GENETIC ANALYSIS OF HEAT TOLERANCE FOR PRODUCTION, REPRODUCTION AND HEALTH TRAITS IN US HOLSTEIN COWS

By
ANIL SIGDEL

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE
UNIVERSITY OF FLORIDA
2018
© 2018 Anil Sigdel
To dairy farmers' around the globe
ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Francisco Peñagaricano for his intellect support, critical analysis of data and for offering valuable ideas and suggestions during my master study. Without his valuable suggestions, meticulous supervision, and constructive criticism, this manuscript would not come up to its present quality. Many thanks also go to members of my committee; Dr. Peter J. Hansen and Dr. Raluca Mateescu for their advice and suggestions.

I am very thankful to Dr. Rostam Abdollahi-Arpanahi for his guidance and expertise that he shared with me during my research. I wish to thank the past and current members of Quantitative Genetics and Genomics group at University of Florida for providing friendly and enthusiastic environment. Special thanks go to Yi Han, Rocio Amorin, Fernanda Rezende, Juan Pablo Nani, Leeher for their friendship and assistance. I would also like to thank Animal Science Graduate Students’ Association here in the Department of Animal Sciences, UF for enriching my personal and professional experiences.

I appreciate the support from South East Milk Check Off and Department of Animal Sciences, UF for funding my study. I would also like to thank Council on Dairy Cattle Breeding and North Florida Holsteins for providing necessary data for this study. I wish to thank my parents Mr. Bishnu Prasad Sigdel and Mrs. Bhagirathi Sigdel, and my sisters and brother for their love and encouragement. Finally, I will like to thank my brother Narendra and friend Bikalpa for their advices and suggestions on my scientific work, future professional career, and personal life. I remain indebted to them.
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS** .................................................................................................................. 4

**LIST OF TABLES** ....................................................................................................................... 7

**LIST OF FIGURES** ...................................................................................................................... 8

**LIST OF ABBREVIATIONS** ......................................................................................................... 9

**ABSTRACT** .................................................................................................................................. 10

## CHAPTER

1  **INTRODUCTION** .................................................................................................................... 12

2  **LITERATURE REVIEW** .......................................................................................................... 14

   Thermoneutral Zone ..................................................................................................................... 15

   Heat Stress .................................................................................................................................... 16

   Responses to Heat Stress ............................................................................................................ 17

   Methods of Heat Transfer .......................................................................................................... 17

      Conduction and Convection ..................................................................................................... 18

      Radiation ................................................................................................................................... 19

      Evaporative Heat Loss ............................................................................................................. 19

   Physiological Responses of Cattle to Heat Stress ...................................................................... 20

      Reduced Feed Intake ................................................................................................................ 20

      Decreased Milk Production .................................................................................................... 21

      Decreased Reproduction ........................................................................................................ 22

      Increased Water Intake ........................................................................................................... 23

      Increased Respiration Rate ...................................................................................................... 23

   Changes in Blood Hormones Concentration ............................................................................. 23

   Methods for Assessment of Heat Stress ..................................................................................... 24

      Rectal Temperature ................................................................................................................ 24

      Milk Temperature .................................................................................................................... 25

      Respiration Rate ...................................................................................................................... 25

      Heat Shock Proteins (HSPs) .................................................................................................... 26

   Strategies for Reduction of Effects of Heat Stress ..................................................................... 27

      Physical Modifications of the Environment ............................................................................ 27

      Nutritional Strategies for Managing Heat Stress in Dairy Cows ........................................... 28

      Genetic Development of Heat Resistant Animals ................................................................... 28

   Genetic Studies on Heat Stress .................................................................................................. 30

3  **GENETIC ANALYSIS OF HEAT TOLERANCE FOR PRODUCTION TRAITS IN US HOLSTEIN COWS** ........................................................................................................... 33

   Background ................................................................................................................................... 33
Materials and Methods................................................................................................................................. 35
  Phenotypic and Genotypic Data .................................................................................................................. 35
  Statistical Model ........................................................................................................................................ 36
  Statistical Analysis ..................................................................................................................................... 37
  Genome Wide Association Mapping using ssGBLUP .................................................................................. 38
Gene Set Analysis ........................................................................................................................................... 39
  Assignment of SNPs to genes ......................................................................................................................... 39
  Assignment of genes to pathways .................................................................................................................. 40
  Pathway-based association analysis ............................................................................................................. 40
Results and Discussion ................................................................................................................................. 41
  Genetic Dissection of Heat Stress for MY, FY and PY ............................................................................. 41
  Whole-Genome Mapping for Thermotolerance Genes affecting MY ....................................................... 42
  Pathway-Based Association Analysis for MY ............................................................................................. 45
Summary ......................................................................................................................................................... 48

4 GENETIC ANALYSIS OF HEAT TOLERANCE FOR REPRODUCTION AND UDDER HEALTH IN US HOLSTEIN COWS ........................................................................................................... 55

  Background .................................................................................................................................................. 55
  Materials and Methods ............................................................................................................................... 58
  Phenotypic and Genotypic Data .................................................................................................................. 58
  Statistical Model ........................................................................................................................................ 59
  Statistical Analysis ..................................................................................................................................... 60
  Genome-wide Association Mapping using ssGBLUP .................................................................................. 61
Gene Set Analysis ........................................................................................................................................... 62
  Assignment of SNPs to genes ......................................................................................................................... 62
  Assignment of genes to pathways .................................................................................................................. 62
  Pathway-based association analysis ............................................................................................................. 62
Results and Discussion ................................................................................................................................. 63
  Genetic Dissection of Heat Stress for Conception and Udder Health ....................................................... 63
  Variance Component Estimates for CI .......................................................................................................... 63
  Variance Component Estimates for SCS ....................................................................................................... 64
  Genomic Scans for Thermotolerance Genes Affecting CI and SCS ......................................................... 64
    Whole Genome association analysis for CI ................................................................................................. 64
    Whole Genome association analysis for SCS ............................................................................................. 66
  Pathway-Based Analysis ............................................................................................................................. 68
    Gene-set analysis for CI ............................................................................................................................. 69
    Gene-set analysis for SCS ........................................................................................................................... 70
Summary ......................................................................................................................................................... 72

5 CONCLUSIONS ............................................................................................................................................ 79

LIST OF REFERENCES ..................................................................................................................................... 81

BIOGRAPHICAL SKETCH ............................................................................................................................... 94
<table>
<thead>
<tr>
<th>Table</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold of 68) for MY (Kg)</td>
</tr>
<tr>
<td>2-2</td>
<td>General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold of 68) for FY (Kg x 100)</td>
</tr>
<tr>
<td>2-3</td>
<td>General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold of 68) for PY (Kg x 100)</td>
</tr>
<tr>
<td>2-4</td>
<td>GO terms significantly enriched with genes associated with MY</td>
</tr>
<tr>
<td>2-5</td>
<td>MeSH terms significantly enriched with genes associated with MY</td>
</tr>
<tr>
<td>3-1</td>
<td>General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold) for CI</td>
</tr>
<tr>
<td>3-2</td>
<td>General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlation and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold) for SCS</td>
</tr>
<tr>
<td>3-3</td>
<td>GO terms significantly enriched with genes associated with CI</td>
</tr>
<tr>
<td>3-4</td>
<td>MeSH terms significantly enriched with genes associated with CI</td>
</tr>
<tr>
<td>3-5</td>
<td>GO terms significantly enriched with genes associated with SCS</td>
</tr>
<tr>
<td>3-6</td>
<td>MeSH terms significantly enriched with genes associated with SCS</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2-1</td>
<td>Manhattan plot showing the results of genome-wide association mapping for MY.</td>
</tr>
<tr>
<td>3-1</td>
<td>Manhattan plot showing the results of genome-wide association mapping for CI.</td>
</tr>
<tr>
<td>3-2</td>
<td>Manhattan plot showing the results of genome-wide association mapping for SCS.</td>
</tr>
</tbody>
</table>
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTA</td>
<td>Bos taurus autosome</td>
</tr>
<tr>
<td>CI</td>
<td>Conception per insemination</td>
</tr>
<tr>
<td>CM</td>
<td>Clinical mastitis</td>
</tr>
<tr>
<td>DEG</td>
<td>Differentially expressed genes</td>
</tr>
<tr>
<td>DIM</td>
<td>Days in milk</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>FY</td>
<td>Fat yield</td>
</tr>
<tr>
<td>MY</td>
<td>Milk yield</td>
</tr>
<tr>
<td>PY</td>
<td>Protein yield</td>
</tr>
<tr>
<td>RR</td>
<td>Respiration rate</td>
</tr>
<tr>
<td>SCS</td>
<td>Somatic cell score</td>
</tr>
<tr>
<td>THI</td>
<td>Temperature humidity index</td>
</tr>
<tr>
<td>TNZ</td>
<td>Thermoneutral zone</td>
</tr>
</tbody>
</table>
Heat stress is an important economic issue in dairy farming, especially in the southern states of the US where the climate is sub-tropical and subject to prolonged periods of high ambient temperature and humidity. Genetic selection for heat tolerance is an attractive alternative for reducing the effects of heat stress on animal performance. Our first goal was to estimate variance components of Milk yield (MY), Fat yield (FY), Protein yield (PY), Conception per insemination (CI) and Somatic cell score (SCS) across lactations considering heat stress. Our second goal was to reveal genes and pathways responsible for thermotolerance. Multi-trait repeatability test-day models with random regressions on THI values were used to estimate variance components. The models included herd-test-day and DIM classes as fixed effects, and regular and heat tolerance additive and permanent environmental as random effects. Interestingly, genetic variance for all traits under study increased across parities, suggesting that cows become more sensitive to heat stress as they age. In addition, our study revealed negative genetic correlations of MY, FY, PY and CI and positive genetic correlation of SCS between general merit and thermotolerance. Whole-genome scans and gene-set analyses were carried out to identify genomic regions, individual genes and pathways
responsible for thermotolerance. Interestingly, some significant regions harbor strong candidate genes for thermoregulation, such as HSF1. The gene-set analysis revealed several functional categories, such as cellular response to heat, response to DNA damage stimulus, protein refolding that are involved in biological processes closely related with thermotolerance. Overall, this study contributes to a better understanding of the genetics underlying heat stress and points out novel opportunities for improving thermotolerance in dairy cattle.
CHAPTER 1
INTRODUCTION

Heat stress alters a variety of physiological functions affecting production, reproduction and health traits in dairy cattle thereby causing huge economic losses to the dairy industry. Heat stress is likely to become more prevalent in years to come as global climate change is expected to increase the frequency and duration of heat-stress related events. Dairy cows are more susceptible to heat stress because of higher metabolic heat production. Heat stress has been a huge issue in dairy farming especially in the southern states of the US where climate is sub-tropical and subject to prolonged periods of high ambient temperature and humidity. As a result, heat load in the cow increases to the point that the body temperature rises, intake declines, milk yield reduces, and health problem increases (Kadzere et al., 2002). In 2010, heat stress estimates the loss of annual milk production for average US dairy by about $39,000 accounting for total $1.2 billion loss in milk production for the US dairy sector (Key and Sneeringer, 2014). Hence, adopting strategies to alleviate heat stress and restore cow’s health, production and reproduction efficiency is necessary for improving profitability of the dairy farm. Over several years, heat stress has been defined as the function of temperature -humidity index. As a result, different approaches like physical modifications of the environment and improved nutritional and management practices have been used to alleviate the effects of heat stress in dairy cows. However, these practices increase production costs, and in general, they cannot eliminate heat stress completely. One complementary strategy for reducing the effects of heat stress on dairy cattle performance is the identification and subsequent selection of animals that are genetically more heat tolerant. Identification of such heat tolerant animal can be made
based on the measurement of their immediate response like rectal temperature, respiration rate among others. However, it is not feasible to use these records in national evaluation system because collection of such records at a national level would be time consuming and labor intensive. Alternatively, decline in performance due to heat stress can be used as an indicator trait of heat tolerance (Ravagnolo et al., 2000). The animal with minimum decline of performance per degree increase of temperature is identified as heat tolerant. Dry bulb temperatures combined with humidity in an index are used to assess the impacts of heat stress on dairy cattle performance. These climatic variables are available from public weather stations. Meteorological data from nearest weather stations can describe environmental conditions of farms (Freitas et al., 2006). Additionally, since this methodology does not require any additional measurements (e.g., body temperature), it can be applied to datasets as large as those utilized for the national evaluation.

The main objective of this research proposal was to perform a comprehensive genomic analysis of heat tolerance in dairy cattle. Production, reproduction and health records, climatic data, and genome-wide dense single nucleotide polymorphism markers were jointly analyzed using test-day models that include a random regression on a heat stress function based on THI values. This strategy allowed us to identify and characterize genomic regions, and individual genes and pathways, responsible for heat tolerance in dairy cattle. The relevant information thus generated in this study will help us better understand genetics underlying heat stress and subsequent use of marker-assisted selection in commercial breeding schemes.
CHAPTER 2
LITERATURE REVIEW

Cattle, being homeothermic animal, try to maintain relatively constant core body temperature regardless of changes in environmental temperature. The normal body temperature of cattle is around 38.6°C. It is important to note that the core body temperature of cattle is generally higher than the peripheral because of extensive metabolic heat production by internal organs. In cows, body temperature also vary with the stage of lactation, amount of milk production, nutritional and health status, stage of the estrous cycle and genetic diversity within a population (Wrenn et al., 1961, Shearer and Beede, 1990). Therefore, to maintain constant core body temperature, cows typically utilize different behavioral, physiological and immunological adaptations.

Several physiological variables such as rectal temperature, respiratory rate, and hormone concentrations have been used to predict body temperature in dairy cattle. For instance, when rectal temperature of the cow is greater than 39°C, it is very likely that the cow undergoes several physiological adaptations that compromises both performance and well-being. Curtis (1983) reported that rectal temperature is not an accurate indicator of body temperature in dairy cattle as there is a time lag between rectal temperature and body temperature. Respiration rate (RR) has also been used as an indicator of body temperature in dairy cows. However, RR varies according to the animal conditions, prior exposure to hot conditions, ambient environment conditions and cooling strategies. Similarly, there is also a time lag of two to four hours between body temperature and RR (Gaughan et al., 2000). High environmental temperature also reduces the plasma concentration of thyroid hormones in the cattle. However, it should be considered that increase in ambient temperature is accompanied with decrease in
feed intake and reduction in hormone concentrations. Therefore, decrease in feed intake should be considered as a confounding factor to correlate high temperature with low hormone concentration (Saber et al., 2009).

Body temperature can also be assessed by measuring meteorological variables such as temperature and humidity. Temperature Humidity Index (THI) formed by combination of temperature and humidity is used to predict potential heat stress conditions. These climatic variables are available from nearest public weather stations which can accurately describe environmental conditions of farms (Freitas et al., 2006). Additionally, this methodology has been used for genetic evaluation of heat tolerance in dairy cattle (Ravagnolo et al., 2000, Aguilar et al., 2009, Sanchez et al., 2009). Since this methodology does not require any additional measurements (e.g., body temperature), it can be applied to datasets as large as those utilized for national evaluations.

**Thermoneutral Zone**

Thermoneutral zone (TNZ) is a range of environmental temperature in which change in ambient temperature does not lead to change in animal heat production. In TNZ, animal is experiencing basal metabolic rate. As a result, animals can regulate temperature by non-evaporative physical process alone. Generally, temperature range of TNZ varies with age, species, feed intake, physiological status, behavioral and immunological conditions of the animal (Yousef, 1985). TNZ is bounded by lower critical temperature (LCT) at the lower end, and upper critical temperature (UCT) at the upper end. LCT is defined as the environmental temperature below which homeotherms must increase metabolic heat production to maintain constant core body temperature. As air temperature drops below the lower critical temperature, animals increase their
metabolic heat production by increasing feed intake, shivering and activating the brown adipose tissue (Carstens, 1994).

The upper end of the thermoneutral zone is the upper critical temperature (UCT) which is defined as the ambient temperature above which animals must decrease heat production to maintain thermal balance. The estimates of UCT for cows producing 30 kg milk per day range from 12 to 24°C (Berman and Meltzer, 1973). When air temperature exceeds the UCT, the heat gradient reduces, as a result, the potential for non-evaporative heat loss is reduced and animals depend on evaporative cooling such as sweating and panting to dissipate any excess heat received from the environment or generated by metabolism (McArthur and Clark, 1988). Above UCT, an increase in body temperature negatively influences animal performance and well-being as the cow enters heat stress conditions.

**Heat Stress**

Stress is the sum total of the forces outside the bodily system that acts to displace the system from the ground state (Yousef, 1985). Heat stress is a combination of air temperature, humidity, wind speed, solar radiation and other environment variables that makes it difficult for the cow to lose bodily heat to environment thereby increasing body temperature. Strain is an internal displacement of the bodily system (Yousef, 1985). Strain from heat stress is in fact, a change in body temperature in the animal that can be easily measured. Heat stress is an important economic issue in dairy farming that negatively impacts the performance and health of livestock. In dairy cows, heat stress decreases milk yield, reduces milk quality, and depresses fertility. Economic losses due to heat stress are estimated to be between $897 and $1500 million per year for the US dairy industry (St-Pierre et al., 2003). Since heat stress is an important
economic issue in dairy farming, it has been extensively reviewed by several authors (Kadzere et al., 2002, St-Pierre et al., 2003, West, 2003, Garcia-Ispierto et al., 2006).

**Responses to Heat Stress**

Adaptation is the animal response to reduce the magnitude of physiological strain caused by heat stress. Most of the adaptations responding to heat stress involve increasing heat transfer to the environment and reducing the production of metabolic heat (Kadzere et al., 2002). However, adaptive response themselves reduce physiological functions thereby compromising performance. For instance, during heat stress, milk production decreases because of the adaptive responses to prevent rise in body temperature. One of the physiological adaptive responses to regulate body temperature during heat stress include reduction in feed intake. When cows eat less, less of the energy will be available for milk production thereby causing reduce in milk production (West, 2003). Another reason why performance suffers during heat stress is failure to successfully adapt to high temperature during heat stress thereby disrupting physiological functions. One example could be reduced fertility during heat stress owing to the effects of elevated temperature on oocyte and embryo development (Hansen, 2009). Thus, maintenance of homeothermy is critical for both performance and well-being of the cow.

**Methods of Heat Transfer**

Homeotherms attempt to balance heat production and heat loss. Heat stress occurs when sum of heat produced from metabolic heat and heat received from the environment exceeds heat lost to the environment. Heat is exchanged between the cow and the environment via two types of mechanisms, viz., sensible heat loss mechanism and latent heat loss mechanism. Sensible heat loss mechanisms, which includes
conduction, convection, and radiation, depend on temperature gradient between animal surface and environment on which animal is exchanging heat. Sensible heat loss reduces as the temperature gradient between environment and cow’s surface temperature decreases (Maia et al., 2005). After sensible heat loss mechanisms become ineffective, the latent heat loss becomes effective. Latent heat loss mechanism, which includes evaporation and condensation, occurs along the vapor gradient between the animal surface and the environmental vapor pressure. When the environmental humidity is high, little heat is lost by evaporative cooling. Therefore, when cows are exposed to high air temperature coupled with high humidity, it becomes increasingly difficult for the cow to lose heat to the environment thereby building increasing heat load in the body.

**Conduction and Convection**

Conduction refers to physical transfer of heat from warm object to cool object without any object moving. Examples are from cow feet to floor. Cows undergo physiological adaptations to regulate conductive heat loss. During heat stress, cows increase blood flow from core to the periphery, thereby increasing surface temperature. This is possible because cows have special adaptations in blood vessels that help increase blood flow to the skin. During periods of heat stress conditions, when body temperature rises, anastomosis between arterioles and venules open thereby allowing rapid flow of blood from core to skin (Walloe, 2016). If exists a temperature gradient between skin surface and environment, then heat lost occurs to the environment.

Convective heat exchange refers to transfer of heat from warm substance to cool substance where the substances are moving past each other. Examples are from cow to wind. Like conductive heat loss, cows can also regulate convective heat loss through
behavioral modification, posture, orientation and by increasing the convective coefficient of wind around animal body (Kadzere et al., 2002).

**Radiation**

The radiant energy from the sun is solar radiation. The heat exchange with solar radiation refers to production and absorption of electromagnetic radiation by animal surface when exposed to sunlight. The amount of heat absorbed by an animal exposed to direct sunlight depends on the absorptivity of the surface. The absorptivity of the white-coat cow is 0.50 while the black coat-cow is 0.90 (Shearer and Beede, 1990) which means net transfer of heat via solar radiation is 1.8 times higher for black coat cows as compared to white coat cows when exposed to sunlight. Radiation is another important medium of heat loss especially during night when the animal radiates heat to the cooler sky. However, high humidity and clouds reduce heat loss by radiation (Fuquay, 1981). When there is no cloud in the sky during night, the temperature gradient between outer sky and the animal surface is greater, allowing large flow of heat from animal surface to outer environment. However, in a cloudy night, a cow outside at night will lose less heat because clouds have a warmer temperature than objects in the outer space.

**Evaporative Heat Loss**

Evaporative heat loss is an important avenue for heat exchange particularly when the temperature gradient between cows’ surface and air temperature falls. Evaporative heat loss occurs when dew point temperature of the air surrounding animal is lower than evaporative surface temperature of the animal (West, 2003). Increased air speed and low humidity facilitates evaporative heat loss. When air temperature is about
40 °C, approximately 84% of the total heat loss is through evaporative heat loss, such as sweating (Yousef, 1985).

**Physiological Responses of Cattle to Heat Stress**

Cows have TNZ within which no additional energy above maintenance is needed to regulate the body temperature. Hyperthermia results from a negative energy balance between heat produced and absorbed by the animal with respect to heat lost to the environment. During heat stress, animals undergo several behavioral and physiological changes from sub-cellular to the whole animal level. All these behavioral and physiological responses involve the expenditure of energy to remove or reduce the impacts of heat stress. Sweating, high respiration rate, vasodilation with increased blood flow to skin surface, reduced metabolic rate, decreased dry matter intake, and altered water metabolism are the physiological responses animals typically undergo during heat stress. Most of these adaptations have in general negative impacts on animal performance, health, and well-being (Shearer and Beede, 1990).

**Reduced Feed Intake**

Feed intake starts to decline at ambient temperature of 25-26 °C and drops more rapidly above 36 °C in dairy cows. Reduced feed intake has been identified as one of the main causes of reduced milk production in dairy cows during heat stress conditions. Heat stress stimulates the medial satiety center of hypothalamus which, in turn, inhibits the lateral appetite center thereby declining both reduced dietary intake and milk production (Albright and Alliston, 1971). Maust et al. (1972) reported a negative genetic correlation between rectal temperature and feed intake which suggests that elevated body temperature reduces feed intake. Similarly, Bouraoui et al. (2002) reported a
9.6% decrease in DMI when THI increases from 69 to 78. According to Baumgard et al. (2013), about 50% decrease in MY is due to reduced feed intake.

**Decreased Milk Production**

Heat stress negatively impacts milk production. Several studies have reported decline in milk production due to high environmental temperature (Ravagnolo and Misztal, 2000, Aguilar et al., 2009, Sanchez et al., 2009). This may be because heat stress has negative effects on energy balance and secretory functions of the udder (Silanikove, 1992). For instance, Ravagnolo et al. (2000) reported loss of milk production per cow per year around 0.2 kg per unit of THI greater than 72. West et al. (2003) found a negative relationship between heat stress and milk production. However, it is important to note that there is a lag effect between the onset of high temperature and humidity on milk production. The lag could be due to reduction in feed intake, a delay between intake and utilization of nutrients, and hormonal changes (West, 2003).

The stage of lactation at which cows’ experience heat stress also affects the effect of heat stress on milk yield. Cows in mid lactation are most adversely affected, cows in early lactation are least affected, while those in late lactation are intermediately affected. Under mild conditions of heat stress, Abeni et al. (2007) reported that heat stress reduces milk yield in early, mid and late stages of lactation by 13%, 24% and 16.5%, respectively. This could be because cows in early lactation produce highest MY and decrease in feed intake due to heat stress is compensated with the catabolism of body fat to replace deficits in energy intake. However, the effects of heat stress are more severe in cows of mid lactation as there are the highest alterations of blood parameters related with energy balance and enzyme activities. High-milk-producing cows are more affected by heat stress compared to low-milk producing cows (Kadzere
et al., 2002). Multiple-parity cows are known to be more susceptible to heat stress than first-parity cows (Aguilar et al., 2009). Similarly, heat stress during the dry period affects milk production in the subsequent lactation, as it was reported by several authors (West, 2003, Avendano-Reyes et al., 2006, Tao et al., 2011).

Decreased Reproduction

Heat stress negatively impacts most aspects of the reproductive functions in mammals. Heat stress depresses estrous behavior, delays oocyte maturation, reduces conception rate, and retards early embryonic development and growth. Badinga et al. (1993) reported that heat stress on the day of ovulation reduces the size of dominant follicle during estrous cycle. Hansen and Arechiga (1999) reported that heat stress reduces expression of estrous, resulting in less number of cows eligible for embryo transfer during summer season. Ravagnolo and Misztal (2002) reported a negative genetic correlation of -0.35 between non-return rate at 90 days (NR90) and THI, suggesting an unfavorable relationship between heat tolerance and reproduction.

The deleterious effects of heat stress on reproductive functions could be due to the results of hyperthermia affecting the functions or physiological adjustments made by the heat-stressed animal to regulate body temperature (Hansen, 2009) Several studies have reported that heat stress decreases oocyte competence for fertilization. Even if high temperature does not affect oocyte competence for fertilization, the resultant embryo develops very slowly and abnormally (Hansen, 2013). Many of the effects of heat stress on oocyte and gametes are closely related with the production of reactive oxygen species which affect the quality of the oocytes and embryos, fertilization and pregnancy success.
Increased Water Intake

Heat stress conditions alter water intake in lactating dairy cows (Murphy et al., 1983). Heat stressed cows increase water intake above the minimal need for maintenance and metabolism. This is because cows must compensate for losses of water from evaporative heat loss through sweating and panting during the periods of heat stress. Also, cows lose large amount of potassium through sweating. So, mineral salts containing potassium are recommended to be included in the feed of heat stressed cows (Shearer and Beede, 1990).

Increased Respiration Rate

Cows undergo physiological adjustments during high ambient temperature, including increased respiration rate (Coppock et al., 1982). In a study involving high-producing dairy cows in heat-stress conditions, Berman et al. (1985) found that respiratory rate is 50-60 breaths per minute higher than in a thermo-neutral environment. Increase in respiration rate is associated with increase in evaporative heat loss (e.g. panting) which helps regulate animal body temperature.

Changes in Blood Hormones Concentration

Secretion of hormones associated with metabolism and water balance are altered during heat stress. The hormones associated with adaptation to heat stress include prolactin (PRL), growth hormone (GH), thyroid hormones, glucocorticoids, mineralocorticoids, acetacholamines and antidiuretic hormone (Farooq et al., 2010). Giesecke (1985) reported an increase in the concentration of plasma PRL during thermal stress in dairy cows. This is because PRL is involved in regulating water and electrolyte demands of heat stressed cows. Growth hormone, a calorigenic hormone released from anterior pituitary, is also altered during heat stress. The decreased
concentration of GH leads to less heat production, which helps regulate body
temperature during heat stress. The thyroid hormones triiodothyronine (T3) and
thyroxine (T4) are primary hormones involved in basal metabolic rate and their
concentrations are found to be reduced during heat stress (Purwanto et al., 1991). The
decline in thyroid hormones along with decline in GH has synergic effect to reduce heat
production and regulate thermal temperature in dairy cattle. There is an increase in the
concentration of glucocorticoid hormone due to activation of adrenocorticotropic
hormone releasing factor in the hypothalamus during heat stress (Farooq et al., 2010).
The glucocorticoids work as vasodilators, thus increasing blood flow from core to
periphery helping regulate body temperature. There is increased concentration of
catecholamine during heat stress (Farooq et al., 2010). Catecholamine is involved in
regulating sweat gland activity. Similarly, there is increased concentration of antidiuretic
hormone (ADH) during heat stress conditions in dairy cattle. Increased water loss
through evaporative cooling in the respiratory tract and skin invokes increases secretion
of ADH, which in turn reduces water loss and increases water intake.

Methods for Assessment of Heat Stress

Rectal Temperature
Rectal temperature is commonly used as indicator trait to predict core body
temperature in dairy cattle. During heat stress condition, rectal temperature rises above
However, there is a limitation in the use of rectal temperature for quantification of heat
stress in dairy cattle because rectal temperature rises slowly in response to rise in
temperature (Curtis, 1983). Moreover, the process of measuring rectal temperature
through thermometer is labor intensive, time consuming, and prone to errors such as
variability probe, and variation in animal handlers (West et al., 2003). Additionally, restraining animals to collect rectal temperature may subject animals to further stress that can alter temperature, therefore rectal temperature might not provide a very accurate measure of body temperature in dairy cattle (Hahn et al., 1990, Prendiville et al., 2003).

**Milk Temperature**

Milk temperature has also been used as an indicator of body temperature in dairy cattle. Several studies have used milk temperature to monitor body temperature and have reported correlations between milk temperature and body temperature between 0.78 and 0.99 (Fordham et al., 1984, Igono et al., 1987, West et al., 1990). However, West (2003) reported that milk temperature is very sensitive to climatic variables at the time of milk collection. Indeed, West (2003) found that the climatic variable having the greatest influence on cow morning milk temperature is the current day minimum air temperature, while evening milk temperature is most influenced by current day mean air temperature. Moreover, it seems milk temperature is also sensitive to other variables. For instance, West et al. (1990) reported that milk temperature was greater for cows administered bST compared with controls, and Igono et al. (1990) observed higher milk temperature in high-producing cows compared to low-producing cows.

**Respiration Rate**

Respiration Rate (RR) has also be used as an indicator of heat stress in cattle. RR can be measured by counting flank movement. Cows exposed to hot weather conditions show increased respiration rate, a physiological adjustment to high ambient temperature (Coppock et al., 1982). Berman et al. (1985) reported that respiratory frequency increases above 50-60 breaths/minute at ambient temperature higher than 25
°C in dairy cows. Increase in RR during hot humid conditions is mainly due to increase in relative humidity which reduces respiratory and surface evaporation thereby increasing body temperature. It should be noted that RR can vary according to animal conditions, prior exposures to hot conditions, any cooling strategies applied and other environmental factors. Since there is a time lag between RR and body temperature, it is recommended that RR measurements are made two to three hours prior to the hottest time of the day (Gaughan et al., 2000).

**Heat Shock Proteins (HSPs)**

Heat shock proteins (HSPs) have been implicated in thermotolerance in cells. In fact, HSPs are molecular chaperones that protect cells from thermal damage, prevent protein denaturation and block apoptosis (Kampinga, 2006). Based on molecular weight, HSPs are classified into five families viz., Hsp100, Hsp90, Hsp70, Hsp60 and the small HSPs. These proteins are found to be produced in increased amount in response to heat shock and are involved in folding, unfolding and refolding of proteins (Sorensen et al., 2003). In livestock, the Hsp70 family has been extensively studied in relation to its involvement in thermo-tolerance. The cellular response to heat stress is regulated at the transcription level and controlled by inducible expression of Hsp70 genes. However, the biosynthesis and chaperonin activity of Hsp70 requires a critical threshold level of temperature above which Hsp70 gene will be induced (Somero, 2002). Hsp70 genes prevent aggregation of misfolded proteins and facilitate folding of misfolded protein to the native state (Mayer and Bukau, 2005).
Strategies for Reduction of Effects of Heat Stress

Several strategies have been commonly used to reduce the effects of heat stress in dairy cattle, including physical modifications of the environment, nutritional strategies and genetic development of heat resistant animals.

Physical Modifications of the Environment

Physical modification of the environment is a commonly used strategy to reduce the effects of heat stress in dairy cows. This strategy aims to either modify the environment to reduce the exposure of heat stress on cows (e.g. shading) or to enhance heat exchange between cows and surrounding (e.g. fans, sprinkles). However, the degree of abatement of heat stress varies with the climate and production level of the cows. In a hot and dry climate, shading is a cost-effective solution for minimizing heat stress in dairy cattle. Indeed, shading reduces direct solar radiation and helps reduce heat load on the animals. Several studies have indicated that cows with access to shade had lower increase in rectal temperature than those without access to shade in hot climate (Fisher et al., 2008, Collier and Gebremedhin, 2015, Fournel et al., 2017).

Shaded environment also impacts production and reproduction performance of dairy cows. Cows in shade generally produce 0.7 kg more milk than cows which are not in shade and the conception rate generally increases from 25.3% in an unshaded cows to 44.4% in a shaded cows (Romanpence et al., 1977).

The use of sprinklers combined with fans is another effective method to promote heat loss during high temperature. Sprinkles produce droplets that wet the cow’s surface. Fans, when used with sprinkles, cause evaporative cooling to take place on the surface. The impact of sprinkles combined with fans was assessed on dairy cows in hot conditions. In a study conducted by Strickland (1989) in Florida, sprinkles and fans
reduced the RR from 95 to 57, increased DMI from 17.8 to 19.1 kg, and increased milk yield from 18.1 to 20.2 kg.

**Nutritional Strategies for Managing Heat Stress in Dairy Cows**

Nutritional management is another commonly used strategy to reduce the effects of heat stress in dairy cows. Nutritional management includes changes in feeding schedule, and ration composition. High quality forages with less fibers produce less heat than high fiber forages during digestion (West, 1999). Therefore, it is necessary to provide good quality forages during heat stress as cows during hot weather will reduce feed intake. Additionally, it is also necessary to increase concentrate of certain minerals supplements to compensate for the loss from the body during heat stress. For instance, it is recommended that cows should be supplemented with 1.5-1.6% DM of potassium and 0.5-0.6% DM of sodium to improve milk yield when cows are under heat stress conditions. Feeding 150–200 g/cow/day of sodium bicarbonate during hot conditions helps buffer rumen and improves the appetite of the dairy cows. Additionally, to minimize heat generation, it is a good strategy that the greater part of the cows’ ration should be fed during cooler period of the day (e.g., between 4.00 and 6.00 am, and 9.00 and 11.00 pm). Another strategy could be, instead of feeding large bulk of feed, provide smaller amount of feed and increase the number of feedings per day (Beede and Collier, 1986).

**Genetic Development of Heat Resistant Animals**

Improving milk production has been a major breeding goal in dairy cattle over several decades. As a result, over the last fifty years, average milk production in US dairy cattle increased by 5,997 kg with 56% of this increase due to genetic selection (VanRaden, 2004). However, several studies have reported that intense selection for
production traits has led to decline in thermo-regulatory physiology of cows (Ravagnolo et al 2002; Kadzere et al., 2002; Dikmen et al. 2012). High producing cows are characterized by larger body size, consumption of more feed and production of more metabolic heat. Moreover, thermoneutral zone of high producing cows has shifted to lower temperature thereby making high producing cows more susceptible to heat stress.

Several studies have reported that heat tolerance is a heritable trait (Ravagnolo and Misztal, 2000, Dikmen et al., 2012, Garner et al., 2017). So, genetic selection can be utilized to increase heat tolerance in dairy cattle. Some studies have estimated the heritability of indicator traits of heat tolerance like body temperature and respiration rate. For instance, Dikmen et al. (2013) estimated the heritability of rectal temperature for US Holsteins around 0.13 to 0.17, which allows selection for lower rectal temperature under heat stress. Body temperature heritability estimates range from 0.11 to 0.68 (Mackinnon et al., 1991, Burrow, 2001, Howard et al., 2014). Heritability estimates of respiration rate range from 0.76 to 0.84 (Seath and Miller, 1947). The estimates of these indicator traits for heat stress show that there is enough genetic variability to allow selection for heat tolerance. There are also studies indicating genes underlying heat tolerance in dairy cattle. For instance, Olson et al. (2003) reported slick hair gene in the bovine genome. This gene can maintain body temperature at lower rates. Several studies have reported the presence of heat shock genes as candidate genes for thermotolerance in dairy cattle (Lee et al., 2006, Page et al., 2006, Xiong et al., 2013). These are molecular chaperones that protects cells from thermal damage, prevents protein denaturation and blocks apoptosis (Kampinga, 2006).
Using random regression models, Ravagnolo et al. (2000) estimated genetic components of heat tolerance in dairy cattle at different values of THI. At THI ≥72, the study estimated the drop of 0.2 kg/unit of THI for milk. The study also reported a negative correlation \( r = -0.35 \) between milk production and thermodurability which means selection for milk yield will make cows more susceptible to heat stress. The study also estimated the breeding value (BV) under no heat stress condition and another BV for performance under heat stress conditions. The application of this model helps calculate simultaneously the genetic merit for performance and the genetic merit for heat tolerance, and therefore help predict rankings of animals in various environments with different climatic conditions.

**Genetic Studies on Heat Stress**

Different approaches including cooling, shading and nutrition are commonly used to mitigate the effects of heat stress. However, these practices increase production costs and in general, they cannot eliminate heat stress completely. One of the complementary strategies of reducing the effects of heat stress is the selection of heat tolerant sires. Several studies have reported the existence of additive genetic variability for heat stress in dairy cattle (Ravagnolo et al., 2000, Aguilar et al., 2009, Sanchez et al., 2009, Nguyen et al., 2016, Macciotta et al., 2017). Nguyen et al. (2017) estimated breeding value for heat tolerance in Australian dairy cattle which provides opportunity to breed cows that are more tolerant to heat with less impact on milk production. Note that selective breeding of dairy cattle for thermodurability is permanent, cumulative, and hence, the most cost-effective approach for mitigating the effects of heat stress in dairy cattle.
Test-day models have been used for genetic evaluation of thermotolerance in dairy cattle. Nguyen et al. (2017) used random regression test day models to estimate genomic estimated breeding values (GEBV) for heat tolerance for milk, fat and protein yield in Australian dairy cattle. The use of test-day models provide an opportunity to accommodate different recording schemes, utilize incomplete lactation records and account for time-dependent effects for each test-day (Swalve, 2000). Many studies in dairy cattle have used repeatability test-day models in which the shape of the lactation curve has been modeled either using fixed regression with fixed effects (Ptak and Schaeffer, 1993) or with function of days in milk.

The actual measurements of heat load on an animal could be performed by measuring rectal temperature, milk temperature or respiratory rate, but this information is expensive and/or difficult to measure in large populations. On the other hand, a linear regression of regularly recorded traits such as milk production, health traits and reproductive traits on climatic data can be fitted to predict the relationship between performance and weather conditions. Ravagnolo et al. (2000) proposed a model to study genetic components of heat stress in dairy cattle by using performance data augmented with public weather information. This model assumes that cow performance is unaffected until a certain level of THI, and above that level the performance declines linearly with increasing THI. Once THI for heat stress is estimated, a heat load function can be calculated as \( HL = \max(0, \text{THI} - \text{THI}_T) \), where THI is the observed THI value for a given day, \( \text{THI}_T \) is a threshold of THI value, and HL is the heat load value. Once the heat load is defined, it can be incorporated in reaction norm models for evaluating genetic tolerance to heat stress. In reaction norm analysis, the weather data (THI) is
treated as continuous, and performance data are regressed on this continuous scale. However, the decline in performance as indicated by the slope is different for different cows. If the variation in the slope contains sizeable additive genetic components, then selection for heat tolerant animals is possible.
CHAPTER 3
GENETIC ANALYSIS OF HEAT TOLERANCE FOR PRODUCTION TRAITS IN US HOLSTEIN COWS

Background

Improving animal’s ability to cope with adverse environmental condition has been a challenge in animal industry. Adverse environmental conditions, especially heat stress, are important economic issues in dairy farming as they negatively impact the production of dairy cows (Aguilar et al., 2010b, Nardone et al., 2010, Biffani et al., 2016). During heat stress, dairy cows undergo several behavioral and physiological changes from sub-cellular to whole animal level, thereby affecting production traits (West, 2003, Nardone et al., 2010). Annual economic losses due to heat stress are estimated between $897 and $1,500 million for the US dairy industry (St-Pierre et al., 2003). Since heat stress is an important economic issue in dairy farming, different approaches including cooling, shading and nutrition are commonly used to mitigate the effects of heat stress. However, these practices increase production costs and in general, they cannot eliminate heat stress completely. One of the complementary strategies of reducing the effects of heat stress is the selection of heat tolerant animals. Some indicator traits of response to heat stress such as rectal temperature and respiration rate exhibit a sizable genetic component. For instance, Dikmen et al. (2013) estimated heritability of rectal temperature in US Holsteins around 0.13 to 0.17. Similarly, heritability estimates of RR range from 0.76 to 0.84 (Seath and Miller, 1947). Heritability estimates of these indicator traits suggest that genetic selection for thermotolerance is feasible. However, the inclusion of these indicator traits in a national genetic evaluation is both expensive and time consuming.
An alternative methodology to evaluate heat tolerance is to examine changes in performance over environmental variables (Ravagnolo et al., 2000). In genetic evaluation of heat stress, the environmental variable is often expressed as Temperature-Humidity Index (THI), a combination of temperature and humidity, obtained from nearest public weather stations. In this model, a linear regression of regularly recorded traits such as MY, FY and PY on climatic data is fitted to predict the relationship between production and weather conditions. This model assumes that production is unaffected until a certain level of THI, and above that level the production declines linearly with increasing THI. For genetic evaluation of heat stress, several studies have assumed that all animals have a common threshold above which production declines linearly (Ravagnolo and Misztal, 2000, Aguilar et al., 2009, Macciotta et al., 2017). Sanchez et al. (2009) estimated however, a separate threshold and slope for each animal and reported a strong genetic correlation of 0.95 between these two components, indicating that animals with higher threshold of heat stress would also have lower rate of performance decline and vice-versa. Therefore, a simple model assuming constant threshold and different slopes would suffice for genetic evaluation of heat stress in dairy cattle.

The identification of genomic regions, and preferably individual genes and pathways, responsible for genetic variation of thermodtolerance will enhance understanding of biological mechanisms and point out novel opportunities for improving thermodtolerance through selective breeding. However, there is a limited information on genes and genomic regions associated with thermodtolerance. Dikmen et al. (2013) reported a genomic region on BTA24 harbors potential candidate genes for regulating
rectal temperature in dairy cows. Macciotta et al. (2017) reported a genomic region on BTA26 that harbors a potential candidate gene \( BTRC \) known to be associated with thermotolerance for MY.

The first objective of this study was to estimate variance components of MY, FY and PY across lactations using multi trait repeatability test-day models considering heat stress. Test-day records of MY, FY and PY were merged with daily THI values, and were jointly analyzed using test-day models that include random regressions on a function of THI values. The second objective of this study was to identify and characterize genomic regions, and preferably individual genes and pathways, underlying thermotolerance. For this, test-day records of MY, weather data, and genome-wide SNP markers were jointly analyzed. As a result, this study will contribute to better understanding of the genetics underlying heat stress and subsequent use of marker-assisted selection in breeding program.

**Materials and Methods**

**Phenotypic and Genotypic Data**

Data comprised 254,215 MY, FY and PY test-day records on 17,522 Holstein cows calved from 2006 through 2016 on two dairy farms viz., North Florida Holsteins and Dairy Research Unit of University of Florida, USA. Lactation records were required to have at least 8 test-days and be between 5 and 305 DIM. Pedigree file was created by tracing the pedigree of cows back to five generations available at the Council on Dairy Cattle Breeding website. The pedigree file included 35,006 animals.

Genotype data for 60,671 single nucleotide polymorphism (SNP) markers were available for 4,776 cows with MY test-day records and 1,592 sires in the pedigree. The SNP data were provided by the Council on Dairy Cattle Breeding and the Cooperative
Dairy DNA Repository. Those SNP markers that mapped to sex chromosomes or were monomorphic or had minor allele frequency less than 1% were removed from the dataset. After data editing, a total of 57,954 single nucleotide polymorphisms (SNPs) were retained for subsequent genomic analyses.

Weather data were obtained from Florida Automated Weather Network for Alachua county and hourly THI was calculated as proposed by Ravagnolo and Misztal (2000) as:

$$\text{THI} = (1.8 \times \text{temp} + 32) - (0.55 - 0.55 \times \text{rh}) \times (1.8 \times \text{temp} - 26)$$  \hspace{1cm} (2-1)

where temp is the temperature in degree Celsius and rh is the relative humidity in percentage. Mean daily THI of 3 days prior the test day was assigned to each test-day record as suggested by Bohmanova et al. (2007).

A function of THI ($f(\text{THI})$) was created to estimate reduction in MY due to heat stress. Therefore,

$$f(\text{THI}) = \begin{cases} 0 & \text{if } \text{THI} \leq \text{THI}_{\text{thr}}. \\ \text{THI} - \text{THI}_{\text{thr}}. & \text{if } \text{THI} > \text{THI}_{\text{thr}}. \end{cases} $$  \hspace{1cm} (2-2)

where the value of THIthr. was set to 68 and thus $f(\text{THI})$ was equal to max (0, THI - THIthr.)

**Statistical Model**

The following multi-trait repeatability test-day model was used to estimate the variance components of production traits, considering multiple lactations as different traits:

$$y_{\text{klmn}} = \text{HTD}_{\text{kl}} + \text{DIM}_{\text{m}} + a_{\text{nl}} + p_{\text{nl}} + v_{\text{nl}} [f(\text{THI})] + q_{\text{nl}} [f(\text{THI})] + e_{\text{klmn}}$$  \hspace{1cm} (2-3)
where $y_{klmn}$ is the record for MY, FY, and PY, $\text{HTD}_{kl}$ is herd-test-day $k$ within parity $l$ ($l = 1, 2, 3$), $\text{DIM}_m$ is the $m^{th}$ DIM class with classes defined every 20 days, $a_{nl}$ is the general random additive genetic effect of animal $n$ in parity $l$, $f(\text{THI})$ is the heat stress function for herd test day $k$, $v_{nl}$ is the random additive genetic effect of heat tolerance for the animal $n$ in parity $l$, $p_{nl}$ is the random permanent environmental effect of cow $n$ in parity $l$, $q_{nl}$ is the random permanent environmental effect of heat tolerance of the cow $n$ in parity $l$ and $e_{klmn}$ is the random residual effect.

Let $a = [a'nl \ v'nl]$ be vector of random additive genetic effects and $p = [p'nl \ q'nl]$ be a vector of random permanent environment effects for parities $n = 1$ to $3$.

The variance-covariance structure was

$$V = \begin{bmatrix} A \otimes G & 0 & 0 \\ 0 & I \otimes P & 0 \\ 0 & 0 & I \otimes R \end{bmatrix}$$

where $A$ is the numerator relationship matrix, and $G$ and $P$ are 6x6 matrices of (co)variances for additive and permanent environment effects respectively. $R$ is a 3x3 diagonal matrix of residual variances corresponding to each trait.

**Statistical Analysis**

Initial estimates for multiple trait analyses were obtained from Aguilar et al. (2009). Variance component for MY, FY and PY using multi-trait repeatability test-day models were estimated in Bayesian framework using GIBBS2F90 (Misztal et al., 2002). Genomic data were not included for variance components estimation. Of a total of 500,000 samples, first 100,000 were discarded as burn-in, and every 100th sample was retained to calculate variance components and posterior standard deviations.

Heritability for the generic merit of MY was estimated as
\[ h^2 = \frac{\sigma_a^2}{\sigma_{\text{total}}^2} \]  \hspace{1cm} (2-5)

Where

\[ \sigma_{\text{total}}^2 = \sigma_a^2 + \sigma_p^2 + \sigma_e^2 \]  \hspace{1cm} (2-6)

And heritability for MY at heat stress level \( f(i) \) was calculated as

\[ h^2 = \frac{\sigma_a^2 + f(i)^2 \sigma_v^2 + 2f(i)\sigma_{av}}{\sigma_a^2 + f(i)^2 \sigma_v^2 + 2f(i)\sigma_{av} + \sigma_p^2 + f(i)^2 \sigma_q^2 + 2f(i)\sigma_{pq} + \sigma_e^2} \]  \hspace{1cm} (2-7)

The genetic correlation between the generic and heat tolerance additive effects was calculated as

\[ \text{corr} [a,f(i)v] = \frac{f(i)\sigma_{av}}{\sqrt{\sigma_a^2 * f(i)^2 \sigma_v^2}} \]  \hspace{1cm} (2-8)

**Genome Wide Association Mapping using ssGBLUP**

The whole-genome association mapping was performed using single-step genomic BLUP methodology (ssGBLUP). The ssGBLUP is indeed a classical BLUP but it replaces the inverse of the pedigree relationship matrix \( (A^{-1}) \) with the inverse of the realized relationship matrix \( (H^{-1}) \) that combines both pedigree and genomic information (Aguilar et al., 2010a). The combined pedigree genomic relationship matrix \( H^{-1} \) was calculated as follows

\[ H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{zz}^{-1} \end{bmatrix} \]  \hspace{1cm} (2-9)

where \( G^{-1} \) is the inverse of the genomic relationship matrix, and \( A_{zz}^{-1} \) is the inverse of the numerator relationship matrix for genotyped animals. Here, \( G^{-1} \) has the dimension of 6,368 x 6,368 which includes 4,776 cows with test-day service records and 1,592
sires in the pedigree. Similarly, A matrix has a dimension of 35,006 X 35,006 which is based on a five generation pedigree.

Candidate genomic regions associated with general merit and thermotolerance additive merit for MY were identified based on the amount of genetic variance explained by 2.0 Mb moving windows of adjacent SNPs. Given the genomic estimated breeding values (GEBVs), the SNP effects can be estimated as $\hat{s} = DZ'[DZ]^{-1} \hat{a}g$, where $\hat{s}$ is the vector of SNP marker effects, D is a diagonal matrix of weights of SNPs, and $\hat{a}g$ is the vector of GEBVs (Wang et al., 2012). The percentage of genetic variance explained by a given 2.0 Mb of moving window of adjacent SNPs was then calculated as

$$\frac{\text{Var}(u_i)}{\sigma^2_i} \times 100 = \frac{\sum_j (z_j)^2 s_j^2}{\sigma^2_i} \times 100 \tag{2-10}$$

where $u_i$ is the genetic value of the $i^{th}$ genomic region under consideration, B is the total number of adjacent SNPs within 2.0 Mb region, and $s_j$ is the marker effect of the $j^{th}$ SNP within the $i^{th}$ region. All the ssGBLUP calculations were performed using BLUPF90 (Misztal et al., 2002).

**Gene Set Analysis**

The gene set analysis includes basically three different steps: i) assignment of SNPs to genes ii): assignment of genes to pathways iii) Pathway-based association analysis

**Assignment of SNPs to genes**

The UMD 3.1 bovine genome sequence assembly was used for SNP assignment using the Bio Conductor R package biomaRt (Zimin et al., 2009). SNPs were assigned to genes if they were located within the genomic sequence of the gene or at most 15 kb either upstream or downstream the gene. If a SNP was found to be located within the
gene or at most 15kb either upstream or downstream the gene, that gene will be included for further analysis. Finally, a gene was considered significant gene for MY, if the gene is flagged by a top 1% SNP that explained the highest genetic variances for MY.

**Assignment of genes to pathways**

Gene Ontology (GO) (Ashburner et al., 2000) and Medical Subject Headings (MeSH) (Nelson et al., 2004) databases were used to define functional sets of genes. Genes assigned to the same functional category are considered more closely related rather than random sets of genes.

**Pathway-based association analysis**

The association of a given pathway with MY was analyzed using Fisher’s exact test. This test, based on cumulative hypergeometric distribution, enables to search for an overrepresentation of significant genes among all the genes in the pathway. The P-value of observing $k$ significant genes in the pathway was calculated by

$$\text{P-value} = 1 - \sum_{i=0}^{k-1} \frac{\binom{S}{i} \binom{N-S}{m-i}}{\binom{N}{m}}$$

(2-11)

where $S$ is the total number of genes that were significantly associated with MY, $N$ is the total number of genes that were analyzed in the study, and $m$ is the total number of genes in the pathway (Abdalla et al., 2016). The GO gene set enrichment analysis was performed using the R-package goseq (Young et al., 2010) whereas MeSH enrichment analysis was carried out using the R-package meshr (Tsuyuzaki et al., 2015).
Results and Discussion

Genetic Dissection of Heat Stress for MY, FY and PY

Variance components of general production level (intercept) and animal ability to respond to heat stress (slope) for MY, FY and PY using multi-trait repeatability test day models (MREP) were estimated respectively (Table 2-1, Table 2-2, Table 2-3). Similarly, additive genetic variances for thermotolerance and relevant genetic parameters, namely heritability and genetic correlation at heat stress level $\phi$ (THI) = 78 (i.e. 10 units above THI threshold of 68) were also estimated.

Estimates of variance components for both generic and heat stress additive effects were comparable to those estimated by Aguilar et al. (2009) for first three parities in US Holsteins, however, the estimates reported in this study are higher than those reported by Ravagnolo and Misztal (2000). Both generic and heat tolerance additive genetic variances for MY, FY and PY increased across parity. Genetic variances for MY under-heat stress increased by 8.31% from parity one to parity two and 5.18 % from parity two to parity three suggesting that cows become more sensitive to heat stress as they age. Additive genetic variances for thermotolerance increased from first to second parity in both FY and PY. Increase from first to second parity were more than 100% for FY and more than 16% for PY.

This is in agreement with the findings of Bernabucci et al. (2014) who also reported that multiple-parity cows are more susceptible to heat stress than first parity cows. Genetic correlations between generic and heat tolerance effects for MY in all three lactations were negative, ranging from -0.30 to -0.55. Genetic correlations between generic and heat tolerance effects for FY and PY in all three lactations were also negative, ranging from -0.18 to -0.68. This is in agreement with the findings of
Aguilar et al. (2009) and Ravagnolo and Misztal (2000) in US Holsteins. Our findings suggest that there is unfavorable genetic correlation between MY, FY and PY and cow thermotolerance, and that continued selection for MY, FY and PY without consideration for thermotolerance will result in greater susceptibility to heat stress.

Heritability estimates for MY at THI 78 were between 0.17 to 0.32 across lactations, which is comparable with the heritability estimates reported by Ravagnolo et al. (2000). Similarly, heritability estimates for FY for first three lactations were between 13 to 20 % whereas heritability estimates for PY across first three lactations were between 18 to 26%. The estimates of heritability decreased across the parity. This could be because phenotypic variances increase across the parity as cows become more sensitive to heat stress in later lactations. Genetic correlations between parities for general additive effect for MY were positive and high (≥0.82) whereas genetic correlations between parities for additive thermotolerance were also positive but slightly lower ranging from 0.61 to 0.78. For FY and PY, genetic correlations between parities for general additive effect were also positive and high (≥0.76) whereas genetic correlations between parities for heat tolerance additive effects were positive but slightly lower, ranging between 0.34 to 0.78. The lower genetic correlation between parities for heat tolerance additive effects could be due to interaction between parity and THI. However, it is important to note that variances and genetic parameters estimated in this study are in lower bounds as MREP ignore the dynamics of heat stress with respect to function of DIM.

**Whole-Genome Mapping for Thermotolerance Genes affecting MY**

Single-step genomic BLUP methodology was utilized to perform GWAS of thermotolerance for MY. The ssGBLUP methodology allows to identify genomic regions
that explain genetic variance for a trait of interest under consideration. The advantage of ssGBLUP is it captures less noise and yields well-defined peaks across the entire genome, however, with ssGBLUP, there is no statistical test to test the significance of a genetic marker for a trait of interest.

The association analysis identified several regions in the bovine genome strongly associated with general production level and animal ability to respond to heat stress. Figure 2-1 displays Manhattan plots for MY across the three lactations under study; the left plots show the results for general merit (intercept) while the right plots show the results for thermotolerance (slope). The results are presented in terms of the proportion of the genetic variance explained by 2.0 Mb SNP windows. For MY, as expected, two different regions on BTA14 (14:1379063-3371507 and 14:3513577-5494654) harbor genes with known relevant biological functions such as \textit{DGAT1}, \textit{CPSF1}, \textit{FABP4} across all three parities. These two QTL regions on BTA14 explained together about 8.28, 5.28 and 4.47\% of genetic variances across the first three parities, respectively. Gene \textit{DGAT1} is a strong candidate gene for MY that encodes a key enzyme involved in the synthesis of milk triglycerides (Cases et al., 1998). Knockdown of \textit{DGAT1} gene in mouse reduces MY (Smith et al., 2000). Cochran et al. (2013) reported SNP located within \textit{CPSF1} to be significantly associated with MY which is involved in 3’ end-processing of pre-messenger RNAs into messenger RNAs. \textit{FABP4} regulates both medium and long chain fatty acids in milk and is upregulated during lactation (Zhou et al., 2015). Another 2.0 Mb SNP window on BTA20 (20:31051302-33048635) explained substantial number of genetic variances for MY (1.07\%). Interestingly, this region harbors \textit{GHR}, a growth hormone receptor known to have a major effect on milk yield.
and milk composition (Blott et al., 2003, Viitala et al., 2006). GHR is implicated in lipid and carbohydrate metabolism and plays a pivotal role in the growth hormone axis by initiating and maintaining lactation (Parmentier et al., 1999).

The association study also identified several regions associated with thermotolerance. Indeed, one region on BTA15 (15:75595790-77595636) was strongly associated with thermotolerance across all three parities. This region harbors PEX16, MAPK8IP1, CREB3L1, CRY2; all genes implicated in cellular response to heat stress. MAPK8IP1 is involved in controlling cellular response to external stimuli such as heat shock, which in turn increases transcription activity of many heat-stress target genes including cell survival, cell proliferation, apoptosis and morphogenesis (Thompson et al., 2001). Therefore, MAPK8IP1 helps repair DNA replication error and damage occurred at high temperature. PEX16 is involved in cell membrane biosynthesis (Farr et al., 2016). PEX16 thus plays an important role in protection against heat shock. CREB3L1 is implicated in endoplasmic reticulum (ER) stress response caused due to the accumulation of misfolded protein and promotes cell survival during heat stress (Liu and Chang, 2008, Greenwood et al., 2015). CRY2 is involved in thermal tolerance and knockdown of CRY2 increased sensitivity to temperature (Sanchez-Bermejo et al., 2015).

The QTL region on BTA22 (22:16637720-18628331) identified in the slope of first lactation harbors potential candidate genes for heat stress, such as FANCD2. Gene FANCD2 is involved in Hsp90-mediated regulation of DNA damage response (Oda et al., 2007).
Another SNP window on BTA14 (14:1651311-3649589) harbors genes $HSF1$, $EEF1D$, $VPS28$ and $TONSL$ which are involved in heat stress response. $HSF1$ is involved in regulating HSP gene expression during heat stress. Upon heat stimulus, $HSF1$ binds promoters containing heat shock elements (HSE) and activates heat stress target gene transcription (Calderwood et al., 2010). $EEF1D$ directs DNA binding at heat shock promoter elements (HSE) and regulates heat shock responsive genes through association with heat shock transcription factors (Cui et al., 2016). $VPS28$ is a vacuolar protein sorting gene involved in heat shock resistance and any mutation in $VPS28$ will lead to heat-shock sensitivity (Jarolim et al., 2013). $TONSL$ is involved in DNA repair and maintenance of genome stability in the presence of DNA damaging stimulus such as heat shock (Saredi et al., 2016).

Another QTL region on BTA15 (15:82479789-84450812) harbors potential candidate genes for heat stress, such as $GLYAT$. Gene $GLYAT$ activates HSE signaling pathway and is known to involve in detoxification of xenobiotic acyl-CoA’s in mammalian cells (Zhang et al., 2007).

**Pathway-Based Association Analysis for MY**

In this study, we complemented the whole-genome scan with a gene set enrichment analysis to detect potential functional categories and molecular mechanisms associated with heat tolerance. The first step of the pathway-based analysis was to assign SNPs to genes. Of the 58,046 SNP markers evaluated in the analysis, a total of 27,488 SNPs was located within annotated genes or within 15kb upstream or downstream from annotated genes. This set of SNPs defined a total of 17,238 genes annotated in the UMD 3.1 bovine genome sequence assembly, which in turn were evaluated for pathway analysis. Additionally, a subset of 344 of these 17,238 genes had
been flagged by at least one SNP whose effect was in the top 1% distribution and were significantly associated with heat tolerance for MY. The second step in the analysis was to assign 17,238 genes into pathways. We tested GO and MeSH categories using a hypergeometric test (Fisher's exact test).

Several GO terms showed significant overrepresentation of genes statistically associated with thermotolerance for MY. Noticeably, some of these terms are directly associated with heat tolerance, such as cellular response to heat (GO:00346), response to temperature stimulus (GO:0009266), and cellular response to stress (GO:0080135). These three categories are highly related in the GO hierarchy and had two significant genes in common, \textit{HSF1} and \textit{MAPT}. There is a well-known connection between heat tolerance and \textit{HSF1}. Upon heat stimulus, \textit{HSF1} activates heat stress target gene transcription including molecular chaperones essential for recovery from cellular damage. Molecular chaperones are involved in protein folding and protect cell against heat stress (Luft et al., 2001). \textit{MAPT} is involved in regulating and coordinating diverse cellular responses to stress. Heat stress leads to protein damage thereby giving rise to misfolded protein and cell death. \textit{MAPT} initiates mechanisms to repress translation and ameliorate misfolded protein accumulation thereby promoting cell survival under stressful conditions (Parker et al., 2014). Furthermore, GO:0080135, cellular response to stress, also includes gene \textit{GAB1}, a gene involved in angiogenesis through mediating angiogenic and survival signaling (Zhao et al., 2011). Development of peripheral blood vessels will enhance blood flow from core to periphery thereby removing heat from core and exchanging with the environment (Charkoudian, 2010). Two significant GO terms are related to DNA repair mechanism during heat stress. These GO terms showed a
very close relationship in the GO hierarchy: *response to DNA damage stimulus* (GO:2001020) and *double strand break repair* (GO:0006302). These highly related GO terms had two genes in common, *HSF1* and *BABAM2*. *BABAM2* is anti-apoptotic and positive regulator of DNA repair which promotes cell survival during heat stress. Furthermore, many significant GO terms were associated with signal transduction, ion transport and homeostasis, including regulation of *calcium-mediated signaling* (GO:0050848), *ion channel regulator activity* (GO:0099106), *glutamine family amino acid catabolic process* (GO:0009065) and *alpha-amino acid catabolic process* (GO:1901606). The regulation of calcium and heat tolerance is well-documented. Calcium can reverse the effects of heat implying protection against heat damage. The term *GO:0050848* is significantly enriched with genes *FKBP1B* and *RCAN2*. Gene *FKBP1B* plays a role in protein folding thereby acting as a molecular chaperone and preventing protein damage during heat stress (Calderwood, 2018). *RCAN2* acts as a molecular chaperone and regulates protein folding. Ion channel regulator activity (GO:0099106) is significantly enriched with at least two genes *PRKG1* and *PRKCB*. These genes are key mediators of NO signaling pathway. NO is implicated in maintaining cellular homeostasis during heat stress by acting as an antioxidant, protecting membrane from damage and decreasing ROS levels (Parankusam et al., 2017).

The term GO:0009065, *Glutamine family amino acid catabolic process*, is significantly enriched with a gene called *GLUD1* which plays a key role in cellular protection during heat stress through enhancement of heat shock protein 70 (HSP70) (Hamiel et al., 2009). *Alpha-amino acid catabolic process* (GO:1901606) is enriched
with a significant gene \textit{GCAT} which has a role in production of glutathione synthesis. Glutathione protects cells from death by removing free oxygen radicals produced during heat shock (Yao et al., 2011).

Table 2-5 shows a panel of MeSH terms significantly enriched with genes associated with heat tolerance of MY. Many of these terms are closely related with heat stress such as \textit{HSP20 Heat-Shock Proteins (D050886), Heat-Shock Response (D018869), Glutathione Transferase (D005982), Nitroso Compounds (D009603). Additionally, functional categories including Inositol 1,4,5-Trisphosphate Receptors (D053496), Calcium-Calmodulin-Dependent Protein Kinase Type (D054732) were also detected as significant in the MeSH-informed enrichment analysis.}

**Summary**

The results of this study reinforce the idea that there is a negative relationship between production traits and thermotolerance. Therefore, continued selection for MY, FY and PY without consideration of thermotolerance will result in greater susceptibility to heat stress. In this study, we also performed GWAS and pathway-association analysis with the purpose of understanding the genetic architecture underlying thermotolerance. Genomic regions in BTA14, BTA15 and BTA22 were associated with thermotolerance. Some of these genomic regions harbor genes that have been implicated in cellular response to heat stress, such as \textit{HSF1, GLYAT, FANCD2}. Moreover, gene set analysis revealed several functional significant terms such as response to heat conditions, response to DNA damage stimulus, and calcium-mediated signaling. Most of these terms are directly implicated in biological process related with thermotolerance. Overall, this study contributes to better understanding of the genetics
underlying heat stress and points out novel opportunities for improving thermotolerance in dairy cattle.
Table 2-1. General ($\sigma^2_a$) and heat tolerance ($100 \sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold of 68) for MY(Kg)$^2$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parity 1</th>
<th>Parity 2</th>
<th>Parity 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_a$</td>
<td>9.26</td>
<td>10.03</td>
<td>10.55</td>
</tr>
<tr>
<td>$100 \sigma^2_v$</td>
<td>0.94</td>
<td>1.56</td>
<td>1.62</td>
</tr>
<tr>
<td>$10\sigma_{a,v}$</td>
<td>-1.21</td>
<td>-1.17</td>
<td>-2.31</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>7.31</td>
<td>12.97</td>
<td>15.65</td>
</tr>
<tr>
<td>$h^2_{f(10)}$</td>
<td>0.32</td>
<td>0.24</td>
<td>0.17</td>
</tr>
<tr>
<td>$r^2(a, v)$</td>
<td>-0.41</td>
<td>-0.30</td>
<td>-0.55</td>
</tr>
<tr>
<td>Cor-ht (par1, parj)</td>
<td></td>
<td>0.78</td>
<td>0.65</td>
</tr>
<tr>
<td>Cor-ht (par2, parj)</td>
<td></td>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td>Cor-gen(par1,parj)</td>
<td></td>
<td>0.82</td>
<td>0.85</td>
</tr>
<tr>
<td>Cor-gen(par2,parj)</td>
<td></td>
<td></td>
<td>0.92</td>
</tr>
</tbody>
</table>

$\sigma^2_a$ - general additive genetic variances, $100\sigma^2_v$ = heat tolerance additive variances at 78 (10 units above a THI threshold of 68); $10\sigma_{a,v}$ - additive genetic covariance between general and heat tolerance effect; $r^2(a, v)$ = genetic correlation between general and heat tolerance effect; Cor-ht = heat tolerance additive genetic correlations; Cor-gen = regular additive genetic correlations
Table 2-2. General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{(10)}$ at THI of 78 (10 units above THI threshold of 68) for FY (Kg x 100)$^2$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parity 1</th>
<th>Parity 2</th>
<th>Parity 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_a$</td>
<td>119.76</td>
<td>205.77</td>
<td>252.33</td>
</tr>
<tr>
<td>$100\sigma^2_v$</td>
<td>18.189</td>
<td>48.614</td>
<td>33.78</td>
</tr>
<tr>
<td>$10\sigma_{a,v}$</td>
<td>-11.44</td>
<td>-37.90</td>
<td>-63.03</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>351.15</td>
<td>666.04</td>
<td>840.74</td>
</tr>
<tr>
<td>$h^2_{(10)}$</td>
<td>0.20</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td>$r^3(a, v)$</td>
<td>-0.25</td>
<td>-0.38</td>
<td>-0.68</td>
</tr>
<tr>
<td>Cor-ht (par1, parj)</td>
<td>0.46</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Cor-ht (par2, parj)</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cor-gen (par1, parj)</td>
<td>0.91</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Cor-gen (par2, parj)</td>
<td></td>
<td></td>
<td>0.95</td>
</tr>
</tbody>
</table>

$\sigma^2_a$. General additive genetic variances, $100\sigma^2_v$ = heat tolerance additive variances at 78 (10 units above a THI threshold of 68); $10\sigma_{a,v}$ = additive genetic covariance between general and heat tolerance effect; $r^3(a, v)$ = genetic correlation between general and heat tolerance effect; Cor-ht = heat tolerance additive genetic correlations; Cor-gen = regular additive genetic correlations
Table 2-3. General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{(10)}$ at THI of 78 (10 units above THI threshold of 68) for PY(Kg x 100)$^2$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parity 1</th>
<th>Parity 2</th>
<th>Parity 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_a$</td>
<td>55.65</td>
<td>63.80</td>
<td>76.01</td>
</tr>
<tr>
<td>$100\sigma^2_v$</td>
<td>8.57</td>
<td>9.98</td>
<td>13.30</td>
</tr>
<tr>
<td>$10\sigma_{a,v}$</td>
<td>-6.31</td>
<td>-4.58</td>
<td>-12.81</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>79.92</td>
<td>127.98</td>
<td>154.51</td>
</tr>
<tr>
<td>$h^2_{(10)}$</td>
<td>0.26</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>$r^\alpha(a, v)$</td>
<td>-0.29</td>
<td>-0.18</td>
<td>-0.40</td>
</tr>
<tr>
<td>Cor-ht (par1, parj)</td>
<td></td>
<td>0.36</td>
<td>0.55</td>
</tr>
<tr>
<td>Cor-ht (par2, parj)</td>
<td></td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td>Cor-gen(par1,parj)</td>
<td></td>
<td>0.78</td>
<td>0.76</td>
</tr>
<tr>
<td>Cor-gen(par2,parj)</td>
<td></td>
<td></td>
<td>0.96</td>
</tr>
</tbody>
</table>

$\sigma^2_a$ - General additive genetic variances, $100\sigma^2_v$ = heat tolerance additive variances at 78 (10 units above a THI threshold of 68); $10\sigma_{a,v}$ = additive genetic covariance between general and heat tolerance effect; $r^\alpha(a, v)$ = genetic correlation between general and heat tolerance effect; Cor-ht = heat tolerance additive genetic correlations; Cor-gen = regular additive genetic correlations
Table 2-4. GO terms significantly enriched with genes associated with MY

<table>
<thead>
<tr>
<th>GO ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0034605</td>
<td>Cellular response to heat</td>
<td>15</td>
<td>3</td>
<td>0.003</td>
</tr>
<tr>
<td>0009266</td>
<td>Response to temperature stimulus</td>
<td>33</td>
<td>3</td>
<td>0.027</td>
</tr>
<tr>
<td>0080135</td>
<td>Cellular response to stress</td>
<td>96</td>
<td>5</td>
<td>0.043</td>
</tr>
<tr>
<td>2001020</td>
<td>Response to DNA damage stimulus</td>
<td>35</td>
<td>3</td>
<td>0.032</td>
</tr>
<tr>
<td>0050848</td>
<td>Calcium-mediated signaling</td>
<td>15</td>
<td>3</td>
<td>0.003</td>
</tr>
<tr>
<td>0009065</td>
<td>Glutamine family amino acid catabolic process</td>
<td>9</td>
<td>2</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Table 2-5. MeSH terms significantly enriched with genes associated with MY

<table>
<thead>
<tr>
<th>MeSH ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>18869</td>
<td>Heat-Shock Response</td>
<td>9</td>
<td>2</td>
<td>0.015</td>
</tr>
<tr>
<td>50886</td>
<td>HSP20 Heat-Shock Proteins</td>
<td>2</td>
<td>1</td>
<td>0.004</td>
</tr>
<tr>
<td>05982</td>
<td>Glutathione Transferase</td>
<td>50</td>
<td>4</td>
<td>0.002</td>
</tr>
<tr>
<td>09603</td>
<td>Nitroso Compounds</td>
<td>2</td>
<td>1</td>
<td>0.004</td>
</tr>
<tr>
<td>53496</td>
<td>Inositol 1,4,5-Trisphosphate Receptors</td>
<td>17</td>
<td>2</td>
<td>0.040</td>
</tr>
<tr>
<td>54732</td>
<td>Calcium-Calmodulin-Dependent Protein Kinase Type 2</td>
<td>13</td>
<td>2</td>
<td>0.002</td>
</tr>
<tr>
<td>51116</td>
<td>Receptors, Scavenger</td>
<td>2</td>
<td>1</td>
<td>0.004</td>
</tr>
</tbody>
</table>

53
Figure 2-1. Manhattan plot showing the results of genome-wide association mapping for Milk yield (MY): the left plots show the result of general merit and the right plots show the results of thermotolerance across lactations numbered vertically as lac1, lac2 and lac3.
CHAPTER 4
GENETIC ANALYSIS OF HEAT TOLERANCE FOR REPRODUCTION AND UDDER HEALTH IN US HOLSTEIN COWS

Background

Heat stress depresses fertility and increases incidence of health disorders in dairy cows. The effects of heat stress on fertility are mainly because of the failure of the inseminated cows to establish and maintain pregnancy. Several studies have reported that there is a marked decrease in conception per insemination during summer (Hansen and Arechiga, 1999, de Vries and Risco, 2005, Huang et al., 2009). Indeed, de Vries (2004) reported that the reduction in conception rate is around 53% during summer in Florida. Reduction in conception rate is mainly because heat stress damages oocyte and early embryo (Hansen, 2013). Embryo loss is 3.7 times more likely during summer season as compared to winter season (Thatcher et al., 2001). Heat stress also has negative effects on udder health, particularly the count of somatic cells is elevated in response to high temperature. Also, high level of circulating stress hormones interferes with immune system thereby increasing susceptibility of cows to clinical mastitis (Hagiya et al., 2017). It is not surprise that in Florida, where more than 257 days have Temperature-Humidity Index (THI) values greater than 68, heat stress has a huge toll in both reproduction performance and health traits of dairy cattle (Ferreira et al., 2016). Conception per insemination (CI) can be found as low as 10% to 20 % during summer season (Hansen and Arechiga, 1999). Even with heat abatement systems, CI is found to be lower during heat stress conditions (Hansen, 2009). Embryo transfer has been considered as one of the most viable tools for improving fertility during heat stress. However, farmers need to spend additional cost of $60 per cow for timed embryo transfer as compared to $20 for conventional AI during summer months (de Vries,
Moreover, it is known that during heat stress, the immune system is compromised, which in turn makes cows more susceptible to infection, increasing markedly the number of cases of clinical mastitis during summer (Kadzere et al., 2002). Economic loss estimates about $179 per case of clinical mastitis. Since heat stress is a huge economic problem in the dairy industry, especially in the southeast, various strategies such as cooling, shading and nutrition have been used to alleviate effects of heat stress and improve performance and well-being in dairy cows. Nonetheless, there is a clear decline in fertility and udder health in dairy cows in hot and tropical climates. In this context, genetic selection of heat tolerance for CI and SCS is an attractive alternative for reducing effects of heat stress and subsequently improving fertility and udder health in dairy cattle.

A growing body of literature suggests that there exists an additive genetic components of heat tolerance for reproduction and udder health traits in dairy cattle (Ravagnolo and Misztal, 2002, Aguilar et al., 2009). Indicator traits of heat stress like respiration rate, body temperature have moderate to strong heritability. Estimates of heritability for rectal temperature during heat stress is 0.17 which suggests genetic improvement of heat tolerance through genetic selection is possible (Dikmen et al., 2012). Ravagnolo and Misztal (2002) estimated heritability for non-return rate of estrous during heat stress condition (THI = 70) and reported heritability estimate increased considerably from 0.06% at 45 days to 5.3% at 90 days. However, there exists a negative genetic correlation between heat tolerance and fertility trait. In the same study, Ravagnolo and Misztal (2002) reported a negative genetic correlation of -0.25, -0.77 and -0.95 between general merit and heat tolerance estimated at 45, 60
and 90 days respectively after insemination. Regarding udder health, Hagiya et al. (2017) reported heritability estimates for SCS under heat stress conditions in the range of 0.18 to 0.20 across lactations. Ravagnolo and Misztal (2002) presented a methodology for genetic evaluation of heat tolerance by combining test day records with public weather station records. Freitas et al. (2006) reported that information from weather stations can be as accurate as that of recorded in farm, and hence, it can be used to dissect the genetics underlying thermotolerance. In this methodology, test day records and daily THI values are jointly analyzed using test-day models that include random regressions on THI values.

Although traditional selection methods could be used to identify and select thermotolerant cows, genetic gains for fertility and health traits could be accelerated if the genes and pathways responsible for thermotolerance are identified and then used in marker-assisted selection in commercial breeding schemes. As such, the first objective of this study was to estimate variance components for CI and SCS across lactations using multi trait repeatability test-day models considering heat stress. The second objective of the study was to identify and characterize genomic regions, individual genes and biological processes responsible for the variation in thermotolerance underlying fertility and udder health. Pregnancy and SCS test-day records, weather data, and genome-wide SNP markers were jointly analyzed. This relevant information can contribute to a better understanding of the genetics underlying heat stress and provide novel opportunities for use of marker-assisted selection in dairy cattle breeding schemes.
Materials and Methods

Phenotypic and Genotypic Data

Conception per Insemination (CI) is one of the important fertility traits in dairy cattle, which is defined as the ability of cow to establish and maintain pregnancy provided cow is inseminated at the time of ovulation (Averill et al., 2004). Somatic cell score (SCS) is an indicator trait for clinical mastitis, and somatic cell counts increase with the severity of udder infection incurring huge economic loss to the dairy industry (Sahana et al., 2013). Our data comprised 74,221 CI records on 13,704 Holstein cows. All available service records maintained between 5 and 400 DIM were used. For SCS, data comprised 355,546 test-day records on 18,975 cows obtained between 5 and 305 DIM. Cows calved from 2006 through 2016 on two dairy farms in Florida, USA. A complete pedigree file was created by tracing the pedigree of cows back to five generations using information available at the Council on Dairy Cattle Breeding website.

Genotype data for 60,671 single nucleotide polymorphism (SNP) markers were available for 4,700 cows with phenotype records, and 1,592 sires in the pedigree. The SNP data were provided by the Council on Dairy Cattle Breeding and the Cooperative Dairy DNA Repository. Those SNP markers that mapped to sex chromosomes or were monomorphic or had minor allele frequency less than 1% were removed from the dataset. After data editing, a total of 57,954 single nucleotide polymorphisms (SNPs) were retained for subsequent genomic analyses.

Weather data were obtained from Florida Automated Weather Network for Alachua county and hourly THI values were calculated as proposed by Ravagnolo and Misztal (2002) as
\[
THI = (1.8 \times \text{temp} + 32) - (0.55 - 0.55 \times \text{rh}) \times (1.8 \times \text{temp} - 26) \tag{3-1}
\]
where temp is the temperature in degree Celsius and rh is the relative humidity in percentage.

A function of THI (f(THI)) was created to estimate effects of heat stress on CI and SCS. Therefore,
\[
f(\text{THI}) = \begin{cases} 
\text{0} & \text{if THI} \leq \text{THI}_{\text{thr}} \\
\text{THI} - \text{THI}_{\text{thr}} & \text{if THI} > \text{THI}_{\text{thr}}
\end{cases} \tag{3-2}
\]
where the value of THI_{thr} was set to 68 and thus f(THI) was equal to max (0, THI – THI_{thr}).

**Statistical Model**

The following multi-trait repeatability test day model was used to estimate the variance components of CI and SCS for heat tolerance, considering multiple lactations as different traits.

\[
y_{klmn} = \text{HTD}_{kl} + \text{DIM}_m + a_{nl} + p_{nl} + v_{nl}[f(\text{THI})] + q_{nl}[f(\text{THI})] + e_{klmn} \tag{3-3}
\]
where \(y_{klmn}\) is the record for CI and SCS, \(\text{HTD}_{kl}\) is herd-test-day \(k\) within parity \(l\) (\(l = 1, 2, 3\)), \(\text{DIM}_m\) is the \(m^{th}\) DIM class (\(m = 1\) to \(20\) for CI and \(m = 1\) to \(16\) for SCS), with classes defined every 20 days, \(a_{nl}\) is the general random additive genetic effect of animal \(n\) in parity \(l\), \(f(\text{THI})\) is the heat stress function for herd test day \(k\), \(v_{nl}\) is the random additive genetic effect of heat tolerance for the animal \(n\) in parity \(l\), \(p_{nl}\) is the random permanent environmental effect of cow \(n\) in parity \(l\), \(q_{nl}\) is the random permanent environmental effect of heat tolerance of the cow \(n\) in parity \(l\) and \(e_{klmn}\) is the random residual effect.

Let \(a = [a'_{nl}, v'_{nl}]\) be vector of random additive genetic effects and \(p = [p'_{nl}, q'_{nl}]\) be a vector of random permanent environment effects for parities \(n = 1\) to \(3\).
The variance-covariance structure was

\[ V \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} A \otimes G & 0 & 0 \\ 0 & I \otimes P & 0 \\ 0 & 0 & I \otimes R \end{bmatrix} \]  \hspace{1cm} (3-4)\]

where \( A \) is the numerator relationship matrix and \( G \) and \( P \) are 6x6 matrices of (co)variances for additive and permanent environment effects respectively. \( R \) is a 3x3 diagonal matrix of residual variances corresponding to each trait.

**Statistical Analysis**

Variance components were estimated in a Bayesian framework using GIBBS2F90 (Misztal et al., 2002). Genomic data were not included for variance component estimation. Of a total of 500,000 samples, first 100,000 were discarded as burn-in, and every 100th sample was retained to calculate variance components and posterior standard deviations of relevant genetic parameters.

Heritability was estimated as

\[ h^2 = \frac{\sigma_a^2}{\sigma_{total}^2} \]  \hspace{1cm} (3-5)\]

Where

\[ \sigma_{total}^2 = \sigma_a^2 + \sigma_p^2 + \sigma_e^2 \]  \hspace{1cm} (3-6)\]

and heritability at heat stress level \( f(i) \) was calculated as

\[ h^2 = \frac{\sigma_a^2 + f(i)^2 \sigma_v^2 + 2f(i)\sigma_{av}}{\sigma_a^2 + f(i)^2 \sigma_v^2 + 2f(i)\sigma_{av} + \sigma_p^2 + f(i)^2 \sigma_q^2 + 2f(i)\sigma_{pq} + \sigma_e^2} \]  \hspace{1cm} (3-7)\]

The genetic correlation between the generic and heat tolerance additive effects was calculated as
\[
\text{corr} \ [a, f(i) v] = \frac{f(i) \sigma_{av}}{\sqrt{\sigma_a^2 + f(i)^2 \sigma_v^2}}
\]  

(3-8)

**Genome-wide Association Mapping using ssGBLUP**

The whole-genome association mapping was performed using single-step genomic BLUP methodology (ssGBLUP). ssGBLUP is indeed a classical BLUP but here pedigree relationship matrix \(A^{-1}\) is replaced with the inverse of the realized relationship matrix \(H^{-1}\) that combines both pedigree and genomic information (Aguilar et al., 2010a). The combined pedigree genomic relationship matrix \(H^{-1}\) is calculated as follows,

\[
H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}
\]  

(3-9)

where \(G^{-1}\) is the inverse of the genomic relationship matrix, and \(A_{22}^{-1}\) is the inverse of the numerator relationship matrix for genotyped animals. Here, \(G^{-1}\) has the dimension of \(6,362 \times 6,362\) which includes 4,700 cows with phenotypic records and 1,592 sires in the pedigree. Similarly, \(A\) matrix has a dimension of \(28,620 \times 28,620\) for CI and \(31,412 \times 31,412\), based on a five generation pedigree.

Candidate genomic regions associated with general merit and heat tolerance additive merit for CI and SCS were identified based on the amount of genetic variance explained by 2.0 Mb windows of SNPs. Given the genomic estimated breeding values (GEBVs), the SNP effects can be estimated as \(\hat{s} = DZ'[ZDZ]^{-1} \hat{a}_g\), where \(\hat{s}\) is the vector of SNP marker effects, \(D\) is a diagonal matrix of weights of SNPs, and \(\hat{a}_g\) is the vector of GEBVs (Wang et al., 2012). The percentage of genetic variance explained by a given 2.0 Mb genomic region was then calculated as

\[
\frac{\text{var} u_i}{\sigma_u^2} \times 100 = \frac{\text{var} \sum_{j=1}^{p} Z_j S_j}{\sigma_u^2} \times 100
\]  

(3-10)
where $u_i$ is the genetic value of the $i^{th}$ genomic region under consideration, $B$ is the total number of adjacent SNPs within the 2.0 Mb region, and $s_j$ is the marker effect of the $j^{th}$ SNP within the $i^{th}$ region. All the ssGBLUP calculations were performed using BLUPF90 family of programs (Misztal et al., 2002)

**Gene Set Analysis**

The gene set analysis includes basically three different steps: i) assignment of SNPs to genes ii): assignment of genes to pathways iii) pathway-based association analysis

**Assignment of SNPs to genes**

The UMD 3.1 bovine genome sequence assembly was used for SNP assignment using the Bio Conductor R package biomart (Zimin et al., 2009). SNPs were assigned to genes if they were located within the genomic sequence of the gene or at most 15 kb either upstream or downstream the gene. If a SNP was found to be located within the gene or at most 15kb either upstream or downstream the gene, that gene will be included for further analysis. Finally, a gene was considered significantly associated with a trait, if the gene is flagged by a SNP in top 1% distribution.

**Assignment of genes to pathways**

Gene Ontology (GO) (Ashburner et al., 2000)and Medical Subject Headings (MeSH) (Nelson et al., 2004) databases were used to define functional sets of genes. Genes assigned to the same functional category are considered more closely related in terms of biology rather than random sets of genes.

**Pathway-based association analysis**

The association of a give pathway with either CI or SCS was analyzed using Fisher’s exact test which is based on the cumulative hypergeometric distribution. This
test enables to search for an overrepresentation of significantly associated genes. The $P$-value of observing $k$ significant genes in the pathway was calculated by

$$P\text{-value} = 1 - \sum_{i=0}^{k-1} \binom{S}{i} \binom{N-S}{m-i} \binom{N}{m}$$

(3-11)

where $S$ is the total number of genes that were significantly associated with MY, $N$ is the total number of genes that were analyzed in the study, and $m$ is the total number of genes in the pathway (Abdalla et al., 2016). The GO gene set enrichment analysis was performed using the R package goseq (Young et al., 2010) whereas MeSH enrichment analysis was carried out using the R package meshr (Tsuyuzaki et al., 2015).

**Results and Discussion**

**Genetic Dissection of Heat Stress for Conception and Udder Health**

Variance components of general performance level (intercept) and animal ability to respond under heat stress conditions (slope) for CI and SCS were estimated using multi-trait repeatability test day models (MREP) (Table 3-1, 3-2). Similarly, heritability and genetic correlations at heat stress level $\phi$ (THI) = 78 (i.e. 10 units above THI threshold of 68) for the first three lactations were also estimated.

**Variance Component Estimates for CI**

Both generic and heat tolerance additive genetic variances for CI increased across parity. Generic additive genetic variances ranged from 0.2% to 0.4% in first three lactations. Genetic variances for CI under-heat stress increased between 0.5% to 0.15% in consecutive parities suggesting that heat stress has greater effects on conception rates in later lactations. This finding agrees with the results of Schueller et al. (2014) who reported that multiple-parity cows are more affected by heat stress than monoparous cows. Genetic correlations between generic and heat tolerance effects for
CI were negative in all three lactations, ranging from -0.35 to -0.82. This agrees with the findings of Dikmen et al. (2012) who also reported a negative genetic correlation between rectal temperature and fertility. These results suggest that there is unfavorable genetic correlation between CI and heat tolerance, and that continued selection for successful pregnancy in thermoneutral conditions will depress fertility under heat stress condition. Moreover, heritability estimates for CI at THI 78 were between 2 to 3%, which is comparable with the heritability estimates reported by Dikmen et al. (2012). Genetic correlations between parities for general additive effect for CI were positive and high (≥0.58), whereas genetic correlations between parities for additive heat tolerances were also positive but slightly lower ranging from 0.17 to 0.49.

**Variance Component Estimates for SCS**

Both general and thermotolerance additive genetic variances increased with parity. This is in agreement with the previous study (Hammami et al., 2015). Heritability estimates at THI 78 ranged from 0.10 to 0.15 across three parities, which are in the range of heritability estimates reported by Santana et al. (2017) in Brazilian Holsteins. Interestingly, genetic correlations between generic and thermotolerance were positive in all three lactations, ranging from +0.10 to +0.43. This suggests that continued selection for lower SCS, will also yield a positive selection response in the cow’s ability to respond under heat stress conditions.

**Genomic Scans for Thermotolerance Genes Affecting CI and SCS**

**Whole Genome association analysis for CI**

The association analysis identified several regions in the bovine genome strongly associated with CI under heat stress. Figure 3-1 displays Manhattan plots across three lactations under study; the left plots show the results for general merit (intercept), while
the right plots show the results for CI under heat conditions (slope). The results are presented in terms of the proportion of the genetic variance explained by non-overlapping 2.0 Mb SNP windows.

Two different regions on BTA23 (23:3626219-5618724 and 23:10620999-12579560) were found to be strongly associated with general merit for CI across all three parities. These regions together explained almost 3% of the additive genetic variances. Interestingly, these strongly associated regions harbor several putative genes closely related to reproduction, including PADI6 and HCRTR2. PADI6 is a maternal-effect gene that regulates mitochondrial activity in oocyte. Mitochondrial activity in oocyte influences the competence of oocyte to fertilize and early embryonic development. Mouse oocyte lacking PADI6 triggers intrinsic apoptotic pathway resulting in early embryonic death following ovulation and fertilization (Fernandes et al., 2012). Gene HCRTR2 binds the hypothalamic neuropeptides orexin A and orexin B which are implicated in normal pregnancy. This orexin system is expressed in uterus, conceptus and trophoblast and is involved in maternal recognition of pregnancy, implantation and proliferation of endometrial epithelial cells (Smolinska et al., 2017). Two genomic regions on BTA2 (2:134837645-136697236) and BTA12(12:14654029-16641931) were also found associated with CI, particularly for first lactation cows. The region on BTA2 harbors gene RPA2 which is involved in DNA damage response in oocyte and embryos. DNA damage leads to cell cycle arrest, which in turn compromises the chances of embryo survival. RPA2 is involved in DNA replication, recombination and repair and maintain cellular and long-term embryo viability through effects on genomic integrity (Zheng et al., 2005). The region on BTA12, which explained more than 1% of the
genetic variance for CI, harbors several putative genes, including \textit{TPT1} which is expressed in placenta and is implicated in calcium binding and homeostasis of trophoblast cells. It is known that placental transfer of calcium is important for growth and development of fetus (Arcuri et al., 2005). The association study also identified several regions associated with thermotolerance. Indeed, one region on BTA10 (10: 80756063-82748913) explained 1.00% of genetic variance for CI under heat stress in first lactation cows. This region harbors gene \textit{RGS6} which is involved in cellular response to heat stress. Gene \textit{RGS6} is implicated in subcellular trafficking of proteins to nucleoli in response to heat stress and prevents further damage of proteins from heat stress (Chatterjee and Fisher, 2003). Another region on BTA11(11:27185742 - 29174959) explained 1.40% of the genetic variance for CI under heat stress in second lactation cows. This region harbors \textit{PRKCE} which is also involved in heat shock response. The heat shock response through the activation of HSFs protects proteins by stabilizing and refolding protein-folding intermediates (Morimoto et al., 1997).

Another region on BTA21 (21:15031948- 17028796) explained 1% of the genetic variance for thermotolerance associated with CI. This region harbors \textit{KLHL25} which is involved in protein ubiquitination (Zhang et al., 2016).

\textbf{Whole Genome association analysis for SCS}

The whole-genome scan identified several regions strongly associated with SCS in both thermoneutral conditions and under heat stress. Figure 3-2 displays Manhattan plots for SCS across the three lactations under study; the left plots show the results for general merit (intercept), while the right plots show the results for SCS thermotolerance (slope). The results are presented in terms of the proportion of the genetic variance explained by 2.0 Mb SNP windows.
Two different regions on BTA29 (29:2294521-4291391 and 29:41755825-43726898) were found to be strongly associated with SCS across all three parities. These regions harbor FAT3 and OTUB1. Gene FAT3 is a calcium ion binding gene that reduces calcium in blood. Reduction in blood calcium level impairs immunity, reduces smooth muscle function in teat sphincter, leading to partial or incomplete closure of teat canals. Lack of teat sphincter closure leads to entry of pathogens into teat canal and hence increases risk of CM (Mulligan et al., 2006). Gene OTUB1 serves as a marker of tumor and therefore could play a role in the development of tumors in mammary gland (Liu et al., 2014). A 2.0 Mb SNP window on BTA6 (6:92648735-94627787) explained 0.54, 0.73 and 0.66% of the genetic variance for SCS for first, second, and third lactation, respectively. This region harbors several genes, including vitamin D-binding protein precursor (GC) and neuropeptide FF receptor 2 (NPFFR2), which are strong candidates for udder health. The GC gene encodes Gc-globulin which transports vitamin D to target cells such as monocytes. In presence of microbes in the mammary gland, vitamin D stimulates monocyte antibacterial activity by increasing production of antimicrobial peptides and enhances mechanism associated with autophagy (Hewison, 2011). Neuropeptide FF receptor 2 is a membrane protein which suppresses the production of nitric oxide in the inflammation process and exhibit anti-inflammatory activity (Sun et al., 2013). Besides, it plays a role in the proliferation of T cells and modulation of immune responses (Minault et al., 1995).

Similarly, one region on BTA24 (24:60394194-62332550) explained 0.83, 0.64 and 0.69% of the genetic variances for SCS for parity one, two, and three, respectively. This region harbors several genes including BCL2 which is an apoptotic gene.
significantly upregulated in the mammary tissues challenged with *Staphylococcus aureus* (Long et al., 2001). Another QTL region on BTA4 (4:60094523-62089794) explained 0.52 and 0.54% of the genetic variances for parity two and three, respectively. This region harbors a potential candidate gene, namely *AOAH*. This gene is expressed in the neutrophils, and its activity decreases when dairy cows are challenged with endotoxin and *E. coli*, resulting in the appearances of lipo-polysaccharides in the blood which causes inflammation in mammary tissues (Mehrzad et al., 2007).

The association study also identified regions associated with SCS under heat stress conditions. One genomic region on BTA2 (2:24639094-26617396) was strongly associated with thermotolerance in first two lactations. This region harbors two candidate genes, *DLX1* and *DLX2* which downregulate cytokine signaling pathway and prevent inflammation (Dulken et al., 2017). Another QTL region on BTA19 (19:42020856-44001678) explained 0.61% of thermotolerance genetic variances for SCS in first lactation cows. This region harbors genes *STAT5B* and *STAT5A* which are directly implicated in the heat stress response. Indeed, *STAT5B* and *STAT5A* are involved in activating immune responses during heat stress. Knockout of *Stat5a<sup>−/−</sup>* and *Stat5b<sup>−/−</sup>* in mouse exhibit decreased immunity thereby suggesting their role in development of immune functions (Lin et al., 2012).

**Pathway-Based Analysis**

We complemented the whole-genome scans with alternative gene set enrichment analyses to detect potential functional categories and molecular mechanism associated with thermotolerance for conception and udder health. Of the 58,046 SNP markers evaluated in the analysis, a total of 27,488 SNPs was located within annotated
genes or within 15kb upstream or downstream from annotated genes. This set of SNPs defined a total of 17,235 genes annotated in the UMD 3.1 bovine genome sequence assembly. A subset of 345 and 361 of these 17,235 genes were significantly associated with CI and SCS, respectively. A gene was declared as significant if it had at least one SNP in the top 1% of the SNP effect distribution.

**Gene-set analysis for CI**

Several GO terms showed significant overrepresentation of genes statistically associated with CI under heat stress. Noticeably, some of these terms are directly associated with heat response such as protein refolding (GO:0042026), inter-strand cross link repair (GO:0036297), and regulation of intra-cellular protein transport (GO:0033157). There is well-known connection between heat tolerance and protein refolding. Interestingly, protein refolding (GO:0042026) is significantly enriched with a gene HSPA1A which acts as a molecular chaperone and prevents unfolding and degradation of protein during heat stress (Calderwood, 2018). Inter-strand cross link repair is implicated in DNA replication, maintenance and repair which promotes cell-cycle, growth and survival during heat stress (McMahon et al., 2016). Regulation of intra-cellular protein transport is implicated in protein trafficking under cellular stress. In response to heat stress, the nucleolus responds to cellular stress by sequestering and releasing a variety of proteins that affect cell cycle and DNA repair (Nalabothula et al., 2010).

Two significant GO terms are directly related with reproduction process: steroid hormone mediated signaling pathway (GO:0043401) and lipid modification (GO:0030258). Lipid modification greatly occurs during pregnancy which is important for fetal growth and development. Two important lipid modifications viz., maternal
accumulation of fat and hyperlipidemia greatly benefits fetal growth particularly under conditions of dietary intake (Herrera and Ortega-Senovilla, 2010).

Two relevant GO terms, namely germ cell development (GO:0007281) and post-embryonic development (GO:0009791) also showed significant overrepresentation of genes associated with conception under heat stress. Interestingly, post-embryonic development includes the gene GATA3 which is selectively expressed in trophectoderm of peri-implantation embryo and regulates morula to blastocyst transformation. Knockdown of GATA3 in pre-implantation mouse embryo inhibits embryonic development (Home et al., 2009).

**Gene-set analysis for SCS**

Several GO terms showed significant enrichment of genes statistically associated with SCS under heat stress. Some of these terms are directly associated with heat response such as chaperone mediated protein folding (GO:0061077). This term had three significant genes, FKBP10, FKBP11 and TCP1. Genes FKBP10 and FKBP11 are chaperones involved in protein folding and prevent the aggregation of unfolded protein in ER at high temperature (Jeffries et al., 2014). TCP1 is another chaperone gene which interacts with Hsp70 to assist protein folding and prevents aggregation of misfolded protein during heat stress (Young et al., 2004). Another significant GO term for SCS is DNA replication (GO:0006260). During heat stress conditions, nucleolin normally found in the nucleolus is translocated to nucleoplasm and interacts with recombinant replication protein, a key DNA replication factor. It is the contribution of these regulatory processes that helps in DNA replication during heat stress (Wang et al., 2001).

Interestingly, four significant GO terms are related to mastitis resistance, namely positive regulation of lymphocyte proliferation (GO:0050671), phagocytosis.
regulation of inflammatory response (GO:0050727) and defense response to bacterium (GO:0046903). The relationship between lymphocyte proliferation and mastitis resistance is well documented. Lymphocytes are able to secrete certain cytokines which based on the kind of cytokines facilitate either a cell-mediated or humoral response and produce antibodies against invading pathogens thereby developing resistance against mastitis (Sordillo and Streicher, 2002). In addition, heat stress is known to induce death signals that lead to apoptosis of mammary cells. During inflammatory reactions, phagocytosis of apoptotic cells is to be recognized and engulfed by macrophages to prevent inflammation of mammary tissues (Kitamura et al., 2005). It should be noted that the cellular stress response, including heat shock response and DNA damage response, and their cross talk with inflammatory pathways regulate host defense against bacterial infections in the udder (Muralidharan and Mandrekar, 2013).

*Intercalated disc* (GO:0014704) is implicated in placental function and development. Several biological processes are related with this term including diffusion of molecules such as cAMP, cGMP, inositol triphosphate (IP3), Ca++ between mother and fetus thereby allowing growth and development of fetus (Malassine and Cronier, 2005).

*Cellular response to tumor necrosis factor* (GO:0071356) showed significant overrepresentation of genes associated with successful pregnancy. TNF-α is also essential for early event in pregnancy such as implantation. Besides, TNF-α is also involved in secreting placental hormone such as chorionic gonadotrophin which is crucial for gestation and successful pregnancy outcome (Haider and Knoefler, 2009).
Summary

The results of this study suggest that there is a negative genetic correlation for conception between thermoneutral and heat stress conditions. Therefore, the continued selection for greater fertility ignoring heat stress conditions will result in increasing even more the susceptibility to heat stress. On the other hand, we identified a positive genetic correlation for SCS between neutral and heat stress, and hence continued selection for lower SCS will also yield a favorable selection response for SCS under heat stress.

Whole-genome scan identified significant regions on BTA10, BTA11 and BTA21 associated with CI under heat stress. Some of the putative genes are RGS6, PRKCE, KLHL25. Moreover, gene set analyses revealed several functional significant terms such as protein refolding, inter-strand cross link repair, regulation of intra-cellular protein transport, and intercalated disc. For SCS, genomic regions on BTA2, BTA3, BTA19 and BTA23 were associated with thermotolerance for SCS. These genomic regions harbor genes such as DLX1, STAT5B, NR2F6 CNPY3 and LCP1 which are potential candidate genes for response of SCS to heat stress. Additionally, gene set analysis revealed several functional significant terms such as chaperone mediated protein folding, DNA replication, positive regulation of lymphocyte proliferation, phagocytosis, regulation of inflammatory response and defense response to bacterium are associated with thermotolerance. Overall, this study contributes to better understanding of the genetics underlying heat stress and points out novel opportunities for improving thermotolerance in dairy cattle.
Table 3-1. General ($\sigma^2_a$) and heat tolerance ($100\sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{f(10)}$ at THI of 78 (10 units above THI threshold) for CI

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parity 1</th>
<th>Parity 2</th>
<th>Parity 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_a$</td>
<td>0.002</td>
<td>0.004</td>
<td>0.008</td>
</tr>
<tr>
<td>$100 \sigma^2_v$</td>
<td>0.005</td>
<td>0.009</td>
<td>0.015</td>
</tr>
<tr>
<td>$10\sigma_{a.v}$</td>
<td>0.001</td>
<td>0.003</td>
<td>0.008</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>0.20</td>
<td>0.19</td>
<td>0.18</td>
</tr>
<tr>
<td>$h^2_{f(10)}$</td>
<td>0.024</td>
<td>0.032</td>
<td>0.024</td>
</tr>
<tr>
<td>$r^G(a, v)$</td>
<td>-0.35</td>
<td>-0.58</td>
<td>-0.82</td>
</tr>
<tr>
<td>Cor-ht (par1, parj)</td>
<td>0.17</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Cor-ht (par2, parj)</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cor-gen(par1, parj)</td>
<td>0.58</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Cor-gen(par2, parj)</td>
<td>0.88</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\sigma^2_a$ = General additive genetic variances, $100 \sigma^2_v$ = heat tolerance additive variances at 78 (10 units above a THI threshold of 68); $10\sigma_{a.v}$ = additive genetic covariance between general and heat tolerance effect; $r^G(a, v)$ = genetic correlation between general and heat tolerance effect; Cor-ht = heat tolerance additive genetic correlations; Cor-gen = regular additive genetic correlations.
Table 3-2. General ($\sigma^2_a$) and heat tolerance ($100 \sigma^2_v$) additive variances, genetic correlations and estimates of heritability $h^2_{i(10)}$ at THI of 78 (10 units above THI threshold) for SCS

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parity 1</th>
<th>Parity 2</th>
<th>Parity 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_a$</td>
<td>0.27</td>
<td>0.28</td>
<td>0.38</td>
</tr>
<tr>
<td>$100 \sigma^2_v$</td>
<td>0.042</td>
<td>0.049</td>
<td>0.091</td>
</tr>
<tr>
<td>$10\sigma_{a,v}$</td>
<td>0.026</td>
<td>0.011</td>
<td>0.081</td>
</tr>
<tr>
<td>$h^2_{i(10)}$</td>
<td>0.11</td>
<td>0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>$\sigma^2_e$</td>
<td>1.34</td>
<td>1.35</td>
<td>1.55</td>
</tr>
<tr>
<td>$r^2(a, v)$</td>
<td>0.24</td>
<td>0.10</td>
<td>0.43</td>
</tr>
<tr>
<td>Cor- ht (par1, parj)</td>
<td>0.18</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Cor- ht (par2, parj)</td>
<td></td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Cor- gen(par1, parj)</td>
<td>0.77</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>Cor- gen(par2, parj)</td>
<td></td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

$\sigma^2_a$ - General additive genetic variances, $100 \sigma^2_v$ = heat tolerance additive variances at 78 (10 units above a THI threshold of 68); $10\sigma_{a,v}$ - additive genetic covariance between general and heat tolerance effect; $r^2(a, v)$ = genetic correlation between general and heat tolerance effect; Cor-ht = heat tolerance additive genetic correlations; Cor-gen = regular additive genetic correlations
### Table 3-3. GO terms significantly enriched with genes associated with CI

<table>
<thead>
<tr>
<th>GO ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0042026</td>
<td>Protein folding</td>
<td>5</td>
<td>2</td>
<td>0.002</td>
</tr>
<tr>
<td>0036297</td>
<td>Inter-strand cross link repair</td>
<td>7</td>
<td>2</td>
<td>0.005</td>
</tr>
<tr>
<td>0003157</td>
<td>Regulation of intra cellular protein transport</td>
<td>66</td>
<td>5</td>
<td>0.010</td>
</tr>
<tr>
<td>0043401</td>
<td>Steroid hormone mediated signaling pathway</td>
<td>34</td>
<td>3</td>
<td>0.030</td>
</tr>
<tr>
<td>0030258</td>
<td>Lipid modification</td>
<td>48</td>
<td>4</td>
<td>0.015</td>
</tr>
<tr>
<td>0007281</td>
<td>Germ cell development</td>
<td>40</td>
<td>3</td>
<td>0.045</td>
</tr>
<tr>
<td>0009791</td>
<td>Post embryonic development</td>
<td>14</td>
<td>2</td>
<td>0.031</td>
</tr>
</tbody>
</table>

### Table 3-4. MeSH terms significantly enriched with genes associated with CI

<table>
<thead>
<tr>
<th>MeSH ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D050884</td>
<td>HSP72 Heat-Shock Proteins</td>
<td>2</td>
<td>1</td>
<td>0.032</td>
</tr>
<tr>
<td>D018840</td>
<td>HSP70 Heat-Shock Proteins</td>
<td>26</td>
<td>3</td>
<td>0.008</td>
</tr>
<tr>
<td>D050883</td>
<td>HSC70 Heat-Shock Proteins</td>
<td>10</td>
<td>2</td>
<td>0.011</td>
</tr>
<tr>
<td>D005982</td>
<td>Glutathione Transferase</td>
<td>50</td>
<td>3</td>
<td>0.048</td>
</tr>
<tr>
<td>D017871</td>
<td>Calcium-Calmodulin-Dependent Protein Kinases</td>
<td>16</td>
<td>2</td>
<td>0.028</td>
</tr>
<tr>
<td>D052250</td>
<td>Nitric Oxide Synthase Type III</td>
<td>73</td>
<td>4</td>
<td>0.043</td>
</tr>
<tr>
<td>D005306</td>
<td>Fertilization</td>
<td>32</td>
<td>3</td>
<td>0.029</td>
</tr>
</tbody>
</table>
Table 3-5. GO terms significantly enriched with genes associated with SCS

<table>
<thead>
<tr>
<th>GO ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0061077</td>
<td>Chaperone mediated protein folding</td>
<td>26</td>
<td>3</td>
<td>0.016</td>
</tr>
<tr>
<td>0006260</td>
<td>DNA replication</td>
<td>57</td>
<td>4</td>
<td>0.031</td>
</tr>
<tr>
<td>0050671</td>
<td>Positive regulation of lymphocyte proliferation</td>
<td>27</td>
<td>4</td>
<td>0.002</td>
</tr>
<tr>
<td>0006909</td>
<td>Phagocytosis</td>
<td>31</td>
<td>4</td>
<td>0.003</td>
</tr>
<tr>
<td>0050727</td>
<td>Regulation of inflammatory response</td>
<td>56</td>
<td>4</td>
<td>0.029</td>
</tr>
<tr>
<td>0042742</td>
<td>Defense response to bacterium</td>
<td>41</td>
<td>4</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Table 3-6. MeSH terms significantly enriched with genes associated with SCS

<table>
<thead>
<tr>
<th>MeSH ID</th>
<th>Term</th>
<th>No. genes</th>
<th>No. sig. genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D0061077</td>
<td>Major Histocompatibility complex</td>
<td>14</td>
<td>2</td>
<td>0.036</td>
</tr>
<tr>
<td>D0006260</td>
<td>Immunity Innate</td>
<td>35</td>
<td>4</td>
<td>0.004</td>
</tr>
<tr>
<td>D0050671</td>
<td>Interleukin-1</td>
<td>29</td>
<td>3</td>
<td>0.012</td>
</tr>
<tr>
<td>D0015850</td>
<td>Interleukin-6</td>
<td>32</td>
<td>3</td>
<td>0.016</td>
</tr>
<tr>
<td>D0054732</td>
<td>Calcium-Calmodulin-Dependent Protein Kinase Type 2</td>
<td>13</td>
<td>2</td>
<td>0.019</td>
</tr>
<tr>
<td>D0053496</td>
<td>Inositol 1,4,5-Triphosphate Receptors</td>
<td>17</td>
<td>2</td>
<td>0.032</td>
</tr>
</tbody>
</table>
Figure 3-1. Manhattan plot showing the results of genome-wide association mapping for CI: the left plots show the result of general merit and the right plots show the results of thermotolerance across lactations numbered vertically as lac1, lac2 and lac3.
Figure 3-2. Manhattan plot showing the results of genome-wide association mapping for SCS: the left plots show the result of general merit and the right plots show the results of thermotolerance across lactations numbered vertically as lac1, lac2 and lac3.
CHAPTER 4
CONCLUSIONS

The results from this study indicate that test-day records combined with meteorological data from public weather stations can be used for genetic-genomic evaluations of heat tolerances. These studies also indicate that heat tolerances additive genetic variances for production, reproduction and health traits increased with parity suggesting that cows become more sensitive to heat stress in later lactations. The genetic correlations between regular and heat stress effects for production and reproduction traits were negative and ranged between -0.30 and -0.55 for MY traits and between -0.35 and -0.82 for reproduction trait. Continued selection for production and reproduction trait without heat stress can result in animals with lower heat stress tolerance and will also deteriorate genetic merit for these traits under heat stress condition. However, the study estimated positive genetic correlations between regular and heat stress effects for SCS ranging between +0.10 to +0.43. Continued selection for lower somatic cell score will also yield positive selection response for heat tolerances in cows.

These studies also have identified genomic regions, and a list of candidate genes and pathways that show significant association with responses of production, reproduction and health traits to heat stress. Our findings reveal that thermotolerance is a quantitative trait influenced by many regions of the genome. Several potential candidate genes and pathways with known roles in thermal regulation, immune response, autophagy, anti-apoptosis among others were identified in these genomic regions. These findings contribute to better understanding of the genetics underlying heat stress and in addition these results can be useful for future fine mapping studies,
identification of causative mutations, and functional studies. Finally, the information provided in this study provides novel opportunities for improving thermotolerance in dairy cattle by means of selective breeding.

Heat stress in our study explained only a small portion of genotype by environment interaction. Moreover, the model in our study identifies only acute response to heat stress and do not consider chronic response to heat stress. It also fails to consider the differences between test days. Therefore, the model could be improved by modeling test days as random effects and by modeling the genotype and environment interactions, but this will increase the computational cost and may raise the issues of convergence of parameters.
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Anil was born and raised in Kathmandu, Nepal. After completing his high school, Anil joined Tribhuvan University where he completed his undergraduate degree of veterinary science and animal husbandry (BVSc & AH). In spring of 2016, he came to the United States to pursue his master's degree in animal sciences and joined University of Florida (UF). He conducted research under the supervision of Dr. Francisco Peñagaricano in dairy cattle genetics and genomics. At UF, he has been conducting his master research to understand genetic aspects of thermotolerance in dairy cattle. His research contributes to a better understanding of the genes and biological mechanisms underlying heat stress and points out novel genomic tools for improving thermotolerance in dairy cattle. His research interests include development and application of statistical and computational models for the analysis of phenotypic and molecular data in dairy cattle.