

## Breeding and Genetics

**T13 Using logistic regression model to analyse some environmental factors affecting mastitis incidence of primiparous Iranian Holsteins.** H. Farhangfar\*, A. Abedini, H. Naeemipour, M. R. Asghari, and M. H. Fathi Nasri, *Birjand University, Birjand, Iran.*

To analyze the effects of some environmental factors such as herd, year, season of calving, stage of lactation on mastitis incidence, a total of 101147 monthly test day somatic cell counts collected from 13977 Iranian Holstein heifers calved between 2002 and 2006 in 183 herds were used. The event of mastitis in cows was coded as 1 if their somatic cell count at specific test day was greater than 250000 cells per milliliter of milk yield sample. Although somatic cell score could be also analyzed as a continuous trait, this study aimed to evaluate mastitis incidence in a new approach. Logistic regression statistical method was applied to model the probability of mastitis incidence. In the model, fixed environmental factors of herd, year, season of calving, stage of lactation, milk (linear covariate) and age at first calving (linear covariate) were included. The model was fitted to the data using logistic procedure of SAS programme. The results obtained in this study indicated that all environmental factors had highly statistical significant ( $p < 0.001$ ) effects on mastitis incidence. Spring calvers (April-June) were found to be more susceptible to mastitis than the other cows calving at different seasons suggesting that calving in the seasons other than spring could result in decreasing mastitis in herds. The results also revealed that the probability of mastitis incidence increased (with odds ratio of 1.029) as the age of cow at first calving increased while it decreased when the amount of milk yield (odds ratio of 0.960) during the lactation period increased. Genetic evaluation of animals for somatic cell score could decrease the level of mastitis incidence.

**Key Words:** Mastitis, Logistic Regression, Iranian Holsteins

**T14 Genetic parameters estimation of somatic cell score in Iranian Holstein heifers using a random regression test day model.** H. Farhangfar\*<sup>1</sup>, A. Abedini<sup>1</sup>, H. Naeemipour<sup>1</sup>, M. Alipanah<sup>2</sup>, K. Shojaeian<sup>2</sup>, and B. Mohammad Nazari<sup>3</sup>, <sup>1</sup>*Birjand University, Birjand, Iran,* <sup>2</sup>*Zabol University, Zabol, Iran,* <sup>3</sup>*Animal Breeding Centre, Karaj, Iran.*

A total of 101,147 monthly test day somatic cell scores (SCS, defined as  $\ln(\text{SCC} \times 10^{-3})$ ) belonging to 13,977 Iranian Holstein heifers calving during 2002-2006 and distributed in 183 herds was used to estimate genetic parameters. A random regression test day model was used to estimate genetic and environmental variance and covariance components. In the model, fixed environmental factors of herd-year-season of production-province, age of cow at recording, Holstein gene percentage, stage of lactation, and random direct additive genetic and environmental effects were included. Orthogonal legendre polynomials of order four (cubic) was applied to take account of the variation of somatic cell score during the lactation course at both genetic and environmental levels. The results found in the present research showed that heritability estimates of SCS ranged from 0.03 (month 2) to 0.07 (month 10) and that the second half of the lactation curve was more heritable than the first half. Phenotypic correlations among SCC at different lactation months were generally lower than the genetic correlations. Genetic correlations among adjacent lactation months were always greater than those obtained among months apart from. The same patterns were also observed for phenotypic and environmental correlations.

**Key Words:** Somatic Cell Score, Test Day Model, Iranian Holsteins

**T15 Genetic parameters and trend estimation for milk and fat yields and fat percentage for primiparous Holstein population of Golestan and Mazandaran provinces of Iran using a univariate animal model.** H. Naeemipour\*<sup>1</sup>, H. Farhangfar<sup>1</sup>, I. Tahmasbi<sup>2</sup>, and M. Bashtani<sup>1</sup>, <sup>1</sup>*Birjand University, Birjand, Iran,* <sup>2</sup>*Zabol University, Zabol, Iran.*

The main objective of the present study was genetic analysis of milk production traits for Holstein heifers in northern provinces of Iran. A total of 7844 first lactation 305-day and 2X milk and fat yields and fat percentage records from 7844 Holstein cows distributed in 139 herd and calved between 1987 and 2003 were used to estimate variance components implementing univariate animal models using DFREML approach. In the model, fixed effect of herd-year-season of calving, age at first calving as well as random effect of additive genetic effect were included. The results obtained in this study indicated that heritabilities for milk yield, fat yield and fat percentage were 0.20, 0.09 and 0.07 respectively. low estimates of fat yield and fat percentage shows a high environmental variance indicating that little genetic gain could be obtained as the genetic selection is based on these traits. Annual genetic trends were estimated to be 3.4, 0.05 Kg/year and 0.0003% respectively. Positive genetic gain obtained for fat percentage indicates that the average predicted breeding value of animals increased for fat percentage along with milk yield over the period of time.

**Key Words:** Genetic Parameters, Holstein, Golestan and Mazandaran

**T16 Genetic relationships between linear type traits, somatic cell score and longevity in Holstein cows of Iran.** M.R. Bakhtiarizadeh\*, M. Moradi Shahr Babak, and A. Pakdel, *University of Tehran, Tehran, Iran.*

The objective of the present study was to estimate the genetic parameters for 13 linear type traits, somatic cell score (SCS) and longevity in Holstein population of Iran. Two set of data including 3000 records (for SCS) and 12226 observations (for longevity) for first lactation records on cows distributed across 219 and 1500 herd-year-season groups respectively were investigated. The genetic parameters were estimated by ASREML software and the fix effects of models were included herd-year-season (in calving year), age in calving, age\*age, days in milk and milk (for SCS). Heritability estimates for the type traits ranged from 0.033 to 0.29. The genetic correlation among type traits ranged from 0.01 to 0.83; among type traits and longevity ranged from 0.41 to -0.5 and among type traits and SCS ranged from 0.85 to -0.6. The genetic correlation between longevity and SCS were negative (-0.38) so that cows with more longevity had a lower SCS. The genetic correlation between body size traits and SCS were positive and between body size traits and longevity were negative. Therefore from genetic point of view we can conclude that cows with higher body size have higher SCS and lower longevity. The results of current study showed more attention should be paid to these type traits in the breeding programs for better animal welfare and to get more profit.

**Key Words:** Type Traits, Longevity, SCS

### **T17 Breed composition of the United States dairy cattle herd.**

R. L. Powell\*, H. D. Norman, and J. L. Hutchison, *Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD.*

Breed composition of the gene pool of all cows (purebred and crossbred) with pedigree data in the USDA national dairy database was summarized by birth year of cow. Partial breed contributions were assigned for individual cows. For cows born in 2005, 1.1% of all genes and 35.1% of genes in crossbreds traced to a female ancestor with breed reported as unknown; i.e., a dam without reported identification information. When a cow and her sire had the same reported breed but her dam's breed was unknown, the sire's breed was assigned to the dam, which decreased the percentage of unknown genes to 0.7 for all cows born in 2005 and to 6.7 for crossbreds. The percentage of the national herd that was crossbred increased from 0.4 for cows born in 1990 to 0.7 in 2000 and 1.6 in 2005. Since 2000, the proportion of genes from Brown Swiss, Jersey, Milking Shorthorn, and nontraditional US dairy breeds has increased, while the proportion from Ayrshire, Guernsey, and Holstein breeds decreased. For cows born in 2005, genetic composition was 0.4% Ayrshire, 1.0% Brown Swiss, 0.4% Guernsey, 90.8% Holstein, 6.5% Jersey, 0.2% Milking Shorthorn, 0.1% other breeds, and 0.7% unknown. Corresponding composition for crossbreds was 2.4, 9.2, 1.1, 44.0, 25.9, 9.3, 1.4, and 6.7%. The most frequent sire breed for crossbreds was Holstein until birth year 1999 and Jersey since then. Frequency of sire breeds for crossbreds born in 2005 was 42% Jersey, 27% Holstein, 13% each for Brown Swiss and Milking Shorthorn, and 5% for all other breeds. About 95% of all first-generation crossbred cows were mated to bulls of one of the crossbred's parental breeds, most frequently the sire breed. Tracing an animal's genetic background rather than relying on its coded breed provides a more complete and accurate representation of the extent of crossbreeding and changes in the genetic composition of the national dairy herd.

**Key Words:** Dairy Cattle Breed, Crossbred, National Dairy Herd

### **T18 Reproductive trends of dairy herds in the United States.**

H. D. Norman, J. R. Wright, S. M. Hubbard\*, M. T. Kuhn, and R. H. Miller, *Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD.*

Trends for reproductive traits were examined for U.S. Holstein and Jersey herds with records in the USDA national dairy database. Traits were days from calving to first service (DFS) and to last service, 70-d nonreturn rates (NRR) and conception rates (CR) for first through fifth services, days open, gestation length, and services per lactation for breedings from 1996 through 2005. Holstein DFS increased from 89 in 1996 to 92 in 2001 and then declined to 86 in 2005; no trend was observed for Jersey DFS. First-service NRR declined from 54 and 57% in 1996 to 46 and 52% in 2005 for Holsteins and Jerseys, respectively; first-service CR declined from 36 and 39% to 30 and 35%. As expected from the NRR decline, services per lactation increased from 2.1 in 1996 to 2.6 in 2005 for Holsteins and from 2.0 to 2.4 for Jerseys. Days to last service also increased by 16 d for both breeds. Days open increased during early years but have stabilized for both breeds. Gestation length showed no change across time for either breed. Data from 2005 breedings were examined for the same traits by parity for both breeds and by geographic region for Holsteins. Later parities (>5) were associated with longer DFS, lower first-service NRR, and increased first-service CR compared with early parities (1 and 2) for both breeds; services per lacta-

tion remained fairly constant for Holsteins and increased slightly with parity for Jerseys. For both breeds, NRR for first and second services and CR for each service through fourth declined across parities. Southwest Holsteins had the fewest DFS (73 d) and lowest first-service NRR (36%); the Mountain region had the greatest DFS (93 d) and the highest NRR (48%). Southeast Holsteins had the lowest first-service CR (23%) and the most services per lactation (3.0); the Northeast had the highest CR (31%) and the fewest services (2.6). Trends likely were impacted by producer preference for animal age for herd retention and increased use of estrous synchronization and timed artificial insemination.

**Key Words:** Reproductive Trend, Conception Rate, Nonreturn Rate

### **T19 Impact of selection for increased daughter fertility on productive life and culling for reproduction.**

H. D. Norman, J. R. Wright\*, and R. H. Miller, *Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD.*

Selection for increased daughter pregnancy rate (DPR) over 2 generations was examined to determine if such selection had affected cow fertility and productive life (PL). Holstein artificial-insemination bulls with a predicted transmitting ability (PTA) for DPR based on  $\geq 35$  daughters were grouped by quintile based on PTA DPR. Then 25 cow groups were formed based on sire and maternal grandsire (MGS) quintiles. Cows had birth dates from 1988 through 1999 and calving dates from 1990 through 2005. Cows that changed herds or had unreported lactations for their first 5 parities were excluded as were herds with <10 cows. Data were available from 4,380,300 cows in 31,759 herds. Mean cow PL was 27.0 mo; time opportunity was a restricting factor. Mean PTA DPR was 2.0 for sires and 2.1% for MGS for the cow group with highest sire and MGS quintiles and -2.1% for both sires and MGS for the group with lowest sire and MGS quintiles. Least squares difference in PL was examined on a within-herd basis with cow birth year in the model. Cows from the highest sire-MGS quintile group had 4.2 mo longer PL than those from the lowest sire-MGS quintile group and were less likely to be culled for reproductive problems (10.0 versus 13.3%) based on reported reason for record termination. Difference in PL between cow groups with highest and lowest sire quintiles for PTA DPR ranged from 2.8 to 3.3 mo; corresponding difference for MGS quintiles ranged from 0.8 to 1.4 mo. Because each month of additional PTA PL is valued at \$29 in the current USDA lifetime net merit index, a 200-cow herd from the highest sire-MGS group for PTA DPR would be worth about \$7,500 more annually than a 200-cow herd from the lowest sire-MGS group without considering any additional income or expense associated directly with DPR. Selection for increased DPR across generations is expected to produce cows with longer herd life because they are less likely to be culled for reproductive problems.

**Key Words:** Daughter Pregnancy Rate, Productive Life, Fertility

### **T20 Modeling nuisance variables for phenotypic evaluation of bull fertility.**

M. T. Kuhn\*, J. L. Hutchison, and H. D. Norman, *USDA-ARS-AIPL, Beltsville, MD.*

This research determined which (available) nuisance variables should be included in a model for phenotypic evaluation of US service sire conception rate (CR), based on DHI data. Models were compared by

splitting data into records for estimation (n=3,613,907) and set-aside data (n=2,025,884), computing predictions using the estimation data, and then comparing predictions to bulls' average CR in set-aside data. There were 803 bulls used for comparison, after requiring a minimum of 50 records for estimation and 100 breedings in at least 30 herds in set-aside data. Correlations and mean differences were used to compare models. Nuisance variables considered were management groups based on herd-yr-season-parity-registry (HYSPR) classes, yr-state-mo, cow age, DIM, a short-cycle variable to account for lower CR for matings preceded by a breeding only 10 to 17 days prior, and various combinations of lactation, service number, and milk yield. Preliminary analysis led to selection of 305d-2x-ME milk yield as the production variable for consideration, and also showed that for each quantitative independent variable, categorical factors provided better bull fertility evaluations than modeling the effects as covariates. Two strategies for management groups were tested, one where HYSPR groups were required to have an absolute minimum number of records and a second where groups were combined across registry, season, and parity subclasses until a minimum group size was achieved. Combining groups to a target size of 20 but still including herd-years with at least 10 breedings maximized correlation with CR in set-aside data. Combining groups implies that some groups have multiple seasons and parities, hence consideration of yr-state-mo and lactation as additional factors. The final variables selected for inclusion were, in addition to HYSPR, yr-state-mo, lactation, service number, milk yield, cow age, short-cycle (yes/no), and the cow effect, partitioned as permanent environment and breeding value. This model maximized correlation with CR in set-aside (55.2%), minimized mean square error (3.25), and mean difference was 0.

**Key Words:** Bull Fertility

**T21 Factors that affect abortion frequency in dairy herds in the United States.** R. H. Miller\*, M. T. Kuhn, H. D. Norman, and J. R. Wright, *Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD.*

Frequency of abortions was studied for lactating cows with 2,980,527 records in the USDA national dairy database. Lactations had been terminated between 1995 and 2005, and cows had been >151 d pregnant at lactation termination. Records without breeding dates were excluded. Abortion frequency for 2005 was 1.6%. Analyses were conducted with PROC GLM to determine effects of herd, year, month, parity (1, 2, ..., 7, ≥8), and gestation stage (152 to 181 d, 182 to 211 d, 212 to 241 d, and ≥242 d pregnant) at lactation termination as well as effects of breed (Holstein, Jersey, and other breeds including crossbreds) and 305-d standardized milk yield. Supplemental analysis examined effects of herd size (50 to 99, 100 to 199, ..., 900 to 999, ≥1,000 cow-years) and location (state). Abortions were most frequent in July and least frequent in December (difference of 0.38%). Abortion frequency decreased from parity 1 to parities ≥8 (difference of 0.52%). Gestation stage had the greatest impact on abortion frequency. Abortions were most frequent at 152 to 181 d pregnant, and trend in abortion frequency was nonlinear across gestation stage. Compared with ≥242 d pregnant, abortion frequency was 3.14% higher at 152 to 181 d pregnant and 1.07% lower at 212 to 241 d pregnant. Abortions for Holsteins and Jerseys occurred more frequently (0.38 and 0.07%, respectively) than for other breeds. Regression of abortion frequency on lactation milk yield was 0.26%, 0.26%, and 0.19% per 1,000 kg of milk for Holsteins, Jerseys, and other breeds, respectively. Regression differences primarily reflect breed differences. Abortion frequency was 0.95% higher for herds with 700 to

799 cow-years than for herds with 50 to 99 cow-years. California herds had the highest abortion frequency (2.18% higher than North Dakota, the state with the lowest frequency). Increased knowledge of factors that affect abortion frequency can aid in the development of management practices for reducing the incidence of abortion, e.g., more intense monitoring of cows 2 to 6 wk after mid-gestation.

**Key Words:** Abortion Frequency, Gestation Stage, Lactation Termination

**T22 Heritability of dairy cow mortality and relationships between mortality and sire genetic evaluations for yield, somatic cell score, productive life and daughter pregnancy rate.** G. W. Rogers\*<sup>1</sup>, J. B. Cooper<sup>1</sup>, and J. S. Clay<sup>2</sup>, <sup>1</sup>*The University of Tennessee, Knoxville,* <sup>2</sup>*Dairy Records Management Systems, Raleigh, NC.*

Lactation records from 1239 Holstein herds in 9 Southeastern states processed by DRMS were utilized to estimate heritability of dairy cow mortality and the relationships between mortality and sire genetic evaluations for milk, fat, protein, SCS, productive life (PL) and daughter pregnancy rate (DPR). Herds were required to have a minimum of 10 years of continuous recording between 1982 and 2005. Binary mortality traits (1=CAR code 6 indicating lactation ending in death vs. 0=all other CAR codes) were developed separately for lactations 1, 2 and 3 or later (3+). To estimate heritability, binary traits were analyzed using a logistic model where the model included sire (random), herd, year of calving, season of calving and age at calving. Number of cows (sires) ranged from 280,000 (1090) to 614,000 (1913). In addition, the binary mortality traits were analyzed using logistic regression with sire replaced by sire PTA (one trait per analysis). Sire PTA were categorized into quartiles and odds ratios for each quartile compared with quartile 4 were determined. Heritability for mortality on the underlying scale was low and ranged from .03 to .04. Odds ratios indicated that mortality tended to be lower for cows sired by bulls with lower PTA milk, protein and SCS, especially in lactations 3+. Odds ratios for fat percent were not consistent across lactations, but higher sire PTA for fat percent were favorable for mortality in lactation 2 and lactations 3+. Odds ratios for PL and DPR indicated that mortality was significantly higher for cows sired by bulls with lower PTA PL and DPR. Daughters of sires with PTA for PL or DPR in the lowest quartile were 22 to 31% more likely to end the lactation with death compared with daughters of sires in highest quartile. Selection for increased milk yield may have increased cow mortality but intense selection for increased PL and DPR will likely reduce cow mortality in the future.

**Key Words:** Dairy Cow Mortality, Heritability, Daughter Pregnancy Rate (DPR)

**T23 Lactation patterns for dairy cattle in a multibreed dairy population in Central Thailand.** A. Seangjun<sup>1</sup>, S. Koonawootrittriron<sup>1</sup>, and M. A. Elzo\*<sup>2</sup>, <sup>1</sup>*Kasetsart University, Bangkok, Thailand,* <sup>2</sup>*University of Florida, Gainesville.*

Milk production in Thailand is based on a dairy population composed of Holstein (H) and fractions of various other *Bos taurus* and *Bos indicus* breeds (O). This population structure is the result of a national effort encouraged by the Thai government to increase milk production. Genetic

evaluations have been conducted by Kasetsart University and the Dairy Promotion Organization for milk yield, fat yield, and fat percentage since 1996. To further improve management and the ability to select animals under Thai tropical conditions, a more complete understanding of lactation characteristics is needed. The objective of this research was to study the lactation patterns of 5 lactation traits: initial yield (IY), peak yield (PY), days to peak (DTP), persistency (PST), and 305-d milk yield (MY) in 5 breed groups in a Holstein-Other breeds (HO) multibreed population in Central Thailand. The dataset included 5,713 monthly test-day records from 520 dairy cows raised in 108 farms from 1998 to 2006. Breeds groups were BG1 (purebred H), BG2 ( $0.9687 \leq H < 1.00$ ), BG3 ( $0.9375 \leq H < 0.9687$ ), BG4 ( $0.875 \leq H < 0.9375$ ), and BG5 ( $0.50 \leq H < 0.875$ ). Seasons were 1 = winter (November to February), 2 = summer (March to June), and 3 = rainy (July to October). The model for each trait included the fixed effects of herd-year-season, calving age, and breed group. Random effects were sire of cow and residual. Procedure Mixed of SAS was used for computations. Least squares means ranged from  $14.16 \pm 0.71$  kg (BG5) to  $16.04 \pm 0.66$  kg (BG4) for IY, from  $18.23 \pm 0.55$  kg (BG5) to  $19.84 \pm 0.51$  kg (BG4) for PY, from  $36.51 \pm 3.78$  d (BG4) to  $52.04 \pm 9.74$  d (BG1) for DTP, from  $6.45 \pm 0.24$  (BG1) to  $6.71 \pm 0.11$  (BG3) for PST, and from  $4,083.74 \pm 103.50$  kg (BG5) to  $4,317.15 \pm 111.56$  kg (BG3) for MY. Breed group differences were non-significant. However, BG3 performed better than BG1 and BG2 suggesting that nutrition, management, and tropical conditions in Thailand may have prevented them from reaching their production potential.

**Key Words:** Lactation, Multibreed, Tropical

**T24 Relationships between reproductive traits of heifers and cows and yield traits for Holsteins in Japan.** H. Abe\*, Y. Masuda, and M. Suzuki, *Obihiro University of A & VM., Obihiro, Japan.*

The objective of this study was to investigate relationships between reproductive traits in heifers and cows and yield traits for Holsteins in Japan. Insemination and lactation records for cows calved between 1990 and 2003 in Hokkaido were obtained from Hokkaido Dairy Milk Recording and Testing Association. Age at conception (AC) and first service conception rates (CR) were calculated for heifers. Days open (DO) and CR were calculated for cows in first and second parities. The yield traits used were 305-d milk, fat and protein yields. A threshold animal model was applied for CR, and a linear animal model was applied for the other traits. For heifers, the model included herd-year and month of the first insemination as fixed effects. For cows, the model included herd-year of calving, month of calving, and age class of calving as fixed effects. Both models included the random animal additive genetic effect. The numbers of records were around 260,000, 780,000 and 640,000 for heifers, first parity, and second parity cows, respectively. Subsets of records were extracted for the two traits by the random sampling of herds, and genetic parameters were estimated either by the AI-REML method or the Bayesian method using Gibbs sampling. Heritability estimates were 13%, 2% for AC and CR in heifers, 3% for CR in cows, 6% and 4% for DO in first and second parities. Genetic correlations between AC and DO in first and second parities were 0.29 and 0.09, respectively. Genetic correlations among CR ranged from 0.33 to 0.93. Genetic correlations of yield traits with AC and CR in heifers ranged from 0.19 to 0.37, and -0.06 to 0.12, respectively. Genetic correlations between yield and reproductive traits in cows were antagonistic. As a result, reproductive traits in heifers and cows should be considered as separate traits. The relationship between yield and reproductive traits

in heifers was slightly preferable, suggesting that genetic selection of reproductive traits could be performed for heifers in Japan.

**Key Words:** Reproductive Traits in Heifers and Cows, Genetic Parameters, Holsteins in Japan

**T25 Genetic aspects of the somatic cells count in dairy buffaloes reared in Sao Paulo state, Brazil.** H. Tonhati\*<sup>1,2</sup>, G. M. Sanches<sup>1</sup>, M. F. Ceron Munoz<sup>1</sup>, L. G. de Albuquerque<sup>1,2</sup>, R. R. A. B. Borquis<sup>1,3</sup>, R. Sesana<sup>1,3</sup>, and L. El Faro<sup>1</sup>, <sup>1</sup>Sao Paulo University, Jaboticabal, Sao Paulo, Brazil, <sup>2</sup>Conselho Nacional de Desenvolvimento Cientifico e Tecnologico, Brasilia, DF, Brazil, <sup>3</sup>Fundacao de Amparo a Pesquisa do Estado de Sao Paulo, Sao Paulo, Brazil.

The goal of this work was to study the relationship between somatic cell count (SCC) and milk yield (MY). Were analyzed 9404 test-day records for SCC and MY obtained from 2198 lactations of 1052 Murrah buffaloes between 1997 and 2005. To quantify the decreases of MY in relation to SCC, the model included a random animal effect and the fixed effects of farm, calving order, year and season of calving and Somatic Cell Score (SCS) as covariate. For estimating genetic parameters, test day models were used for SCC transformed in SCSt ( $SCSt = [\log_2(SCC/1000000)]+3$ ). For average of somatic cells count in the lactation (SCCt270) and milk yield to 270 days (MY270) the (co) variance components were estimated. SCC of every month of lactation were considered as different traits. The model included additive genetic, permanent environmental (for SCCt270 and for MY270) and residual random effects. Other fixed effects were: contemporary group; test-day and age of cow at calving as a covariate (linear and quadratic effects). For SCSt, contemporary groups were defined as herds-year-month of the control, and for SCCt270 and MY270 as herd-year-season of the calving. It was found that all effect influenced the expression of SCSt. For first parity females, there no relation between MY and SCC was found. The largest decreases of milk production were observed in female with more than one calving. This category should receive a special attention in relation to udder health. The farm, year and calving order effects should be considered in the comparison among animals for genetic evaluations. Heritability estimates obtained from single trait analyses ranged, among the months, from 0.06 to 0.50 for SCSt and 0.28 for SCCt270. Heritability estimates in double trait analyses oscillated between 0.65 and 0.28 for SCSt and up 0.66 for SCCt270. All correlations between SCSt and SCCt270 were positive, ranging from 0.50 to 0.91 (genetic) and from 0.59 to 0.82 (phenotypic). The genetic correlations between SCSt and MY270 ranged from 0.52 to 0.10 and the phenotypic correlation ranged from 0.37 to 0.0. The genetic correlation between SCCt270 and MY270 was 0.11 and the phenotypic correlation was 0.15.

**Key Words:** Genetic Evaluation, Variance Components, Longitudinal Data

**T26 Computing options for multiple-trait test-day random-regression models.** I. Aguilar\*<sup>1,2</sup>, S. Tsuruta<sup>1</sup>, and I. Misztal<sup>1</sup>, <sup>1</sup>University of Georgia, Athens, GA, <sup>2</sup>Instituto Nacional de Investigación Agropecuaria, Las Brujas, Uruguay.

Data included 90,242,799 test day records from 5,402,484 Holstein cows in 3 parities. The total number of animals in the pedigree file was

9,326,754. Additionally, daily temperature humidity indexes (THI) from 202 weather stations were available. The effects of herd test day, age at calving, milking frequency and days in milk classes (DIM) were made fixed, and the effects of additive, permanent environment and herd-year were made random. Random effects were fit as random regression. Covariates included linear splines with four knots at 5, 50, 200, 305 DIM, and a function of THI of the 3rd day before the test day from a weather station closest to the farm. The first three lactations were used as separate traits, resulting in 15 by 15 (co)variance matrices for each random effect. The mixed model equations were solved using an iteration on data program with a preconditioned conjugate gradient (PCG) algorithm. Several preconditioners were used: diagonal (D), block diagonal due to traits (BT), and block diagonal due to traits and correlated effects (BTCORR). One run included BT with a "diagonalized" model in which the random effects were reparameterized for diagonal (co)variance matrices among traits (BTDIAG). Memory requirements were 8.7 Gbytes for D, 10.4 Gbytes for BT and BTDIAG, and 24.3 Gbytes for BTCORR. Computing times (rounds) were 14 d (952) for D, 10.7 d (706) for BT, 7.7 d (494) for BTDIAG and 4.6 d (289) for BTCORR. The convergence criterion for BTCORR showed high fluctuation that required either a moving average or a strict stopping criterion. The convergence pattern was strongly influenced by the choice of fixed effects. When sufficient memory is available, the option BTCORR is the fastest and simplest to implement; the next efficient method, BTDIAG, requires additional steps to diagonalization and back-diagonalization.

**Key Words:** Heat Stress, Multiple Trait Random Regression Model, Genetic Evaluation

**T27 One predominant and preeminent common ancestor in Bell family.** R. D. Shanks\* and K. Boesche, *University of Illinois, Urbana*.

Purpose was to improve interpretation of inbreeding. As a means to increase interpretation, inbreeding was partitioned relative to the common ancestor that was the source of the inbreeding. To focus pedigrees, the family of Carlin-M Ivanhoe Bell (Bell) was used as a test sample. Forty sons with high inbreeding coefficients and nine grandsons of Bell were evaluated for their source of inbreeding. All but one evaluated son and all evaluated grandsons of Bell had the opportunity to share alleles identical by descent from Osborndale Ivanhoe (Ivanhoe). Ivanhoe was the grandsire of Bell. For twenty-one of the inbred sons, Ivanhoe was the only common ancestor as the source of inbreeding. Ivanhoe as a source of inbreeding was responsible for inbreeding ranging from .4% to 3.1% for the sons and from .2% to .4% for the grandsons. Additionally, Penstate Ivanhoe Star, a son of Ivanhoe, was a common ancestor for eight inbred sons of Bell. Penstate Ivanhoe Star was responsible for inbreeding of 6.25% or 3.125%. On the maternal side of the pedigrees, Ivanhoe was found within 3 to 6 generations of the sons and 5 or 6 generations for the grandsons. More diversity existed among the grandsons of Bell as only 2 grandsons had all of their inbreeding from Ivanhoe as the common ancestor. Inbreeding of the sample of sons ranged from .4% to 7.8%. Inbreeding of the nine grandsons ranged from .2% to 2.6%. Osborndale Ivanhoe was an Excellent bull and a legend in dairy cattle genetics. In 1992, 40 years after his birth, he was still related to more than 5% of the Holstein population. The vast majority of inbreeding within the Bell family has Ivanhoe as the source. Although the purpose of this study was to improve interpretation of inbreeding, results of the study were that the number of common ancestors was extremely limited. Inbreeding depression would be a function of the alleles available in the common ancestor. If few, but different, common ancestors would be

found in other sire families, the need to determine inbreeding depression within family may be warranted.

**Key Words:** Inbreeding, Holstein, Common Ancestors

**T28 Heritability of genetic tolerance to Johne's disease.** R. Zanella\*<sup>1</sup>, M. Settles<sup>1</sup>, T. Fyock<sup>2</sup>, R. Whitlock<sup>2</sup>, Y. Schukken<sup>3</sup>, J. Van Kessel<sup>5</sup>, J. Karns<sup>5</sup>, E. Hoving<sup>4</sup>, J. Smith<sup>6</sup>, C. Van Tassel<sup>5</sup>, C. Gaskins<sup>1</sup>, and H. Neibergs<sup>1</sup>, <sup>1</sup>Washington State University, Pullman, <sup>2</sup>University of Pennsylvania, Kennett Square, <sup>3</sup>Cornell University, Ithaca, NY, <sup>4</sup>Penn State University, University Park, <sup>5</sup>USDA, ARS, Beltsville, MD, <sup>6</sup>University of Vermont, Burlington, VT.

Johne's disease, also known as paratuberculosis, is a prevalent and economically important disease in cattle caused by bacterial infection of *Mycobacterium avium* subsp. *paratuberculosis*(MAP). Johne's disease often results in weight loss, lowered milk production, and death. MAP is excreted in feces and milk of infected animals which helps propagate the disease. Infection generally occurs in neo-natal calves, but clinical disease and detection are often delayed for 2 to up to 6 years, prolonging the exposure of the herd to MAP. Disease prevalence is estimated to be present in 75% of US dairy herds and is increasing. Current vaccination and treatments are ineffective and diagnostic methods detect less than 25% of infected animals, at one time. Cattle exposed to MAP may have three distinct responses: failure to develop disease (resistant animals), disease with little fecal shedding of MAP (tolerant), or disease with significant fecal shedding (intolerant-susceptible animals). The objective of this study was to determine the heritability of genetic tolerance to Johne's disease in 4 Holstein dairy herds. Genetic tolerance was determined by sequential diagnostic testing of 260 animals every 3-6 months followed by post mortem tissue examination. This testing regime represents the gold standard for Johne's diagnosis. MAP testing was conducted on a minimum of 4 tissue samples per animal. Comparison of fecal and tissue testing results was used to determine if cows were classified as resistant, tolerant or intolerant-susceptible. Heritability was calculated using an animal method. Selection of animals with genetic tolerance to Johne's would be advantageous by lowering infection pressure and extending the period between infection and clinical symptoms.

**Key Words:** Johne's Disease, Heritability, Tolerance

**T29 Detection of polymorphism in bovine polymeric immunoglobulin receptor gene promoter region and association with milk IgA and IgM concentration.** C. G. Zhang<sup>1,2</sup>, J. Q. Wang\*<sup>1</sup>, D. P. Bu<sup>1</sup>, G. L. Liu<sup>1</sup>, J. B. Cheng<sup>1</sup>, X. L. Dong<sup>1,2</sup>, H. Y. Wei<sup>1</sup>, L. Y. Zhou<sup>1</sup>, G. Q. Zhao<sup>2</sup>, and K. L. Liu<sup>1</sup>, <sup>1</sup>State Key Laboratory of Animal Nutrition, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China, <sup>2</sup>Yangzhou University, Yangzhou, China.

The polymeric immunoglobulin receptor (pIgR) transports dimeric immunoglobulin A (dIgA) and tetramer IgM across the epithelial cell layers into the secretions of mammary gland. The object of this study was to detect bovine pIgR gene promoter polymorphism and its association with milk IgA and IgM concentration. Blood and milk samples (n = 189) from each cow were collected randomly from more than 1,600 animals in Beijing. Determination of IgA and IgM in milk samples were performed by a commercial sandwich enzyme-linked immunosorbent assay (ELISA) using the Bovine IgA and IgM ELISA Quantitation Kit.

SNPs were found by 25 dairy cows' blood genome DNA samples by PCR amplifying and sequencing. The SNPstream system (Beckman Coulter, USA) was used for genotyping for 189 individuals. The results indicated that SNP1 (-3128) and SNP2 (-3072) were the G/A variation and SNP3 (-2834), SNP4 (-2348) and SNP5 (-515) were the T/C variation. Seven haplotype block and ten main diplotypes were inferred by phase 2.1. Statistical analysis indicated that diplotypes had no effect on milk IgA and IgM concentration ( $P > 0.05$ ). Duncan's multiple-range test showed the least square mean for milk IgA concentration (mg/ml) of diplotype H1H1 ( $0.323 \pm 0.042$ ), H1H5 ( $0.421 \pm 0.106$ ) and H3H6 ( $0.486 \pm 0.128$ ) was significantly higher than those of H4H4 ( $0.206 \pm 0.041$ ) and H5H6 ( $0.051 \pm 0.128$ ) diplotypes ( $P < 0.05$ ), whereas other diplotypes had no significant differences on the least square mean for milk IgA and IgM concentration ( $P > 0.05$ ). Therefore, our findings implied that pIgR gene promoter SNPs had effect on milk IgA and IgM concentration.

Acknowledgement: Research funded by Ministry of Science and Technology (2006BAD12B03).

**Key Words:** pIgR Gene Promoter, IgA and IgM, SNPs

**T30 Effect of marker-assisted preselection in Japanese dairy population.** H. Ohmiya\* and M. Suzuki, *Obihiro University of Agriculture & Veterinary Medicine, Obihiro, Hokkaido, Japan.*

The utility of the "modified bottom-up" approach to marker-assisted selections in a conventional progeny test scheme was examined on dairy breeding in Japan. In this scheme, genotyping the daughters to decide the genotypes for the quantitative trait locus (QTL) of their sires was not needed as in a usual bottom-up approach, which is the preselection scheme that uses marker information before a traditional progeny test. Because the cows are not tested for markers to be closely linked to a QTL in Japan, the development of selection schemes without genotyping the daughters for markers may benefit the Japanese dairy population. By computing the differences between the deviations of the breeding values and the genetic variance (G) in a population, it can be decided whether their sire is homozygous for locus controlling milk product. If the sires were identified as homozygous with allelic contrast smaller than G, the mean breeding value of the daughters from a sire is more favorable than that in a population used for generating candidate young bulls. Stochastic simulations of dairy population structures, methods of evaluation, and parameters similar to Japanese dairy population were used to evaluate each rate of the additional genetic gains of a traditional scheme using only progeny testing and BLUP with an animal model, a conventional bottom-up approach, and the modified bottom-up approach suggested in this study. These simulations under the 2 QTLs and 40 polygenes each with 4 alleles were replicated 20 times. It was found that these rates were increased by an average of 11.5% and 5% compared with a traditional progeny test scheme when conventional bottom-up and modified bottom-up approaches were used. Additionally, with the modified bottom-up approach, the dairy population improved more with only progeny testing but less than with the conventional bottom-up approach. When only dairy bulls are genotyped in Japan, this approach would be effective.

**Key Words:** Marker-Assisted Selection, QTL, Bottom-Up Approach

**T31 Optimized measured genotype analysis for genome-wide quantitative trait loci mapping using dense SNP chips.** J. R. O'Connell\*, *University of Maryland School of Medicine.*

The availability of low cost, dense fixed content single nucleotide polymorphism (SNP) chips has changed the landscape of genetics. These chips provide sufficient density for linkage disequilibrium mapping to localize alleles that increase susceptibility to disease, determine variation of quantitative phenotypes or predict genetic merit for economically important traits.

A standard quantitative trait analysis for genome-wide SNP analysis is the mixed model measured genotype approach, which treats the genotype as a measured covariate while controlling for residual familial correlation through a polygenic component. The model easily incorporates dominant, recessive or additive models of single SNPs and joint, conditional or interaction analysis of multiple SNPs and genomic prediction using estimated regression parameters. Mixed model maximum likelihood estimation, however, is computationally challenging due to dense matrix operations required for a single analysis, making genome-wide analysis often infeasible. A computationally efficient approach is presented based on the diagonalization of the covariance-variance structure that replaces operations on dense matrices with diagonal matrices in the maximization.

A genome-wide association analysis of yield and conformation traits measured on 2602 Holstein bulls using the BovineSNP50 chip was performed to evaluate and validate the method. In total 38,000 SNPs with a minor allele frequency  $> 0.02$  were analyzed using an additive model. As expected a SNP located at ~300kb on chromosome 14 near DGAT provided the most significant association for milk yield with a  $p$ -value  $< 10e-41$ . In addition a cluster of SNPs located around 50Mb on chromosome 15 reached genome-wide significance with  $p$ -values  $< 10e-09$ .

**Key Words:** Pedigree Analysis, Genomic Selection

**T32 Effect of varying degree of relatedness in family designs on estimation of IBD probability and other parameters in QTL mapping based on variance component analysis.** G. Freyer<sup>2</sup> and N. Vukasinovic\*<sup>1</sup>, <sup>1</sup>Newsham Choice Genetics, Saint Louis, MO, <sup>2</sup>Research Institute for the Biology of Farm Animals, Dummerstorf, Germany.

In variance component approach the QTL effect is modeled as a function of probabilities that two alleles in the same or in different animals at a particular genomic position are identical by descent (IBD probabilities). Estimates of IBD probabilities and therefore, proper modeling and estimation of QTL variances depend on the number and informativeness of markers, the strength of linkage and linkage disequilibrium of markers and QTL, and the relatedness of animals in the pedigree. In this simulation study we investigate how the level of relatedness of animals in a pedigree influences IBD probabilities, their correlations to the IBD probability at the true QTL-position, and the quality of test statistic profiles. Four multi-generational pedigrees (FS) resembling real dairy populations were simulated. All pedigrees comprised 850 individuals. The final offspring of 9 sires originated from 2 founder (great-grand) sires. FS0 was non-inbred; FS1 contained an inbred sire from an aunt-nephew mating; FS2 contained a 25% inbred sire from a half-sib mating. In both pedigrees, the inbred sires had 78 final offspring each. FS3 was the same as FS2, except for increasing the number of offspring of the

inbred sire to 138. FS4 contained a sire originating from a mother by (inbred) son mating with 73 offspring. This design was used to demonstrate how extreme levels of inbreeding affect QTL mapping. Animals were assumed genotyped for 11 markers within a 55cM long putative QTL region, with one QTL explaining 15% of the phenotypic variance. The number of alleles was 2, 4, or 6, at unevenly spaced markers. Twenty replicates were run for each parameter combination. IBD probabilities for each cM in the segment were calculated using the nearest informative marker bracket. Estimation of QTL parameters was performed using maximum likelihood approach within the ASReml software. The results indicated that the existence of inbred animals in a pedigree may lead to more precise estimates of IBD probabilities and QTL parameters that are less sensitive to variations in simulation parameters.

**Key Words:** QTL Mapping, Inbreeding, Variance Components

**T33 Spermatozoal transcriptome profile as marker for bull fertility and sperm motility: A potential tool to evaluate semen quality.** N. Bissonnette<sup>\*1,2</sup>, J.-P. Lévesque-Sergerie<sup>1,2</sup>, and G. Boissonneault<sup>2</sup>, <sup>1</sup>*Agriculture and Agri-Food Canada, Sherbrooke, Québec, Canada*, <sup>2</sup>*University of Sherbrooke, Sherbrooke, Québec, Canada*.

Fertility represents significant costs to the dairy industry and is still a major concern for dairy producers. Efforts are made to evaluate bull fertility in order to offer semen of the highest quality value. In a previous study, we had analyzed the sperm transcript content and found that bulls with different non-return rate, a measure of field fertility, contain a different profile (Lalancette et al. 2007, *Biology of Reproduction*, in press). This transcript panel comprising 3216 EST has been used to construct a microarray by automated spotting. Since motility is a prerequisite of high quality semen, we verified if spermatozoa collected from fresh semen and presenting different motility indexes could also display different transcriptome profiles. The transcript content of fresh spermatozoa isolated at the bottom fraction 90% (highly motile) of a Percoll step gradient was compared by microarray hybridization to the transcript extracted from the spermatozoa harvested at the 70%–90% interface (less motile). Spermatozoa derived from the same ejaculated bull semen, but displaying subtle yet different motility, contain different transcript abundance ( $p < 0.001$ ). Among the genes of which expression was confirmed by real-time PCR, we validated some well-known transcript, notably that of protamine 1, whose encoded proteins are architectural and required for an adequate genome packaging. Uncharacterized gene such as a gene encoding a protein similar to human metalloproteinase was among the differentially expressed genes. Interestingly, expression of this candidate was predicted in December 2006 by automated computational analysis and was found and for which no bovine ortholog sequence has yet been reported. Analyzing the spermatozoal transcriptome would not only be helpful in determining bull fertility but could also be used in a semen quality analysis. Whereas this specialized cell has long been considered a vehicle that contains only a half-genome, it also contains remnant transcripts of spermatogenesis, whose profile can be used as a signature for semen quality.

**Key Words:** Fertility, Sperm Quality, Genetic Marker

**T34 Molecular characterization of the bovine DDX3Y gene.** W.-S. Liu<sup>\*1</sup>, A. Wang<sup>2</sup>, Y. Yang<sup>1</sup>, E. Landrito<sup>3</sup>, and H. Yasue<sup>4</sup>, <sup>1</sup>*The Pennsylvania State University, University Park*, <sup>2</sup>*Virginia Polytechnic Institute and State University, Blacksburg*, <sup>3</sup>*University of Nevada, Reno, NV*, <sup>4</sup>*National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan*.

The DEAD box polypeptide 3, Y-linked gene (DDX3Y, also known as DBY, DEAD box gene on the Y) encodes a putative ATP-dependent RNA helicase. This gene belongs to the DEAD box protein family, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD). In the human, DDX3Y is located in the AZFa (azoospermia factor a) interval in the non-recombining region of the Y chromosome. Deletion of the AZFa region has been showed to disrupt spermatogenesis, causing subfertility and infertility in otherwise healthy men. We have characterized the bovine DDX3Y (bDDX3Y) gene. We found two transcripts for bDDX3Y, bDDX3Y-L and bDDX3Y-S, which correspond to the long and short transcripts of the human DDX3Y and mouse Ddx3y gene. The two bDDX3Y transcripts are identical except for a three-base-pair (AGT) insertion in the position of nt 2027 and an expanded 3'UTR (nt 2155-2790) in bDDX3Y-L. The bDDX3Y-S encodes a peptide of 660 amino acids (aa), while the bDDX3Y-L encodes a peptide of 661 aa as a result of the insertion of a serine (S) in the position of aa 634 in the bDDX3Y-L peptide. Both bDDX3Y isoforms contain the conserved motifs of DEAD-box RNA helicases. Expression analysis of the bDDX3Y gene in 12 bovine tissues by RT-PCR showed that both transcripts are predominantly expressed in the bovine testis. This project was supported by grants from USDA-CSREES-NRI to WSL (No. 2005-35205-15455 and No. No. 2005-35205-18653).

**Key Words:** DDX3Y Gene, Y Chromosome, Bovine

**T35 A gene frequency model to map QTL using bayesian inference.** W. He<sup>\*1</sup>, R. L. Fernando<sup>1</sup>, J. C. M. Dekkers<sup>1</sup>, and D. Gianola<sup>2</sup>, <sup>1</sup>*Iowa State University, Ames*, <sup>2</sup>*University of Wisconsin, Madison*.

Information for mapping of quantitative trait loci (QTL) comes from two sources: linkage disequilibrium (non-random association of allele states) and cosegregation (non-random association of allele origin). Information from LD can be captured by modeling conditional means and variances at the QTL given marker information. Similarly, information from cosegregation can be captured by modeling conditional covariances. Here, we consider a model where both conditional means and variances are modeled as a function of the conditional gene frequencies at the QTL. The parameters in this model include these gene frequencies, additive effect of the QTL, its location, and the residual variance. Bayesian methodology was used to estimate these parameters. The priors used were: logit-normal for gene frequencies, normal for the additive effect, uniform for location, and inverse chi-square for the residual variance. Computer simulation was used to compare the power to detect and accuracy to map QTL by this method with least squares using a regression model (LSR). LD was simulated in a chromosomal segment of 1cM with one QTL by random mating for 1000 generations in a population of size 500 and for 50 generations in a population of size 100. The comparison was studied under a range of conditions, which included SNP density of 0.1, 0.05 or 0.02 cM, sample size of 500 or

1000, and phenotypic variance explained by QTL of 2 or 5%. Both 1 and 2-SNP models were considered. Power of LSR ranged from 0.66 to 1.0, and was always higher than power of the Bayesian method (BM), which ranged from 0.56 to 1.0. The accuracy of BM to map QTL position, quantified by the root mean squared error, ranged from 0.107 to 0.189 cM, and was always better than the accuracy of LSR, which ranged from 0.134 to 0.260 cM. Results support that given a high SNP density, the gene frequency model can be used to map QTL with considerable accuracy even within a 1 cM region.

Supported by National Research Initiative Grant no. 2007-35205-17862

**Key Words:** Bayesian Inference, Gene Frequency Model, Map QTL

**T36 The study of gene delivery system in avian species using human adenoviral vector.** D. R. Bae\*, J. H. Shin, J. M. Reddish, J. D. Latshaw, M. P. Wick, and K. Lee, *The Ohio State University, Columbus.*

Adenovirus has been used in vivo and in vitro as a vector to carry a foreign gene for gene transfer. Several adenoviral vectors have been widely used to transfect a target gene into various cell types and tissues of various animal origins for studying gene function and also clinical application. However, the feasibility of the human recombinant adenoviral vectors system has not been evaluated in avian species. The objective of the present study was to evaluate this system in poultry. The primary muscle cells obtained from chicken, turkey, and quail were cultured and infected with recombinant adenovirus. Over 50% of chicken and turkey cells were infected by adenovirus containing green fluorescence protein (GFP) gene with less infection efficiency in quail. These cells were further tested with recombinant adenovirus containing GFP gene as a tracking marker and 3-hydroxyisobutryl-CoA hydrolase (HIBCH) gene as a target gene. Immunofluorescence analysis revealed that cells infected with adenovirus expressed both GFP and HIBCH protein, showing a successful tracking of the target gene by GFP expression. In addition, our western blot analysis showed that the expression of HIBCH protein in cultured primary muscle cells infected with the recombinant adenovirus. These data demonstrate the feasibility of using human recombinant adenoviral vectors as a tool to express foreign genes in the muscle cells of avian species. Depending

on the target gene, the adenoviral vector system will provide a new approach to study the role of target genes in muscle development and metabolism in avian species.

**Key Words:** Human Recombinant Adenovirus, Gene Delivery, Avian Species

**T37 The use of machine learning techniques for analysis of high-dimension gene expression data sets.** K. R. Robbins, W. Zhang, J. K. Bertrand, and R. Rekaya\*, *University of Georgia, Athens.*

The analysis of microarray data has become common place in the field of animal science; however, due to the high dimensions and complex structure of many expression data sets, traditional statistical models may be inadequate for the analysis of such data. To address issues associated with commonly used methods for the identification of predictive genes sets, the ant colony algorithm (ACA) is proposed for use on data sets with large numbers of features and complex structures. The ACA is an optimization algorithm capable of modeling complex data structures without the need for explicit parameterization. The incorporation of prior information and communication between simulated ants allows the ACA to search the sample space more efficiently than other optimization methods. When applied to simulated data, as well as, a high-dimensional cancer microarray data set, the ACA was able to identify small subsets of highly predictive and biologically relevant genes without the need for simplifying assumptions. For simulated data, genes selected by the ACA to train a latent variable model yielded increases in prediction accuracy of 24.4%, 6.7%, and 10.3% when compared to genes sets selected by fold change (FC), t-test (T) and penalized t-test (PT), respectively. For the cancer data set, the ACA yielded increases in accuracy of 16.6% and 6.5% over the best performing test statistics and other optimization models. Furthermore, the ACA was able to converge to good solutions without the need for significant truncation of the data, as required by other optimization algorithms. The ACA was also able to achieve higher prediction accuracies using fewer selected genes when compared the test statistics. This was attributed to ability of ACA to model the complex gene interactions, reducing the collinearity in selected genes when compared with FC, T and PT.

**Key Words:** Ant Colony Algorithm, Machine Learning, Gene Expression