Abstract 35, Midwest Animal Science Meeting, Des Moines.

Yu Y. H. Zhang, L. Bacon, F. Tian, J. Song 2008 Methylation Pattern Variations of Global ALVE and TVB gene in a Neoplastic Disease Susceptible and Resistant Chicken Model: A Quantitative Evaluation by Pyrosequencing PLoS ONE 3(3): e1731. doi:10.1371/journal.pone.0001731

Yu Y. H. Zhang, F. Tian, W. Zhang, J. Song 2008 An Integrated Epigenetic and Genetic Analysis of DNA Methyltransferase Genes (DNMTs) in Tumor Resistant and Susceptible Chicken Lines. PLoS ONE. 2008; 3(7): e2672. Published online 2008 July 16. doi: 10.1371/journal.pone.0002672.

Zhang, G., A.M. Sebolt, S.S. Sooriyapathirana, D. Wang, M.C. Bink, J.W. Olmstead, and A.F. Iezzoni. 2009. Fruit size QTL analysis of an F_1 population derived from a cross between a domesticated sweet cherry cultivar and a wild forest sweet cherry. Tree Genetics and Genomes:(in review).

Zhang, G., C. Gu, and D. Wang. 2009. A novel locus for soybean aphid resistance. Theor. Appl. Genet.:(in review).

Zhang, G., C. Gu, and D. Wang. 2009. Molecular mapping of soybean aphid resistance genes in PI 567541B. Theor. Appl. Genet. 118:473-482.

Zhang, B., P. Chen, A. Shi, A. Hou, T. Ishibashi, and D. Wang. 2009. Putative Quantitative trait loci (QTL) associated with calcium content in soybean seed. J. Hered. 100: 263-269.

Zhang, B., P. Chen, C.Y. Chen, D. Wang, A. Shi, and A. Hou. 2008. Mapping quantitative trait loci for seed hardness in soybean. Crop Sci. Crop Sci. 48:1341-1349.

Zumbach, B., I. Misztal, S. Tsuruta, J. P. Sanchez, M. Azain, W. Herring, J. Holl, T. Long, and M. Culbertson. 2008. Genetic components of heat stress in finishing pigs: Development of a heat load function. J. Anim. Sci. 86: 2082-2088.

Zumbach, B., I. Misztal, S. Tsuruta, J. P. Sanchez, M. Azain, W. Herring, J. Holl, T. Long, and M. Culbertson. 2008. Genetic components of heat stress in finishing pigs: Parameter estimation. J. Anim. Sci. 86: 2076-2081.

End of NCCC204 Report.▫