THESIS

PREDICTIVE POTENTIAL OF GENOTYPES WITHIN THE PROLACTIN, GROWTH HORMONE AND INSULIN -LIKE GROWTH FACTOR-I PATHWAYS IN GENETIC EVALUATION OF 305 DAYS MILK YIELD IN HOLSTEIN COWS IN SONORA, MEXICO

Submitted by

Ana Isabel Hernandez Cordero

Department of Animal Sciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2015

Master's Committee:

Advisor: Milton G. Thomas
Co-Advisor: Richard M. Enns

Scott Speidel
Craig McConnel

Copyright by Ana Isabel Hernandez Cordero 2015

All Rights Reserved

ABSTRACT

PREDICTIVE POTENTIAL OF GENOTYPES WITHIN THE PROLACTIN, GROWTH HORMONE AND INSULIN -LIKE GROWTH FACTOR-I PATHWAYS IN GENETIC EVALUATION OF 305 DAYS MILK YIELD IN HOLSTEIN COWS IN SONORA, MEXICO

The objective of this study was to calculate a molecular breeding value (MBV) using single nucleotide polymorphims (SNP) within genes of the prolactin (PRL) and growth hormone and insulin-like growth factor (GH-IGF1) pathways associated with milk production traits and evaluate their effectiveness in genetic prediction in Holstein cows in Sonora, Mexico. We hypothesized that MBV constructed using DNA markers within the PRL and GH-IGF1 pathways have the potential to predict milk production traits in heat-stressed lactating Holstein cows.

The data contained observations of 659 Holstein dairy cows collected during 2012 from the city of Obregón, Sonora, Mexico. Milk yield observations were recorded monthly and 305 d milk yield was calculated. Cows were genotyped for 179 tag SNP within 43 genes in the PRL and GH-IGF1 pathways. Eight SNP within 5 genes were associated with 305d milk yield ($P \leq 0.05$). No previous research reported these associations. Their effects were used to estimate a MBV. The linear correlation of the MBV and 305 d milk yield was 0.21 and the adjusted R^2 was 4.5%. Genetic parameters were estimated in ASREML for 305 d milk yield ($h^2 = 0.39 \pm 0.11$).

A training and predicting exercise, was performed using SAS 9.4 with the same data set. The SNP effects and association were estimated and used to calculate an MBV. The MBV was estimated and evaluated by comparing estimates from a 5-fold strategy of random clustering. This procedure was repeated five times, resulting in five MBV. To evaluate the effectiveness of these

MBV, correlations and adjusted R^2 were estimated between MBV and 305 d milk yield. One MBV (MBV5) was correlated (-0.27) and had an adjusted R^2 of 6.37%.

The MBV estimated from SNP within the PRL and GH-IGF1 pathways genes was positive but weakly associated with 305 d milk yield. In the training-predicting exercise, only 1 of the 5 MBV explained a portion of the variation in 305 d milk yield. The small amount of phenotypic variation may be due to the small numbers of SNP used to calculate the MBV and the polygenic nature of the trait under heat stress conditions. The quality of the data, could also affect the results. We accept our hypothesis, the MBV was capable of predicting a portion of the phenotypic variation in 305 d milk yield in lactating Holstein cows in Sonora, MX. Nevertheless, the accuracy and amount of variability explained was not enough to be feasible for use in genetic selection procedures.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation for the support and time of my advisor, Dr. Milt Thomas, who has guided me and patiently helped me to improve with each challenge during my Master's program. I would like to thank and acknowledge my committee members, Dr. Mark Enns, Dr. Scott Speidel, and Dr. Craig McConnel for their commitment and support in each of their areas of expertise, which made it possible to accomplish this thesis.

I am also grateful to Dr. Pablo Luna, from the Instituto Tecnológico de Sonora in Mexico, and his students for their collaboration and continuous help. Additionally, I would like to thank Dr. Ann Hess, from the Statistics department for her help with SAS coding to analyze data. I also want to acknowledge Colorado Ag Experimental Station and the department of Animal Science of Colorado State University for providing the tools to analyze the data, and UC MEXUS program for funding this research. Finally, I would like to thank my graduate student cohort, whose support, encouragement and help were priceless.

DEDICATION

I dedicate this thesis to God, my parents, and brothers. My biggest support in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

One of the challenges of dairy production in tropical and hot climates, as well the summer season in non-tropical ecosystems, is heat stress. This stress reduces both milk production and reproduction efficiency of Holstein dairy cattle (Jordan, 2003). With high temperatures and intense radiant energy, lactating dairy cows accumulate metabolic heat, increasing body temperature and subsequently decreasing feed intake and therefore milk production (West, 2003). The decline in milk production due to heat stress has an adverse economic effect estimated in the US dairy industry at approximately $900 million in 2006 (Collier et al., 2006). In September of 2014, United States Department of Agriculture (USDA) reported losses of approximately 1.2 billion dollars in the dairy industry in 2010 because of the heat stress (Key et al., 2014).

To deal with heat stress conditions, most management emphasis has focused on altering and improving the environment of the lactating Holstein. These management strategies such as, providing fresh alfalfa, focus on reducing heat during digestion (Dunshea et al., 2013). Additionally, fans, shade, and showers are used to decrease the negative effect of high temperatures and persistent solar energy (Dunshea et al., 2013). However, most of the genetic selection pressure in Holstein dairy cattle has focused on improving milk yield rather than improving to heat-stress environment (Collier et al., 2006).

With exposure to high temperatures, a heat stress response is initiated in lactating dairy cattle. This response appears to be a highly conserved cascade of gene expression and protein activation atypical to the non-stressed cow (De Rensis and Scaramuzzi, 2003). When heat stress persists, gene expression changes, leading to the alteration of the physiological state, a process referred to as “acclimation” and one largely controlled by the endocrine system (Collier et al.,

2008). Two hormones associated with up-regulation of heat shock proteins (HSP70 and HSP90) are known to increase in plasma in response to thermal stress (prolactin and glucocorticoids). These proteins protect against protein denaturalization (Collier et al., 2008).

For many traits DNA technologies have been used to develop genomic selection methodologies applicable to dairy breeding. The DNA markers (SNP) associated with a quantitative trait loci (QTL), contribute to variations in phenotype. These DNA markers are used to construct genomic or molecular breeding values (MBV) to make selection decisions, especially for phenotypes that are complex and difficult to measure. With this perspective, milk production and reproductive traits in Holstein dairy cattle under heat stress are complex traits and might be suitable for use of DNA technologies in genetic improvement.

Genotype data in dairy cattle, in particular genes of the prolactin pathway, have the potential to be used as tools to improve milk production. Prolactin plays a key role in the initiation and maintenance of lactation in mammals. Genotypes in this pathway are associated with milk production (Lü et al., 2010) and heat stress response in Holstein dairy cattle (Collier et al., 2008). We hypothesized that MBV constructed with DNA markers within the prolactin pathway have the potential to predict milk production traits in heat stressed Holstein cows in Sonora, Mexico; where high ambient temperature is common. The aim of this study is to estimate a molecular breeding value using molecular markers within the prolactin (PRL), growth hormone (GH) and insulin-like growth factor I (IGF1) pathways and evaluate its predictive potential for 305 d milk yield.

CHAPTER 2: LITERATURE REVIEW

Genetic evaluation of dairy cattle

Selection of livestock animals started over 5,000 years ago. The appearance of the animals or their phenotypes based the initial artificial selection (i.e.; milk and beef production). These initial selections used the existing natural variation within a species, within a breed and (or) the population. Traditional selection was made without molecular information of the genes affecting phenotypes of interest. The selection of superior animals for mating thru time has enhanced the breeding values by combining phenotype recording of individual performance with genealogical information (Silva et al., 2014). In the case of milk production, sons of high production cows were retained for breeding. Milk yield has been increasing by 110 kg per animal per year in Holstein cattle (Eggen, 2012), which is a powerful example of the results that can be achieved with breeding methods.

Over time, new methods and technology have been used to create breeding programs via the needs of the production systems and industries. Estimation of genetic merit of dairy cattle using quantitative approaches has been in place for more than half a century. One of the first methods used to achieve rapid genetic improvement was the use of index selection methodology introduced in 1942 (Hazel and Lush, 1942). Selection index uses the correlations between phenotypic measures as well as the genetic relationship between animals and phenotypes to combine several sources of information into a single breeding objective (Silva et al., 2014). Selection index was the first methodology that used pedigree information. The resulting properties of this methodology include decreasing prediction error variance and maximizing the correlation between prediction values and true values, therefore increasing the accuracy of the estimates.

Mixed model and Best Linear Unbiased Prediction (BLUP)

The mixed model methodology developed and calculated more accurate estimates of breeding values as per inclusion of the sire-progeny relationship (Henderson, 1950). This mixed model methodology allowed to estimate fixed (BLUE- best linear unbiased estimator) and random effects (BLUP- best linear unbiased prediction) at the same time (Silva et al., 2014). The BLUP methodology was first applied to genetic evaluation of dairy sires in the northeastern United State in 1970. The BLUP allows an efficient use of all the information available for each individual and it relatives, while adjusting for biases such as age, calving interval, sex, and farm management improving accuracy of the predictions (Parnell, 2004).

The BLUP described as a linear model includes:

$$y = X\beta + Zu + e$$

where y is a vector of n observable random variables, β is a vector of p unknown parameters having fixed values (fixed effects), X and Z are known matrices, and u and e are vectors of q and n, respectively.(Henderson, 1975, Robinson, 1991).

The BLUP estimates of true values of random variables are linear functions of the data and unbiased. The average value of the estimates is equal to the average value of the true value while minimizing the mean squared error. The mixed model equations (MME) we used to calculate an animal model predictor's (BLUP) is defined as:

$$\begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \dfrac{\sigma_e}{\sigma_a} A^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

in which σ_e is the residual variance, σ_a is the additive variance and A is the numerator relationship matrix. In the numerator relationship matrix, additive relationships are a measure of the proportion

of genes, which are identical by descent. Therefore the *A* is necessary to account for the additive genetic covariance between records of related individuals. The construction of *A* is relatively easy, however with a large quantity of animals, computing *A* can be time consuming and inverting a large *A* matrix very time consuming (Henderson, 1975), to deal with this the *A* matrix is generized.

Dairy research and industry evaluations use BLUP methodology for genetic evaluation of milk and fertility traits. It is well known that high milk-producing cows have low fertility (Pryce et al., 2004). Most dairy genetic selection objectives have focused on milk production, with minimal attention to fertility. Given the negative genetic correlation between milk yield and fertility, a negative genetic trend in fertility is expected. In 2004, it was documented a decrease of 1% per year in pregnancy rates occurring at first service (Pryce et al., 2004).

A multi-trait model could help to improve fertility. According to the principles of BLUP, the information in milk production traits will give a benefit to accuracy of EBV for fertility traits as they have moderate, yet negative, genetic correlations (Sun et al., 2010). Thus, a multiple-trait model including milk production will reduce the bias (due to indirect selection) in genetic evaluation of fertility traits. It has been found by comparing several multi-trait models to a single trait model that a single trait model underestimated genetic trend of fertility traits (Sun et al., 2010).

Marker assisted selection (MAS)

The MAS was introduced in the 1990´s. This methodology was initially based on a relative small number of DNA marker such as single nucleotide polymorphisms (SNP). The MAS method was used to eliminate deleterious gene alleles or select for favorable conditions based on some marker information (Eggen, 2012). This methodology requires the prior knowledge of markers association with a given trait with quantitative estimates of these associations in the population of interest. Moreover, MAS, only explains a small portion of the genetic variation (Eggen, 2012).

The MAS method typically consists of a SNP marker locus (ML) closely linked to a QTL. An individual typically possess a two alleles at each locus($M^p\ M^m$), one for each allele inherited from its paternal (p) and maternal (m) parent, and assumes that this marker is linked with a QTL ($Q^p\ Q^m$) (Fernando and Grossman, 1989). This marker linked to the QTL, will tend to be inherited together in recombination events during gamete formation. More than one QTL can affect a given trait and the additive effect of the alleles at the remaining QTLs unlinked to the marker locus will be denoted as a residual effect (Fernando and Grossman, 1989).

When BLUP is used to obtain additive effects, the numerator relationship matrix is constructed using familial relationship information. When MAS is performed using BLUP, the numerator relationship matrix will change because marker information is available (Fernando and Grossman, 1989). For example, with only relationship information, the covariance between half-sibs will be 0.25 or 25 %. When marker information is available, covariance between half-sibs that receive the same marker allele from their common parent is higher tha the covariance between half-sibs that receive different marker alleles.

Genomic selection (GS)

The efficiency of traditional methods of selection decreases when traits are complex, hard to measure, and (or) have low heritability. Examples of such trait categories include: fertility, longevity, feed efficiency, environmental tolerance, and (or) disease resistance. Traditional methods of selection or genetic evaluation use only phenotypic data and probabilities assuming that genes are identical by descendent using pedigree information (Forni et al., 2011). Genomic selection use molecular data and pedigree data to construct a genomic relationship matrix that potentially improves the accuracy of the predictions.

Principles of genomic selection in dairy cattle

In the last decade, quantitative dairy traits have been selected and studied with the aid of DNA markers. With the large number of available DNA markers, such as SNP, automated methods for SNP genotyping were developed and are commercially available. The use of SNP arrays that cover the bovine genome and explain a large portion (45% in extreme casess) of the genetic variation in economically relevant traits has been the tool to develop genomic selection (Schefers and Weigel, 2012, Silva et al., 2014). The GS is based on the principle that information from a large number of DNA markers (SNP) can be used to estimate breeding values without knowledge of the causative gene locations in the genome for the trait of interest (Eggen, 2012).

The GS uses similar methods as marker-assisted selection, but with large SNP panels across the entire bovine genome. When large numbers of SNP are use to analyze quantitative traits, most markers will be indirectly associated with the causative gene mutation and most likely in linkage with the causal mutation (Silva et al., 2014). This suggested the quantitative trait nucleotide (QTN) and QTL are inherited together. Nonrandom association of alleles at different loci is called linkage disequilibrium (LD) and is affected by recombination events (Nordborg and Tavaré, 2002). Recombination take place during the formation of gametes, and involve cross over events among homologous chromosomes; the pair of chromosomes exchange genetic material in a random fashion. Since the recombination rate of two loci depends on the physical distance between them on a chromosome, the smaller the distance between loci, the slower the frequency of the locus will get to equilibrium under generations of random mating (Silva et al., 2014) and closely linked loci will tend to be highly correlated as well (Nordborg and Tavaré, 2002).

The LD is calculated by using statistics of association between two allele loci. To calculate LD, consider two loci (A and B) with alleles A_1 A_2 and B_1 B_2 and alleles frequency p_{A_i} and p_{B_j},

respectively. Let $p_{A_iB_j}$ stand for the frequency of the haplotype A_iB_j. $|D|'$ is the absolute value of

$$D = p_{A_1B_1} - p_{A_1}p_{B_1}$$

normalized to take values between 0 and 1 regardless of the allele frequency;

$$r^2 = \frac{D^2}{p_{A_1}p_{A_2}p_{B_1}p_{B_2}}$$

the squared correlation in allelic state as they occur in haplotypes. Both of these measures are symmetric since in these two scenarios is not relevant alleles associations (Nordborg and Tavaré, 2002), and therefore there is equlibrium between the alleles. A different equation for pairwise association

$$d^2 = \left(\frac{p_{A_2B_1}}{\frac{p_{A_2} - p_{A_1B_1}}{p_{A_1}}} \right)$$

can measure the association between two alleles (Nordborg and Tavaré, 2002), d^2 is the prefered value to estimate LD. The common threshold used for d^2 is 0.05, which mean, a pair alleles of different genes with values of d^2 above 0.05 are consider to be in LD. Nevertheless, researchers have access to multi locus data and this approach would not be efficient measuring multiple alleles within a locus at the same time.

Genome-wide association studies (GWAS)

The GWAS is used to analyze and understand the variation in complex traits and detect QTL. This is achived by associating phenotypes with genotype of a large number of molecular markers (SNP) covering the whole genome and phenotypes (Gondro et al., 2013). The commercialization of SNP-chips that cover the entire bovine genome have made it possible to obtain large amounts of genomic data on individuals and conduct GWAS. This methodology exploits linkage disequilibrium between markers and causative mutations (Gondro et al., 2013).

The simplest GWAS methodology involves marker regression on the phenotypes with the following equation:

$$y = Wb + Xg + e$$

where y is vector of phenotypes, W an incidence matrix relating fixed effects to records, b is the fixed effect vector, X is the matrix assigning records to the marker effect, g is the marker effect and e is the error term or residual (Gondro et al., 2013). The marker effect is treated as a fixed effect and the model is additive, consequently two copies of the same allele have twice the effect of one copy assuming that the marker will affect the trait if it is in LD with the QTL. The LD between the allele and the QTL decrease base in the recombination rate, in this case the allele will be in equilibrium with the QTL and the assumption would not be correct.

In the case of single marker regression, the mean and the SNP effect can be calculated as:

$$\begin{bmatrix} \hat{\mu} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} W'W & W'X \\ X'W & X'X \end{bmatrix}^{-1} \begin{bmatrix} W'y \\ X'y \end{bmatrix}$$

in which y is the number of animals or vector of phenotypes.

Whole genome association studies can also use haplotypes. The use of haplotypes has an advantage as they may be in greater LD with QTL alleles than single markers. Assuming that this is true, the r^2 between the haplotypes and the QTL would increase, as a result the power of the GWAS would increase (Gondro et al., 2013).

Identical haplotypes can be generated through two different methodologies, they can be derived from the same common ancestor (identical by descent (IBD)), or the same marker haplotypes can be generated by recombination (identical by state (IBS)). If the haplotype contains only a single SNP, the possibility of being identical by state depends on the homozygosity,which mean identical alleles of a gene are in both chromosme . Therefore, the chance of identical haplotypes by recombination is reduced as more markers are included in the haplotypes hence the

chance of identical haplotypes by state (IBS) decreased. Chromosome segments with haplotypes identical by descent will carry the same QTL alleles, and as the haplotypes increase with number of markers, the variance proportion explained by the QTL will increase; thus, the haplotype is more likely to be associated with the QTL (Gondro et al., 2013).

Haplotype frequency can be calculated, if there are two alleles for a QTL and their frequencies (q_1 q_2) are estimated, then the surrounding markers can be classified into n haplotypes with there on frequencies, where p_i is the frequency of the surranding markers. Consequently the haplotype frequency for the QTL allele 1 and QTL allele 2 would be $p_i q_1 - D_i$ and $p_i q_2 - D_i$ respectively, where i represents a particular haplotype. The disequilibrium is calculated as $D_i = p_i(q_1) - p_i q_1$, where $p_i(q_i)$ is the proportion of haplotypes i from the QTL allele1 (Gondro et al., 2013). The variance of the QTL explained by the haplotypes can be estimated as:

$$r^2(h,q) = \frac{\sum_{i=1}^{n} \frac{D^2}{p_i}}{q_1 q_2}$$

The model that tests the association of haplotypes is described as

$$y = 1_n'\mu + Xg + Zu + e$$

but g is the haplotype effect rather that a single marker effect and can be estimated from the equations:

$$\begin{bmatrix} \hat{\mu} \\ \hat{u} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 1_n{}'1_n & l_n{}'Z & 1_n{}'X \\ Z'1_n & Z'Z + A^{-1}\lambda_1 & Z'X \\ X'1_n & X'Z & X'X + I\lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} 1_n{}'y \\ Z'y \\ X'y \end{bmatrix}$$

where $\lambda_1 = \frac{\sigma_e^2}{\sigma_a^2}$ and $\lambda_2 = \frac{\sigma_e^2}{\sigma_h^2}$ and σ_h^2 is the haplotype variance that in practice is estimated since is unlikely to be known.

Genomic best linear unbiased prediction (gBLUP)

The gBLUP is a modification of BLUP methodology that involves the use of genomic relationships to estimate genetic merit of an individual. The DNA marker (SNP) information is used to construct a genomic relationship matrix that establishes the covariance between animals at a genomic level (Gondro et al., 2013). The gBLUP has been analyzed in various research studies and has shown to be as accurate or more accurate that pedigree-based BLUP; it has been reported to increase EBV accuracy 20 to 50% (VanRaden et al., 2009). The gBLUP is being used to predict genetic merit in livestock breeding (genomic estimated breeding value-GEBV), and it has also been used to study complex traits like disease resistance and low heritable traits. This type of analysis is the most commonly used for genetic prediction in Holstein cattle (Pryce et al., 2004, Gondro et al., 2013).

To execute GS, it is necessary to incorporate DNA marker information into the relationship matrix used in BLUP. Genomic relationships can better estimate the proportion of genetic information shared by individuals. High-density genotyping identifies loci identical in state that may be shared through common ancestors and not recorded on the pedigree (Forni et al., 2011).

To incorporate genomic information into genetic evaluation, there are different methods that can be used. One method is the ridge-regression BLUP (RR-BLUP), which assumes that SNP effects are assumed to be random, in which the function relating genotype $g(x_i)$ to EBV and can be considered as a molecular breeding value (MBV) having the form

$$g(x_i) = \sum_{k=1}^{p} x_{ik}\, \beta_k ,$$

where β_k is the effect of each SNP, x_k is the SNP genotype (0, 1 or 2) at locus k (Moser et al., 2009, Gondro et al., 2013). The regression coefficient can be found by solving the equation

$$\hat{\beta} = (X'X + I\lambda)(X'X + I\lambda)^{-1} X'y$$

in which $\lambda = \frac{\sigma_e}{\sigma_a}$ is constant for all SNP. Another method that can be used to incorporate genomic information in BLUP is by substituting the numerator relationship matrix for a genomic relationship matrix (gBLUP). The linear model used in this practice is:

$$y = Xb + Zg + e$$

where y is a vector of phenotypes, X is a known matrix relating the fixed effects to each animal, b is a vector of fixed effects, Z is a design matrix assigning records to the markers effects, g is a vector of additive genetic effects for an individual, and e is a vector of residual of error terms. In addition, $\text{var}(g) = G\sigma_g^2$ in which G is the genomic relationship matrix and σ_g^2 is the genetic variance for the model (Gondro et al., 2013).

The mixed model equation for an animal model using gBLUP looks like:

$$\begin{bmatrix} X'X & X'Z & 0 \\ Z'Z & Z'Z + G^{11} & G^{12} \\ 0 & G^{21} & G^{22} \end{bmatrix} \begin{bmatrix} b \\ g1 \\ g2 \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ 0 \end{bmatrix}$$

where G^{11} G^{12} G^{21} G^{22} are the positions into the genomic relationship matrix; G^{11} is the subgroup of individuals having phenotypic and genotypic information; G^{12} G^{21} are the relationship between the animals with phenotypic information and without phenotypic information, and G^{22} represent the animals without phenotypic data.

Other methods, in addition to RR-BLUP and gBLUP that can be used to incorporate genomic information, include Bayes regression (Bayes-R), support vector regression (SVR), and partial least squares regression (PLSR). All of these methods have been reported to generate similar accuracy for MBV estimates (Moser et al., 2009). Nevertheless, the incorporation of genomic information using gBLUP methodology has several desirable characteristics. The gBLUP is computationally convenient and efficient since the dimension of the genetic effect are

maintained as $m \times m$ and accuracy can be calculated in the same way it is done in pedigree-based BLUP estimates. This genomic information can be integrated with pedigree information in a single step method.

Genomic relationship matrix (G-matrix)

The idea to integrate genomic information with the numerator relationship matrix (A) was suggested in 1997 (Nejati-Javaremi et al., 1997) and several methods have been proposed and applied to livestock production (Forni et al., 2011). A BLUP evaluation using a single matrix (H) that contains the combined information from the numerator relationship matrix (A) and the genomic relationship matrix (G) in a single step method has been effectively applied to dairy cattle (Forni et al., 2011). The formula for H calculates the difference between genomic and pedigree based relationships($G - A$).

VanRaden et al. (2009) presented these equations to calculate a genomic relationship matrix:

$$G = \frac{(M - P)(M - P)'}{2\sum_{i=1}^{m} p_i\,(1 - p_i)}$$

or

$$G = \frac{WW'}{2\sum_{i=1}^{m} p_i\,(1 - p_i)}$$

he defined an incidence matrix M in an allele-sharing matrix with m columns (m=total number of markers) coded as -1, 0, 1 (-1=homozygote (AA), 0=heterozygote (AB), 1=homozygote (BB)) that represented the alleles for each individuals and n rows (n=numer of individuals). A P matrix that contains the frequency of the second allele (p_i) or minor frequency allele, expressed as a difference from 0.5 and multiplied by 2, such that i of P is $2(p_j - 0.5)$. W is therefore the subtraction of P from M.

Yang et al. (2010) proposed another genomic matrix approach. Specifically, the information of all the SNP (i) coded as 0, 1, 2, to calculate the relationship among the individuals j and k into the genomic relationship matrix (G_{ijk}) . The scheme used to construct the genomic relationship matrix is based on allele frequency similar to VanRaden et al. (2009) but weighted the off-diagonal and diagonal elements differently, when ($j \neq k$) then:

$$G_{ijk} = \frac{1}{N}\sum_{i} G_{ijk} = \frac{(w_{ij} - 2p_i)(w_{ik} - 2p_i)'}{2p_i(1-p_i)}$$

and when j is equal to k then:

$$G_{ijk} = \frac{1}{N}\sum_{i} G_{ijk} = \frac{1}{N}\sum_{i} \frac{w_{ij}^2 - (1-2p_i)w_{ij} + 2p_{ij}^2)'}{2p_i(1-p_i)}$$

where w_{ij} is the element of W relating to marker i and individual j. These estimates of relationships between individuals were all relative to a base population in which the average relationship between individuals were zero.

Likewise, Goddard et al. (2011) described another approach constructing a genomic relationship matrix, where G_m matrix that can be constructed as $G_m = WW'/M$ in which the elements of matrix W is constructed using the Yang et al. (2010) method:

$$M = \sum 2p_j\left(1-p_j\right)$$

then $\hat{G}$ can be calculated as:

$$\hat{G} = [A + b(G_m - A)]$$

where A is the numerator relationship matrix base on pedigree information and $b = \sigma_g^2/\sigma_a^2$ (σ_g^2= variance of each marker effect; σ_a^2= additive genetic variance). With the regression of $\hat{G}$ back

towards *A*, some of the error associated with the estimation of *G* from a finite number of markers is removed and consequently $\hat{G}$ is an estimate of the true genomic relationship.

Accuracy of Genomic breeding values from dairy cattle breeding programs

With the progress that has come from GS, predictions have become more accurate using dense marker genotypes and phenotypes. Using a large quantity of markers across the entire bovine genome, allowed use of the effect of all loci in the prediction, even if the effect of each locus was very small. Simulation studies found accuracy of 85% of predicting breeding values using genome-wide markers (Meuwissen et al., 2001) in animals (bovines). Meuwissen et al. (2001) concluded that selection on genetic values predicted from markers could substantially increase the rate of genetic gain in animals and plants, especially if combined with reproductive techniques to shorten the generation interval.

The rate of genetic gain can potentially be doubled by using GS (Hayes et al., 2009). Traditional progeny testing takes a long time as the generation interval is approximately 63 months (Schefers and Weigel, 2012). The GS allows AI companies developing young sires to make selection decisions based on genomic breeding values (GEBV) in very young bulls. The GS makes possible the use of bulls at approximately 12 months of age (Schefers and Weigel, 2012).

GS in dairy cattle breeding programs from Australia, New Zealand, and the United States have proven that GEBV are significantly greater breeding values than breeding values from parent averages . In these countries, the GEBV were calculated by combining the parental average breeding value with the genomic information using selection index theory (Hayes et al., 2009).

The primary goal of GS is to improve accuracy of the genetic predictions. The accuracy of the GEBV depends on four parameters :

1. Level of LD among the markers and the QTL

2. Number of animals with genotypes and phenotypes
3. Heritability of the trait
4. The distribution of the QTL effect

The first two parameters are under the control of the experimenter hence there is an opportunity to increase the accuracy by improving these two. For example, it has been reported that as r^2 between the markers and the QTL increases, the accuracy of the GEBV will do so as well (Hayes et al., 2009).

Implications of GS and genotyping

The implementation of GS has great advantages: it can be applied from an early age, is not limited by sex, and can be used in any trait that can be measured in a reference population. The GS is particularly useful for traits that have a low heritability or are difficult to improve (complex), since GS allows the use of more information to estimate breeding values and polymorphism may be associated with different phenotypes of a trait. The use of DNA markers in genetic evaluation and selection can improve accuracy, decrease generation intervals and increase intensity of selection (Eggen, 2012). Consequently, GS explains a greater amount of genetic variation than traditional selection (no DNA markers (SNP)).

Since the application of genomics has been of great use in many fields, not just animal, but also human health, the cost for sequencing and genotyping have been decreasing with recent adoption of the technology. If this trend continues, the next big step could be the integration of whole genome sequence information into the genetic evaluations (Eggen, 2012). Currently genomic information is helping improve accuracy of predictions in dairy and beef cattle national evaluations (Hayes et al., 2009).

The GS provides a great opportunity for genetic improvement in developing countries. Often, recording of phenotypes in pedigreed animals, and genetic evaluations for herd improvement are absent in developing countries. In these cases, a genomic approach can help to identify animals that are well adapted to a particular weather or environmental conditions. These animals could be selected to breed or crossbreed, and improve the quality of the animals (Eggen, 2012). However it should be noted that the training and prediction process requires substantial numbers of animals and genotypes (Boddhireddy et al., 2014)

Genomic prediction validation

The validation of genomic selection involves using a training population with genotypes and phenotypes to simultaneously estimate SNP effects (Meuwissen et al., 2001). The SNP effects can be combined with EBV to improve predictions. Several factors influence genomic prediction accuracy including the sample size for the trained population, and the relationship between the discovery and the target validation population (Habier et al., 2007, Clark et al., 2012), the type of phenotypic variable used for estimating the SNP effects, and the methodology used for grouping the data for cross-validation (Saatchi et al., 2011). Some of the methods that can be used for clustering for cross-validation are: K-means (Saatchi et al., 2011), IBS-based (Aulchenko et al., 2007), random clustering, and IBS-based with unequal cluster size (Boddhireddy et al., 2014). Other factors include: extend of LD, number of QTL contributing to the phenotypes, the heritability of the trait, and the accuracy of the measurements of the phenotypes.

Traits of dairy breeding programs using genomic selection

Milk yield is the primary trait in the dairy industry although reproductive traits have been recently garnering more attention, since high producing Holstein have low fertility. Seykora and McDaniel (1983) reported heritability for milk yield adjusted to 305 days in milk and day-open in

Holstein dairy cattle to be 0.32 ± 0.04 and 0.13 ± 0.04 respectively. Van Dorp et al. (1998) report a heritability for milk yield of 0.26 in Holstein dairy cattle. Research reported that the heritability for milk production ranged from 0.30 to 0.37 and heritability of fertility from 0.08 to 0.02 in Holstein dairy cattle using GS (Karoui et al., 2012).

The effectiveness of GS depends on the use of a large reference or training population with animals that have both phenotype and genotype information.. In the United States, more than 30 traits are evaluated for dairy production and they are related to health, milk yield, and fertility. The traditional categories include: milk yield, protein yield, fat yield, protein percentage, fat percentage, cow productive life, somatic cell score and daughter pregnancy rate (Silva et al., 2014).

Milk production is a highly specialized and competitive trait. Selection objectives need to include traits for profitability and animal efficiency. Consequently new traits, not traditionally measured, are being evaluated and included in genetic evaluations. Some of these traits are feed efficiency, energy balance, diseases resistance, novel fertility traits, resistance to heat stress and calf birth weight (Silva et al., 2014). The limitation with new traits is that there are not large reference populations, and often these traits have low heritability. Therefore, the accuracy of GEBV for these traits is low.

Lactation curve

"Accurate description of a lactation curve is relevant to activities such as conducting feeding trials with lactating cattle, estimating total lactational yield from incomplete records, and forecasting herd performance on a monthly or individual cow basis" (Sauvant, 1988). Dairy cows with a flat lactation curve are considered to have more persistent lactations than those with the same lactation yield but a steep lactation curve (Tekerli et al., 2000).

To estimate a lactation curve there are several equations, some of those are presented in Table 2.1 (Val-Arreola et al., 2004). The different models to represent the lactation curve are an essential research tool to explain the main features of the milk production pattern in terms of the known biology of the mammary gland during pregnancy and lactation (Macciota et al., 2005), since the lactation curve is influence for two factors, the processes of cell growth and death

Table 2.1 Equations used to describe the lactation curve of dairy cows managed under small-scale and intensive systems in central Mexico.

Equation	Functional form[1]
Ganines	$Y = \alpha e^{-bt}$
Wood	$Y = \alpha t^{b} e^{-ct}$
Rook	$Y = \alpha\{\frac{1}{\left[1 + \frac{b}{(c+t)}\right]}\}e^{-dt}$
Dijkstra	$Y = \alpha \exp[\frac{b(1-e^{-dt})}{c} - dt]$

[1]Y is milk yield (Kg/d), t is time of lactation (d), and a, b, c, d (all >0) are parameter that define the scale and shape of the curve.

Genes involved in milk production traits and milk composition in Holstein dairy cattle

Quantitative genetic approaches have resulted in tremendous genetic improvement in milk yield, and the use of genotype information in genetic evaluation made the predictions even more accurate. Many researchers have studied the genes associated with milk production and composition (Yao et al., 1996, Cobanoglu et al., 2006), as well as genes related with novel traits, such as feed intake and energy balance (Liefers et al., 2002) that could improve the accuracy of predictions for milk yield. Important genes are discussed below.

Signal Transducer and Activator of Transcription 1 (STAT-1)

This gene is part of a family of signal transducer and activator of transcription factors (i.e: Janus kinase (JAK)/STAT) that play a role on the activation on cytokine, growth factors, and

hormones (Sengupta et al., 1995). Seven bovine genes have been identified (Ihle, 2001, Cobanoglu et al., 2006) to be part of this family of genes. The *STAT1* gene is expressed during pregnancy and lactation (Darnell, 1997) in endometrial tissue (Carvalho et al., 2014) and other tissues such as liver, which play a key role during lactation as well. There is some evidence that suggest that *STAT1* is involved in the development and differentiation of mammary gland development (Cobanoglu et al., 2006). This gene is located on chromosome 2, and whole genome scans have shown an association between milk yield and DNA markers near the *STAT1* gene. Cobanoglu et al. (2006) identified an SNP (C/T) in the *STAT1* gene in Holstein dairy cattle. The allele combination of this SNP (CC and CT) were associated with and increased milk yield, fat, and protein.

Growth hormone (GH)

This gene plays a key role in nutrient utilization, mammary development, growth, lactation, intermediary metabolism, reproduction and several other important endocrine physiological processes (Le Gac et al., 1993, Yao et al., 1996). The somatotropin is synthesized and stored by somatotroph cells within the anterior pituitary gland (Kim, 2014). Six polymorphisms were found by Yao et al. (1996) in 128 Holstein bulls, from which two of them (T/C and A/C) were located in the third intron and fifth exon of the GH gene. These polymorphisms were found to be associated with milk production traits. The leucine to valine non-synonymous SNP caused by a C to G nucleotide change in the fifth exon of the GH gene is related to milk yield (Yao et al., 1996); Holstein cows injected with the valine recombinant variant of GH had a higher milk yield than those cows that were injected with the Leucine variant (Eppard et al., 1992). However, other researcher, have found a decrease in milk yield with the valine variant of GH in Holstein cows (Lee et al., 1993, Lucy et al., 1993).

Leptin

Adipose tissue synthesizes this protein, which regulates feed intake, fertility, and immune response (Frühbeck et al., 1998). Leptin plasma concetration is known to increase in response to lipids and glucose stimulation (Chelikani et al., 2003). Liefers et al. (2002) found that a polymorphism in the leptin gene in Holsteins cattle was associated with the initiation of luteal activity, energy balance, milk yield and mean live weight. They reported that heifers with the favorable allele (144bp), situated in exon 3 which causes an amino acid change from Alanine to Valine, produced 1.32 kg/d more milk and consumed 0.73 kg/d less feed in comparison than cows with the non-favorable allele (136bp) located in the intron between the two exons of the Leptin gene. More evidence of leptin genotype associated to milk yield and fertility have been report as well by Clempson et al. (2011). This study investigated associations of SNP in the leptin and leptin receptor with milk and fertility traits. Mixed model analyses revealed that leptin SNP were associated with early skeletal growth, fertility, and milk production (Clempson et al., 2011).

Prolactin (PRL)

Prolactin is a peptide hormone released by the anterior pituitary gland; however, its secretion is also attributed to central nervous system, immunological system, and mammary gland (Bole-Feysot et al., 1998). This hormone is pleiotropic as, more than 300 biological functions have been attributed to PRL (Bole-Feysot et al., 1998). Biological functions of the PRL can be categorized in five groups: reproduction, osmoregulation, growth, integument (natural cover of an organ or organism, i.e: skin), and synergism with steroids. It has been suggested that PRL is involved in parental behavior in mammals. This protein is codified by the PRL gene located on chromosome 23 and it has 5 exons and 4 introns (Freeman et al., 2000, Rincón et al., 2012). Most importantly, the PRL protein is known for its action on the mamary gland. During pregnancy, the

growth of this gland is regulated by hormones such as estrogen, progesterone, insulin, GH, and PRL (Freeman et al., 2000). Although hormonal requirements for the maintenance of milk production change in different species, the common factor is that PRL in the primary hormone secreted from the anterior pituitary lactotrophs (Ben-Jonathan and Liu, 1992) responsible for milk synthesis of proteins, lactose and lipids. The PRL hormone is also involved in adaptive response to stress in mammals, as it exhibits analgesical effects that can be mimicked by a number of neurotransmitters (Bole-Feysot et al., 1998).

Other genes

Researchers have constructed a database that contain 934 gene loci involved in mammary gland development, milk production traits and resistance or susceptibility to mastitis (Ogorevc et al., 2009). Several studies have shown other genes associated with milk production, milk composition and other traits such as fertility, and diseases resistance. Included in this list are ATP-Binding Cassete Sub-family G Member 2 (ABCG2), which transports various molecules across extra and intra cellular membrane and is funcionally active in mamary gland tissue (Cohen-Zinder et al., 2005), PRL and STAT-5A (Schennink et al., 2009).

Genes involved in heat stress response

Heat stress response is a highly conserved alteration of gene expression and proteins across the body (Collier et al., 2008). Gene expression dynamics due to heat stress is regulated by the transcription factor Heat Stress Factor (HSF). The HSF promotes the transcription of heat shock proteins (HSP) that have a key role of cytoprotection during heat stress response.

A primary research focus of heat tolerance are the genes that regulate hair coat. In high temperatures and in dairy cattle, cutaneous evaporative heat loss (EVHL) is active. Factors such as sweat gland density and function, hair coat density and thickness, hair length and color, skin

color, and regulation of epidermal vascular supply effect EVHL (Collier et al., 2008). A major gene (slick hair gene) found to be responsible for short and sleek hair coat, is located on the bovine chromosome 20 (Mariasegaram et al., 2007, Flori et al., 2012). Flori et al. (2012) found one positional candidate gene, namely "Retinoic Acid induced 14 gene" (RAI14 or NORPEG). This region was narrowed even further, to a 0.8 Mb (37.7-38.5 Mb) consensus region for the slick hair gene locus on the bovine chromosome 20. This region contains the S-phase kinase-associated protein 2, E3 ubiquitin protein ligase (SKP2) and sperm flagellar 2 (SPEF2), which are possible candidate genes for heat tolerance (Huson et al., 2014). Polymorphisms in these genes are often found in heat tolerant *Bos taurus* breeds such as Senepol and Romosinuano of Central and South America (Olson et al., 2003). Slick hair has shown to decrease rectal temperature in crossbred animals such as a Holstein × Carora. This effect on rectal temperature is not only dependent on the slick hair gene but also in the degree of heat stress, age of the animals, and lactation status (Olson et al., 2003).

Additionally, in 2014 was found a region of recent selection in chromosome 20 that contain the Slick gene in *Bos taurus* cattle (Flori et al., 2012). The SPEF2 (38.4 Mb) and prolactin receptor (PRLR) (38.0 Mb) loci are within this region and correspond to reproduction and milk production (Huson et al., 2014). Therefore, the PRLR may play a roll not only in lactation but also other physiological functions. These candidate genes could be selection targets to improve thermal adaptation in lactating *Bos taurus* populations exposed to heat stress.

In cattle, PRL-hormone interacts with target cells by binding to the PRL receptor located in the membrane (Żukiewicz et al., 2012). Two isoforms of PRLR have been described, resulting from alternative splicing: a long form, with a length of 557 amino acids, and a short one, with a

length of 272 amino acids. The PRLR gene is mapped on the bovine chromosome 20, from the 38.9 Mb to 39.1 Mb position (Żukiewicz et al., 2012).

Candidate gene criteria for GS

Most candidate genes are related to a trait of interest or physically near the gene associated with a trait. Candidate genes by definition are those that can positively help improve relevant traits since they exhibit gene polymorphism associated with phenotypes being recorded. A common approach in dairy research is to study genes with sequence variations that show allele-phenotype interactions associated with milk production or mastitis. In addition, they can be also a genetic marker associated with an animal trait (QTL), which encompass candidate genes (Ogorevc et al., 2009).

Influence of the environment on Holstein dairy cattle

The environment (weather/climate and management) has both direct and indirect influence on health and production of dairy animals. The degree of environmental influence depends on the stage of the life cycle and adaptation of breeds (Collier et al., 1982). Cold stress in lactating Holsteins has little effect on reproduction; in contrast, heat stress reduces libido, fertility and embryonic survival. Under heat stress in late gestation, fetal growth is reduced, and endocrine status of the dam is altered. Blood hormone concentrations that are altered include increases in glucorticosteroids and prolactin, and decreases in gonadotropins (Collier et al., 2008). Additionally, intrauterine blood flow decreases and uterine temperature increases. These changes increase the probability of embryo death and inhibit embryo development (De Rensis and Scaramuzzi, 2003).

Environmental factors that influence dairy cattle are temperature of the air, humidity, wind speed, solar radiation, management and breed. These factors, as well as the influence of feed

quality and quantity, make an even greater impact on the animal health, fertility and production. In 2003, West reported that in high temperatures, body temperature increased leading to health problems in Holstein cows. Some of these health problems included mastitis and metritis (West, 2003). Other researchers have reported that high humidity and temperature caused higher cow mortality due to heat stress (Vitali et al., 2009).

Climate effects

Climate is a combination of elements that include temperature, humidity, rainfall, air movement, radiation, barometric pressure and ionization (West, 2003). Changes in climatic conditions are known to affect dairy cattle. When dairy cattle are exposed to climate conditions that are out of their zone of thermoneutrality (thermoneutrality: rate of heat production = rate of heat loss), the cows start to make metabolic adjustments to maintain homeostasis. Temperatures that are below or above of critical temperatures (-25ºC to 25ºC) have a direct negative effect on a lactating cow milk yield. Lactating cattle in general are more heat sensitive and cold resistance (Collier et al., 1982). Lactation has a significant physiological impact on dairy cattle. There are strong effects on conception rates between heifers and lactating cows when the environmental temperature increases. Therefore, reproductive performance of lactating cows is greatly affected, in contrast, non-lactating heifers show little response to increased ambient temperature even in humid environments (Collier et al., 2006). Specific to reproduction, conception rate decreases in cows in environmental temperatures above 30ºC in comparison with heifers conception rate that are negatively affected, when temperatures reach 35ºC in subtropical environments (Badinga et al., 1985).

Effect of climate on reproduction

Bading et al. (1985) reported a negative relationship between conception rate and maximum air temperature on the day after insemination. They reported an increase in ambient temperature from 23.9ºC (March) to 32.2ºC (July) yielded a decrease in conception rate from 52% to 32% (Badinga et al., 1985).

Heat stress is associated with a reduction in duration of estrus and can lead to a poor detection of estrus. High temperatures also affect gamete formation since the production of spermatozoa requires temperatures below the body temperature (De Rensis and Scaramuzzi, 2003). Follicular development in cows is influenced by the alteration of hormones. Follicle stimulating hormone (FSH) increases while lutenizing hormone (LH) decreases due to high temperature subsequently affecting synthesis of estradiol (De Rensis and Scaramuzzi, 2003). Embryo survival is negatively affected by heat stress, since high environmental temperature is associated with a detrimental effect on developing embryos (Badinga et al., 1985). High temperatures will reduce prostaglandin secretion by the endometrium leading to early luteolysis, increasing early embryonic loss and reducing successful inseminations (De Rensis and Scaramuzzi, 2003).

Management and climate

The understanding of the effects that climate has on animal production has increased over the years. Crossbreeding Holstein dairy cattle with breeds that are more tolerant to high temperature tend to decrease milk production, nevertheless is an opportunity for heterosis (Jordan, 2003). Management strategies like fans, shade, showers, fresh water and hay, have been used and proven to decrease the detrimental of heat stress effect on milk production (Jordan, 2003, Collier et al., 2006). Management of the animals accounts for a portion of the phenotypic variation in dairy

cattle traits (i.e.; producer breed preference, time and number of inseminations, year of service and estrous detection method, etc.).

A breed-to-breed differences in response to the heat stress has been documented (Badinga et al., 1985). When the cows were managed in a hot-stressful enviroment Jersey cows had higher conception rates (45%) than Holsteins (39%) and Brown Swiss (41%) cows. In addition, breed differences in relation to the number of services necessary for conception were also documented (Badinga et al., 1985). Jersey cows had an average of 1.7 services per conception compare to Holsteins with 1.9 services and Brown Swiss with two services. These differences are most likely due to differences in milk production and the thermoregulatory physiology of each the breeds.

Estrus detection techniques have a significant effect on pregnancy rates due to the precision of each method. The research conducted by Badinga et al. (1985) revealed that cows and heifers presented for insemination after observed standing estrus were more fertile (46%) than the females whose estrus was detected by heat mount detectors (41%) and mounting activity (39%). It is probable that lower pregnancy rates are affected by false positive and not by infertility of the cows or heifers. Previews research have shown that lactating Holstein cows under heat stress have the same estrus length as that of control Holstein population ($P > 0.10$) (Trout et al., 1998). Trout et al. (1998) report the progesterone levels increased under heat stress conditions.

Mastitis

Management practices and overall production trait levels are related to the incidence of diseases. Height temperatures and summer season of the year have a negative effect on the health of dairy cattle. For example incidence of Mastitis increase during summer season. Mastitis is the most costly and frequent disease in the dairy industry and can be limited by improving management practices, such as hygiene. Mastitis is an inflammation of the mammary gland due to

pathogens infection like *Staphylococcus aureus*, yet this infection can be caused by a wide range of microorganisms (Heringstad et al., 2000, Halasa et al., 2007).

Mastitis has a complex nature and a result of multiple factors. Some of these factors are hygiene, bedding, milking frequency, age, and exposure to microorganism. Improvement in the incidence of mastitis consists of taking management action on the treatment of the disease, dry cow therapy, and prevention the transmission of the infection (Halasa et al., 2007). Many antimicrobial drugs have been used as a treatment, including compounds that do not readily penetrate the mammary gland; such as, penicillins and sulfonamides (Barkema et al., 2006).

Given the high frequency and cost of treating lactating cows for mastitis, genetic selection is a valid option. Breeding for increased resistance or less susceptibility can be performed by direct selection using clinical records, or indirectly using traits genetically correlated to mastitis. The commonly used indirect measure is somatic cell count (SCC). The SCC consists of many types of cells including, neutrophil leukocytes, macrophages, lymphocytes, eosinophils, and epithelial cells types of the mammary gland. Neutrophils in particular are found in more than the 95% of diseased udders. The correlation between SCC and clinical mastitis indicated that both are related to udder health. Reducing mastitis should always be included in breeding programs, using either direct or indirect measures (Heringstad et al., 2000).

Heat stress in Holstein dairy cattle

Season of the year has an impact on dairy cattle milk production, growth and reproduction (Jordan, 2003). Under constant exposure to high temperatures and intense radiant sunlight, heat stress is initiated. These conditions are observed in subtropical and tropical climates as well in the summer season around the world. Heat stress can by simply defined as a condition that occurs when an animal cannot dissipate an adequate quantity of heat, whether it is produced or absorbed

by the body, to maintain body thermal balance (Bernabucci et al., 2014). The external environmental conditions that contribute to heat stress, trigger physiological and behavioral responses that negatively influences farm animal performance. Heat stress is caused by a combination of factors such as temperature, humidity, wind and direct and indirect radiations (Bernabucci et al., 2010).

The effect of the environment on the performance of cattle (beef and dairy) has been measured by combining temperature and relative humidity into the temperature-humidity index (THI), which can be measured as:

$$THI = (1.8 \times AT + 32) - (0.55 - 0.55 \times RH) \times [(1.8 \times AT + 32) - 58]$$

where AT is a the ambient temperature expressed in degrees Celsius, and RH is the relative humidity (Bernabucci et al., 2014). This index is commonly used to predict the degree of heat stress in dairy cattle. It is traditionally thought that milk yield will start to decrease when the THI reaches 72 (Bernabucci et al., 2014). Milk yield declined by 0.2 kg per unit increase in THI when THI exceeded 72 (West, 2003, Hayes et al., 2009). It has been reported that Holstein, Jersey and Brown Swiss milk yield will be normal when ambient temperature is 29ºC and 40% relative humidity. These breeds under these conditions have 97%, 93% and 98% of normal milk yield, but when the relative humidity reaches 90%, milk yield is reduced to 69%, 75% and 83% of normal production (West, 2003).

Metabolic heat production

Metabolism accounts for a significant portion of a cow´s heat production. The increased physical activity, feed intake, and therefore digestive requirement, cause body heat production to

increase as well. In consequence, when cows are under heat stress they accumulate heat not only associated with the animal metabolic processes, but also from the environment. Failing to dissipate this heat loads lead to decrease milk yield (West, 2003).

Heat stress induces metabolic changes in dairy cattle, leading to low fertility and milk production, but regardless of the massive economic impact, there is not enough information on how heat stress causes alteration in metabolism. It is known that an increase in heat load suppressed nutrient uptake. It is assumed that a decrease in dry matter intake is primarily responsible for a reduction on milk yield (Collier et al., 2008), but recent research showed that poor nutrient intake accounts for approximately ~40% to 50% of the reduction in milk yield (Dunshea et al., 2013).

Physiological response to heat stress

To maintain body temperature in an optimal condition, mammals use four mechanisms to exchange heat: conduction, convection, radiation, and evaporation (Collier et al., 2006, Hayes et al., 2009). Three of these routes (conduction, convection, and radiation) are referred to as sensible routes of heat loss and require a thermal gradient to operate. The fourth: evaporation, works on a vapor/pressure gradient and is defined as insensible heat loss (Collier et al., 2006, Dunshea et al., 2013). When ambient temperatures approach body temperature, the only viable route of heat loss is evaporation. If ambient conditions exceed body temperature, heat flow will reverse and the animal becomes a heat sink (Collier et al., 2006, Hayes et al., 2009). In cattle, evaporative heat loss via sweating, is the major mode of heat loss, and passive evaporation can account of ~15% of the core heat loss of a cow under heat load (Dunshea et al., 2013).

Although THI index has been used to estimate the impact of the environment or as a measurement heat load on dairy cattle, this is only an estimation of the outside ambient conditions. This index is not a precise estimation of the housing structure. Housing temperatures are affected

by the cooling systems locations relative to the animals position, consequently THI values do not always account for the microenvironments. However, it has been shown that infrared skin temperature is highly correlated with respiration rates, which a good measure of the microenvironment of animals in closed structures (Collier et al., 2006, Hayes et al., 2009). Infrared skin temperature is typically measured to grade the severity of the heat stress as being either low (20 to 60 breaths per minute), medium (60 to 80 breaths per minute) high (80 to 120 breaths per minute) or severe heat stress (>150 breaths per minutes) (Dunshea et al., 2013).

Heat stress alters the nutritional needs of the dairy cattle. Changes include decreased dry matter intake requiring increased nutrient density, altered mineral and water requirements, and altered digestive tract function (Collier et al., 2006). About 50% of the milk yield decrease that is due to heat stress, is due to low feed intake (Dunshea et al., 2013). These alteration of the nutritional needs are also associated with a lower rumen pH and acidosis (Hayes et al., 2009).

Endocrine alterations during heat stress

Acclimation is a physiological change that can reduce the effect of stressful conditions on cows. An example of these conditions are climate changes and thermal stress. Endocrine systems are implicated in the acclimatory response to heat stress, and they primarily include thyroid hormones, PRL, GH, glucocorticoids and mineralcorticoids (Bernabucci et al., 2010). The adrenal gland reduces aldosterone and glucocorticoid secretion while increasing epinephrine and progesterone secretion. Adipose tissue increases leptin secretion and the anterior pituitary gland increases PRL synthesis and secretion, and decreass somatotropin secretion. Under hyperthermia IGF1 synthesis and secretion by the liver decreses due to the low concentrations of GH (Bernabucci et al., 2010). The thyroid decreases thyroxine secretion and the placenta decreases estrone sulfate secretion (Bernabucci et al., 2010). Figure 2.1 is a schematic description of the

possible mechanisms for the effect of heat stress on reproduction in the lactating dairy cow (De Rensis and Scaramuzzi, 2003). Figure 2.1 shows that heat stress will affect the participaring key hormones in fertility resulting in a negative effect in cattle fertility.

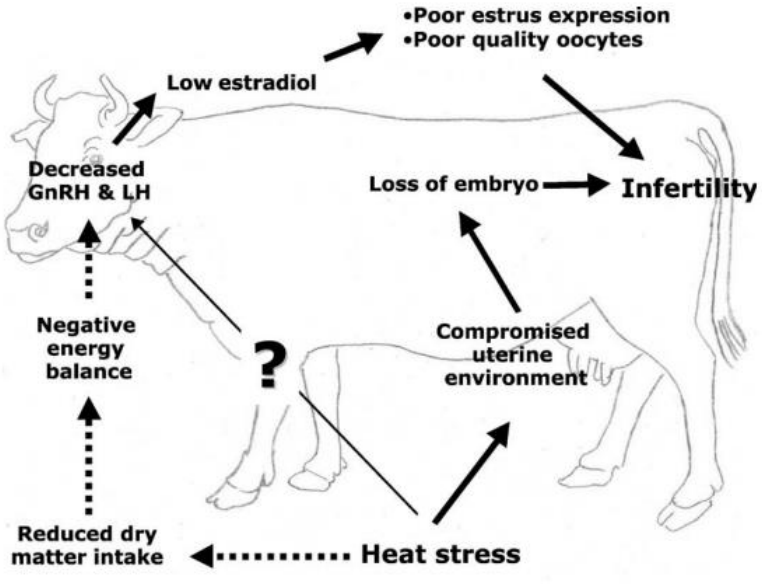

Figure 2.1. Effect of heat stress in dairy cattle. Note: reduced dry matter intake indirectly inhibits GnRH and LH secretion from the hypothalamo-pituitary system (dashed lines). However, it is not clear if heat stress can also directly influence the hypothalamo-pituitary system (thin solid line) to reduce GnRH and LH secretion. Heat stress can directly compromise the uterine environment (solid lines) to cause embryo loss and infertility (De Rensis and Scaramuzzi, 2003).

The PRL, as a metabolic hormone, is sensitive to change in temperature; and therefore expression levels increase during the summer (Collier et al., 1982, De Rensis and Scaramuzzi, 2003). Since PRL also maintains galactopoiesis and lactogenesis during lactation in ruminants, the increase in PRL expression levels would seem to cause an increase in galactopoiesis and lactogenesis. The PRL improves insensible heat loss and sweat gland function during acclimation (Beede and Collier, 1986). Other consequences of temperature alterations can influence the follicular development and suckling-induce PRL secretion leading to a postpartum anestrus (De Rensis and Scaramuzzi, 2003).

In addition to prolactin other molecules are altered in response to heat stress due to the decrease in feed intake, such as: glucose, GH and IGF1 (Bernabucci et al., 2010) and non-esterified acids decreased under these conditions (De Rensis and Scaramuzzi, 2003). Both IGF-I and glucose are generally stimulatory to follicular growth and implantation and glucose is the primary nutritional metabolite for the ovary (Rabiee et al., 1997). Glucose availability is also directly involved in modulating LH secretion (Bucholtz et al., 1996), and severe hypoglycemia inhibits pulsatile LH secretion and prevents ovulation and fertility in cows (Jolly et al., 1994).

The GH and IGF1 plays a key role in lactation. The GH secreate by the anterior pituitary gland in response to the acction of the growth hormone release hormone (Hadsell et al., 2008) and is negatuvely regulated by somatotastin (Muralidhar and Lee, 3013). The GH binds to its receptors in the liver causing the systeisis and secretion of IGF1. The IGF1 protein is involve on the development of the mammary gland during lactation in lactating Holstien cattle (Hadsell et al., 2003). This protein is regulated by the insulin-like growth factor binding proteins family (IGFBP) (Accorsi et al., 2002). For example, it has been reported that IGFBP5 acction on the IGF1 results in the apoptosis of the mammary glan tissue, nevertheless, the IGFBP5 acction is downregulated by the PRL (Sakamoto et al., 2007). A large number of proteins are involve in this pathway, for example: the STAT proteins family, which, participate in the transcription of milk proteins (Zhang et al., 2010).

Conclusion

Genetic evaluation of Holstein dairy cattle has been executed in the United States since the 1950's proving to be a valuable tool for herd improvement. Over the years, most selection pressure was placed on milk production traits. However, most livestock production traits are polygenic. Complex traits can be better evaluated using GS since genetic markers can be used to improve

accuracy of predictions. One situation in which candidate gene markers could be highly useful in genetic improvement is in Holstein dairy programs were cows are under heat stress. Heat stress can have a detrimental effect on health, milk production, and reproduction and it is known that heat stress influences the PRL and GH-IGF1 endocrine pathways. The challenge in these conditions is to keep improving milk yield while selecting for other traits such as high fertility.

CHAPTER 3: GENOTYPES WITHIN THE PRL AND GH-IGF1 PATHWAYS ASSOCIATED WITH 305 D MILK YIELD IN LACTATING HOLSTEIN CATTLE

Introduction

One of the challenges of dairy production in tropical and hot climates, as well the summer season in non-tropical ecosystems, is heat stress. This stress reduces both milk production and reproductive efficiency of Holstein dairy cattle (Jordan, 2003). Heat stress is a condition that occurs when an animal is not capable of dissipating an adequate amount of heat, and therefore, they cannot maintain thermal balance. Heat stress occurs via exposure to ambient conditions of high temperature and humidity. The persistent exposure to these conditions will alter physiological status and initiate a behavioral response. The physiological changes increase water intake and decrease feed intake decreasing reproduction and milk yield (Bernabucci et al., 2010). Additionally, components of milk such as fat, solid, lactose and protein content are reduced by heat stress (Collier et al., 2006).

A heat stress acclimation response involves changes in the endocrine system, which includes the thyroid hormones, PRL, GH, IGF1, glucocorticoids and mineralcorticoids (Bernabucci et al., 2010). For instance, the anterior pituitary gland synthesis and secretion of PRL is sensitive to changes in temperature and level of this hormone increases during summer (Collier et al., 1982, De Rensis and Scaramuzzi, 2003). This hormone may also have a role in acclimation through heat loss and sweat gland function (Bernabucci et al., 2010).

In addition to PRL, other biological molecules are altered due to the decrease in feed intake, these include: glucose, due to a lack of energy sources, also the GH, IGF1 (Bernabucci et al., 2010) and non-esterified acids (De Rensis and Scaramuzzi, 2003). Both IGF1 and glucose are generally stimulatory to follicular growth and embryo implantation, and glucose is a primary nutritional

metabolite for the ovary (Rabiee et al., 1997). Glucose availability is also directly involved in modulating pituitary gland secretion of LH (Bucholtz et al., 1996), and severe hypoglycemia inhibits LH secretion and prevents ovulation in cows (Jolly et al., 1994). These would result in a negative effect in reproduction.

The hypothesis of this research is that the molecular breeding values (MBV) constructed with DNA markers within the PRL and GH-IGF1 pathways have the potential to predict milk production traits in heat stressed Holstein cows in Sonora. The objective of this research was to calculate an MBV using SNP within genes of the PRL and GH-IGF1 pathways and their association with milk production traits in Holstein cows of Sonora, Mexico.

Materials and Methods

Data

Data were collected from 659 Holstein dairy cows. Observations were collected during 2012 from three dairy farms located in Block 910 and 1114 of the Yaqui Valley city of Obregón, Sonora, Mexico. Farm 1 (n = 298) was located at the coordinates 27°21'N 109°54W, farm 2 (n = 106), which is part of the "Instituto Tecnologico de Sonora", located at coordinates 27°21'N 109°54'48W, and farm 3 (n = 255) located at coordinates 27°19'N 109°52'W. Cows used in the study had body condition scores ranging from 2.5 to 3.5 (1 = very thin, 5 = very fat) (Kadarmideen and Wegmann, 2003) and were ~150 d of lactation at the beginning of the study. Cows with missing observation were not used in the analyses.

Milk yield observations were recorded monthly and used to calculate total 305 d milk yield and daily average 305 d milk yield per cow in kg. The 305 d milk yield was calculated multiplying the milk production observations by an adjustment factor for days in milk and age of the cow. These factors were obtained from the National Cooperative Dairy Herd Improvement Program

Handbook (1985). In addition to 305 d milk yield and total milk production, the variables available for analyses are listed in Table 3.1.

Table 3.1. Summary statistics for variables used to study lactating Holstein cows in Sonora, MX.

Variable	N	Mean	SD	Minimum	Maximum
305d MY[1], kg	589	6308	1468	636	10787
Years of age	596	5.26	1.97	1	13
Lactation number	596	3.06	1.83	1	11
Number AI[2]	596	1.96	1.23	1	10
DIM[3]	595	308	70	21	716

[1]MY = Milk yield
[2]AI = Service per conception from artificial insemination in 2012
[3]DIM = Days in milk

Management and health status

Cows were cooled with fans, and fresh water showers (cooling system) to minimize the hot environmental effects (Figure 3.1). The cooling system consisted of: showers during the warmer hours of the day for 5 minutes followed by 10 minutes of ventilation (~11,000 horse power). The cooling system had 16 sprinklers (water distribution of: ~15L per cow/series) and three electrical fans of half horse power, located 2.73 meters above the floor. Cows had free access to fresh water and shade (8.5 m^2 per cow). Although cooling systems were used to alleviate heat stress, it is reported that dairy cattle production decreases by 10 to 15% in operations with cooling systems (Dunshea et al., 2013). Cows were fed twice a day with a ration of 75% alfalfa hay and 25% corn silage that supplied their nutritional needs according to the requirements established by the NRC (2001) for lactating Holstein cows with an average weight of 650 kg and producing ~30 kg/d of milk with average composition of 3.5% fat and 3.2% true protein.

Figure 3.1. Holstein cows being showered in shade to alleviate in Sonora, Mexico.

The housing for the cows in the three farms had open pens, which allows free air flow. The flooring of the pens were made from soil, shades were provided ($8.5m^2$ per cow). Figure 3.2, Figure 3.3 and Figure 3.4 shows the structural differences among the three farms. Each holding pen in farm 2 contain 25 to 30 cows. Pens in farm 1 and farm 3 contained approximately 50 cows.

Figure 3.2. Aerial view of the housing structure of Farm1.

Figure 3.3. Aerial view of the housing structure of Farm 2.

Figure 3.4. Aerial view of the housing structure of Farm 3.

Cows had free access to water, which was located on the areas with shade. Additionally, the food was also located in shade areas in farm 2 and farm 3. Food was also located in a different far from the water.

Cows were palpated 20 to 25 d after parturition to check for signs of uterine infection and had a voluntary waiting period between 40 to 50 d postpartum to prepare the cows for breeding season. All cows started a hormonal treatment to synchronize ovulation and received a fixed-time artificial insemination (AI) at day 70 postpartum. Records of mastitis and metritis diagnosis and treatment were obtained and recorded for the cows on the three dairy farms. Subclinical mastitis

diagnosis was based on the qualitative California mastitis test (Laboratorios Sanfer, S.A. de C.V., Obregon, Mexico). The scoring results were:

- no infection: no precipitation
- type 1: light precipitation
- type 2: light precipitation with granulose filaments
- type 3: gel formation
- type 4: fast gel formation and coagulation of the sample

Cows with clinical mastitis were excluded from the project. Ultrasound (SonoSite, Inc., Bothell, WA) scans were performed 40 d postpartum, and metritis was diagnosed if the ultrasound revealed dense fluid from placental retention or uterine infection. Health records were compiled into one variable known as health status. Health status was used as a categorical variable and coded as 0 for no disease diagnosis and 1 for any disease diagnosis. Summary statistics (Table 3.2) were calculated for each health status group using PROC MEANS (SAS 9.4). All other analyses were performed with this version of SAS, unless otherwise stated. Table 3.2 shows that overall healthy cows performed better and had lower variability during 2012, as opposed to the cows in health group 1 ($P \leq 0.05$).

Table 3.2. Summary statistics per health status group in lactating Holstein cows in Sonora, MX.

Health	Variable	N	Mean	SD	Minimum	Maximum
0	305d MY[1], kg	554	6372	1432	636	10787
	Years of age	561	5.27	1.98	1	13
	Number of lactation	561	3.09	1.85	1	11
	Number AI[2]	563	1.87	1.05	1	8
	DIM[3]	560	305	61	21	716
1	305d MY[1], kg	35	5296	1678	1787	8029
	Years of age	35	5.11	1.85	2	10
	Number of lactation	35	2.71	1.54	1	8
	Number AI[2]	33	3.55	2.51	1	10
	DIM[3]	35	354	150	42	696

[1]MY= 305 d milk yield. [2]AI= Service per conception from artificial insemination in 2012. [3]DIM=Days in milk

Temperature and humidity index (THI)

The THI was calculated and provided by the "Instituto Tecnologico de Sonora" in 2011 and 2012, since the cows used in this research calving in two years. The climatic records were obtained thru Sonora, Mexico Agro-climatic Station Network available through www.agroson.org .Mexico. This calculation was based on the equation:

$$THI = 0.8\,(T°) + RH/100\,(T° - 14.4) + 46.4$$

where THI was the temperature and humidity index, $T°$ was the temperature in Celsius degrees and RH was the relative humidity in decimals (Mader et al., 2006). Index values were calculated each hour of each day (24), and monthly average and standards deviations were estimated.

SNP discovery and genotype

Forty-three candidate genes (Table 3.3) within the PRL and GH-IGF1 pathways (Etherton, 2003, Chagas et al., 2007, Lucy, 2008) were studied. These genes were selected based on their physiological function and involvement in milk production (Etherton, 2003, Chagas et al., 2007, Lucy, 2008).

For DNA extraction, 3 ml of blood were collected via venipuncture of the median caudal tail vein or artery for each cow. This sample was spotted in fast analysis of nucleic acid cards (GeneSeek, Inc., Lincon, NE). Later, DNA was extracted from the cards and quantified for the following analyses. Genotyping was completed using several multiplex SNP assays within the Sequenom MassArray platform (GeneSeek, Inc., Lincon, NE). Polymorphisms were analyzed and regions of disequilibrium (linkage disequilibrium) were identified using the software Haploview (Barrett et al., 2005). A range from 2 to 50 SNP were found within each gene and yielded a panel of 179 tag SNP (Supplementary table 1).

Table 3.3. Candidate genes within the PRL and GH-IGF1 pathways used for SNP discovery and previously reported to be involved in lactating Holstein Cows.

Gene	Definition
AVP	Arginine Vasopressin
AVPR1A	Arginine Vasopressin Receptor 1A
CISH	Cytokine inducible SH2-containing Protein
FURIN	FURIN (paired basic amino acid cleaving enzyme)
GH1	Growth Hormone
GHRH	Growth Hormone Releasing Hormone
GHRHR	Growth Hormone Releasing Hormone Receptor
GHSR	Growth Hormone Segretagogue Receptor
IGF1	Insulin-Like Growth Factor-1
IGF1R	Insulin-Like Growth Factor-1 Receptor
IGFBP2	Insulin-Like Growth Factor Binding Protein-2
IGFBP3	Insulin-Like Growth Factor Binding Protein-3
IGFBP4	Insulin-Like Growth Factor Binding Protein-4
IGFBP5	Insulin-Like Growth Factor Binding Protein-5
IGFBP6	Insulin-Like Growth Factor Binding Protein-6
IGFBP7	Insulin-Like Growth Factor Binding Protein-7
OXTR	Oxytocin receptor
OXT	Oxytocin
PAPPA1	Pregnancy Associated Plasmatic Protein A1
PAPPA2	Pregnancy Associated Plasmatic Protein A2
PCSK2	Proprotein Convertase K2
PIAS	Protein Inhibitor of Activated STAT-1
PMCH	Pro-Melanin Concentrating Hormone
PRL	Prolactin
PRLR	Prolactin Receptor
SCGV	Secretogranin V
SOCS1	Supressor of Cytokine Signalling-1
SOCS2	Supressor of Cytokine Signalling-2
SOCS3	Supressor of Cytokine Signalling-3
SOCS4	Supressor of Cytokine Signalling-4
SOCS5	Supressor of Cytokine Signalling-5
SOCS6	Supressor of Cytokine Signalling-6
SOCS7	Supressor of Cytokine Signalling-7
SST	Somatostatin
SSTR2	Somatostatin Receptor-2
SSTR3	Somatostatin Receptor-3
SSTR5	Somatostatin Receptor-5
STAT1	Signal Transducer and Activator of Transcription-1
STAT3	Signal Transducer and Activator of Transcription-3
STAT4	Signal Transducer and Activator of Transcription-4

STAT5A	Signal Transducer and Activator of Transcription-5A
STAT5B	Signal Transducer and Activator of Transcription-5B
STAT6	Signal Transducer and Activator of Transcription-6

Note: Genes biological functions provided in Supplementary Table 2.

SNP effects

The associative analysis between genotypes and phenotypes for 305 d milk yield was performed using PROC MIXED. The statistical model used for was:

Model 1

$$y = Xb + Za + e$$

where y was the vector of milk yield to 305 d of lactation, b was the vector of fixed effects, and a was the vector of random effects, which include random sire effect. Fixed effects included the lactation number, the genotype (SNP effect), the contemporary group (farm), days in milk, the health status, and calving month. The genotype (SNP effect) and days in milk were used as a covariant. X and Z were incidence matrices relating records of fixed effects to random effects and e was the error vector. Associations between each SNP and 305 d milk yield were reported based on their significance ($P \leq 0.05$).

Each single SNP effect was estimated individually using two different approaches where both have been used in other association studies (Cochran et al., 2013). In the first method the genotypes (SNP effect) were included as a covariant to determine the allele substitution effect using PROC MIXED. In the second method, the genotypes were included in the model as a categorical variable and orthogonal contrasts were used to estimate additive effects. The false discovery rate (FDR) or Q-value was calculated to control for false positives using PROC MULTTEST (Benjamini and Hochberg, 1995), no other adjustments were performed on the initial *P* - values. Cows with missing observations were excluded from the analysis.

Molecular breeding value (MBV)

The MBV for 305 d milk yield was calculated for each of the Holstein cows used by summing the additive genotype effect at each locus ($P \leq 0.05$). The calculation of the MBV was performed using the Animal Breeder Tool Kit (ABTK) (Colorado State University, Fort Collins, CO). Pearson's correlation between 305 d milk yield and the MBV was calculated using PROC CORR.

Results and Discussion

Temperature humidity index (THI)

The THI calculated for this region in Mexico for the year 2011 and 2012 is presented in (Figure 3.5 and Figure 3.6). Cows under constant environmental conditions such as high humidity and ambient temperature, with THI above 72, are known to be heat stressed (West, 2003, Hayes et al., 2009). Nevertheless, it has been suggested a THI threshold for lactating dairy cows producing more than 35 kg/day should be 68 since the threshold of 72 was estimated in the 1950s. Global temperatures are higher and cows have much higher milk yield at the present (Zimbelman et al., 2009).

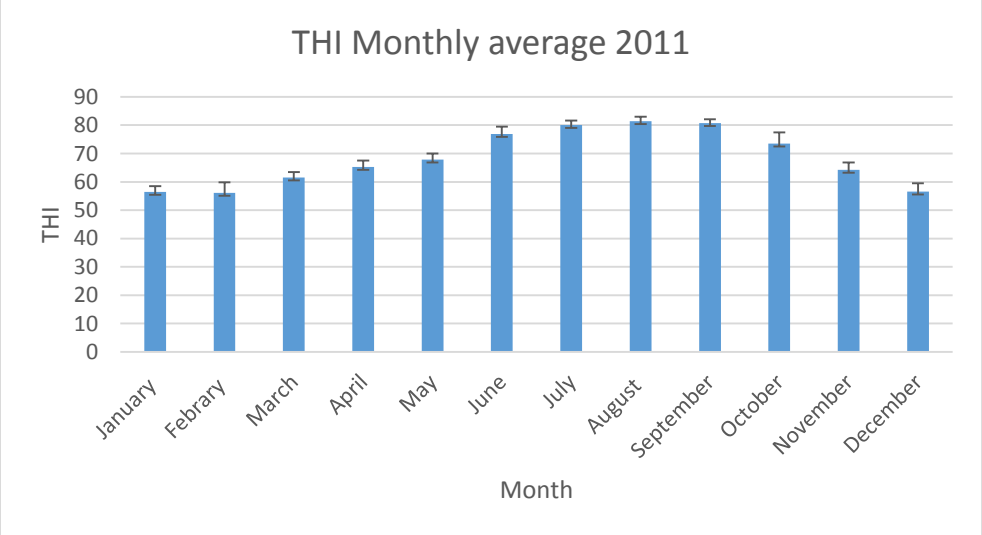

Figure 3.5. Mean ± SD temperature and humidity index in 2011 in Obregón, Sonora, MX.

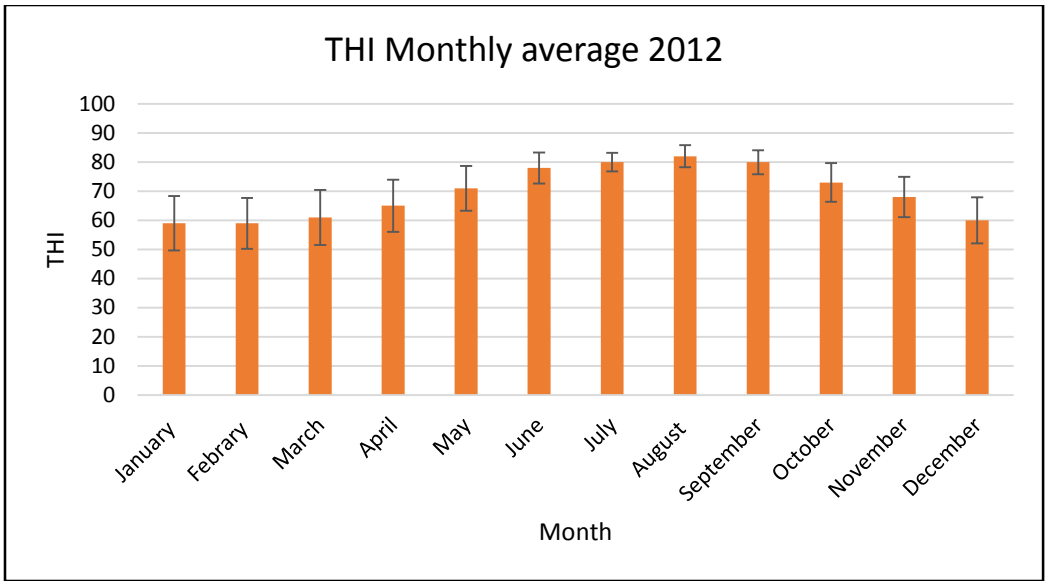

Figure 3.6. Mean ± SD temperature and humidity index in 2012 in Obregón, Sonora, MX.

Based on the monthly average (Figure 3.5 and 3.6) THI and standard deviations, the cows were potentially heat stressed from March through November in 2011 and 2012. These index values varied from light stress (72 - 79) to moderate stress (80 - 89) (Armstrong, 1994). The average annual THI was 69.6 ± 2.6 for the year 2012. The months with highest TH I were May, June, July, August, and September. The calving month distribution for this population is shown in Figure 3.7. The majority of the cows calved from August 2011 to March 2012, which is the end of the heat stress season. This management decision aims to minimize potential negative environmental effect on milk yield.

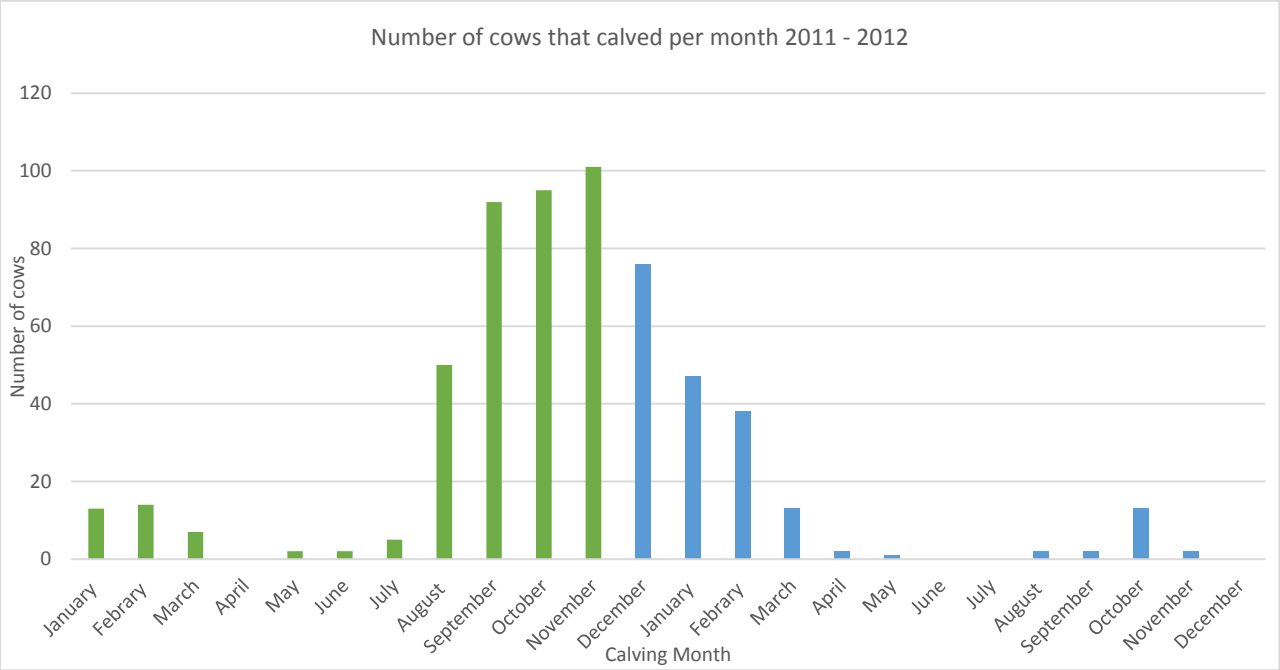

Figure 3.7. Number of cows that calving each month for the year 2011 and 2012. Green bards are for cows that calf in 2011 and blue bards were for cows that calf in 2011.

SNP effects

Eight SNP in 5 genes were associated ($P \leq 0.05$) with 305 d milk yield (Table 3.4). These genes were *AVPR1A*, *FURIN*, *IGFBP6*, *PMCH*, *PRLR*. These genes possess different functions in the various biological processes influencing milk production.

Table 3.4. SNP within the PRL and GH-IGF1 pathways associated with 305 d milk yield in heat stressed lactating Holstein cows in Sonora, MX.

Gene	Chr[1]	Location (Mb)	SNP	*P*	FDR[2]	Alleles	Additive SNP effect (Kg)
AVPR1A	5	50.5	rs207971189	0.05	0.05	**C** / G	272.58
AVPR1A	5	50.5	rs210011420	0.03	0.04	**T** / C	276.39
AVPR1A	5	50.5	rs209300854	0.04	0.05	**C** / C	257.11
IGFBP6	5	27.0	rs211039223	0.00	0.01	**C** / T	388.42
PMCH	5	65.3	rs14197280	0.00	0.01	**A** / T	30.58
PRLR	20	39.1	rs135164815	0.00	0.01	**A** / G	196.30
PRLR	20	39.1	rs136247583	0.04	0.02	**C** / T	194.22
FURIN	21	22.2	rs381099643	0.04	0.05	**G** / A	251.92

Favorable allele are bolded.[1]Chr = chromosome. [2]FDR = false discovery rate. ($P < 0.05$).

The *AVPR1A* gene was reported to change levels of expression in endometrial and myometrium tissues during estrus and early pregnancy in cattle (Fuchs et al., 1990). "Endometrial

receptor (AVPR1A) levels varied significantly during the cycle; it was lowest on days 7 and 14, rose significantly on day 17, and peaked on day 21. Myometrial receptor levels decreased from levels at estrus on days 7 and 14, but the changes were not significant" (Fuchs et al., 1990). This gene, located on bovine chromosome 5, is also involved in regulation of systemic arterial pressure (Gozdz et al., 2002). The *AVP* gene codes for the arginine vasopressin protein, which is a diuretic hormone involved in the secretion and ejection of milk during lactation in cattle (Nussey et al., 1987). To our knowledge, no previous research showed SNP association within this gene and 305 d milk yield. However, our findings reveled two intra-gene SNP (rs209300854, rs210011420) associated with 305 d milk yield.

Additionally, *IGF1* and *IGFBP* participate in the lactation process. The activity of IGF1 in milk production and cell proliferation during lactation is regulated by the IGFBP protein family (IGFBP1, IGFBP2, IGFBP3, IGFBP4, IGFBP5, and IGFBP6) (Accorsi et al., 2002). The *IGFBP6*, shows a decrease in expression during lactation (Fenwick et al., 2008). This gene is located on bovine chromosome 5 and regulates the actions of IGF1 (i.e., free vs bound form in circulation). Past research has not reported SNP within this gene associated with milk production traits; however, we found one variant (rs211039223) associated with 305 d milk yield.

The *FURIN* gene is located on bovine chromosome 21 and is involved in the activation of precursor proteins through the cleavage of a single or paired basic amino acid residue (Khatib and Sfaxi, 2012, Maruotti et al., 2012). Previous research studied *FURIN* in lactating cows (Cánovas et al., 2010); but an association with milk production traits was never reported. Our research showed association of one SNP (rs381099643) in this gene with a significant effect on 305 d milk yield. The association found here may be explained by *FURIN* involvement in the posttranslational

processing of *GHRH* and, indirectly, in the synthesis and secretion of GH (Posner et al., 2004), which could affect nutrient mobilization and cell proliferation during lactation.

The *PMCH* gene located on bovine chromosome 5 is involved in regulation of energy homoeostasis and could be a defense mechanism against energy deficiency (Beerda et al., 2008). It is reported that cows with high gene expression levels of *PMCH* show reduced estrus behavior (Beerda et al., 2008). Past research has not been able to establish SNP within this gene to be associated with milk production traits. However, herein one SNP (rs14197280) was identified to be associated with 305 d milk yield. This association may be explained by the effect of this gene on the energy status, which could lead to a negative change in energy balance in heat-stressed cows. It should be noted that energy balance decreases rapidly within the first 100 days in milk in high yield lactating Holstein cows (Huttmann et al., 2009).

An important candidate gene identified as being associated with 305 d milk yield was *PRLR*. The relevance of *PRLR* is due to its role in milk production and stress response (Collier et al., 2008, Lü et al., 2010). The SNP (rs135164815, rs136247583) within *PRLR* associated with milk yield identified in our research were located in exon 2, position 39.1 Mb on bovine chromosome 20. In contrast, other studies found a *PRLR* mutation in exon 10 that introduces a premature stop codon and is considered a candidate for Slick coat genotype and heat tolerance in Senepol cattle (Littlejohn et al., 2014). It was identified a novel SNP (39.1 Mb) in the *PRLR* gene consisting of a single base deletion in exon 10 that introduces a premature stop codon (p.Leu462) and loss of 120 C-terminal amino acids from the long isoform of the receptor (Littlejohn et al., 2014).

Previous research "identified a phenotype characterized by development of a very short, sleek hair coat that is inherited as if controlled by a single dominant gene", the slick gene (Olson

et al., 2003). Olson et al. (2003) reported that Holstein dairy cattle with the slick haplotype exhibit higher milk yields than non-slick contemporaries. This is of particular interest in dairy farming contexts, where most selection has occurred in heat-intolerant *Bos taurus* breeds (Littlejohn et al., 2014).

Additionally, The *PRLR* is found at the same locus, bovine chromosome 20, as other DNA markers used to map the Slick gene in *Bos taurus* cattle. A long range of homozygosity extending over 5 Mb of the Slick gene in slick haired cattle was reported in 2012 (Flori et al., 2012). Within this region two genes, *PRLR* (38.0 Mb) and sperm flagellar 2 (SPEF2) (38.4 Mb) were found. The *SPEF2* and *PRLR* genes are also involved in reproduction and milk production (Huson et al., 2014). Additionally, values of integrated haplotype scores indicated that the region between *PRLR* and *SPEF2* is a target of recent selection (Huson et al., 2014).

MBV and correlations

The molecular breeding value was calculated for 546 cows that were genotype for all the SNP that were found to have a positive association with 305 d milk yield. The summary statistics for the MBV is show in table 3.

Table 3.5. Summary statistic for the molecular breeding value

N	Mean	SD	Minimum	Maximum
546	2780.82	656.30	251.92	3735.04

Pearson's correlations were estimated from the MBV that was previously calculated from the SNP effects within the PRL and GH-IGF1 pathways associated with 305 d milk yield with the variables used in this study. Under heat stress, the MBV had a weak but positive linear correlation with 305 d milk yield (Figure 3.8). Other correlated variables and their levels of significance are also shown in Table 3.5.

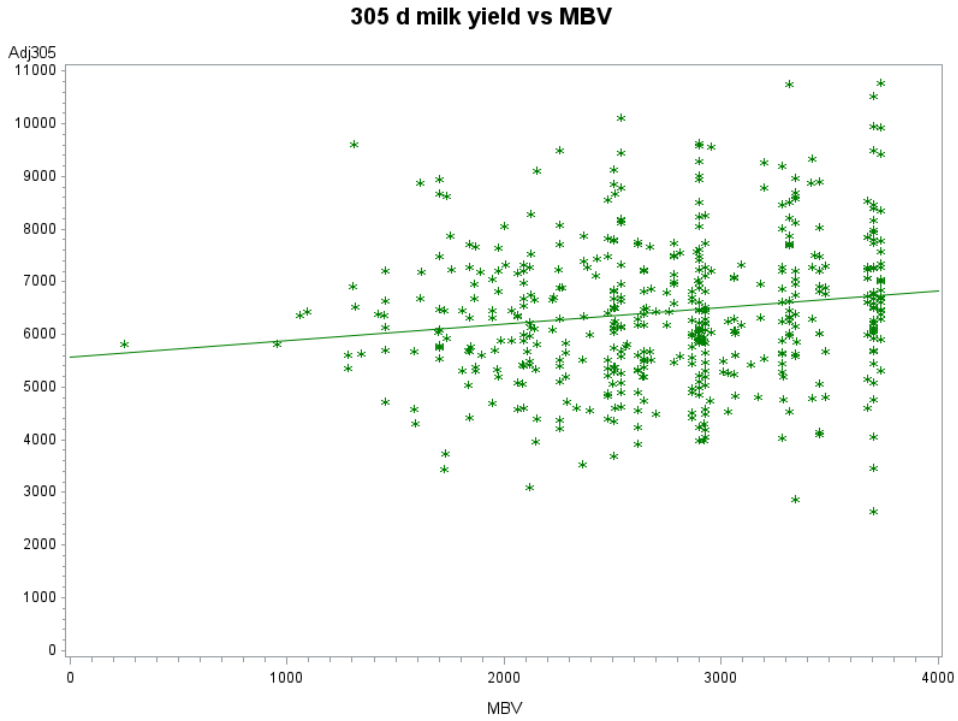

Figure 3.8. Scatter plot and regression of 305 d milk yield versus MBV in lactating Holstein cows in Sonora, MX. Slope=0.54 kg.

Table 3.6. Correlations between variables used to study lactating Holstein cows in Sonora, MX.

Variables	305d MY[1]	Age	Lactation number	Number AI[2]	DIM[3]	MBV[4]
305d MY[1]	1	0.41*	0.43*	-0.08*	0.09*	0.15*
Age		1	0.95*	0.02	0.14*	-0.04
Lactation			1	-0.02	0.03	-0.02
Number AI[2]				1	0.54*	-0.005
DIM[3]					1	-0.006
MBV[4]						1

*($P \leq 0.05$). [1]MY = Milk yield, kg. [2]AI = Artificial insemination services per conception.[3]DIM = Days in milk. MBV[4] = molecular breeding value.

There was a strong linear correlation between age and lactation number (Table 3.6), based on these correlations, age was not included in either of the models since it was auto-correlated. Additionally, not all cows in this study began lactating at the same age.

Conclusion

We were able to find a positive association between 8 SNP and 305 d milk yield and no previews authors have reported these finding before. The MBV estimated from SNP within the PRL and GH-IGF1 pathways genes was weakly associated with 305 d milk yield in Holstein cows from Sonora, MX.

CHAPTER 4: MOLECULAR BREEDING VALUE ASSOCIATION WITH 305 D MILK YIELD IN HOLSTEIN CATTLE IN SONORA, MX.

Introduction

Climate change in the last decade has resulted in higher average temperatures, hotter daily maximums, and frequent heat waves (Key et al., 2014). Climate model predictions reported by the USDA on September of 2014 suggested that annual temperatures will increase between 0.8°C and 1.3°C by the year 2030. Current reports by the USDA estimated financial losses due to heat stress in the dairy industry were $1.2 billion in the year 2010 (Key et al., 2014) and are likely to increase if current climate model are correct.

In addition to the United States, other countries, such as Mexico, have a growing interest in improving dairy cattle breeds, in particular Holstein. Some regions in Mexico, like the state of Sonora and especially the area surrounding Obregón, have annual average temperatures of ~ 36°C; resulting in heat stressed dairy cattle. Given the economic importance of heat stress in the dairy industry globally, developing genotyping strategies for dairy cattlein those areas has the potential to assist in selection of tolerant animals while also taking into account other traits such as milk production and fertility. The objective of this chapter is to evaluated the association between the MBV prevously estimated and 305 d milk yield.

Materials and Methods

Models and parameter estimation

A regression model was calculated using PROC MIXED. For this model, we included the MBV previously calculated as a continuous variable. The model was:

Model 2

$$Y = \mu + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 + X_6\beta_6 + e$$

where Y was defined as the dependent variable of 305 d milk yield or vector of observations. μ was the population mean, X_1 was the covariate for MBV, β_1 was the slope for the MBV, X_2 was the covariate for days in milk, β_2 was the slope for the variable days in milk, $X_3\beta_3, X_4\beta_4, X_5\beta_5$ and $X_6\beta_6$ were the incidence matrices for the categorical variables which included: the contemporary group (farm), the lactation number, calving month and the health status with their vector for fixed effects respectively. e was the vector of residual effect or error term. No relationship matrix was fixed for this model.

An additional model was constructed. The model was:

Model 3

$$Y = \mu + X_1\beta_1 + e$$

where Y was defined as the dependent variable of 305 d milk yield or vector of observations. μ was the population mean, X_1 was used as a covariant term for the MBV and β_1 was the slope for the MBV, and e was the vector of residual effect or error term. No relationship matrix was fixed for this model.

Mixed model equation

A mixed model equation was developed and calculated using ASREML 3.0 (Gilmour et al., 2009). The general equation was:

Model 4

$$y = Xb + Za + e$$

were y was the vector of 305 d milk yield, b was the vector of fixed effects, and a was the vector of random effects which included the individual cow. Fixed effects included lactation number,

the farm, days in milk, calving month, MBV, and health status. The MBV and days in milk were included as covariant. The X and Z were incidence matrices relating records of the animal's fixed effects to random effects and e was the vector of random residual effects. One generation of the pedigree of the cows was also fixed (relationship matrix) in the model to estimate the variance components. An additional model (**Model 5**) was also constructed using the same variables as **Model 4** without the MBV and executed with ASREML 3.0 (Gilmour et al., 2009). A relationship matrix was fixed for this model for one generation.

Results and Discussion

Models and parameter estimation

Model 2 (linear regression of 305 d milk yield on fixed effects of MBV, days in milk, contemporary group, lactation numbers, calving month, and health status) had an adjusted R^2 of 42.92%. All variables included in the model resulted in a statistical significance ($P \leq 0.05$) (Table 4.1). For comparison, in Model 3 (the regression of 305 d milk yield on MBV), an adjusted R^2 of 2.18% was estimated. The small amount of variation explain by the MBV led us to postulate that this may be due to the small number of SNP used, therefore the trait may be highly polygenic (Nussey et al., 1987, Sakamoto et al., 2007, Fenwick et al., 2008, Khatib et al., 2008, Zhang et al., 2010).

Table 4.1. *P* values for independent variables in lactating Holstein cows in Sonora, MX.

Fixed Effects	*P*
DIM[1]	<.0001
Lactation	<.0001
MBV[2]	<.0001
Farm effect	<.0001
Health	0.0022
Calving month	0.0003

[1]DIM = days in milk. ($P \leq 0.05$). [2]MBV = molecular breeding value

Genetic parameters

Additive and phenotypic variances, and heritabilities (h^2) were estimated and reported in kilograms. Our study revealed h^2 of 0.39 ± 0.11 for 305 d milk yield (Table 4.2). Other studies reported similar heritabilities ranging between 0.29 to 0.37 for milk yield (Carlén et al., 2004, Cohen-Zinder et al., 2005, Raven et al., 2013).

Table 4.2. Genetic parameters for 305 d milk yield in lactating Holstein cows in Sonora, MX. without using the MBV as a fixed effect.

	305 d MY (Kg)	SE (Kg)
σ^2_a	482758	151580
σ^2_e	736300	136310
σ^2_p	1219100	74169
h^2	0.39	0.11

[1]MY = milk yield

When genetic parameters were calculated without the MBV (Model 5), the heritability was 0.42 ± 0.11 for 305 d milk yield. These results showed that the MBV was subtracting genetic variance resulting is a smaller estimation of heritability for 305 d milk yield. This could be due to the fact that the MBV was estimated from the same population that was used to estimate the heritability.

Conclusions

This MBV accounted for 2.18% of the phenotypic variation in a 305 d milk yield. The small amount of phenotypic variation may be due to the small numbers of SNP used to calculate the MBV. An additional explanation could be the polygenic nature of the trait under heat stress conditions.

Moreover, an important contribution of this study is that eleven SNP within nine genes were found associated with 305 d milk yield for the first time. Our genetic parameter estimations were

consistent with previous research for production traits in Holstein dairy cattle. Nevertheless, the MBV did not influence heritability estimates.

CHAPTER 5: CROSS-VALIDATION OF MBV ESTIMATED FROM SNP WITHIN THE PRL AND GH-IGF1 PATHWAYS IN LACTATING HOLSTEIN COWS IN SONORA, MEXICO

Introduction

A breeding objective in heat stressed populations of dairy cattle include selection for tolerance as well as the improvement of milk production. Molecular markers, in particular SNP associated with a quantitative trait loci (QTL) may contribute to phenotype variations in a given trait. These DNA markers can be used to construct genomic breeding values or MBV especially for complex traits such as environmental tolerance, disease resistance, and reproduction. The success of genetic improvement programs based on MBV can be accomplished by training and predicting independent populations (Saatchi et al., 2012, Mateescu et al., 2013, Boddhireddy et al., 2014).

Genomic selection validation involves using a training populations with genotypes and phenotypes to simultaneously estimate SNP effects (Meuwissen et al., 2001). The SNP effects can be combined with EBV to improve predictions. Several factors influence genomic prediction accuracy including the sample size of the training population and the relationship between the discovery and the validation population (Habier et al., 2007, Clark et al., 2012), the type of phenotypic variable (continuous, categorical, hard to measure) used for estimating the SNP effects, and the methodology used for grouping the data for cross-validation (Saatchi et al., 2011). Other factors include: extend of LD, number of QTL contributing to the phenotypes, the heritability of the trait, and the accuracy of the measurements of the phenotypes (Saatchi et al., 2011). The objective of this chapter was to validate an MBV by conducting a training and prediction exercise.

Materials and Methods

Data

The data used in this study was described in the Materials and Methods section of Chapter 3.

Management and health status

The management practices and health of the cows were described in the Materials and Methods section of Chapter 3.

Cross-validation

The MBV were estimated and evaluated by comparing estimates from a 5-fold strategy of random clustering using the MACRO statement and PROC SURVEY. This procedure divided the cows (n = 659) into five groups (5 folds). Individuals were randomly assigned to each group. Previous research have used different clustering methods, such as: K-means, identical-based clustering with equal and unequal sample size to validate genomic breeding values (GEBV), or direct genetic values (DGV) (Saatchi et al., 2011, Saatchi et al., 2012, Mateescu et al., 2013, Boddhireddy et al., 2014). The success of such methods depends on the amount of relation between individuals to minimize relatedness within each group. However, random clustering was used in our study due to the small number of cows per sire in the data set. Additionally, a large number of cows (138) did not have pedigree data and progeny numbers were highly variable. Sire's progeny mean and standard error was 3.4 ± 0.37 (Table 5.1). Training and predicting exercises in crossbreeds or across breeds are more problematic because different breeds may exhibit different QTL and LD (Hayes et al., 2009, Garrick, 2011). Only information available on sire was their name on each farm.

Table 5.1. Sires' progeny number of the lactating Holstein cows from Sonora, Mx used in an training and predicting exercise

Sire name	PN[1]	Sire name	PN[1]	Sire name	PN[1]
Unkown sires	138	JE3241	5	7H6745	4
WRANGLER	7	JE3214	1	7H6682	4
Testyfysex	3	Japelou	2	7H6250	3
Terminator	5	Income	1	7H6168	1
Taboosex	6	Igniter	1	7H5708	2
SORBY91716	2	Icepack	13	7H5687	2
SEMENTAL	1	Harry	2	7H5435	1
Silver	1	GUNNER	1	72H1758	2
SHOWTIME	1	GALLEON	1	29HO9899	1
SCORE	2	FOREVER	2	29HO9635	3
Saylor	7	FIRED	7	29HO10889	1
SAILOR	32	EMERSON	2	29HO10644	1
RUEBEN	2	ELWAY	1	29HO10641	1
RSVP	1	DUSTMAN	2	29HO10615	1
Rebel	1	DREVIL	8	29HO10493	10
RANGER	3	DOZIT	3	29HO10461	2
Predictor	1	DOUG	6	29HO10181	1
POTTER	8	DIFFERENCE	3	29H9899	1
PAT	1	DEMAND	1	29H10808	1
PARADISE	3	DECKER	13	29BS3781	1
OSMIUM	1	DAN	9	25HO803	1
ONIX	9	Damion	1	200HO5127	1
NZFRESIAN	1	Cumulus	1	200H0040	1
NZEU	2	CROPPER	3	200H00232	1
NZDESC	1	Criollo	1	1BS560	1
NZBMASTER	1	CORONATION	1	154BN513	1
NZAMBIENCE	3	Champion	1	14JE406	1
NZ2006	28	Cevis	2	14JE0406	1
NUCLEAR	3	BS3781	5	14JE0366	2
Mystique	1	BRIDGE	1	14H4400	1
Mr.Sam	2	BRANGUS	1	14H4360	1
Morty	3	BOSS	1	14H4099	2
MONTA	5	Blitz	4	14H4056	2
MITCHELL	1	BLASTOFF	7	14H4026	2
MicMac	2	Billion	5	14H3913	3
MICH	2	BEAVER	1	14H3597	23
MATT	6	Baxter	11	14H2586	6
MARTINI	2	Bambam	3	14H2288	1
Marion	8	Arthur	2	14/HO3913	1
MANASSA	11	Armstead	2	122H1410	3

MAGNA	14	AN603ANGUS	1	122H1286	1
LUCKY	1	AMBIENCE	2	122H0137	1
LON	17	AMBAR	5	11H5534	2
Lheros	8	Airraid	1	11H5286	1
Letterman	1	ADAN	1	11H5240	3
KENNETH	1	9H2704	1	11H5009	1
KARAT	4	7JE0570	1	11H4712	4
JE3643	1	7H8425	4	11H4631	1
JE3346	4	7H6960	2	11H4623	2
JE3307	2	7H6834	1	11H4338	1
11H08730	3	7H6758	2	11H4131	1
				11H3754	2

[1]PN= sires' progeny number

Genotyped cows were first divided into five mutually exclusive groups (5 folds). In each training analysis, the data excluded one group (20% or 1 fold) as to train with the other four groups (80% or 4 folds) to estimate marker effects, which were then used to estimate MBV of individuals from the omitted group (validation set) (Saatchi et al., 2011). This procedure was repeated five times (Figure 4.1).

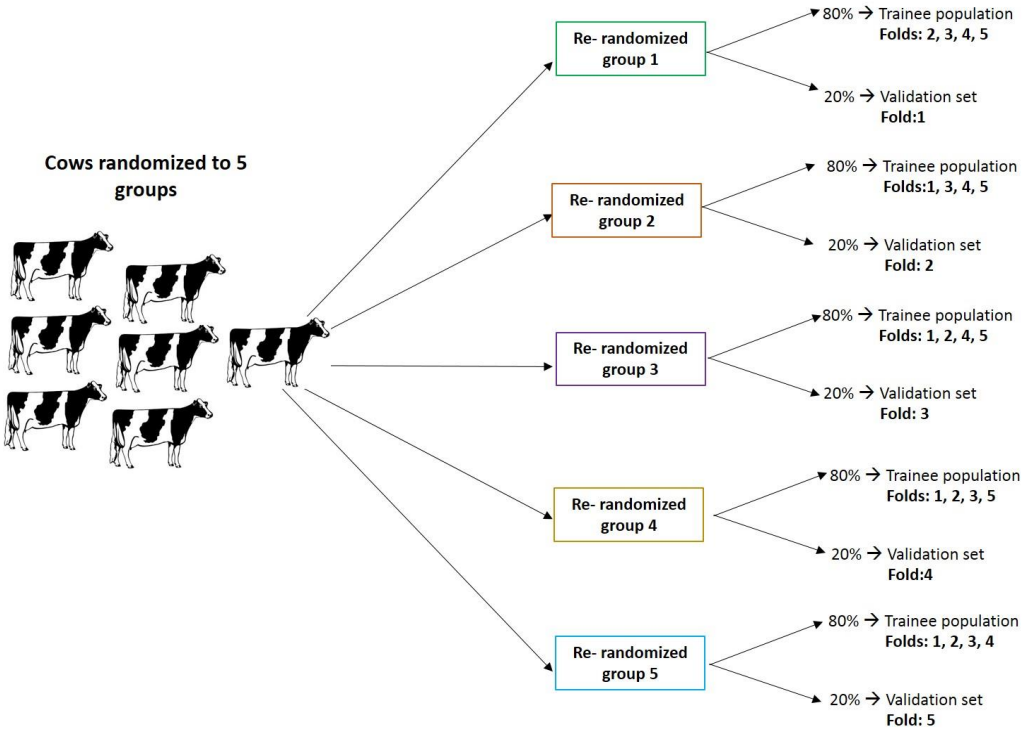

Figure 5.1. Diagram of the cross-validation scheme used to study the effectiveness of an MBV-training and predicting exercise in lactating Holstein cows (n = 659) from Sonora, MX.

Molecular breeding value (MBV)

The MBV calculation procedure was described in the Materials and Methods section of Chapter 3 and was used in the training population of each re-randomized group. Pearson's correlation between the MBV calculated from each re-randomize group and 305 d milk yield was estimated using PROC CORR.

Additionally the adjusted R^2 was calculated to evaluate the amount of variability explained by the MBV calculated from each trainee population, which we called MBV 1, MBV 2, MBV 3, MBV 4, and MBV 5 (Figure 4.1).. This calculation was performed using PROC MIXED. The model used was described as Model 3 in the Materials and Methods section of Chapter 3. The SNP effects used to estimate each MBV is reported in Supplementary table 3.

Estimations for 305 d milk yield was performed on the validation set of each re-randomized group using PROC MIXED. The model used for this analysis was described in the Materials and Methods section of Chapter 3 as a Model 2.

Results and Discussion

Cross-validation and MBV accuracy

To evaluate the efficiency of the MBV in a training and predicting exercise, correlations were estimated between the MBV 1, MBV 2, MBV 3, MBV 4, and MBV 5 and 305 d milk yield. This approach has been used by several authors describing genetic relationships in Holstein cows (Moser et al., 2010, Brøndum et al., 2011), beef cattle breeds (Saatchi et al., 2011, Mateescu et al., 2013, Boddhireddy et al., 2014) and plant species (Resende et al., 2012). Other reports of cross-validation strategies (IBS-based and K-means), reported strong correlations (0.96) among the methodologies (Boddhireddy et al., 2014); however, these MBV and results were generated with high density SNP-chip data and massive amounts of data from animals with pedigrees. In the

current study, the combination of SNP used to calculate MBV do not account for variation in these level of effects. It is worthy to mention that our data had missing observations and incomplete pedigrees.

Previous research on direct genomic value (DGV) cross-validation have reported moderate correlations between DGV with milk production traits (milk yield, fat, and protein) (Moser et al., 2010) using data from BovineSNP50 in Holstein bulls and cows. The average the reliability of the cows in that study was 0.57 (Moser et al., 2010). This previous report is not supported by our results. One reason for such lower MBV correlations could be the clustering method used to produce the 5 re-randomized groups, in which pedigree contribution was not taken into consideration. Additionally, Brøndum et al. (2011) reported low correlations between milk yield and DGV calculated from different cattle breeds (Danish Jersey, Nordic Holstein, Finnish Red, Danish Red, Swedish Red, and combined Red).

Table 5.2. Correlations between the five MBV and 305 d milk yield used to study lactating Holstein dairy cows in Sonora, MX.

	MY[1]	MBV[2] 1	MBV[2] 2	MBV[2] 3	MBV[2] 4	MBV[2] 5
MY[1]	1	0.17	-0.01	-0.08	-0.08	-0.27*
MBV[2] 1		1	0.24*	0.23*	0.36*	0.11*
MBV[2] 2			1	0.13*	0.38*	-0.02
MBV[2] 3				1	0.10*	0.29*
MBV[2] 4					1	-0.08
MBV[2] 5						1

*($P < 0.05$). [1]MY = 305 d milk yield. [2]MBV = molecular breeding value.

The adjusted R^2 for the five MBV (Table 5.3) revealed that MBV 5 ($P \leq 0.05$) explained 6.37 % of the 305 d milk yield variability in this Holstein population. In contrast, the other MBV (MBV 1, MBV 2, MBV 3 and MBV4), did not explain variability ($P \leq 0.05$) in milk yield. The MBV 5, which explained a portion of the variation in 305 d milk yield, had the highest correlation

with 305 d milk yield. Nevertheless, this correlation was negative. Previous research reported highest accuracies of prediction with subsets SNP from a high-density assay (Moser et al., 2010).

Table 5.3. Coefficient of determination of the MBV in relation to 305 d milk yield in lactating Holstein cows in Sonora, MX.

MBV[1]	Coefficient of determination. (%)
1	0
2	0
3	0
4	0
5	6.37

[1]MBV= molecular breeding value

A limitation to the current study that may be affecting our results was the data structure and quality. These data were from a relatively small sample size (n = 659) in comparison with other published studies that used a larger amounts cows and genotypes (SNP) to estimate an MBV. Some of the cows included in the dataset had no records for 305 d milk yield. Additionally, the pedigree was incomplete for a big portion of the cows.

Conclusions

The five MBV calculated from SNP within the PRL and GH-IGF1 pathways were found to be correlated with 305 d milk yield. Nevertheless, only one of the five MBV explained a portion of the variation in 305 d milk yield. The small amount of variation explained may be due to management and environmental conditions, which could be masking the positive effect of these SNP in this population. Additionally, the quality of the data, which had missing observations and incomplete pedigrees, could also affect the results. Another feasible explanation may be the polygenic nature of the trait under heat stress conditions. Finally, we accept our hypothesis, the MBV was capable of predicting a portion of the phenotypic variation in 305 d milk yield in lactating Holstein cows in Sonora, MX. Nevertheless, the accuracy and amount of variability explained was not enough to be feasible for use in genetic selection procedures.

REFERENCE

Accorsi, P., B. Pacioni, C. Pezzi, M. Forni, D. Flint, and E. Seren. 2002. Role of prolactin, growth hormone and insulin-like growth factor 1 in mammary gland involution in the dairy cow. J. Dairy Sci. 85:507-513.

Armstrong, D. 1994. Heat stress interaction with shade and cooling. J. Dairy Sci. 77:2044-2050.

Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. Van Duijn. 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23:1294-1296.

Badinga, L., R. Collier, W. Thatcher, and C. Wilcox. 1985. Effects of climatic and management factors on conception rate of dairy cattle in subtropical environment. J. Dairy Sci. 68:78-85.

Barendse, W. 2005. The transition from quantitative trait loci to diagnostic test in cattle and other livestock. Anim. Prod. Sci. 45:831-836.

Barkema, H., Y. Schukken, and R. Zadoks. 2006. Invited Review: The Role of Cow, Pathogen, and Treatment Regimen in the Therapeutic Success of Bovine Staphylococcus aureus Mastitis. J. Dairy Sci. 89:1877-1895.

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265.

Beede, D. and R. Collier. 1986. Potential nutritional strategies for intensively managed cattle during thermal stress. J. Anim. Sci. 62:543-554.

Beerda, B., J. Wyszynska-Koko, M. te Pas, A. de Wit, and R. Veerkamp. 2008. Expression profiles of genes regulating dairy cow fertility: recent findings, ongoing activities and future possibilities. Animal 2:1158–1167.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B. Met. 57:289-300.

Ben-Jonathan, N. and J. W. Liu. 1992. Pituitary lactotrophs: endocrine, paracrine, juxtacrine, and autocrine interactions. Trends. Endocrin. Met. 3:254-258.

Bernabucci, U., N. Lacetera, L. H. Baumgard, R. P. Rhoads, B. Ronchi, and A. Nardone. 2010. Metabolic and hormonal acclimation to heat stress in domesticated ruminants. Animal 4:1167-1183.

Bernabucci, U., S. Biffani, L. Buggiotti, A. Vitali, N. Lacetera, and A. Nardone. 2014. The effects of heat stress in Italian Holstein dairy cattle. J. Dairy Sci. 97:471-486.

Blum, J., T. Elsasser, D. Greger, S. Wittenberg, F. De Vries, and O. Distl. 2007. Insulin-like growth factor type-1 receptor down-regulation associated with dwarfism in Holstein calves. Domes. Anim. Endocrin. 33:245-268.

Boddhireddy, P., M. Kelly, S. Northcutt, K. Prayaga, J. Rumph, and S. Denise. 2014. Genomic predictions in Angus cattle: comparisons of sample size, response variables, and clustering methods for cross-validation. J. Anim. Sci. 92:485-497.

Bole-Feysot, C., V. Goffin, M. Edery, N. Binart, and P. A. Kelly. 1998. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocr. Rev. 19:225-268.

Bonakdar, E., H. Rahmani, M. Edriss, and B. S. Tabatabaei. 2010. IGF-I gene polymorphism, but not its blood concentration, is associated with milk fat and protein in Holstein dairy cows. Genet. Mol. Res. 9:1726-1734.

Brøndum, R. F., E. Rius-Vilarrasa, I. Strandén, G. Su, B. Guldbrandtsen, W. Fikse, and M. S. Lund. 2011. Reliabilities of genomic prediction using combined reference data of the Nordic Red dairy cattle populations. J. Dairy Sci. 94:4700-4707.

Bucholtz, D. C., N. M. Vidwans, C. G. Herbosa, K. K. Schillo, and D. L. Foster. 1996. Metabolic interfaces between growth and reproduction. V. Pulsatile luteinizing hormone secretion is dependent on glucose availability. Endocrinology 137:601-607.

Cánovas, A., G. Rincon, A. Islas-Trejo, S. Wickramasinghe, and J. F. Medrano. 2010. SNP discovery in the bovine milk transcriptome using RNA-Seq technology. Mamm. Genome 21:592-598.

Carlén, E., E. Strandberg, and A. Roth. 2004. Genetic Parameters for Clinical Mastitis, Somatic Cell Score, and Production in the First Three Lactations of Swedish Holstein Cows. J. Dairy Sci. 87:3062-3070.

Carvalho, A. V., C. Eozenou, G. Healey, N. Forde, P. Reinaud, M. Chebrout, L. Gall, N. Rodde, A. L. Padilla, and C. G. Delville. 2014. Analysis of STAT-1 expression and biological activity reveals interferon-tau-dependent STAT-1-regulated SOCS genes in the bovine endometrium. Reprod. Fertil. Dev. http://dx.doi.org/10.1071/RD14034.

Chagas, L., J. Bass, D. Blache, C. Burke, J. Kay, D. Lindsay, M. Lucy, G. Martin, S. Meier, and F. Rhodes. 2007. Invited review: New perspectives on the roles of nutrition and metabolic priorities in the subfertility of high-producing dairy cows. J. Dairy Sci. 90(9):4022-4032.

Chelikani, P. K., D. H. Keisler, and J. J. Kennelly. 2003. Response of plasma leptin concentration to jugular infusion of glucose or lipid is dependent on the stage of lactation of Holstein cows. J. Nutr. 133:4163-4171.

Clark, S. A., J. M. Hickey, H. D. Daetwyler, and J. H. van der Werf. 2012. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. Genet. Sel. Evol. 44:4.

Clempson, A., G. Pollott, J. Brickell, and D. Wathes. 2012. Associations between bovine IGFBP-2 polymorphisms with fertility, milk production, and metabolic status in UK dairy cows. Anim. Biotechnol. 23:101-113.

Clempson, A., G. Pollott, J. Brickell, N. Bourne, N. Munce, and D. Wathes. 2011. Evidence that leptin genotype is associated with fertility, growth, and milk production in Holstein cows. J. Dairy Sci. 94:3618-3628.

Cobanoglu, O., I. Zaitoun, Y. M. Chang, G. E. Shook, and H. Khatib. 2006. Effects of the signal transducer and activator of transcription 1 (STAT-1) gene on milk production traits in Holstein dairy cattle. J. Dairy Sci. 89:4433-4437.

Cochran, S. D., J. B. Cole, D. J. Null, and P. J. Hansen. 2013. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genet. 14:49.

Cohen-Zinder, M., E. Seroussi, D. M. Larkin, J. J. Loor, A. Everts-van der Wind, J. H. Lee, J. K. Drackley, M. R. Band, A. G. Hernandez, M. Shani, H. A. Lewin, J. I. Weller, and M. Ron. 2005. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. 15:936-944.

Collier, R., D. Beede, W. Thatcher, L. Israel, and C. Wilcox. 1982. Influences of environment and its modification on dairy animal health and production. J. Dairy Sci. 65:2213-2227.

Collier, R., G. Dahl, and M. VanBaale. 2006. Major advances associated with environmental effects on dairy cattle. J. Dairy Sci. 89:1244-1253.

Collier, R., J. Collier, R. Rhoads, and L. Baumgard. 2008. Invited Review: Genes Involved in the Bovine Heat Stress Response. J. Dairy Sci. 91:445-454.

Darnell, J. E. 1997. STATs and gene regulation. Science 277:1630-1635.

De Rensis, F. and R. J. Scaramuzzi. 2003. Heat stress and seasonal effects on reproduction in the dairy cow--a review. Theriogenology 60:1139-1151.

Do Amaral, B., E. Connor, S. Tao, J. Hayen, J. Bubolz, and G. Dahl. 2009. Heat-stress abatement during the dry period: Does cooling improve transition into lactation? J. Dairy Sci. 92:5988-5999.

Dunshea, F. R., B. J. Leury, F. Fahri, K. DiGiacomo, A. Hung, S. Chauhan, I. J. Clarke, R. Collier, S. Little, and L. Baumgard. 2013. Amelioration of thermal stress impacts in dairy cows. Anim. Prod. Sci. 53:965-975.

Eggen, A. 2012. The develpment and application of genomic selection as a new breeding paradigm. Anim. Frontiers 2:10-15.

Eppard, P., L. Bentle, B. Violand, S. Ganguli, R. Hintz, L. Kung, G. Krivi, and G. Lanza. 1992. Comparison of the galactopoietic response to pituitary-derived and recombinant-derived variants of bovine growth hormone. J. Endocrinol. 132:47-56.

Etherton, T. 2003. Somatotropic function: the somatomedin hypothesis revisited. J. Anim. Sci. 82:E239-244.

Fenwick, M., S. Llewellyn, R. Fitzpatrick, D. Kenny, J. Murphy, J. Patton, and D. Wathes. 2008. Negative energy balance in dairy cows is associated with specific changes in IGF-binding protein expression in the oviduct. Reproduction 135:63-75.

Fernando R.L. and M. Grossman.1989. Marker assiteed selection using best linear unbiased prediction Genet. Sel. Evol. 21: 467-477.

Flori, L., M. I. Gonzatti, S. Thevenon, I. Chantal, J. Pinto, D. Berthier, P. M. Aso, and M. Gautier. 2012. A quasi-exclusive European ancestry in the Senepol tropical cattle breed highlights the importance of the slick locus in tropical adaptation. PLoS ONE 7:e36133.

Forni, S., I. Aguilar, and I. Misztal. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Genet. Sel. Evol. 43:1.

Freeman, M. E., B. Kanyicska, A. Lerant, and G. Nagy. 2000. Prolactin: structure, function, and regulation of secretion. Physiol. Rev. 80:1523-1631.

Frühbeck, G., S. Jebb, and A. Prentice. 1998. Leptin: physiology and pathophysiology. Clin. Physiol. 18:399-419.

Fuchs, A., O. Behrens, H. Helmer, C.H. Liu, C. Barros, and M. Fields. 1990. Oxytocin and Vasopressin Receptors in Bovine Endometrium and Myometrium during the Estrous Cycle and Early Pregnancy. Endocrinology 127:629-636.

Garrick, D. J. 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. Genet. Sel. Evol. 43:10.1186.

Gilmour, A. R., B. Gogel, B. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK.

Goddard M.E., B. J. Hayes, T. H. Meuwissen. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. J. Anim. Breed. Genet. doi:10.1111/j.1439-0388.2011.00964.x

Gondro, C., J. Van der Werf, and B. J. Hayes. 2013. Genome-wide Association Studies and Genomic Prediction. Chapters: 8,6,13,15. Springer.

Gozdz, A., E. Szczepanska-Sadowska, K. Szczepanska, W. Maslinski, and B. Luszczyk. 2002. Vasopressin V1a, V1b and V2 receptors mRNA in the kidney and heart of the renin transgenic TGR(mRen2)27 and Sprague Dawley rats. J. Physiol. Pharmacol. 53:349-357.

Habier, D., R. Fernando, and J. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. Genetics 177:2389-2397.

Hadsell, D. L., A. F. Parlow, D. Torres, J. George, and W. Olea. (2008). Enhancement of maternal lactation performance during prolonged lactation in the mouse by mouse GH and long-R3-IGF-I is linked to changes in mammary signaling and gene expression. J. Endocrinol. 198: 61-70.

Halasa, T., K. Huijps, O. Østerås, and H. Hogeveen. 2007. Economic effects of bovine mastitis and mastitis management: a review. Vet. Q. 29:18-31.

Hayes, B., P. Bowman, A. Chamberlain, K. Verbyla, and M. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genet. Sel. Evol. 41:51.

Hazel, L. and J. L. Lush. 1942. The efficiency of three methods of selection. J. Hered. 33:393-399.

Henderson, C. R. 1950. Estimation of genetic parameters Ann. Math. Stat. 21:309–310.

Henderson, C. R. 1975. Best Linear Unbiased Estimation and Prediction Under a Selection Model. Biometrics 31:423-447.

Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: a review with focus on the situation in the Nordic countries. Livest. Prod. Sci. 64(2–3):95-106.

Huson, H. J., E. S. Kim, R. W. Godfrey, T. A. Olson, M. C. McClure, C. C. Chase, R. Rizzi, A. M. O'Brien, C. P. Van Tassell, J. F. Garcia, and T. S. Sonstegard. 2014. Genome-wide

association study and ancestral origins of the slick-hair coat in tropically adapted cattle. Front. Genet. 5:101.

Huttmann, H., E. Stamer, W. Junge, G. Thaller, and E. Kalm. 2009. Analysis of feed intake and energy balance of high-yielding first lactating Holstein cows with fixed and random regression models. Animal 3:181-188.

Ihle, J. N. 2001. The Stat family in cytokine signaling. Curr. Opin. Cell Biol. 13:211-217.

Jakobsen, J. H., P. Madsen, J. Jensen, J. Pedersen, L. Christensen, and D. Sorensen. 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. J. Dairy Sci. 85:1607-1616.

Jamrozik, J. and L. Schaeffer. 1997. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. J. Dairy Sci. 80:762-770.

Jolly, P., S. McDougall, L. Fitzpatrick, K. Macmillan, and K. Entwistle. 1994. Physiological effects of undernutrition on postpartum anoestrus in cows. J. Reprod. Fertil. 49:477-492.

Jordan, E. 2003. Effects of heat stress on reproduction. J. Dairy Sci. 86:E104-E114.

Kadarmideen, H. N., and Wegmann, S. 2003. Genetic parameters for body condition score and its relationship with type and production traits in Swiss Holsteins. J. Dairy Sci. 86:3685-3693.

Karoui, S., M. J. Carabaño, C. Díaz, and A. Legarra. 2012. Joint genomic evaluation of French dairy cattle breeds using multiple-trait models. Genet. Sel. Evol. 44:39.

Key, N., S. Sneeringer, and D. Marquardt. 2014. Climate Change, Heat Stress, and US Dairy Production. USDA-ERS Economic Research Report (175).

Khatib, A.M. and F. Sfaxi. 2012. FURIN (FURIN (paired basic amino acid cleaving enzyme)). Atlas Genet. Cytogenet. Oncol. Hermatol. 16:639-643.

Khatib, H., R. Monson, V. Schutzkus, D. Kohl, G. Rosa, and J. Rutledge. 2008. Mutations in the STAT-5A gene are associated with embryonic survival and milk composition in cattle. J. Dairy Sci. 91:784-793.

Kim, J. W. 2014. Modulation of the Somatotropic Axis in Periparturient Dairy Cows. Asian-Australas J. Anim. Sci. 27:147-154.

Le Gac, F., O. Blaise, A. Fostier, P.Y. Le Bail, M. Loir, B. Mourot, and C. Weil. 1993. Growth hormone (GH) and reproduction: a review. Fish Physiol. Biochem. 11:219-232.

Lee, B., G. Lin, B. Crooker, M. Murtaugh, L. Hansen, and H. Chester-Jones. 1993. Association of somatotropin (bST) gene polymorphism with selection for milk yield in Holstein cows. J. Dairy Sci. 76:149.

Li, C., D. Sun, S. Zhang, S. Wang, X. Wu, Q. Zhang, L. Liu, Y. Li, and L. Qiao. 2014. Genome Wide Association Study Identifies 20 Novel promising genes associated with milk fatty acid traits in chinese holstein. PloS ONE 9:e96186.

Liefers, S. C., M. F. te Pas, R. F. Veerkamp, and T. van der Lende. 2002. Associations between leptin gene polymorphisms and production, live weight, energy balance, feed intake, and fertility in Holstein heifers. J. Dairy Sci. 85:1633-1638.

Littlejohn, M. D., K. M. Henty, K. Tiplady, T. Johnson, C. Harland, T. Lopdell, R. G. Sherlock, W. Li, S. D. Lukefahr, and B. C. Shanks. 2014. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nat. Commun. doi:10.1038/ncomms6861.

Lü, A., X. Hu, H. Chen, J. Jiang, C. Zhang, H. Xu, and X. Gao. 2010. Single nucleotide polymorphisms in bovine PRL gene and their associations with milk production traits in Chinese Holsteins. Mol. Biol. Rep. 37:547-551.

Lucy, M. 2008. Functional Differences in the Growth Hormone and Insulin-like Growth Factor Axis in Cattle and Pigs: Implications for Post-partum Nutrition and Reproduction. Reprod. Domest. Anim. 43:31-39.

Lucy, M. C., S. D. Hauser, P. J. Eppard, G. G. Krivi, J. H. Clark, D. E. Bauman, and R. J. Collier. 1993. Variants of somatotropin in cattle: gene frequencies in major dairy breeds and associated milk production. Domest. Anim. Endocrinol. 10:325-333.

Macciotta, N. P. P., Vicario, D., and Cappio-Borlino, A. 2005. Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. J. Dairy Sci. 88:1178-1191.

Mader, T. L., M. Davis, and T. Brown-Brandl. 2006. Environmental factors influencing heat stress in feedlot cattle. J. Anim. Sci. 84:712-719.

Mariasegaram, M., C. C. Chase, J. X. Chaparro, T. A. Olson, R. A. Brenneman, and R. P. Niedz. 2007. The slick hair coat locus maps to chromosome 20 in Senepol-derived cattle. Anim. Genet. 38:54-59.

Maruotti, J., M. Muñoz, S. A. Degrelle, E. Gómez, C. Louet, C. D. Monforte, P. H. De Longchamp, V. Brochard, I. Hue, and J. N. Caamaño. 2012. Efficient derivation of bovine embryonic stem cells needs more than active core pluripotency factors. Mol. Reprod. Dev. 79:461-477.

Mateescu, R., D. Garrick, R. Tait, A. Garmyn, Q. Duan, Q. Liu, M. Mayes, A. Van Eenennaam, D. VanOverbeke, and G. Hilton. 2013. Genome-wide association study of concentrations of iron and other minerals in longissimus muscle of Angus cattle. J. Anim. Sci. 91:3593-3600.

Meuwissen, T. H., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Moser, G., B. Tier, R. E. Crump, M. S. Khatkar, and H. W. Raadsma. 2009. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. Genet. Sel. Evol. 41:56.

Moser, G., M. S. Khatkar, B. J. Hayes, and H. W. Raadsma. 2010. Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP markers. Genet. Sel. Evol. 42:37.

Muralidhar, K. and J. Lee. 2013. Prolactin and Angiogenesis: Biological Implications of Microheterogeneity, Prolactin. InTech, DOI: 10.5772/54318.

National Research Council. 2001. Nutrient requirements of dairy cattle. 7th rev. ed. Natl. Acad. Press, Washington, DC.

Nejati-Javaremi, A., C. Smith, and J. P. Gibson. 1997. Effect of total allelic relationship on accuracy of evaluation and response to selection. J. Anim. Sci. 75:1738-1745.

Nordborg, M. and S. Tavaré. 2002. Linkage disequilibrium: what history has to tell us. Trends. Genet. 18:83-90.

Nussey, S., R. Prysor-Jones, A. Taylor, V. Ang, and J. Jenkins. 1987. Arginine vasopressin and oxytocin in the bovine adrenal gland. J. Endocrinol. 115:141-149.

Ogorevc, J., T. Kunej, A. Razpet, and P. Dovc. 2009. Database of cattle candidate genes and genetic markers for milk production and mastitis. Anim. Genet. 40:832-851.

Ojango, J. and G. Pollott. 2001. Genetics of milk yield and fertility traits in Holstein-Friesian cattle on large-scale Kenyan farms. J. Anim. Sci. 79:1742-1750.

Olson, T. A., C. Lucena, C. C. Chase, and A. C. Hammond. 2003. Evidence of a major gene influencing hair length and heat tolerance in Bos taurus cattle. J. Anim. Sci. 81:80-90.

Parnell, P. 2004. Best linear unbiased prediction (BLUP) and the design of breeding programs. P. Aus. S. Ani. 15:513-516.

Pimentel, E., S. Bauersachs, M. Tietze, H. Simianer, J. Tetens, G. Thaller, F. Reinhardt, E. Wolf, and S. König. 2011. Exploration of relationships between production and fertility traits in dairy cattle via association studies of SNP within candidate genes derived by expression profiling. Anim. Genet.. 42:251-262.

Posner, S. F., C. A. Vaslet, M. Jurofcik, A. Lee, N. G. Seidah, and E. A. Nillni. 2004. Stepwise posttranslational processing of progrowth hormone-releasing hormone (proGHRH) polypeptide by furin and PC1. Endocrine 23:199-213.

Pryce, J., M. Royal, P. Garnsworthy, and I. L. Mao. 2004. Fertility in the high-producing dairy cow. Livest. Prod. Sci. 86:125-135.

Rabiee, A., I. Lean, J. Gooden, B. Miller, and R. Scaramuzzi. 1997. An evaluation of transovarian uptake of metabolites using arterio-venous difference methods in dairy cattle. Anim. Reprod. Sci. 48:9-25.

Raven, L.-A., B. G. Cocks, J. E. Pryce, J. J. Cottrell, and B. J. Hayes. 2013. Genes of the RNASE5 pathway contain SNP associated with milk production traits in dairy cattle. Genet. Sel. Evol. 45:25.

Resende, M. F., P. Muñoz, M. D. Resende, D. J. Garrick, R. L. Fernando, J. M. Davis, E. J. Jokela, T. A. Martin, G. F. Peter, and M. Kirst. 2012. Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda L.). Genetics 190:1503-1510.

Rincón, J. C., A. López, and J. Echeverri. 2012. Variabilidad genética del loci RsaI de la prolactina bovina en Ganado Holstein de Antioquia. Rev. Colomb. Cienc. Pecu. 25:191-201.

Robinson, G. K. 1991. That BLUP is a good thing: The estimation of random effects. Statist. Sci. 6:15-32.

Rocheville, M., D. C. Lange, U. Kumar, R. Sasi, R. C. Patel, and Y. C. Patel. 2000. Subtypes of the somatostatin receptor assemble as functional homo-and heterodimers. J. Biol. Chem. 275:7862-7869.

Rodriguez-Martinez, H., J. Hultgren, R. Båge, A. Bergqvist, C. Svensson, C. Bergsten, L. Lidfors, S. Gunnarsson, B. Algers, and U. Emanuelson. 2008. Reproductive performance in high-producing dairy cows: can we sustain it under current practice. IVIS Reviews in Veterinary Medicine, IVIS (Ed.). International Veterinary Information Service, Ithaca NY (www. ivis. org). Last updated: 12-Dec-2008.

Saatchi, M., M. C. McClure, S. D. McKay, M. M. Rolf, J. Kim, J. E. Decker, T. M. Taxis, R. H. Chapple, H. R. Ramey, and S. L. Northcutt. 2011. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genet. Sel. Evol. 43:40.

Saatchi, M., R. D. Schnabel, M. M. Rolf, J. F. Taylor, and D. J. Garrick. 2012. Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle. Genet. Sel. Evol. 44:38.

Sakamoto, K., T. Yano, T. Kobayashi, A. Hagino, H. Aso, and Y. Obara. 2007. Growth hormone suppresses the expression of IGFBP-5, and promotes the IGF-I-induced phosphorylation of Akt in bovine mammary epithelial cells. Domest. Anim. Endocrinol. 32:260-272.

Schefers, J. M. and K. A. Weigel. 2012. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. Anim. Frontiers 2:4-9.

Schennink, A., H. Bovenhuis, K. M. Léon-Kloosterziel, J. A. van Arendonk, and M. H. Visker. 2009. Effect of polymorphisms in the FASN, OLR1, PPARGC1A, PRL and STAT-5A genes on bovine milk-fat composition. Anim. Genet. 40:909-916.

Sauvant, D. 1988. Modelling of lactation and nutrition. Pages 149– 155 in Korver S., and J. A. M. Van Arendonk, ed. Modelling of Livestock Production Systems. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Sengupta, T. K., A. Chen, Z. Zhong, J. E. Darnell, and L. B. Ivashkiv. 1995. Activation of monocyte effector genes and STAT family transcription factors by inflammatory synovial fluid is independent of interferon gamma. J. Exp. Med. 181:1015-1025.

Seykora, A. J. and B. T. McDaniel. 1983. Heritabilities and Correlations of Lactation Yields and Fertility for Holsteins. J. Dairy Sci. 66:1486-1493.

Silva, M. V. B., D. J. A. dos Santos, S. A. Boison, A. T. H. Utsunomiya, A. S. Carmo, T. S. Sonstegard, J. B. Cole, and C. P. Van Tassell. 2014. The development of genomics applied to dairy breeding. Livest. Sci. 166:66-75.

Sun, C., P. Madsen, M. Lund, Y. Zhang, U. Nielsen, and G. Su. 2010. Improvement in genetic evaluation of female fertility in dairy cattle using multiple-trait models including milk production traits. J. Anim. Sci. 88:871-878.

Szewczuk, M., S. Zych, E. Czerniawska-Piatkowska, and J. Wójcik. 2012. Association between IGF1R/i16/TaqI and IGF1/SnaBI polymorphisms and milk production traits in Polish Holstein-Friesian cows. J. Anim. Sci. 30:13-24.

Tekerli, M., Z. Akinci, I. Dogan , and A. Akcan. 2000. Factors affecting the shape of lactation curves of Holstein cows from the Balikesir province of Turkey. J. Dairy Sci. 83:1381-1386.

Trout, J. P., L. R. McDowell, and P. J. Hansen. 1998. Characteristics of the estrous cycle and antioxidant status of lactating Holstein cows exposed to heat stress. J. Dairy Sci. 81:1244-1250.

Val-Arreola, D., E. Kebreab, J. Dijkstra, and J. France. 2004. Study of the lactation curve in dairy cattle on farms in central Mexico. J. Dairy Sci. 87:3789-3799.

Van Dorp, T. E., J. C. M. Dekkers, S. W. Martin, and J. P. T. M. Noordhuizen. 1998. Genetic parameters of health disorders, and relationships with 305-day milk yield and conformation traits of registered Holstein cows. J. Dairy Sci. 81:2264-2270.

VanRaden, P., C. Van Tassell, G. Wiggans, T. Sonstegard, R. Schnabel, J. Taylor, and F. Schenkel. 2009. Invited Review: Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16-24.

Vitali, A., M. Segnalini, L. Bertocchi, U. Bernabucci, A. Nardone, and N. Lacetera. 2009. Seasonal pattern of mortality and relationships between mortality and temperature-humidity index in dairy cows. J. Dairy Sci. 92:3781-3790.

West, J. 2003. Effects of heat-stress on production in dairy cattle. J. Dairy Sci. 86:2131-2144.

Wickramasinghe, S., G. Rincon, and J. Medrano. 2011. Variants in the pregnancy-associated plasma protein-A2 gene on Bos taurus autosome 16 are associated with daughter calving ease and productive life in Holstein cattle. J. Dairy Sci. 94:1552-1558.

Yan, X., R. C. Baxter, and S. M. Firth. 2010. Involvement of pregnancy-associated plasma protein-A2 in insulin-like growth factor (IGF) binding protein-5 proteolysis during pregnancy: a potential mechanism for increasing IGF bioavailability. J. Clin. Endocrinol. Metab. 95:1412-1420.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. E. Goddard, and P. M. Visscher. 2010. Common SNP explain a large proportion of the heritability for human height. Nat. Genet. 42:565-569.

Yao, J., S. E. Aggrey, D. Zadworny, J. F. Hayes, and U. Kühnlein. 1996. Sequence variations in the bovine growth hormone gene characterized by single-strand conformation polymorphism (SSCP) analysis and their association with milk production traits in Holsteins. Genetics 144:1809-1816.

Zhang, F., J. Huang, Q. Li, Z. Ju, J. Li, F. Shi, J. Zhong, and C. Wang. 2010. Novel single nucleotide polymorphisms (SNP) of the bovine STAT-4 gene and their associations with production traits in Chinese Holstein cattle. Afr. J. Biotechnol. 9:4003-4008.

Zimbelman, R., R. Rhoads, M. Rhoads, G. Duff, L. Baumgard, and R. Collier. 2009. A re-evaluation of the impact of temperature humidity index (THI) and black globe humidity index (BGHI) on milk production in high producing dairy cows. Pages 158-168 in Proc. Southwest Nutr. Manag. Conf. Tempe, AZ. ARPAS, Savoy, IL.

Żukiewicz, A., W. Grzesiak, I. Szatkowska, P. Błaszczyk, and A. Dybus. 2012. Genetic factors of milk yield in dairy cattle–advances in the quest for universal markers. Israel J. Vet. Med. 67:2.

APPENDIX

Supplementary table 1. P value from the Hardy-Weinberg equilibrium (HWE) chi-square test and allele frequencies for 179 tag SNP in the PRL and GH-IGF1 pathways used to study heat stressed lactating Holstein cows in Sonora, MX.

Gene	RN[1]	*P*	Allele 1	Allele 1 frequency	Allele 2	Allele 2 frequency
AVP	rs467297442	0.27	A	0.16	G	0.84
AVP	NA	0.01	C	0.25	T	0.75
AVP	rs469243577	0.00	C	0.79	T	0.21
AVPR1A	rs207971189	0.90	C	0.67	G	0.33
AVPR1A	rs210011420	0.68	C	0.39	T	0.61
AVPR1A	rs209300854	0.80	C	0.61	G	0.39
CISH	rs209463645	0.00	A	0.59	G	0.42
CISH	rs208019931	0.27	C	0.81	G	0.19
FURIN	rs382538054	0.31	A	0.23	G	0.77
FURIN	rs134721854	0.00	C	0.13	T	0.87
FURIN	rs210731409	0.91	C	1.00	T	0.00
FURIN	rs466130569	0.75	G	0.99	T	0.01
FURIN	rs463846971	0.36	G	0.94	T	0.06
FURIN	rs381099643	0.06	A	0.15	G	0.85
GHR	rs41639262	0.82	A	0.38	G	0.62
GHRH	rs133786352	0.67	C	0.95	T	0.05
GHRH	rs109663333	0.84	A	0.64	T	0.36
GHRH	rs109912355	0.01	A	0.65	G	0.35
GHRH	rs109981400	0.02	G	0.79	T	0.21
GHRH	rs380969504	0.32	A	0.05	G	0.95
GHRH	rs137760387	0.35	A	0.04	C	0.96
GHRHR	rs43407600	0.91	A	0.17	G	0.83
GHRHR	rs465206110	0.97	A	0.00	G	1.00
GHSR	rs210921858	0.60	C	0.97	T	0.03
GHSR	rs110476783	0.88	A	0.04	C	0.96
GHSR	rs133986528	0.78	A	0.97	G	0.03
GHSR	rs481253740	0.14	G	0.05	T	0.95
GHSR	rs385048010	0.92	A	0.00	G	1.00

GHSR	rs110721203	0.98	C	0.04	T	0.96
GHSR	rs110950555	0.28	A	0.86	C	0.14
GHSR	NA	.	A	1.00	.	.
IGF1	rs109763947	0.29	C	0.39	T	0.61
IGF-2	rs137289661	0.86	C	0.77	T	0.23
IGF1R	rs380909637	0.54	A	0.13	G	0.87
IGF1R	rs41961338	0.31	A	0.43	G	0.57
IGF1R	rs133310242	0.63	A	0.08	G	0.92
IGF1R	rs210778604	0.06	C	0.74	T	0.26
IGF1R	rs211549206	0.07	A	0.33	G	0.67
IGF1R	rs41640706	0.18	A	0.05	G	0.95
IGF1R	rs110343126	0.75	A	0.52	G	0.48
IGF1R	rs470246390	0.74	A	0.01	G	0.99
IGF1R	rs208140993	0.31	C	0.50	T	0.50
IGF1R	rs41960583	0.24	C	0.67	T	0.33
IGF1R	rs109762729	0.07	G	0.32	T	0.68
IGFBP2	rs134739850	0.21	A	0.11	G	0.89
IGFBP2	rs110305498	0.56	C	0.11	T	0.89
IGFBP2	rs134705980	0.87	A	0.56	C	0.44
IGFBP2	rs209576314	0.80	C	0.18	G	0.82
IGFBP2	rs443442023	0.67	A	0.17	G	0.83
IGFBP3	rs17870204	0.01	G	0.57	T	0.43
IGFBP3	rs17870212	0.04	C	0.41	T	0.59
IGFBP4	rs378389402	0.79	G	0.06	T	0.94
IGFBP5	NA	0.98	A	1.00	G	0.00
IGFBP5	rs135457390	0.30	C	0.89	G	0.11
IGFBP5	rs208989155	0.21	A	0.21	G	0.79
IGFBP5	rs134231478	0.89	C	0.89	T	0.11
IGFBP5	rs110668467	0.68	C	0.88	G	0.12
IGFBP6	rs211039223	0.41	C	0.90	T	0.10
IGFBP6	rs135291157	0.06	A	0.96	G	0.04

IGFBP6	NA	.	C	1.00	.	.
IGFBP7	rs43477917	0.54	A	0.29	C	0.71
IGFBP7	rs208148962	0.61	C	0.21	T	0.79
IGFBP7	rs449568329	0.90	A	0.05	G	0.95
IGFBP7	rs43477922	0.76	C	0.07	T	0.93
IGFBP7	rs43477925	0.95	A	0.00	C	1.00
IGFBP7	rs454266321	0.00	C	0.05	T	0.95
IGFBP7	rs443140594	0.01	A	0.94	C	0.06
IGFBP7	rs43469649	0.00	C	0.98	T	0.02
IGFBP7	rs43477915	0.58	A	0.24	G	0.76
OXT	rs135620444	.	T	1.00	.	.
OXT	rs137154444	0.00	A	0.86	G	0.14
OXT	rs137388314	0.00	C	0.25	T	0.75
OXTR	rs42002643	0.67	A	0.50	G	0.50
OXTR	rs42002659	0.00	A	0.63	C	0.37
OXTR	rs42002660	0.00	C	0.93	T	0.07
PAPPA1	rs379196319	0.43	A	0.20	C	0.80
PAPPA1	rs384230354	0.25	A	0.29	G	0.71
PAPPA1	rs209859180	0.00	A	0.29	G	0.71
PAPPA2	rs109779265	0.06	A	0.33	C	0.67
PAPPA2	rs109952914	0.28	A	0.18	T	0.82
PAPPA2	rs42301978	0.17	A	0.07	C	0.93
PAPPA2	rs42301985	0.90	A	0.05	G	0.95
PAPPA2	rs42300479	0.00	C	0.45	T	0.55
PAPPA2	rs109706337	0.54	C	0.02	T	0.98
PAPPA2	NA	.	G	1.00	.	.
PCSK2	rs41688130	0.15	A	0.28	G	0.72
PCSK2	rs137423265	0.12	G	0.66	T	0.34
PCSK2	rs41685759	0.00	A	0.16	G	0.84
PIAS1	rs385447261	.	.	.	.	.
PIAS1	rs383631454	0.53	A	0.84	C	0.16

PIAS1	rs137166453	0.04	A	0.46	G	0.54
PMCH	rs14197280	0.62	A	0.52	T	0.48
PRL	rs110586822	0.09	A	0.86	G	0.14
PRL	rs211032652	0.00	A	0.19	G	0.81
PRL	rs110494133	0.27	A	0.70	G	0.30
PRL	rs110904118	0.00	C	0.12	T	0.88
PRL	rs134028641	0.85	C	0.99	T	0.01
PRLR	rs209364409	0.79	A	0.03	G	0.97
PRLR	rs109428015	0.19	C	0.95	T	0.05
PRLR	rs135164815	0.12	A	0.75	G	0.25
PRLR	rs136247583	0.13	C	0.75	T	0.25
PRLR	rs43158737	.	G	1.00	.	.
SCG5	rs385034220	0.97	C	1.00	T	0.00
SCG5	rs109962791	0.48	A	0.46	G	0.54
SCG5	rs109273675	0.72	A	0.64	G	0.36
SOCS1	rs383043882	0.66	A	0.02	G	0.98
SOCS1	rs210216882	0.17	C	0.42	T	0.58
SOCS1	rs441084041	0.59	A	0.98	G	0.02
SOCS1	rs109183195	0.00	G	0.17	T	0.83
SOCS2	rs136895314	0.02	C	0.51	T	0.49
SOCS2	rs109409520	0.24	A	0.18	G	0.82
SOCS2	rs132661440	0.69	C	0.98	T	0.02
SOCS2	rs137463248	.	C	1.00	.	.
SOCS2	rs136382760	0.78	C	0.99	T	0.01
SOCS3	rs458247445	0.00	A	0.13	G	0.87
SOCS3	NA	0.00	C	0.82	T	0.18
SOCS4	NA	0.96	C	0.00	G	1.00
SOCS4	rs109702177	0.86	A	0.37	G	0.63
SOCS4	rs456481871	0.89	A	0.04	G	0.96
SOCS5	rs210573908	0.49	A	0.37	A	0.63
SOCS6	rs109979250	0.03	A	0.49	G	0.51

SOCS6	rs381761783	0.08	A	0.22	G	0.78
SOCS6	rs110213772	0.30	C	0.65	T	0.35
SOCS7	rs480561519	0.56	A	0.11	G	0.89
SOCS7	rs110136164	0.98	A	0.65	G	0.35
SOCS7	rs211359837	0.02	A	0.46	C	0.54
SOCS7	NA	.	A	1.00	.	.
SOCS7	NA	.	T	1.00	.	.
SOCS7	rs209926244	0.00	C	0.40	C	0.60
SOCS7	rs109563188	0.01	G	0.39	T	0.61
SST	rs17870997	0.76	A	0.33	G	0.68
SST	rs472257957	0.72	C	0.99	T	0.01
SSTR2	rs110053675	0.00	C	0.10	T	0.90
SSTR2	rs110602382	0.15	A	0.40	G	0.60
SSTR2	rs137754010	0.32	C	0.08	T	0.92
SSTR2	rs207769413	0.67	C	0.88	T	0.12
SSTR2	rs381275188	0.94	C	0.00	T	1.00
SSTR3	rs466764839	0.70	C	0.95	T	0.05
SSTR3	rs137314909	0.95	A	0.60	G	0.40
SSTR3	rs109318052	0.03	C	0.30	T	0.70
SSTR3	rs136447809	0.02	C	0.70	G	0.30
SSTR3	rs43438660	0.44	C	0.70	G	0.31
SSTR3	rs43438659	0.32	A	0.30	T	0.70
SSTR3	rs109931679	0.00	A	0.83	G	0.17
SSTR5	rs383554671	0.02	C	0.90	T	0.10
SSTR5	rs132901966	0.58	C	0.94	T	0.06
SSTR5	rs109914110	0.29	C	0.90	T	0.10
STAT1	NA	.	C	1.00	.	.
STAT1	rs209274978	0.65	A	0.45	G	0.55
STAT1	rs471883369	0.84	C	0.99	T	0.01
STAT1	rs134291403	0.69	A	0.55	G	0.45
STAT1	rs43706906	0.99	C	0.55	G	0.45

STAT1	rs134129900	0.80	A	0.45	G	0.55
STAT1	NA	0.04	A	0.73	G	0.27
STAT3	rs137587098	0.00	C	0.40	T	0.60
STAT3	rs110942700	0.77	A	0.51	G	0.49
STAT4	rs134874928	0.18	C	0.80	T	0.20
STAT4	rs385524813	0.91	A	0.00	C	1.00
STAT4	rs110153328	.	T	1.00	.	.
STAT4	rs110344022	0.37	C	0.32	T	0.68
STAT4	rs110893400	0.91	A	0.47	G	0.53
STAT4	NA	.	T	1.00	.	.
STAT4	rs384033065	0.94	A	0.05	G	0.95
STAT4	rs109845537	0.55	C	0.51	G	0.49
STAT4	rs209426968	0.00	C	0.55	T	0.45
STAT5A	rs137182814	0.27	C	0.55	G	0.45
STAT5A	rs379638945	0.43	C	0.96	T	0.04
STAT5B	NA	.	A	1.00	C	.
STAT5B	rs134393319	0.01	A	0.09	G	0.91
STAT5B	rs132929933	0.62	C	0.95	T	0.05
STAT5B	rs441151034	0.00	G	0.62	T	0.38
STAT5B	rs384930401	0.63	A	0.15	G	0.85
STAT5B	rs43706496	0.93	C	0.44	T	0.56
STAT5B	rs41915659	0.88	C	0.44	T	0.56
STAT6	rs109171041	0.01	C	0.65	G	0.35
STAT6	rs110335864	0.55	A	0.11	C	0.89
STAT6	rs109238562	0.47	C	0.14	T	0.86
STAT6	rs110097583	0.21	A	0.16	G	0.84
STAT6	rs109821685	0.48	A	0.14	C	0.86

[1]RN=SNP Reference number on dbSNP (NCBI). January 2015.

Supplementary table 2. Candidate genes within the PRL and GH-IGF1 pathways used to study heat stressed lactating Holstein cows in Sonora, MX. General biological functions

Gene	Definition	Function (Gene ontology annotation)
AVP	Arginine Vasopressin	V1A vasopressin receptor binding, cysteine-type endopeptidase inhibitor activity involved in apoptotic process, neurohypophyseal hormone activity, protein kinase activity, signal transducer activity.
AVPR	Arginine Vasopressin Receptor	Regulation of systemic arterial blood pressure by vasopressin, signal transducer activity, G-protein coupled receptor activity, vasopressin receptor activity, protein binding, signal transduction, G-protein coupled receptor signaling pathway, membrane, integral component of membrane.
CISH	Cytokine inducible SH2-containing Protein	Protein kinase inhibitor activity.
FURIN	FURIN (paired basic amino acid cleaving enzyme)	Metal ion binding, serine-type endopeptidase activity
GH	Growth Hormone	Growth hormone receptor binding, hormone activity, metal ion binding.
GHRH	Growth Hormone Releasing Hormone	Growth hormone-releasing hormone activity, growth Hormone-releasing hormone activity.
GHRHR	Growth Hormone Releasing Hormone Receptor	Growth hormone-releasing hormone receptor binding
GHSR	Growth Hormone Segretagogue Receptor	Growth hormone-releasing hormone receptor activity
IGF1	Insulin-Like Growth Factor-1	Growth factor activity, hormone activity.
IGF1R	Insulin-Like Growth Factor-1 Receptor	ATP binding, identical protein binding, insulin binding, insulin receptor binding, insulin receptor substrate binding, insulin-like growth factor I binding, insulin-like growth factor-activated receptor activity, phosphatidylinositol 3-kinase binding, protein tyrosine kinase activity.
IGFBP2	Insulin-Like Growth Factor Binding Protein-2	Insulin-like growth factor I binding, insulin-like growth factor II binding.

IGFBP3	Insulin-Like Growth Factor Binding Protein-3	Fibronectin binding, insulin-like growth factor I binding, insulin-like growth factor II binding, protein tyrosine phosphatase activator activity.
IGFBP4	Insulin-Like Growth Factor Binding Protein-4	Insulin-like growth factor I binding, insulin-like growth factor II binding.
IGFBP5	Insulin-Like Growth Factor Binding Protein-5	Fibronectin binding, insulin-like growth factor I binding.
IGFBP6	Insulin-Like Growth Factor Binding Protein-6	Insulin-like growth factor binding.
IGFBP7	Insulin-Like Growth Factor Binding Protein-7	Regulation of cell growth.
OXTR	Oxytocin receptor	Oxytocin receptor activity, peptide binding, vasopressin receptor activity.
OXT	Oxytocin	Neurohypophyseal hormone activity.
PAPPA1	Pregnancy Associated Plasmatic Protein A1	Metalloendopeptidase activity, metallopeptidase activity, zinc ion binding.
PAPPA2	Pregnancy Associated Plasmatic Protein A2	Protein binding, proteolysis, metallopeptidase activity, zinc ion binding, membrane, cell differentiation, bone morphogenesis, extracellular vesicular exosome.
PCSK2	Proprotein Convertase K2	Phosphoenolpyruvate carboxykinase activity, phosphoenolpyruvate carboxykinase (GTP) activity, protein binding, GTP binding, mitochondrion, gluconeogenesis,purine nucleotide binding, extracellular vesicular exosome.
PIAS	Protein Inhibitor of Activated STAT-1	Negative regulation of transcription from RNA polymerase II promoter, protein binding transcription factor activity.
PMCH	Pro-Melanin Concentrating Hormone	Melanin-concentrating hormone activity.
PRL	Prolactin	Hormone activity, prolactin receptor binding.
PRLR	Prolactin Receptor	Cytokine receptor activity, metal ion binding.
SCGV	Secretogranin V	Enzyme inhibitor activity, protein binding.
SOCS1	Supressor of Cytokine Signalling-1	Regulation of protein phosphorylation, insulin-like growth factor receptor binding, protein binding.

SOCS2	Supressor of Cytokine Signalling-2	Protein kinase inhibitor activity, insulin-like growth factor receptor binding, JAK pathway signal transduction adaptor activity.
SOCS3	Supressor of Cytokine Signalling-3	Protein kinase inhibitor activity.
SOCS4	Supressor of Cytokine Signalling-4	Protein kinase inhibitor activity.
SOCS5	Supressor of Cytokine Signalling-5	Epidermal growth factor receptor binding, protein binding.
SOCS6	Supressor of Cytokine Signalling-6	Immunological synapse, protein binding.
SOCS7	Supressor of Cytokine Signalling-7	Protein binding.
SST	Somatostatin	Hormone activity.
SSTR2	Somatostain Receptor-2	Somatostatin receptor activity.
SSTR3	Somatostatin Receptor-3	Molecular function, signal transducer activity, somatostatin receptor activity, protein binding, cellular component, cell, cytoplasm, plasma membrane, cilium,signal transduction, spermatogenesis.
SSTR5	Somatostatin Receptor-5	Neuropeptide binding, somatostatin receptor activity.
STAT1	Signal Transducer and Activator of Transcription-1	RNA polymerase II core promoter proximal region sequence-specific DNA binding, RNA polymerase II core promoter sequence-specific DNA binding, RNA polymerase II core promoter sequence-specific DNA binding transcription factor activity, negative regulation of endothelial cell proliferation, DNA binding, double-stranded DNA binding, sequence-specific DNA binding transcription factor activity, signal transducer activity, tumor necrosis factor receptor binding, calcium ion binding, protein binding.
STAT3	Signal Transducer and Activator of Transcription-3	DNA binding, protein dimerization activity, protein kinase binding, sequence-specific DNA binding transcription factor activity, signal transducer activity.

STAT4	Signal Transducer and Activator of Transcription-4	Calcium ion binding, double-stranded DNA binding, RNA polymerase II core promoter proximal region sequence-specific DNA binding, RNA polymerase II core promoter sequence-specific DNA binding transcription factor activity, signal transducer activity.
STAT5A	Signal Transducer and Activator of Transcription-5A	DNA binding, sequence-specific DNA binding transcription factor activity,signal transducer activity.
STAT5B	Signal Transducer and Activator of Transcription-5B	DNA binding, RNA polymerase II core promoter sequence-specific DNA binding, chromatin binding, glucocorticoid receptor binding, protein dimerization activity, protein phosphatase binding, sequence-specific DNA binding transcription factor activity, signal transducer activity.
STAT6	Signal Transducer and Activator of Transcription-6	Sequence-specific DNA binding transcription factor activity, sequence-specific DNA binding.

[1]Sources: http://www.ncbi.nlm.nih.gov/ http://www.ebi.ac.uk/ http://www.uniprot.org/. February 2015.

Supplementary table 3.The SNP within the PRL and GH-IGF1 pathways associated with 305 d milk yield for each re-randomized group of heat stressed lactating Holstein cows in Sonora, MX. used in a training and predicting exercise.

G[1]	Gene	Chr[2]	Location (Mb)	SNP	*P*	FDR[3]	Allele	Additive effect (Kg)
1	AVPR1A	5	50.5	rs210011420	0.03	0.05	**T**/A	308
	FURIN	21	22.5	rs381099643	0.02	0.05	**G**/A	316
	GHSR	1	95.7	rs110950555	0.03	0.05	**A**/G	75
	IGFBP5	2	105.3	rs208989155	0.0	0.05	**A**/G	696
	PAPPA1	8	107.1	rs379196319	0.02	0.05	**C**/A	331
	PMCH**	5	65.3	rs14197280	0.00	0.03	**A**/T	150
	PRLR**	20	39.1	rs135164815	0.01	0.05	**A**/G	187
	PRLR	20	39.1	rs136247583	0.03	0.05	**C**/T	187
	STAT5A	19	43	rs137182814	0.02	0.05	**G**/C	416
2	AVPR1A	5	50.5	rs207971189	0.01	0.05	**C**/G	354
	AVPR1A	5	50.5	rs210011420	0.01	0.05	**T**/C	323
	AVPR1A	5	50.5	rs209300854	0.03	0.05	**C**/G	296
	FURIN	21	22.5	rs382538054	0.05	0.05	**G**/A	27
	GHR	20	32	rs41639262	0.01	0.04	**A**/G	130
	IGFBP6	5	27	rs211039223	0.00	0.03	**C**/T	435
	PMCH**	5	65.3	rs14197280	0.05	0.05	**A**/T	3
	PRLR**	20	39.1	rs135164815	0.05	0.05	**A**/G	172
	SSTR5	25	0.85	rs109914110	0.01	0.04	**T**/C	1299
	STATA5B	19	42.9	rs384930401	0.03	0.05	**A**/G	1057
3	FURIN	21	22.5	rs381099643	0.05	0.05	**G**/A	279
	IGFIR	21	8.2	rs41960583	0.05	0.05	**T**/C	85
	IGFBP2	2	105.3	rs443442023	0.04	0.05	**G**/A	225
	IGFBP6	5	27	rs211039223	0.00	0.02	**C**/T	509
	PMCH**	5	65.3	rs14197280	0.00	0.02	**A**/T	232
	PRLR**	20	39.1	rs135164815	0.00	0.02	**A**/G	227
	PRLR	20	39.1	rs136247583	0.00	0.03	**C**/T	224

	STAT4	2	80	rs134874928	0.05	0.05	**T**/C	795
4	AVPR1A	5	50.5	rs207971189	0.02	0.05	**C**/G	362
	AVPR1A	5	50.5	rs210011420	0.03	0.05	**T**/C	313
	AVPR1A	5	50.5	rs209300854	0.04	0.05	**C**/G	304
	IGFBP6	5	27	rs211039223	0.01	0.05	**C**/T	409
	PAPPA2	16	59.3	rs109952914	0.03	0.05	**T**/A	108
	PMCH**	5	65.3	rs14197280	0.03	0.05	**A**/T	13
	PRLR**	20	39.1	rs135164815	0.04	0.05	**A**/G	141
	SSTR3	5	76	rs43438660	0.04	0.05	**G**/C	487
	SSTR3	5	76	rs43438659	0.03	0.05	**A**/T	508
5	IGFBP6	5	27	rs211039223	0.04	0.05	**T**/C	516
	OXTR	22	17.8	rs42002643	0.03	0.05	**A**/G	114
	PMCH**	5	65.3	rs14197280	0.00	0.03	**A**/T	27
	PRLR**	20	39.1	rs135164815	0.01	0.04	**A**/G	260
	PRLR	20	39.1	rs136247583	0.01	0.04	**C**/T	252
	SOCS1	25	9.9	rs210216882	0.05	0.05	**C**/T	408

Significance ($P \leq 0.05$) of SNP that were found positively associated with milk yield adjusted to 305 d of lactation and their favorable allele with expected effect. Favorable allele are bolded.[1]G = re-randomized group. [2]Chr = Chromosome.[3]FDR = false discovery rate. **= SNP within genes that were repeated in the five re-randomized groups.