PAMELA ITAJARA OTTO

ARCHITECTURE AND DISSECTION OF F2 POPULATION AND GIROLANDO BREED GENOME

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of Doctor Scientiae.

VIÇOSA
MINAS GERAIS - BRASIL
2019

Ficha catalográfica preparada pela Biblioteca Central da Universidade Federal de Viçosa - Câmpus Viçosa

T
O91a
2019

Otto, Pamela Itajara, 1990-
Architecture and dissection of F2 population and girolando breed genome / Pamela Itajara Otto. – Viçosa, MG, 2019.
vii, 196f. : il. (algumas color.) ; 29 cm.

Texto em inglês.
Orientador: Simone Eliza Facioni Guimarães.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Girolando (Bovino) - Melhoramento genético. 2. Alelos. 3. Estudo de Associação Genômica Ampla. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-Graduação em Zootecnia. II. Título.

CDD 22 ed. 636.2142

PAMELA ITAJARA OTTO

ARCHITECTURE AND DISSECTION OF F2 POPULATION AND GIROLANDO BREED GENOME

Thesis submitted to the Animal Science Graduate Program of the Universidade Federal de Viçosa, in partial fulfillment of the requirements for the degree of *Doctor Scientiae.*

APPROVED: February 27th, 2019.

Mario Calus

Lucas Lima Verardo

Marcos Vinicius G. Barbosa da Silva
(Co-advisor)

Joao Claudio Do Carmo Panetto

Simone Eliza Facioni Guimarães
(Advisor)

ACKNOWLEDGMENTS

Firstly, would like to thank God for all protection and strength that allowed me to conclude my studies.

Sincere thanks also to:

Especially to my mother Roseli Otto, who always cared to give a good education to her daughters, encouraging, being always present and doing everything with much love; and to my father Darci Otto, who has always been very picky, mainly related to the studies, showing the importance of the study for a good professional future and for allowing and helping me to achieve my goals;

My sisters Thatiana and Thatiara Otto, who always supported me and believed in me;

My boyfriend Diego Gobo, especially his support, for all love, patience, friendship and everything;

All my family for always cheer with my achievements;

To my adviser, Professor Simone Eliza Facioni Guimarães, for patience and orientation through all these years;

Universidade Federal de Viçosa and Animal Science Department for support and the opportunity of carrying out the course;

To CAPES and CNPq for financial support with scholarships;

Embrapa Dairy Cattle Research Center for the partnership and for providing the data needed to carry out this study;

My co-advisor Marco Antonio Machado for the help, incentive and confidence in my work;

My co-advisor Marcos Vinicius for patience, advices and help with the analyses;

Wageningen University & Research, for allowing me do part of my doctorate degree at the University, providing all support that I needed;

WUR Animal Breeding and Genetic Group for lessons, good discussions and support during my training period;

Mario Calus, my advisor during my training period at WUR, for all help, lessons, good discussions, tips and patience;

My friends from LABTEC (Animal Biotechnology Laboratory) for continuous knowledge, learning exchange and pleasant workplace.

TABLE OF CONTENTS

ABSTRACT

OTTO, Pamela Itajara, D.Sc., Universidade Federal de Viçosa, February, 2019. **Architecture and Dissection of F2 Population and Girolando Breed Genome.** Advisor: Simone Eliza Facioni Guimarães. Co-advisors: Marcos Vinícius Gualberto Barbosa da Silva and Marco Antonio Machado.

Livestock production is one of the most important activities of Brazilian agribusiness. Milk production has taken a prominent position in this sector, with continuous growth over the last decades. Most of the milk in Brazil is produced in systems relying on grazing mainly located in tropical and subtropical regions, which favors parasite infections in dairy cattle. Among the parasites that affect cattle, ticks are the main cause of economic losses in tropical and subtropical regions around the world. The bovine tick, Rhipicephalus (Boophilus) microplus, promotes several losses in bovine production. In parallel, gastrointestinal nematodes infections are the biggest problem on livestock production, especially in ruminants, in which the infections are usually subclinical, what makes difficult the diagnosis of this parasitosis and aggravates the losses caused by this parasite. Such parasite infections and also heat stress, bring, large losses in dairy industry. In these circumstances, an alternative to overcome parasite infections and heat stress in dairy production systems is the use of crossing or synthetic breeds containing some portion of Bos indicus genetics. To meet these requirements, the Girolando breed was created aiming to produce milk sustainably in tropical and subtropical regions. Despite of its relevance in dairy production, Girolando Breed still lack information about its genomic architecture and genomic data regarding important traits for production and adaptive. GWAS studies was the first step in our process to find causal variants in our work. Post-GWAS studies, such as ontology terms enrichment and gene-transcription factors networks, were used to identify most likely candidate genes. By the BOA approach we assigned breed-of-origin of alleles in the crossbred animals. In this context, our overall aims were 1) to identify genetic markers and haplotypes associated with tick counts, variation of rectal temperature and averages of last measurements for fecal egg count (AVFEC) in Gir x Holstein crossbred cattle, as well as for 305-days milk yield in Girolando breed; 2) to identify candidate genes for the four evaluated traits using the ontology terms enrichment and gene-transcription factors networks; 3) To infer the breed origin of marker alleles located on the candidate genes identified for adaptive traits in F2 population, to check if there is a prevalence of alleles from a certain breed in animals groups; 4) To investigate whether estimated SNP-

specific variance for 305MY is affected by the breed of origin. The results of this thesis showed the genetic variability for the evaluated traits, which makes the use of genetic selection promising for the improvement of them. Results from GWAS approaches pinpointed relevant QTL regions for four traits in several Bos taurus chromosomes. The post-GWAS analyses increased the biological understanding of evaluated traits and allowed us to identify the TREM1, TREM2, TREML1, CD83, MYO5A and PRSS16 candidate genes for tick counts, the LIF, OSM, TXNRD2 and DGCR8 candidate genes for variation of rectal temperature, and finally the LEP, CLOCK, CASR, LRRC4, DOCK1, SLC15A2 and SND1 candidate genes for the 305MY. Based in the literature review we identified the MYOD1 and CALCB candidate genes for AVFEC. According to the origin of alleles analysis we observed that most of resistant animals for tick infestation showed two alleles from Gir breed, while the susceptible ones showed alleles from Holstein. Holstein breed alleles could be associated to higher response to heat stress effects and to AVFEC resistance. For the seven candidate genes identified for 305MY in Girolando cattle, the SNP variances were different depending upon the breed-of-origin, while the alleles coming from the Gir breed-of-origin showed the highest genetic variance. The present study was shown a rich information resource to add more pieces to the intricate nature of Girolando Cattle, providing better biological comprehension of its nature and physiology of production under tropical conditions.

RESUMO

OTTO, Pamela Itajara, D.Sc., Universidade Federal de Viçosa, fevereiro de 2019. **Arquitetura e dissecação do genoma da população F2 e da raça Girolando.** Orientador: Simone Eliza Facioni Guimarães. Coorientadores: Marcos Vinícius Gualberto Barbosa da Silva e Marco Antonio Machado.

A produção pecuária é uma das atividades mais importantes do agronegócio brasileiro. A produção de leite assume posição de destaque neste setor, com crescimento contínuo nas últimas décadas. A maior parte do leite no Brasil é produzido em sistemas que dependem de pastagem, principalmente localizados em regiões tropicais e subtropicais, o que favorece infestações por parasitas em gado leiteiro. Entre os parasitas que afetam o gado, os carrapatos são a principal causa de perdas econômicas em regiões tropicais e subtropicais ao redor do mundo. O carrapato bovino, Rhipicephalus (Boophilus) microplus, promove diversas perdas na produção bovina. Paralelamente, infecções por nematódeos gastrintestinais são o maior problema na produção pecuária, especialmente em ruminantes, onde as infecções são geralmente subclínicas, o que dificulta o diagnóstico dessa parasitose e agrava as perdas causadas por esse parasita. Tais infestações parasitárias e também estresse térmico, trazem grandes perdas na indústria de laticínios. Nestas circunstâncias, uma alternativa para superar infestações por parasitas e estresse térmico em sistemas de produção de laticínios é o uso de cruzamentos ou raças sintéticas contendo alguma porção da genética de Bos indicus. Para atender a esses requisitos, foi criada a raça Girolando, que visa a produção leite de forma eficiente nas condições de produção das regiões tropicais e subtropicais. Apesar de sua relevância na produção leiteira, a Raça Girolando ainda carece de informações sobre sua arquitetura genômica e dados genômicos sobre características importantes para produção e adaptação. Os estudos de GWAS foram o primeiro passo em nosso processo para encontrar variantes causais para as características em questão. Estudos pós-GWAS, tais como enriquecimento de termos de ontologia e redes de fatores de transcrição genética, foram usados para identificar os genes candidatos mais prováveis. Por meio da abordagem BOA, atribuímos a raça de origem dos alelos nos animais cruzados. Nesse contexto, nossos objetivos gerais foram 1) identificar marcadores genéticos e haplótipos associados à contagem de carrapatos, variação da temperatura retal e às médias das últimas medições para contagem de ovos fecais (AVFEC) em bovinos mestiços Gir x Holandês, bem como para produção de leite até os 305 dias na raça Girolando; 2) identificar genes candidatos para as quatro características avaliadas utilizando as redes de

enriquecimento de termos de ontologia e fatores de transcrição genética; 3) Inferir a origem da raça de alelos marcadores localizados nos genes candidatos identificados para características adaptativas na população F2, verificando se há uma prevalência de alelos de uma determinada raça em grupos de animais; 4) Investigar se a variância específica estimada de SNPs para a produção de leite até os 305 dias (305MY) é afetada pela raça de origem. Os resultados desta tese mostraram a variabilidade genética para as características avaliadas, o que torna o uso da seleção genética promissor para a melhoria das mesmas. Resultados de abordagens de GWAS identificaram regiões de QTL relevantes para quatro características em vários cromossomos do Bos taurus. As análises pós-GWAS aumentaram o entendimento biológico das características avaliadas e nos permitiram identificar os genes candidatos TREM1, TREM2, TREML1, CD83, MYO5A e PRSS16 para contagem de carrapatos, os genes candidatos LIF, OSM, TXNRD2 e DGCR8 para variação da temperatura retal, e finalmente os genes candidatos LEP, CLOCK, CASR, LRRC4, DOCK1, SLC15A2 e SND1 para o 305MY. Com base na revisão de literatura, identificamos os genes candidatos MYOD1 e CALCB para AVFEC. De acordo com a análise da origem dos alelos, verificamos que a maioria dos animais resistentes para infestação de carrapatos apresentou dois alelos da raça Gir, enquanto os suscetíveis apresentaram alelos da raça Holandesa. Os alelos da raça Holandesa podem estar associados a uma maior resposta aos efeitos do estresse térmico e à resistência AVFEC. Para os sete genes candidatos identificados para 305MY no gado Girolando, as variâncias dos SNPs foram diferentes dependendo da raça de origem, indicando maior variância para os alelos provenientes da raça de origem Gir. O presente estudo mostrou-se um rico recurso de informação para agregar mais peças à natureza intrincada do gado Girolando, proporcionando melhor compreensão biológica de sua natureza e fisiologia de produção em condições tropicais.

CHAPTER 1

GENERAL INTRODUCTION

Girolando breed

Livestock production is one of the most important activities of Brazilian agribusiness. Milk production has taken a prominent position in this sector, with continuous growth over the last decades. In 2017, the milk activity in Brazil generated R$ 37,1 billion for national economy and the milk production was estimated in 33,5 billion liters, while the average production was 1,963 liters/cow, with an increase of 14,7% compared to 2016 (IBGE, 2018).

Brazil has a large climate diversity with most pastures located on the tropical and subtropical regions, which emphasizes the importance of selecting animals adapted to different Brazilian regions for the success of milk production (Ferraz and Felício, 2010). To meet these requirements, Gir cattle (**G**) was crossed with Holstein cattle (**H**) to generate the Girolando breed. This crossbreeding has been used to generate cows that combine the high milk production capacity of Holstein and the robustness, adaptability to high-temperature conditions and resistance to parasites of the Gir breed.

Girolando breed, which was officialized in February, 1996, has a genetic composition ranging from 1/4 H + 3/4 G to 7/8 H + 1/8 G. However, the crossings have been directed to establish the breed's genetic composition at 5/8 H + 3/8 G. Animals resulting from mating between 5/8 H + 3/8 G individuals are considered as Pure Synthetics (Freitas et al., 2002).

The fast growth of Girolando semen production in Brazil contributed to the genetic improvement of the breed, which has an average 305-day milk yield of 5,445 kg (Silva et al., 2018). Girolando cattle are notable for their excellent productivity, high fertility indexes

and good vigor in tropical and subtropical regions (Canaza-Cayo et al., 2017). Because of these advantages, cows crossed between Holstein and Gir cattle became the predominant animals on most of Brazilian dairy farms, being responsible for 80% of the milk produced in the country.

Tick infestations

Among the parasites that affect cattle, ticks are the main cause of economic losses in tropical and subtropical regions around the world (Molento et al., 2013). The bovine tick, Rhipicephalus (Boophilus) microplus, promotes several losses in bovine production due to weight loss, behavioral changes, secondary skin infections and transmission of viral, bacterial and protozoa pathogens that can be lethal to hosts (Léger et al., 2013).

Brazil has the largest commercial cattle herd in the world with more than 215.2 million head in 2015 (IBGE, 2015), and the losses due to tick burdens are approximately of US$ 3.24 billion dollars per year (Grisi et al., 2014). The production decrease corresponds to the amount of 51% of these losses, while 49% are associated with the parasite control (Jonsson et al., 2001). Furlong et al. (1996) observed decrease of 23% in milk production when Gir x Holstein animals were infested with an average of 105 ticks.

Most of the milk in Brazil is produced in systems relying on grazing mainly located in tropical regions (Ferraz and Felício, 2010), which favors tick infestations in dairy cattle. Therefore, R. microplus tick is considered as the main ectoparasite affecting cattle production in Brazil (Shyma et al., 2015). In addition to the environment conditions, the tick burdens can also be influenced by age, physical and nutritional status, breed and individual immunological pattern (Leite et al., 2010).

Tick control is routinely accomplished by acaricides, but their misuse leaves chemical residues in meat, milk and in the environment, besides contributing to the emergence of resistant tick populations. However, these drugs are still essential to control tick infestations. According to Amaral et al. (2011), the farmers have difficulty in understanding the parameters required for implementing strategic controls against cattle ticks, which results in incorrect application. The correct use of control strategy benefits consumers, milk industry and the environment, but mainly avoids selection and proliferation of resistant tick populations (Mendes et al., 2007).

Another alternative against tick infestations is the use of resistant animals. Bos indicus breeds are genetically more resistant to ticks than Bos taurus due to the presence of naturally selected genes throughout their evolutionary process (Ibelli et al., 2012; Ayres et al., 2015). Crosses between both genetic groups have been intensely used in the last years in order to generate animals with high milk production and resistant to parasite infestations and to climate conditions of tropical and subtropical regions (Bianchin et al., 2007).

There are two types of cattle resistant to tick infestation. Innate resistance is the body's first line of defense and can be linked to breed, while the acquired resistance comes from the contact of the host organism with the parasite (Riek, 1962). According to Maharana et al. (2011), Bos indicus cattle exhibits the strongest innate and acquired defenses, with greatest capacity of immune response against tick infestation.

Genetic variability for bovine tick resistance has been reported in literature (Machado et al., 2010; Cardoso et al., 2014), showing potential for the improvement of this trait. The evaluation of this adaptive trait allows the identification of the most tolerant animals and can be used as selection criteria in breeding programs.

Heat Stress

Most of Brazil's territory is located in tropical regions, with prevalence of high temperatures and grazing systems dependent on the rainy season (Cruz et al., 2011). Climate conditions may induce an increase in body temperature and heat stress, and consequently, large losses in the dairy industry.

Stress can be defined as a biological response or a set of reactions that defend the body from a threat against its homeostasis, which is essential for its survival (Azevêdo and Alves, 2009). Thermal stress is induced by cold or heat, depending on the critical temperature reached in the zone of thermoneutrality. Within this zone, the physiological requirement is minimal, there is no sweating, and the feed intake and respiration rate are normal. When the body temperature exceeds the critical upper temperature from the thermoneutral zone, heat stress is installed (West, 2003).

Among the factors influencing heat stress, there are solar radiation, relative humidity, environment temperature, age of animal, homeothermic capacity, breed, production capacity, nutritional condition and breeding system (Cunningham, 2008). Although the environment temperature is often considered as the most important climatic factor, relative humidity is another important factor for animals, that depend of evaporation mechanisms to dissipate the body heat (Pinheiro et al., 2015). The bioclimatic index usually used to evaluate the degree of heat stress affecting dairy cows is the temperature–humidity index (ITU), which represents the combined effects of air temperature and humidity (Armstrong, 1994; Hahn et al., 2003).

Dairy cattle, as any mammal, are homeothermic animals and have the ability to regulate body temperature regardless of the environment temperature. Some examples of thermoregulatory mechanisms are: conduction, which is the heat transfer by tissue contact

with solid or liquid surfaces; radiation, which is the loss of heat by the emission of heat rays; convection, that consists in the energy transfer from the hot bodies to cold bodies by air movement; and finally the evaporation mechanisms that work through sweating and respiration and are responsible for 80% of body heat loss to the environment (Shearer and Beede, 1990; West, 2003; Cunningham, 2008). The respiration rate is an important physiological parameter in the prediction of heat stress in dairy cattle (Mcdowell et al., 1975).

The animals' coat presents important roles on the evaporative heat loss mechanisms. As described by Collier et al. (2008), hair density is directly associated with number of sweat glands, and hair diameter and length regulate airflow at the skin surface. Moreover, hair coat color and hair coat density control the absorption of solar radiation.

Heat stress is one of the main factors affecting reproductive performance in dairy cattle. The adverse effects include a longer service period, negative hormone regulation leading to ovarian inactivity and decrease in conception rate (Dash et al., 2016). In milk production, animals of high genetic merit are the most affected, in which a reduction of 21% in the production was reported by Bouraoui et al. (2002). According to Spiers et al. (2004), milk yield decreases by 0.41 kg/cow/day for each increase of above 69 in THI unit. These losses can be caused by reduction of feed intake (Wheelock et al., 2010) and increase in metabolic maintenance requirements by 7 to 25% (NRC, 2001).

Bos indicus animals have a greater ability to regulate their body temperature since they have undergone natural selection for thousands of years under high temperatures in India. Part of this ability results from lower metabolic rates and higher capacity of heat loss through thermoregulatory mechanisms, which helps in heat stress prevention (Hansen, 2004). Among the thermoregulatory tools larger sweat glands, with a greater number of

layers of cells by epithelium and location close to the skin surface promote the better sweating capacity (Cattelam and Vale, 2013). In contrast, Bos taurus animals are more susceptible to the deleterious effects of heat stress (Renaudeau et al., 2012) due to the large amount of metabolic heat produced by high milk yield.

In tropical areas an alternative to overcome heat stress in dairy production systems is the use of crossing or synthetic breeds, both containing some portion of Bos indicus genetics, such as the Girolando breed which is derived from crossing Gir and Holstein animals.

Gastrointestinal nematodes infections

Nematodes are parasites that live inside the animal and represent the biggest problem on livestock production (Stear et al., 2007). Gastrointestinal nematodes (GIN) infections are present in all cattle herds in Brazil and the animals can be infected by one or more worm species at the same time (Costa and Lima, 2007).

In cattle, Cooperia spp. and Haemonchus spp. are responsible for more than 95% of the total parasitic load in the infection (Costa and Borges, 2010). The Haemonchus spp. is a hematophagous nematode and in its adult life, it consumes approximately 0.5 ml of blood per day. Infections with 10,000 of these worms are normal in cattle, which result in severe anemia (Fortes, 2004). Due to these circumstances, Haemonchus spp. is the most pathogenic nematode in the tropical regions. In addition to blood loss, the gastric lesions induce the decrease of hydrochloric acid and pepsinogen levels, resulting in inactivation of pepsin and interruption of protein digestion (Costa and Borges, 2010).

The GIN infections in ruminants are usually subclinical, which makes the diagnosis difficult and aggravates the losses caused by these parasites (Charlier et al., 2009). In

Brazil, the economic losses were estimated in approximately $ 7.11 billion (Grisi et al., 2014). Strategic controls can be used to deal with the infections and their negative effects. However, as in the tick control, the use of drugs is still essential, and their indiscriminate and incorrect use favors selection and proliferation of resistant populations (Souza et al., 2008).

A long period of host exposure to nematodes is necessary for the development of an effective level of immunological protection (Gasbarre et al., 2001). Young animals are more susceptible to worm infections because they do not have an efficient immune system yet (Nascimento et al., 2010). Furthermore, nutritional condition plays an important role in immunoglobulins production, which are immune cells essential for host defense (Brunetto et al., 2007). According to Gasbarre et al. (1990), the ability to induce immune response is also influenced by host genetic composition.

Fecal egg counts (FEC) per gram of feces is used to evaluate the degree of GIN infestations and can be used to classify the animals in: animal that never demonstrates high FEC values (Type I); animals with increased FEC values through the first 2 months on pasture, followed by decrease and maintenance of levels similar to the Type I animals (Type II); or animals that maintain high FEC levels (Type III). The approximate percentage of these animals in herd is 25:50:25, respectively (Gasbarre et al., 2001).

Bos indicus animals are more resistant to worm infection than Bos taurus, mainly to Haemonchus spp. and Cooperia spp. infections (O'Kelly, 1980). Silva et al. (2012) reported that Holstein animals showed higher significant FEC compared to purebred Gir and crossbred animals (1/2 Holstein x ½ Gir and 3/8 Holstein x 5/8 Gir). Neves et al. (2014) found the same results, in which Simental x Guzerá and Holstein x Gir crossbreed animals showed lower FEC compared to Bos taurus purebred cattle. Thus, crosses between

Bos taurus and Bos indicus breeds can be an effective alternative to decrease the use of drugs in GIN control in Brazilian cattle herds, contributing to control worm populations resistant to anthelmintic and to decrease the contamination of pastures and products of animal origin (Peña et al., 2000; Oliveira et al., 2009).

305-day milk yield

Brazil is one of the world's largest milk producers and part of this success is result of the implementation of dairy cattle breeding programs. In this context the use artificial insemination (AI) and the possibility of diluting a single ejaculate to create many progenies was essential to maximize the impact of elite bulls globally (Robertson and Rendel, 1947; Dejarnette et al., 2004). The selection of the best bulls based on the progeny testing results and, in recent years, using the genomic selection have allowed substantial increases in rates of genetic gain for milk yield in many countries (Hayes et al., 2009; Wiggans et al., 2017).

In Brazil, milk yield control is performed at 30-days intervals, in which individual cow production is measured and recorded in a 24-hour period (Pereira et al., 2010). These records are usually implemented by cooperatives, breeders associations or research centers, and represent the single source of data information to genetic evaluations for dairy cattle.

Several studies using the test-day milk yield data in cattle have been published (Bignardi et al., 2009, 2012; Pereira et al., 2013). However, aggregated test-day milk yield data representing total milk yield, normally standardized for a period of 305 days (305MY), have been used as the breeding goal for production in dairy cattle in the genetic breeding programs (Silva et al., 2017, 2018; Panetto et al., 2018). The accuracy of 305-d estimates depends on the number of tests and the used procedures (Swalve, 1995; Schaeffer and Jamrozik, 1996).

GWAS and Post-GWAS

Genome-wide association studies (GWAS) use high throughput genotyping technologies to assay the single-nucleotide polymorphisms (SNPs) and relate them to traits of interest (Pearson and Manolio, 2013). The classical GWAS is based on the single-marker analysis, which treats each SNP marker as a covariate in the model and tests its statistical association with the evaluated trait (Hirschhorn and Daly, 2005). Afterward, the estimated statistical significance value (P-value) can be used for identification of significant SNPs and visualization of GWAS results using quantile-quantile (Q-Q) and manhattan plots (Turner, 2014). However, this approach does not allow the direct inclusion of phenotypic information from non-genotyped individuals (Wang et al., 2012).

The single-step GWAS (ssGWAS) is a method that considers all markers simultaneously in the model and estimates SNP effects through genomic estimated breeding values (GEBV) resulted from single-step genomic best linear unbiased prediction (ssGBLUP) method (Aguilar et al., 2010). An advantage of the ssGBLUP is that it considers all phenotypic, genotypic and pedigree information (Wang et al., 2012) through the use of the H matrix, which combines pedigree and genomic relationships as in Aguilar et al. (2010). However, this approach does not allow the calculation of P-values (Wang et al., 2012).

GWAS is only the first step in the process to find causal variants, and further studies are needed to validate these results. Post-GWAS studies can be done by exploiting quantitative trait loci (QTL) regions found in GWAS through online database tools (i.e. NCBI, Emsembl) and by gene network analyses, which allow the identification of candidate genes related to these QTL regions and provide a better biological understanding of the evaluated traits. Among the Post-GWAS analyses, gene ontology terms enrichment

network can be used to select genes with enriched biological processes relevant to the evaluated trait. On the other hand, the selection of genes can also be done based on the gene-transcription factors networks, generated from enriched transcription factors (Verardo et al., 2016). Therefore, Post-GWAS analyses are a powerful approach to obtain the best of a GWAS analysis.

Assignment of Breed of Origin of Alleles

Recently, it was developed an approach that enables to assign breed-of-origin of alleles in crossbred animals, so-called BOA approach (Vandenplas et al., 2016). Applications of the BOA approach so far have been limited to 2- or 3-way crosses in pigs and poultry (Sevillano et al., 2016), in which this approach was applied in real data and on average the authors assigned 95.2% of the alleles a breed-of-origin using pedigree information. The BOA results can be used in genomic evaluation models that take breed-specific SNP effects into account, to improve genomic prediction, or GWAS (Sevillano et al., 2017, 2018). According to Sevillano et al. (2017), the inclusion of breed-specific SNP effect in the genomic evaluation model allowed, a better prediction for crossbred pigs performance. In GWAS, the parental breed of origin influenced the estimation of SNP effects and explained variance of SNPs associated with crossbred pigs performance (Sevillano et al., 2018). Therefore, the BOA approach is an important tool in crossbred animals genomic evaluation.

Genomic regions associated with tick, heat stress and nematodes resistance traits and 305-days milk yield in dairy cattle

In last years, the high-throughput genotyping and molecular techniques have provided a valuable source of information to study relationships between genotypes and phenotypes of interest in livestock (Pryce et al., 2010). The access to genotyping platforms with low cost has made the study of GWAS a common practice. These studies, based on molecular markers, can be used to identify genomic regions and genes associated with tick, heat stress and nematodes resistance traits and 305MY, becoming a powerful auxiliary tool in breeding programs.

Markers and QTL associated with cattle resistance traits and milk production in different chromosomes have been described in the literature (Ashwell et al., 2004; Coppieters et al., 2009; Turner et al., 2010; Machado et al., 2010; Silva et al., 2011; Meredith et al., 2012; Dikmen et al., 2013; Howard et al., 2013; Kim et al., 2015; Benavides et al., 2016; Mapholi et al., 2016). Among all genes of the immune system associated with the host resistance to pathogens, the major histocompatibility complex genes (MHC or bovine lymphocyte antigen - BoLA) are the most studied. Studies reported a significant relationship between MHC genes and tick counts (Acosta-Rodríguez et al., 2005; Martinez et al., 2006; Untalan et al., 2007). Turner et al. (2010) also related a few candidate genes for tick burden: TNFSF8 (CD30), SIRPA, CD47, SATB2, MAN2A1, ABCA9. In relation to candidate genes for heat stress traits, the following associations have been reported: ATP1A1A (Liu et al., 2011), HSP90AB1 (Charoensook et al., 2012), HSP70A (Deb et al., 2013) and PPARA (Fang et al., 2014). For the gastrointestinal infections, the ITGAE gene was described by Coppieters et al. (2009). Regarding milk production in cattle,

several genes have been described, among them: DGAT1 (Werner et al., 2002), ABCG2 (Cohen-zinder et al., 2005), GHR and PRLR (Blott et al., 2003; Viitala et al., 2006).

Objectives and outliers

Despite of its relevance in dairy production, Girolando Breed still lack information about its genomic architecture and genomic data regarding such important traits for production and survival. The current work comes to add more pieces to the intricate nature of Girolando Cattle, providing better understanding of its nature and physiology of production under tropical conditions. In Chapter 2, we aimed to present a full association study for tick resistant trait in a in F2 Bos taurus x Bos indicus crossbred cattle using the single-step GWAS. Moreover, we used the gene set to build the gene ontology terms enrichment and gene-transcription factors (TF) networks, generated from enriched TF, identified from the promoter sequences of selected gene sets. In addition, the BOA approach was used to assign the genetic origin of marker alleles located on candidate genes, which allowed us to check if there is a prevalence of alleles from a certain breed in resistant and susceptible animals. In Chapter 3, we also performed GWAS and post-GWAS for the heat stress response in F2 Bos taurus x Bos indicus crossbred cattle. For this study, we used the single-SNP GWAS to identify the significant SNPs. QTL regions were defined based on the location of the significant SNPs and the average linkage disequilibrium. The QTL's coordinates were used to obtain the related genes and generate ontology terms enrichment and gene-transcription factors (TF) networks, generated from enriched TF, to identify the most likely candidate genes. The origin of alleles of selected genes was assigned to help understanding the gene and inheritance involved in heat stress response.

In Chapter 4, we present a single-SNP GWAS for gastrointestinal nematodes trait in a F2 Bos taurus x Bos indicus crossbred cattle. The position of significant SNPs was used to identify the putative candidate genes. Moreover, we used the BOA analysis to infer the breed of origin of SNPs adjacent and located on the candidate genes in the resistant and susceptible animals, which allowed us to evaluate the gametic segregation of these alleles. In Chapter 5, we performed a single-step GBLUP procedure to identify QTL associated with 305MY in Girolando breed commercial population and the most likely candidate genes were identified through follow-up analyses. Genomic breeding values specific for Holstein and Gir were estimated for Girolando animals using a model that uses breed-specific partial relationship matrices and those were converted to SNP effects by breed-of-origin. In addition, differences between breeds-of-origin were evaluated by comparing estimated SNP variances between breeds. In Chapter 6 we aimed to summarize our results and put them together for a better understanding of our findings regarding the mosaic nature of Girolando genome.

REFERENCES

Acosta-Rodríguez, R., R. Alonso-Morales, S. Balladares, H. Flores-Aguilar, Z. García-Vazquez, and C. Gorodezky. 2005. Analysis of BoLA class II microsatellites in cattle infested with Boophilus microplus ticks: Class II is probably associated with susceptibility. Vet. Parasitol. 127:313–321. doi:10.1016/j.vetpar.2004.10.007.

Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta, and T.J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score.. J. Dairy Sci. 93:743–52. doi:10.3168/jds.2009-2730.

Amaral, M.A.Z. do, C.M.B.M. da Rocha, J.L. Faccini, J. control of cattle ticks : milk producers ' perceptionsn Furlong, C.M. de O. Monteiro, and M.C. de A. Prata. 2011. Strategic control of cattle ticks : milk producers ' perceptions. Rev. Bras. Parasitol. Veterinária 20:148–154.

Armstrong, D. V. 1994. Heat Stress Interaction with Shade and Cooling. J. Dairy Sci. 77:2044–2050. doi:10.3168/jds.S0022-0302(94)77149-6.

Ashwell, M.S., D.W. Heyen, T.S. Sonstegard, C.P. Van Tassell, Y. Da, P.M. VanRaden, M. Ron, J.I. Weller, and H.A. Lewin. 2004. Detection of Quantitative Trait Loci Affecting Milk Production, Health, and Reproductive Traits in Holstein Cattle. J. Dairy Sci. 0302:468–475. doi:10.3168/jds.S0022-0302(04)73186-0.

Ayres, D.R., R.J. Pereira, A.A. Boligon, F. Baldi, V.M. Roso, and L.G. Albuquerque. 2015. Genetic parameters and investigation of genotype × environment interactions in Nellore × Hereford crossbred for resistance to cattle ticks in different regions of Brazil. J. Appl. Genet. 56:107–113. doi:10.1007/s13353-014-0238-5.

Azevêdo, D.M.M.R., and A.A. Alves. 2009. Bioclimatologia Aplicada à Produção de Bovinos Leiteiros Nos Trópicos. 1st ed. Documentos / Embrapa Meio-Norte, ISSN 0104-866X, 188, Teresina.

Benavides, M.V., T.S. Sonstegard, and C. Van Tassell. 2016. Genomic Regions Associated with Sheep Resistance to Gastrointestinal Nematodes. Trends Parasitol. 32:470–480. doi:10.1016/j.pt.2016.03.007.

Bianchin, I., J.B. Catto, A.N. Kichel, R.A.A. Torres, and M.R. Honer. 2007. The effect of the control of endo- and ectoparasites on weight gains in crossbred cattle (Bos taurus taurus x Bos taurus indicus) in the central region of Brazil. Trop. Anim. Health Prod. 39:287–296. doi:10.1007/s11250-007-9017-1.

Bignardi, A.B., L. El Faro, V.L. Cardoso, P.F. Machado, and L.G. de Albuquerque. 2009. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. Livest. Sci. 123:1–7. doi:10.1016/j.livsci.2008.09.021.

Bignardi, A.B., L. El Faro, G.J.M. Rosa, V.L. Cardoso, P.F. Machado, and L.G. Albuquerque. 2012. Short communication : Principal components and factor analytic models for test-day milk yield in Brazilian Holstein cattle. J. Dairy Sci. 95:2157–2164. doi:10.3168/jds.2011-4494.

Blott, S., J. Kim, S. Moisio, A. Schmidt-ku, A. Cornet, P. Berzi, N. Cambisano, C. Ford, B. Grisart, D. Johnson, L. Karim, P. Simon, R. Snell, R. Spelman, J. Wong, J. Vilkki, W. Coppieters, and M. Georges. 2003. Molecular Dissection of a Quantitative Trait Locus : A Phenylalanine-to-Tyrosine Substitution in the Transmembrane Domain of the Bovine Growth Hormone Receptor Is Associated With a Major Effect on Milk Yield and Composition. Genet. Soc. Am. Mol. 266:253–266.

Bouraoui, R., M. Lahmar, A. Majdoub, M. Djemali, and R. Belyea. 2002. The relationship of temperature-humidity index with milk production of dairy cows in a Mediterranean climate. Anim. Res. 51:479–491. doi:10.1051/animres.

Brunetto, M.A., M.D.O.S. Gomes, J.T. Jeremias, L.D. Oliveira, and A.C. Carciofi. 2007. Imunonutrição: o papel da dieta no restabelecimento das defesas naturais. Acta Scien Vet. 35:230–232.

Canaza-Cayo, A.W., P.S. Lopes, J.A. Cobuci, M.F. Martins, and M.V.G.B. da Silva. 2017. Genetic parameters of milk production and reproduction traits of Girolando cattle in Brazil. Ital. J. Anim. Sci. 17:22–30. doi:10.1080/1828051X.2017.1335180.

Cardoso, C.P., B.F. Silva, D.S. Gonçalves, N.J. Tagliari, M.E. Saito, and A.F.T. Amarante.

2014. Resistência contra ectoparasitas em bovinos da raça crioula lageana e meio-sangue angus avaliada em condições naturais. Pesqui. Vet. Bras. 34:141–146. doi:10.1590/S0100-736X2014000200008.

Cattelam, J., and M.M. do Vale. 2013. Estresse térmico em bovinos. Rev. Port. Ciências Veterinárias 108:96–102.

Charlier, J., J. Höglund, G. von Samson-Himmelstjerna, P. Dorny, and J. Vercruysse. 2009. Gastrointestinal nematode infections in adult dairy cattle: Impact on production, diagnosis and control. Vet. Parasitol. 164:70–79. doi:10.1016/j.vetpar.2009.04.012.

Charoensook, R., K. Gatphayak, B. Brenig, and C. Knorr. 2012. Polymorphisms in the bovine HSP90AB1 gene are associated with heat tolerance in Thai indigenous cattle. Trop. Anim. Health Prod. 44:921–928. doi:10.1007/s11250-011-9989-8.

Cohen-zinder, M., E. Seroussi, D.M. Larkin, J.J. Loor, A.E. Der Wind, J. 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. doi:10.1101/gr.3806705.5.

Collier, R.J., J.L. Collier, R.P. Rhoads, and L.H. Baumgard. 2008. Invited review: genes involved in the bovine heat stress response.. J. Dairy Sci. 91:445–54. doi:10.3168/jds.2007-0540.

Coppieters, W., T.H.M. Mes, T. Druet, F. Farnir, N. Tamma, C. Schrooten, A.W.C. a Cornelissen, M. Georges, and H.W. Ploeger. 2009. Mapping QTL influencing gastrointestinal nematode burden in Dutch Holstein-Friesian dairy cattle.. BMC Genomics 10:96. doi:10.1186/1471-2164-10-96.

Costa, A.J., and F.A. Borges. 2010. Controle de endoparasitos em bovinos de corte.

FEALQ, ed. Piracicaba.

Costa, M.S.V.L.F., and W.D.S. Lima. 2007. Infecções Helmínticas Em Novilhas Gir Durante O Periparto. Rev. Ceres 54:182–184.

Cruz, L.V. da, D. de S.R. Angrimani, B.R. Rui, and M.A. da Silva. 2011. Efeitos do estresse térmico na produção leiteira: revisão de literatura. Rev. científica eletrônica Med. veterinária. doi:16797353.

Cunningham, J.G. 2008. Tratado de Fisiologia Veterinária. 4th ed. G. Koogan, ed. Rio de Janeiro.

Dash, S., A.K. Chakravarty, A. Singh, A. Upadhyay, M. Singh, and S. Yousuf. 2016. Effect of heat stress on reproductive performances of dairy cattle and buffaloes: A review. Vet. World 9:235–244. doi:10.14202/vetworld.2016.235-244.

Deb, R., B. Sajjanar, U. Singh, S. Kumar, M.P. Brahmane, R. Singh, G. Sengar, and A. Sharma. 2013. Promoter variants at AP2 box region of Hsp70 . 1 affect thermal stress response and milk production traits in Frieswal cross bred cattle. Gene 532:230–235. doi:10.1016/j.gene.2013.09.037.

Dejarnette, J.M., C.E. Marshall, R.W. Lenz, D.R. Monke, W.H. Ayars, and C.G. Sattler. 2004. Sustaining the Fertility of Artificially Inseminated Dairy Cattle : The Role of the Artificial Insemination Industry. J. Dairy Sci. 87:E93–E104. doi:10.3168/jds.S0022-0302(04)70065-X.

Dikmen, S., J.B. Cole, D.J. Null, and P.J. Hansen. 2013. Genome-Wide Association Mapping for Identification of Quantitative Trait Loci for Rectal Temperature during Heat Stress in Holstein Cattle. PLoS One 8:1–7. doi:10.1371/journal.pone.0069202.

Fang, W., J. He, J. Huang, and Z. Ju. 2014. Study on genetic variations of PPAR a gene and its effects on thermal tolerance in Chinese Holstein. Mol. Biol. Rep. 41:1273–1278.

doi:10.1007/s11033-013-2972-x.

Ferraz, J.B.S., and P.E. de Felício. 2010. Production systems – An example from Brazil. Meat Sci. 84:238–243. doi:10.1016/j.meatsci.2009.06.006.

Fortes, E. 2004. Parasitologia Veterinária. 4th ed. Ícone, ed. São Paulo.

Freitas, A.F. de, M.C. Durães, and C.R.A. Menezes. 2002. Girolando: Raça Tropical Desenvolvida No Brasil. Circular T. E.G. de Leite, ed. Juis de Fora, MG.

Furlong, J., F. Deresz, L.L. Matos, and M. V. Balbi. 1996. The effect of cattle tick Boophilus microplus (Acari: Ixodidae) infestation on feed intake and milk yield of Holstein × Zebu crossbred cows. Page in Proccedings of Congresso Panamericano de Veterinária, XV, Campo Grande.

Gasbarre, L.C., E.A. Leighton, and C.J. Davies. 1990. Genetic control of immunity to gastrointestinal nematodes of cattle. Vet. Parasitol. 37:257–272. doi:10.1016/0304-4017(90)90009-Z.

Gasbarre, L.C., E.A. Leighton, and T. Sonstegard. 2001. Role of the bovine immune system and genome in resistance to gastrointestinal nematodes. Vet. Parasitol. 98:51–64. doi:10.1016/S0304-4017(01)00423-X.

Grisi, L., R.C. Leite, J.R.D.S. Martins, A.T.M. De Barros, R. Andreotti, P.H.D. Cançado, A.A.P. De León, J.B. Pereira, and H.S. Villela. 2014. Reassessment of the potential economic impact of cattle parasites in Brazil.. Brazilian J. Vet. Parasitol. 23:150–156. doi:10.1590/S1984-29612014042.

Hahn, G.L., T.L. Mader, and R.A. Eigenberg. 2003. Perspective on development of thermal indices for animal studies and management. EAAP Tech. Ser. 7:31–44.

Hansen, P.J. 2004. Physiological and cellular adaptations of zebu cattle to thermal stress. Anim. Reprod. Sci. 82–83:349–360. doi:10.1016/j.anireprosci.2004.04.011.

Hayes, B.J., P.J. Bowman, A.J. Chamberlain, and M.E. Goddard. 2009. Invited review : Genomic selection in dairy cattle : Progress and challenges. J. Dairy Sci. 92:433–443. doi:10.3168/jds.2008-1646.

Hirschhorn, J.N., and M.J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6:95–108. doi:10.1038/nrg1521.

Howard, J.T., S.D. Kachman, W.M. Snelling, E.J. Pollak, D.C. Ciobanu, L.A. Kuehn, and M.L. Spangler. 2013. Beef cattle body temperature during climatic stress: A genome-wide association study. Int. J. Biometeorol. 58:1665–1672. doi:10.1007/s00484-013-0773-5.

Ibelli, A.M.G., A.R.B. Ribeiro, R. Giglioti, L.C.A. Regitano, M.M. Alencar, A.C.S. Chagas, A.L. Paço, H.N. Oliveira, J.M.S. Duarte, and M.C.S. Oliveira. 2012. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. Vet. Parasitol. 186:425–430. doi:10.1016/j.vetpar.2011.11.019.

IBGE. 2018. Produção Da Pecuária Municipal 2017. Diretoria de Pesquisas, Coordenação de Agropecuária, Pesquisa da Pecuária Municipal 2017, Rio de Janeiro.

Jonsson, N.N., R. Davis, and M. De Witt. 2001. An estimate of the economic effects of cattle tick (Boophilus microplus) infestation on Queensland dairy farms. Aust. Vet. J. 79:826–831. doi:10.1111/j.1751-0813.2001.tb10929.x.

Kim, E.S., T. S.Sonstegard, M.V.G.B. Da Silva, L.C. Gasbarre, and C.P. VanTassell. 2015. Genome-Wide scan of gastrointestinal nematode resistance in closed angus population selected for minimized influence of MHC. PLoS One 10:1–18. doi:10.1371/journal.pone.0119380.

Léger, E., G. Vourc'h, L. Vial, C. Chevillon, and K.D. McCoy. 2013. Changing

distributions of ticks: Causes and consequences. Exp. Appl. Acarol. 59:219–244. doi:10.1007/s10493-012-9615-0.

Leite, R.C., A.P. Cunha, A.C.P.P. Bello, L.N. Domingues, and E. Bastianetto. 2010. Controle de ectoparasitos em bovinocultura de corte. FEALQ, ed. Piracicaba.

Liu, Y., D. Li, H. Li, X. Zhou, and G. Wang. 2011. A novel SNP of the ATP1A1 gene is associated with heat tolerance traits in dairy cows. Mol. Biol. Rep. 38:83–88. doi:10.1007/s11033-010-0080-8.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

Maharana, B.R., R.K. Baithalu, and I.M. Allaie. 2011. Mechanism of immunity to Tick infestation in Livestock. Vet. World 4:131–135. doi:10.5455/vetworld.2011.131-135.

Mapholi, N.O., A. Maiwashe, O. Matika, V. Riggio, S.C. Bishop, M.D. MacNeil, C. Banga, J.F. Taylor, and K. Dzama. 2016. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick. Borne. Dis. 7:487–497. doi:10.1016/j.ttbdis.2016.02.005.

Martinez, M.L., M.A. Machado, C.S. Nascimento, M.V.G.B. Silva, R.L. Teodoro, J. Furlong, M.C.A. Prata, A.L. Campos, M.F.M. Guimarães, A.L.S. Azevedo, M.F.A. Pires, and R.S. Verneque. 2006. Association of BoLA-DRB3.2 alleles with tick (Boophilus microplus) resistance in cattle. Genet. Mol. Res. 5:513–524. doi:0211 [pii].

Mcdowell, R.E., N.W. Hooven, and J.K. Camoens. 1975. Effect of Climate on Performance

of Holsteins in First Lactation. J. Dairy Sci. 59:965–971. doi:10.3168/jds.S0022-0302(76)84305-6.

Mendes, M.C., J.R. Pereira, and A.P. Prado. 2007. Sensitivity of boophilus microplus (acari: ixodidae) to pyrethroids and organophosphate in farms in the vale do Paraíba region, São Paulo, Brazil. Arq. Inst. Biol. (Sao. Paulo). 74:81–85.

Meredith, B.K., F.J. Kearney, E.K. Finlay, D.G. Bradley, A.G. Fahey, D.P. Berry, and D.J. Lynn. 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC Genet. 13:1–11.

Molento, M.B., C.J. Veríssimo, A.T. Amarante, J.A. Van Wyk, A.C.S. Chagas, J. V De Araújo, and F.A. Borges. 2013. Alternativas Para O Controle De Nematoides Gastrintestinais De Pequenos Ruminantes. Arq. Inst. Biol 80:253–263.

Nascimento, G.A.M. do, S. Sossai, G.D. Carvalho, and V.G. Lelis. 2010. Controle do carrapato Rhipicephalus (Boophiluz) microplus e doenças trasnmitidas, por esse ectoparasita, em bovinos da cidade de viçosa, MG, e região. Pages 183–188 in Anais II SIMPAC.

Neves, J.H. das, N. Carvalho, L. Rinaldi, G. Cringoli, and A.F.T. Amarante. 2014. Diagnosis of anthelmintic resistance in cattle in Brazil: A comparison of different methodologies. Vet. Parasitol. 206:216–226. doi:10.1016/j.vetpar.2014.10.015.

NRC. 2001. Requirements of Dairy Cattle. 7th ed. NATIONAL ACADEMY PRESS, ed. Washington, D.C.

O'Kelly, J.C. 1980. Parasitism and blood composition in genetically different types of cattle grazing in a tropical environment. Vet. Parasitol. 6:381–390. doi:10.1016/0304-4017(80)90082-5.

Oliveira, M.C.S., M.M. Alencar, A.C.S. Chagas, R. Giglioti, and H.N. Oliveira. 2009.

Gastrointestinal nematode infection in beef cattle of different genetic groups in Brazil. Vet. Parasitol. 166:249–254. doi:10.1016/j.vetpar.2009.09.006.

Panetto, J.C. do C., M.V.G.B. da Silva, R. da S. Verneque, M.A. Machado, A.R. Fernandes, M.F. Martins, D.R. de L. Reis, W.A. Arbex, R. da S. Maciel, H.T. Ventura, and M.A. Pereira. 2018. Programa Nacional de Melhoramento Do Gir Leiteiro - Sumário Brasileiro de Touros - Resultado Do Teste de Progênie - Abril 2018. 1 (Documen. E.G. de Leite, ed. Juiz de Fora : Embrapa Gado de Leite.

Pearson, T.A., and T.A. Manolio. 2013. How to Interpret a Genome-wide Association Study. JAMA 299:1335–1344.

Peña, M.T., J.E. Miller, W. Wyatt, and M.T. Kearney. 2000. Differences in susceptibility to gastrointestinal nematode infection between Angus and Brangus cattle in south Louisiana. Vet. Parasitol. 89:51–61. doi:10.1016/S0304-4017(99)00229-0.

Pereira, R.J., A.B. Bignardi, L. El Faro, R.S. Verneque, A.E.V. Filho, and L.G. Albuquerque. 2013. Random regression models using Legendre polynomials or linear splines for test-day milk yield of dairy Gyr (Bos indicus) cattle. J. Dairy Sci. 96:565–574. doi:10.3168/jds.2011-5051.

Pereira, R.J., P.S. Lopes, R. da S. Verneque, M.L.S. Júnior, M.R. Lagrotta, and R. de A. Torres. 2010. Funções de covariância para produção de leite no dia do controle em bovinos Gir leiteiro. Pesqui. Agropecu. Bras. 45:1303–1311.

Pinheiro, C., E.P. Saraiva, C. Aparecida, S. Saraiva, V. De França, C. Fonseca, M. Elivania, V. Almeida, G. Gonçalves, M. Leal, C. Menezes, P. José, and R. Neto. 2015. Características anatomofisiológicas de adaptação de bovinos leiteiros ao ambiente tropical. Rev. Agropecuária Técnica 36:280–293.

Pryce, J.E., S. Bolormaa, A.J. Chamberlain, P.J. Bowman, K. Savin, M.E. Goddard, and

B.J. Hayes. 2010. A validated genome-wide association study in 2 dairy cattle breeds for milk production and fertility traits using variable length haplotypes. J. Dairy Sci. 93:3331–3345. doi:10.3168/jds.2009-2893.

Renaudeau, D., A. Collin, S. Yahav, V. de Basilio, J.L. Gourdine, and R.J. Collier. 2012. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. Animal 6:707–728. doi:10.1017/S1751731111002448.

Riek, R. 1962. Studies on the reactions of animals to infestation with ticks. VI. Resistance of cattle to infestation with the tick Boophilus microplus (Canestrini). Aust. J. Agric. Res. 13:532–550. doi:https://doi.org/10.1071/AR9620532.

Robertson, A., and J.M. Rendel. 1947. The use of progeny testing with artificial insemination in dairy cattle. J. Genet. 50:21–31.

Schaeffer, L.R., and J. Jamrozik. 1996. Multiple-Trait Prediction of Lactation Yields for Dairy Cows. J. Dairy Sci. 79:2044–2055. doi:10.3168/jds.S0022-0302(96)76578-5.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, and M.P.L. Calus. 2016. Empirical determination of breed-of-origin of alleles in three-breed cross pigs. Genet. Sel. Evol. 48:55. doi:10.1186/s12711-016-0234-9.

Shearer, J.K., and D.K. Beede. 1990. Thermoregulation and physiological responses of dairy cattle in hot weather. Agri-Practice 11:5–17. doi:0745-452X.

Shyma, K.P., J.P. Gupta, and V. Singh. 2015. Breeding strategies for tick resistance in tropical cattle : a sustainable approach for tick control. J. Parasit. Dis. 39:1–6. doi:10.1007/s12639-013-0294-5.

Silva, A.A., A.L.S. Azevedo, R.S. Verneque, K. Gasparini, M.G.C.D. Peixoto, M.V.G.B. da Silva, P.S. Lopes, S.E.F. Guimarães, and M.A. Machado. 2011. Quantitative trait loci affecting milk production traits on bovine chromosome 6 in zebuine Gyr breed. J. Dairy Sci. 94:971–980. doi:10.3168/jds.2009-2970.

Silva, J.B. Da, C.P. Rangel, B.D.A. Baêta, and A.H. Da Fonseca. 2012. Risk factors relating to helminth infections in cows during the peripartum. Rev. Bras. Parasitol. Veterinária 21:92–96. doi:10.1590/S1984-29612012000200004.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, and B.R. Coelho. 2017. Girolando Breed Genetic Improvement Program – Genetic Evaluation of Cows – June/2017. Juiz de Fora: Embrapa Gado de Leite.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Souza, A.P. De, C.I. Ramos, V. Bellato, A.A. Sartor, and C.A. Schelbauer. 2008. Resistência de helmintos gastrintestinais de bovinos a anti-helmínticos no Planalto Catarinense. Ciência Rural 38:1363–1367. doi:10.1590/S0103-84782008000500026.

Spiers, D.E.Ã., J.D. Sampson, and R.P. Rhoads. 2004. Use of physiological parameters to predict milk yield and feed intake in heat-stressed dairy cows. J. Therm. Biol. 29:759–764. doi:10.1016/j.jtherbio.2004.08.051.

Stear, M.J., M. Doligalska, and K. Donskow-Schmelter. 2007. Alternatives to anthelmintics for the control of nematodes in livestock. Parasitology 134:139–151. doi:10.1017/S0031182006001557.

Swalve, H.H. 1995. The Effect of Test Day Models on the Estimation of Genetic Parameters and Breeding Values for Dairy Yield Traits. J. Dairy Sci. 78:929–938. doi:10.3168/jds.S0022-0302(95)76708-X.

Turner, L.B., B.E. Harrison, R.J. Bunch, L.R.P. Neto, Y. Li, and W. Barendse. 2010. A genome-wide association study of tick burden and milk composition in cattle. Anim. Prod. Sci. 50:235–245. doi:10.1071/AN09135.

Turner, S.D. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv 3–4. doi:10.1086/519795.

Untalan, P.M., J.H. Pruett, and C.D. Steelman. 2007. Association of the bovine leukocyte antigen major histocompatibility complex class II DRB3*4401 allele with host resistance to the Lone Star tick, Amblyomma americanum. Vet. Parasitol. 145:190–195. doi:10.1016/j.vetpar.2006.12.003.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Verardo, L.L., M.S. Lopes, S. Wijga, O. Madsen, F.F. Silva, M.A.M. Groenen, E.F. Knol, P.S. Lopes, and S.E.F. Guimarães. 2016. After genome-wide association studies: Gene networks elucidating candidate genes divergences for number of teats across two pig populations. J. Anim. Sci. 94:1446–1458. doi:10.2527/jas.2015-9917.

Viitala, S., J. Szyda, S. Blott, N. Schulman, M. Lidauer, M. Georges, and J. Vilkki. 2006. The Role of the Bovine Growth Hormone Receptor and Prolactin Receptor Genes in

Milk, Fat and Protein Production in Finnish Ayrshire Dairy Cattle. Genet. Soc. Am. 173:2151–2164. doi:10.1534/genetics.105.046730.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, and W.M. Muir. 2012. Genome-wide association mapping including phenotypes from relatives without genotypes. Genet. Res. (Camb). 94:73–83. doi:10.1017/S0016672312000274.

Werner, F.A.O., S. Kollers, S. Kata, G. Durstewitz, A. Winter, W. Kra, J. Buitkamp, J.E. Womack, G. Thaller, and R. Fries. 2002. Association of a lysine-232¯alanine polymorphism in a bovine gene encoding acyl-CoA : diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. PNAS 99:9300–9305.

West, J.W. 2003. Effects of Heat-Stress on Production in Dairy Cattle. J. Dairy Sci. 86:2131–2144. doi:10.3168/jds.S0022-0302(03)73803-X.

Wheelock, J.B., R.P. Rhoads, M.J. Vanbaale, S.R. Sanders, and L.H. Baumgard. 2010. Effects of heat stress on energetic metabolism in lactating Holstein cows. J. Dairy Sci. 93:644–655. doi:10.3168/jds.2009-2295.

Wiggans, G.R., J.B. Cole, S.M. Hubbard, and T.S. Sonstegard. 2017. Genomic Selection in Dairy Cattle : The USDA Experience. Annu. Rev. Anim. Biosci. 5:309–327. doi:10.1146/annurev-animal-021815-111422.

CHAPTER 2

Genome Wide Association Studies for tick resistance in Bos taurus x Bos indicus crossbred cattle: a deeper look into this intricate mechanism [1]

Pamela I. Otto*, Simone E. F. Guimarães*, Lucas L. Verardo*, Ana Luísa S. Azevedo†, Jeremie Vandenplas‡, Aline C. C. Soares*, Claudia A. Sevillano‡§, Renata Veroneze*, Maria de Fatima A. Pires†, Célio de Freitas†, Márcia Cristina A. Prata†, John Furlong†, Rui S. Verneque†, Marta Fonseca Martins†, João Cláudio C. Panetto†, Wanessa A. Carvalho†, Diego O. R. Gobo*, Marcos Vinícius G. B. da Silva†, Marco A. Machado†[1]

*Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, Brazil; †EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, Brazil; ‡Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands; §Topigs Norsvin Research Center, 6640 AA Beuningen, the Netherlands

Journal of Dairy Science
Received November 30, 2017
Accepted May 29, 2018
doi: https://doi.org/10.3168/jds.2017-14223.

ABSTRACT

Rhipicephalus (Boophilus) microplus, is the main cattle ectoparasite in tropical areas. Gir x Holstein crossbred is well adapted to different production systems in Brazil. In this context, we performed genome wide association study (**GWAS**) and post-GWAS analyses for R. microplus resistance in an experimental Gir x Holstein F2 population. SNPs identified in GWAS were used to build gene networks, and to investigate the breed of origin for its alleles. Tick artificial infestations were performed during dry and rainy seasons. Illumina BovineSNP50 BeadChip and single-step BLUP procedure (ssGBLUP) was used for GWAS. Post-GWAS analyses were performed by gene ontology terms enrichment and gene-transcription factors (**TF**) networks, generated from enriched TF, identified from the promoter sequences of selected gene sets. Also, the genetic origin of marker alleles in the F2 population was assigned using the breed of origin of alleles (**BOA**) approach. Heritability estimates for tick counts were 0.40 ± 0.11 in the rainy season and 0.54 ± 0.11 in dry season. The top 10 "0.5 Mb-windows" with the highest percentage of genetic variance explained by SNP markers were found in chromosomes 10 and 23 for both dry and rainy seasons. Gene network analyses allowed the identification of genes involved with biological processes relevant to immune system functions (TREM1, TREM2 and CD83). Gene-transcription factors network allowed the identification of genes involved with immune functions (MYO5A, TREML1 and PRSS16). In resistant animals, the average proportion of animals showing significant SNPs with paternal and maternal alleles originated from Gir breed was 44.8% whereas the proportion of animals with both paternal and maternal alleles originated from Holstein breed was 11.3%. Susceptible animals showing both paternal and maternal alleles originated from Holstein breed represented 44.6% on average, whereas both paternal and maternal alleles originated from Gir breed

animals represented 9.3%. This study allowed us to identify candidate genes for tick resistance in Gir x Holstein crossbreds in both rainy and dry seasons. According to the origin of alleles analysis, we found that most animals classified as resistant showed two alleles from Gir breed, while the susceptible ones showed alleles from Holstein. Based on these results, the identified genes may be thoroughly investigated in additional experiments aiming to validate their effects on tick resistance phenotype in cattle.

Key words: breed of origin, gene network, genetic variance, Gir x Holstein crossbred

INTRODUCTION

Among the parasites that affect cattle, ticks are the main cause of economic losses in tropical and subtropical regions around the world (Molento et al., 2013). The bovine tick, Rhipicephalus (Boophilus) microplus, promotes several losses in bovine production due to weight loss, behavioral changes, secondary skin infections and transmission of viral, bacterial and protozoa pathogens that can be fatal to the host (Léger et al., 2013).

Indicine breeds are genetically more resistant to ticks than taurine due the presence of naturally selected genes throughout their evolutionary process (Ibelli et al., 2012; Ayres et al., 2015). Genetic variability for bovine tick resistance has been reported (Cardoso et al., 2014), evidencing a potential for the improvement of this trait. According to Walker (2011), individual ability of cattle to acquire tick immunity is hereditary. This author described that European animals showed moderate immunological capacity (approximately 85% of the larvae died due to the animals' immune defense), while indicine cattle exhibited high immune defense capacity (98% of dead larvae).

R. microplus tick is considered as the main ectoparasite affecting cattle production in Brazil, showing high prevalence in tropical and subtropical regions (Shyma et al., 2015). Due to its a large climatic diversities, the search for animals adapted to the different Brazilian regions is very important for the success of milk production. In this sense, Gir cattle (Bos indicus), a breed which is resistant to hot temperatures and tropical diseases, are crossed with Holstein cattle (Bos taurus), to generate the Girolando. Girolando is a synthetic breed which is 5/8 Holstein and 3/8 Gir, by definition. This breed has shown a great adaptation throughout Brazil and accounts for 80% of the milk produced in the country (Ferro et al., 2010). The evaluation of adaptive traits is of great importance to livestock production systems because it allows the identification of the most tolerant

animals to biotic and abiotic stress conditions. Subsequently, these traits can be used as selection criteria in breeding programs.

Genome wide association studies (**GWAS**) have been performed to identify genomic regions associated with resistance to bovine ticks (Gasparin et al., 2007; Machado et al., 2010; Porto Neto et al., 2011). Nevertheless, information about genes underlying this trait and their breed of origin are still scarce. Post-GWAS analyses, including gene networks and gene ontology enrichment analyses, as well as the identification of relevant transcription factors can contribute to a better understanding of the molecular mechanisms involved in the complex pathogen-host interaction during R. microplus infestation in cattle. The identification of the origin of alleles in crossbred animals (Vandenplas et al. 2016, Sevillano et al., 2016) is also important in this context.

Thus, we aimed to perform GWAS and Post-GWAS analyses for R. microplus resistance in cattle, using an experimental Gir x Holstein F2 population. SNPs that were identified in GWAS were used to build gene networks and to investigate the breed of origin for its alleles.

MATERIALS AND METHODS

Experimental Population and Phenotypic Data

Data were obtained from Embrapa Dairy Cattle Research Center, located in southeastern Brazil. An experimental population was generated and evaluated for R. microplus resistance as described by Machado et al. (2010). Tick artificial infestations were performed in each F2 animal during the dry (April to September) and rainy (October to March) seasons, from 2001 to 2007. After 21 days of infestation, the number of ticks was measured in each F2 animal. Animals were evaluated in contemporary groups aged

between 10 and 14 months. Additional traits that might affect the tick resistance evaluations were also assessed: coat color, coat thickness, and length, density and type of hair.

DNA Extraction and SNP Genotyping

Blood samples from parental, F1 and F2 populations were collected using vacuum tubes containing anticoagulant. DNA was extracted from white blood cells using a modified phenol/chloroform method (Machado et al., 2010). DNA quality and concentration were determined by NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA). Subsequently, all samples were genotyped with the Illumina BovineSNP50 BeadChip (Illumina, San Diego, California, USA).

Genotype quality control was implemented using the SNPStats package in R software. Samples with a call rate < 0.90 and heterozygosity of 3.0 standard deviations above or below than the observed mean were removed. For the quality control of mapped SNPs, only autosomal SNPs with call rate > 0.90 and minor allele frequency (**MAF**) > 0.03 were considered. After this quality control, the genotype file contained 40,283 SNP markers and 476 samples, including four Holstein bulls, 23 Gir cows, 65 F1 and 384 F2 animals. Missing genotypes were imputed using FImpute software (Version 2.2) to ensure better statistical analysis effectiveness (Sargolzaei et al., 2011).

Statistical Analyses

Because the tick count data did not follow normal distribution, they were normalized using a natural log transformation: ln (tick count + 1). The bivariate model used was:

$$\mathbf{y = X\beta + Wa + e}$$

where **y** is the observations vector of the transformed tick count in rainy and dry seasons; **β** is the vector of fixed effects (i.e., contemporary group, sex, coat color and hair type, and the covariates, such as age, coat thickness, length and hair density); **a** is the vector of random additive genetic effects, assuming $\mathbf{a} \sim N(0; G_0 \otimes \mathbf{H})$; where **H** is a matrix combining genomic and pedigree information, as proposed by Aguilar et al. (2010) and G_0 is the covariance matrix related to the additive genetic effect; **X** is the incidence matrix of fixed effects; **W** is the incidence matrix of random additive genetic effects; and **e** is the incidence vector of residual effect, assuming $\mathbf{e} \sim N(0; R_0 \otimes \mathbf{I})$, where **I** is the identity matrix and R_0 is the covariance matrix related to the residual effect.

Analyses for estimation of variance components and genetic parameters, as well as the GWAS, were performed using the single-step BLUP procedure (ssGBLUP) (Aguilar et al.; 2010), as implemented in AIREMLF90 from the BLUPF90 family of programs (http://nce.ads.uga.edu/).

In order to map the loci explaining genetic differences in the evaluated trait linkage disequilibrium, blocks were identified in the genome based on linkage disequilibrium, genetic variance was estimated in 0.5 Megabase (**Mb**) overlapping windows using PostGSf90 software (http://nce.ads.uga.edu/) (Aguilar et al., 2011). Percentage of additive genetic variance explained by window has been calculated based of genetic value of the each window of 0.5 Mb and the total additive genetic variance, as described by Wang et al. (2014). According to Habier et al. (2011), windows capture the QTL effect better than a single SNP, since several SNPs are in LD with a QTL, where each of these SNPs explains a part of the QTL effect, allowing more precise estimation of this effect in a given chromosomal region. For the Manhattan plots construction, we selected non-overlapping

windows which exhibit the highest genetic variance, using the mhtplot function of R software (R version 3.3.1) to illustrate the genetic variance explained by each window.

Gene Search and Gene Network Building

According to Wang et al. (2014), the top 10 windows based on genetic variance were selected for each evaluated season (dry and rainy) in order to identify candidate genes from the initial and final coordinates of the each selected windows. Gene database of the National Center for Biotechnology Information (**NCBI**) (http://www.ncbi.nlm.nih.gov) was used to identify candidate genes located in the selected regions based on the UMD 3.1 assembly of the bovine genome.

From the identified candidate gene sets (rainy and dry season), we excluded ncRNA genes and those remaining were used for biological processes enrichment analyses. For that, the ClueGO application in Cytoscape (Bindea et al., 2009) was used based on a unilateral hypergeometric test and Bonferroni correction, which simultaneously analyzes one or more sets of genes and searches for a Functional Gene Ontology (**GO**) term or pathways that establish relationships among genes.

Besides from the gene sets, promoter sequences (FASTA format) were obtained by considering 3000 base pairs upstream and 300 base pairs downstream from genes transcription start site (Verardo et al., 2016) based on the bovine genome. These data were used as input to the TFM-explorer program (http://bioinfo.lifl.fr/TFM/TFME/) for the identification of enriched TF. This program uses weighting matrices from JASPAR database (http://jaspar.binf.ku.dk/) (Sandelin et al., 2004) to detect all potential transcription factor binding sites (**TFBS**) of a set of gene sequences and search for over-represented TFBS. This program extracts significant groups (TFBS regions of the selected

gene sequences associated with a factor) by calculating a scoring function threshold which chosen to give a p-value equal to or less than 10^{-3} for each position in each sequence, as described in Touzet and Varré (2007).

The TF list was then analyzed in Cytoscape software (Shannon et al., 2003) using the Biological Networks Gene Ontology tool (**BiNGO**) (Maere et al., 2005) to determine significantly overrepresented functional gene ontology terms. Based on the over-represented biological processes in Biological Networks Gene Ontology tool and evidence of literature review related to the studied trait, we were able to identify the TF most related to tick resistance, allowing to generate a gene-TF interaction network.

Gene-TF network analysis was performed in Cytoscape software based on the number of TFBS (connections with the selected TF) in each gene to determine the strongest associations with tick counts what will aid in the identification of the most likely candidate genes.

Assignment of Breed of Origin of Alleles

In order to assign the genetic origin of marker alleles in the F2 population, the breed origin of alleles (**BOA**) approach (Vandenplas et al., 2016) was used with the parameter settings recommended by Sevillano et al. (2016). The BOA approach consists of three steps: (1) phasing of pure and crossed animal haplotypes; (2) determining the unique haplotypes for each pure breed; and (3) assigning the breed of origin to alleles observed in F2 animals based on the determined unique haplotypes.

The software AlphaPhase 1.1 (Hickey et al., 2011) was chosen for haplotype phasing. The phasing step was performed using pedigree, and during the second step of the BOA approach a relaxation factor ($\mathbf{f_r}$) of 20% was used, which represents the maximum

percentage of haplotype copies that can be observed in a population of a different breed. If the percentage of haplotype copies identified in a single breed is less than (100 - relaxation factor) %, the breed of origin of this haplotype is defined as unknown.

Crossing between B. taurus and B. indicus animals results in the formation of long chromosome segments in F2 crossed animals, from which are expected to have the same origin. These segments allow a clear classification of origin of the SNPs as taurine or indicine (Bolormaa et al., 2013).

After the origin of alleles analysis for the entire population, F2 animals were classified as resistant (10% of animals with the lowest genomic breeding values for the number of ticks), or susceptible (10% of animals with the highest genomic breeding values). In these selected samples, the adjacent SNPs close to genes in the networks were used to identify the origin of their alleles. The SNP alleles from each group were assigned to their breed of origin and paternity and classified into four types: father H (Holstein), father G (Gir), mother H (Holstein) and mother G (Gir). After this classification, the alleles of each SNP were quantified and classified according to the breed of origin, using the Chi-square test to compare the segregation between resistant and susceptible groups.

RESULTS AND DISCUSSION

Heritability estimates for tick counts were 0.40 ± 0.11 in the rainy season and 0.54 ± 0.11 in dry season. In other studies, heritability for tick resistance ranged from 0.09 to 0.37 (Prayaga et al., 2009; Machado et al., 2010; Turner et al., 2010). In a previous study using 180 microsatellite markers in this same F2 experimental population, Machado et al. (2010) estimated a lower heritability (0.21 ± 0.12). A total of 40,283 SNP markers were used, which allowed identifying a greater number of QTL regions for tick count, contributing to

the estimation of a higher heritability. The different statistical models applied in each experiment may also explain the different estimates of heritability in both studies. Moreover, Machado et al. (2010) analyzed the data using a single-trait model, so the correlation (0.36) between counts made in the same individual in two seasons was estimated using Pearson Product-Moment correlation coefficient. On the other hand, in this current study, we estimated the genetic correlation (0.96) between tick count in each season using bivariate analyses. Independently of the statistical methodologies and molecular tools used to assess animal genotypes, our results indicate the presence of genetic variability for tick counts. According to Turner et al. (2010), the selection for resistance would not affect the milk production in cattle. These authors identified a low correlation between the effects of alleles involved in milk production and tick counts in an association study using SNPs in Bos taurus animals from Australia. Prayaga et al. (2009) also reported low and non-significant genetic correlation between tick counts and several production traits, confirming that selection for this trait does not promote unfavorable effects on other economically important traits.

Based on GWAS results, we identified 7,327 windows with mean density of 9 SNPs per window (Figures 1 and 2). We selected 10 windows with the highest percentage of genetic variance explained distributed in chromosomes 10 and 23. Other studies, which evaluated tick resistance in cattle, also reported several QTLs on the same chromosomes (Machado et al., 2010; Turner et al., 2010; Mapholi et al., 2016). All windows identified in the rainy season were also found in the dry season, differing only by the genetic variance explained by each window (Table 1). The window with the highest genetic variance was located in BTA23 and was the same for both seasons, explaining 1.80% and 1.84% for the rainy and dry seasons, respectively. In the previous study performed by Machado et al.

(2010), the identified QTL showing the highest phenotypic variance was also found on chromosome 23 for both seasons. The other windows followed the same pattern of additive genetic variance between both seasons, a fact that might be due to the high genetic correlation (0.96) between the traits evaluated for dry and rainy seasons. The percentage of genetic variances explained by all selected windows is noticeable, 5.34% for each season, and suggests the existence of a strong relation between these genome regions and the resistance to bovine tick in this population.

From the windows selected in both seasons, 60 genes encoding proteins with different functions were annotated. To better understand the functions of these tick resistance genes, we performed a biological process analyses (Figure 3). In this GO network, we identified genes biological process enrichment involved in relevant immune functions, such as TREM1, TREM2 (window 6, Table 1) and CD83 (window 10, Table 1). Triggering receptor expressed on myeloid cells (TREM) are important regulators of immune response as they are potent amplifiers of inflammatory responses (Genua et al., 2014). TREM1 has been related to monocytes and neutrophils blood circulating, macrophages, dendritic cells (**DCs**) and activation of T-cells, CD86 and major histocompatibility complex (**MHC**) class II (Schenk et al., 2007; Chen et al., 2008; Sharif and Knapp, 2008). The expression of TREM2 was detected in DCs, monocytes and macrophages (Turnbull et al., 2006; Seno et al., 2009; Bosco et al., 2010). DCs, along with CC chemokine receptors 7 (**CCR7**), increase expression in dendritic cells of co-stimulatory molecule (B7) and MHC class I and II (Gao et al., 2017).

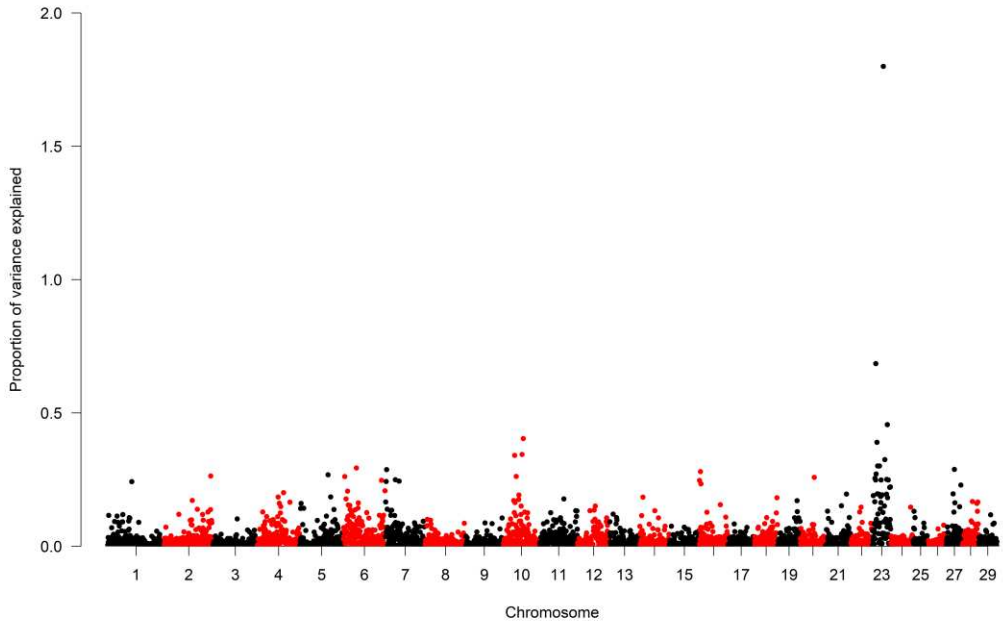


Figure 1. Manhattan Plot of additive genetic variance explained by non-overlapping windows in the experimental Holstein x Gir F2 population evaluated for cattle resistance to tick in the rainy season.

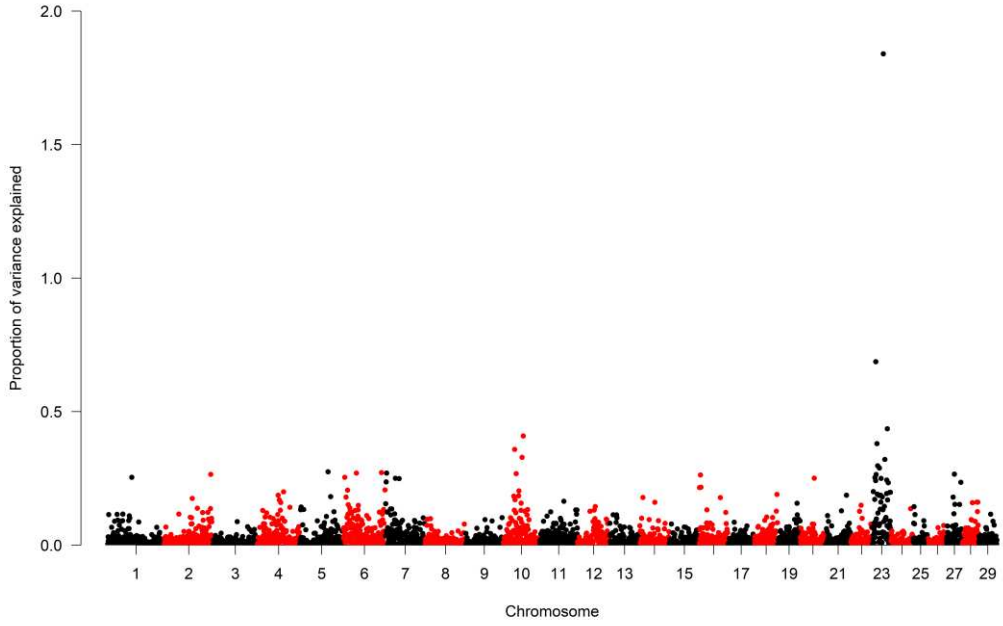


Figure 2. Manhattan Plot of additive genetic variance explained by non-overlapping windows in the experimental Holstein x Gir F2 population evaluated for cattle resistance to tick in the in the dry season.

Table 1. Top10 selected windows for percentage of additive genetic variance for the rainy and dry seasons in an experimental Holstein x Gir F2 population evaluated for cattle resistance to tick

Window	Chromosome	Start[1]	End[2]	Var. Exp. Rainy[3]	Var. Exp. Dry[4]
1	23	30806829	31286064	1.80	1.84
2	23	9712645	10167082	0.68	0.69
3	23	42206374	42669248	0.46	0.44
4	10	57588822	58054825	0.40	0.41
5	23	12997137	13484531	0.39	0.38
6	10	54341869	54787499	0.34	0.33
7	10	33343859	33774666	0.34	0.36
8	23	35156681	35642738	0.32	0.32
9	23	19915536	20402236	0.30	0.29
10	23	14992700	15465413	0.30	0.30

[1]Position in base pair of the first SNP in the window.

[2]Position in base pair of the last SNP in the window.

[3]Percentage of additive genetic variance explained by each window for rainy season.

[4]Percentage of additive genetic variance explained by each window for dry season.

The CD83 gene is an immunoglobulin superfamily member and is regulated during dendritic cells activation, being used to identify the maturation/activation of DCs in peripheral circulation (Breloer and Fleischer, 2008). CD83 is expressed in most immune cells such as B cells (Kretschmer et al., 2007), monocytes and macrophages (Cao et al., 2005), neutrophils (Yamashiro et al., 2000) and a regulatory subset of natural killers (**NK**) cells (Mailliard et al., 2005), and thus showing an essential role in the initiation and regulation of innate and adaptive immune responses. It acts in the stimulation of immature and memory T cells by activation of DCs (Aerts-Toegaert et al., 2007), thymic maturation and activation of $CD4^+$ and $CD8^+$ (Ju et al., 2016), maturation and maintenance of

peripheral lymphocyte B homeostasis (Lüthje et al., 2008) and in the increased expression of MHC class I and CD86 (Tze et al., 2011). These evidences suggest that this gene also has functions potentially involved with the physiological response to tick infestation in cattle, what makes it a candidate gene for this trait along with TREM1 and TREM2. Carvalho et al. (2014) identified potential tick resistance genes from this same family (CD14, CD5L, CD55, CD68, CD84, CD86) while analyzing gene expression microarrays in the first hours of tick infestation in the same experimental population.

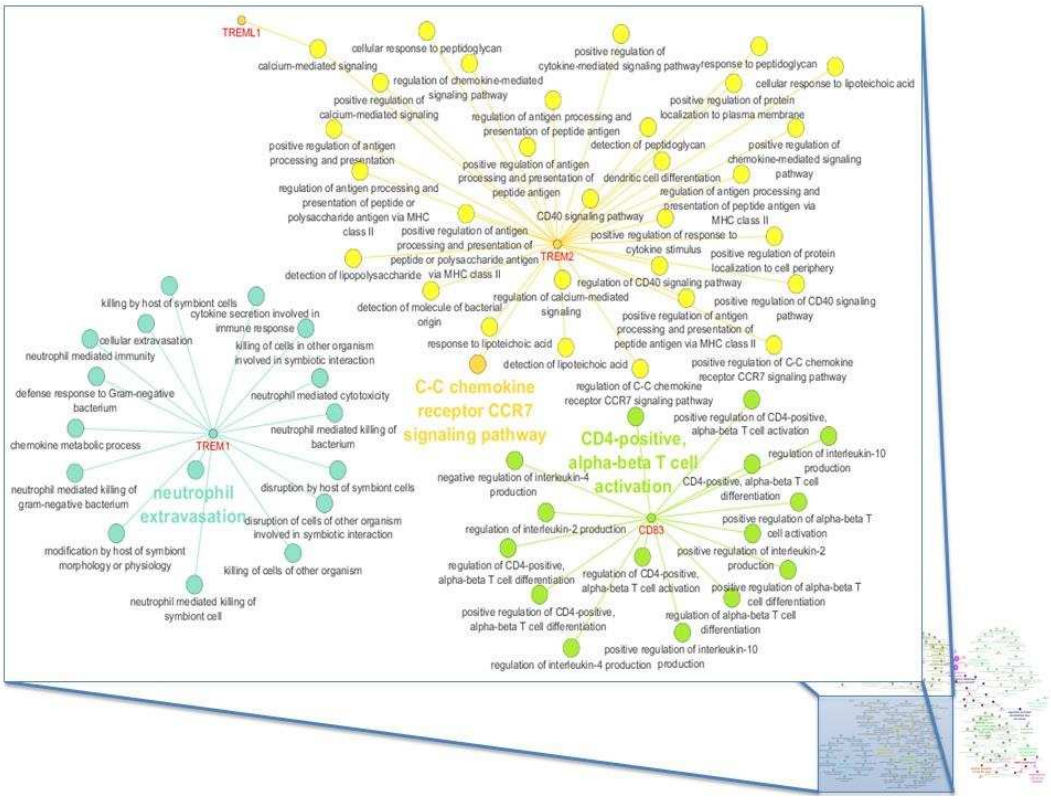


Figure 3. Functional networks showing gene interactions (red label) related to cattle tick resistance in dry and rainy seasons and the relationship across genes and their sub-networks related to immune system biological processes.

We performed in silico analyses of regulatory sequences for all detected genes and found 25 significantly enriched TF ($P < 0.0001$). The TF list was analyzed to determine which functional gene ontology terms were significantly overrepresented. Then, the TF

most associated with tick counts, based on biological processes and literature review (Table 2), were chosen to generate a gene-TF network (Figure 4), which allowed the identification of candidate genes for tick resistance.

Table 2. Enriched transcription factors (TF) associated with genes identified for cattle tick counts in rainy and dry season, based on their biological process and literature review. Data obtained in an experimental Holstein x Gir F2 population evaluated for bovine tick resistance

TF[1]	Biological Process	Literature review[2]
TAL1::TCF3	Immune System Processes	Lymphocyte B regulator (Aspland et al., 2001) and T cell development (Veer et al., 2014)
NFIL3	Immune System Processes	Adaptive immune lines and NK cell development (Seillet et al., 2014)
EGR1	Immune System Processes	Differentiation of thymocytes mediated by T cell pre-receptors (TCR) (Carter et al., 2007)
SOX10	Pigment cell differentiation	Proliferation and survival of melanocytes and melanogenesis (Fufa et al., 2015)
REL	Regulation of interleukin-12 biosynthesis processes	Key regulation of inducible gene expression in the immune system (Hayden et al., 2011; Herrington et al., 2016)
PAX5	Immune System Processes	Entry of common lymphoid progenitors (CLPs) into the B cell line (Kwon et al., 2008)
PBX1	Immune System Processes	Transformation of pre-B cells (Aspland et al., 2001)

[1]Transcription factor; TAL1 = T-Cell Acute Lymphocytic Leukemia 1; TCF3 = Transcription Factor 3; NFIL3 = Nuclear Factor, Interleukin 3 Regulated; EGR1 = Early Growth Response 1; SOX10 = SRY (Sex Determining Region Y)-Box 10; REL = REL proto-oncogene, NF-kB subunit; PAX5 = Paired Box 5; PBX1 = Pre-B-Cell Leukemia Transcription Factor 1

[2]The cited literature studies are just a sample from the vast available literature.

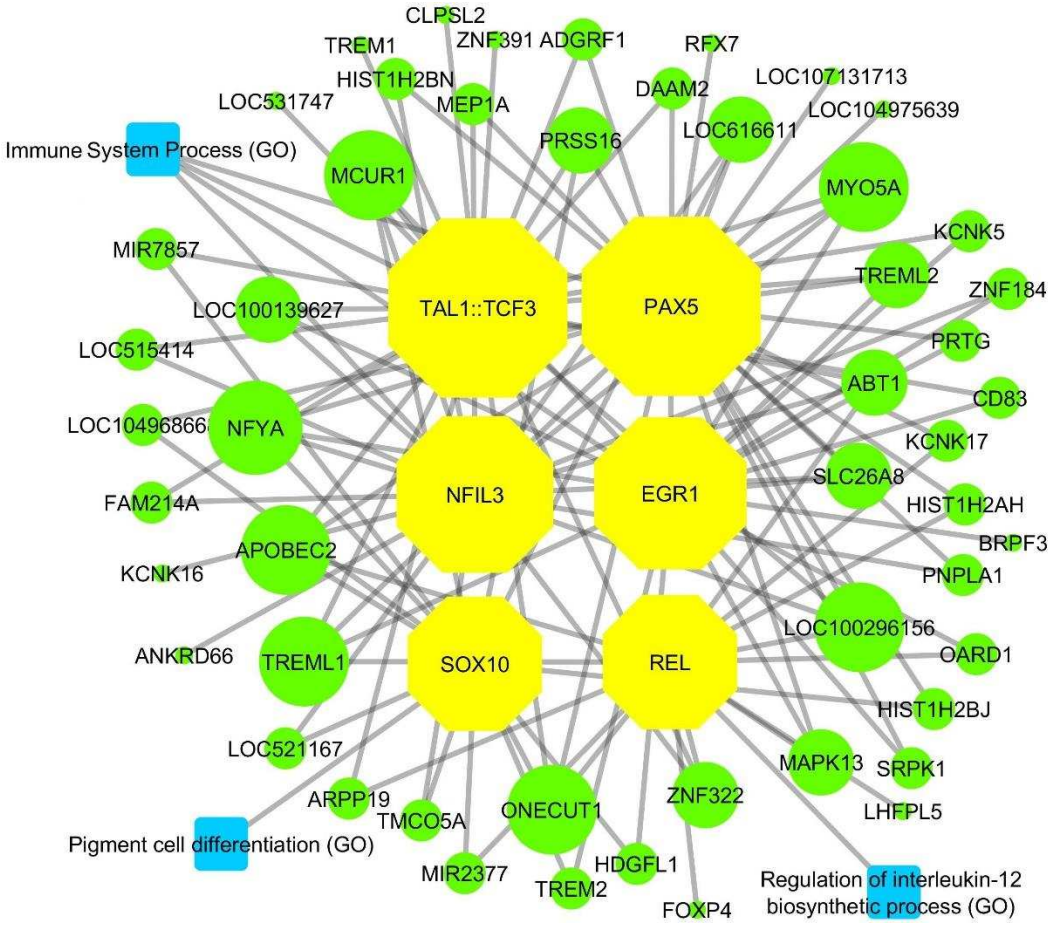


Figure 4. Gene-transcription factor (TF) network: genes located in the top 10 windows for tick counts in dry and rainy seasons (green circle nodes) and their associated TF (yellow octagon nodes). The node size corresponds to the network analyses (Cytoscape, Shannon et al., 2003) in which the larger nodes corresponds to a higher edge density associated with the number of TF binding sites. Blue square nodes are the biological processes related to TF (Gene Ontology Terms). Data obtained in an experimental Holstein x Gir F2 population evaluated for cattle tick resistance.

Among the analyzed TF in the network, two nodes were linked to a high number of genes (TAL1::TCF3 and PAX5). TCF3 transcription factor is a regulator of lymphopoiesis B and acts in the first detectable stages of B-cell involvement, development and maturation (Veer et al., 2014). It is also an important regulator of T cell development and its deficiency in mice leads to defects in the selection and differentiation of thymocytes (Aspland et al.,

2001). The TAL1 is the master regulator of hematopoietic stem/progenitor cells (**HSPCs**) in humans (Benyoucef et al., 2015) and it is essential to formation but not involved in the maintenance of hematopoietic stem cells in adults (Capron et al., 2006). PAX5 is involved in the entry of common lymphoid cells (**CLPs**) into B cell line. In the absence of TCF3, PAX5 is responsible for in vivo activation of B lymphocyte transcription program (Souroullas et al., 2009).

We also found four TF showing a high number of connections in the gene-TF network: NFIL3, EGR1, SOX10 and REL. NFIL3 is believed to be associated with the function of adaptive immune lines and natural killers cell development (Seillet et al., 2014). EGR1 (Carter et al., 2007), SOX10 (Fufa et al., 2015) and REL (Hayden and Ghosh, 2011; Herrington and Nibbs, 2016) were found to be associated with pre-receptor-mediated differentiation of thymocytes, proliferation and survival of melanocytes and melanogenesis, and key regulation of inducible gene expression in the immune system, respectively.

From the gene-TF network, candidate genes for tick counts were highlighted based on TFBS. The MYO5A gene (window 3, Table 1) showed connections to a set of TF (TALL3::TCF3, NFIL3, EGR1 and PAX5). This gene encodes the myosin Va protein found in melanocytes and, along with other proteins, assists in the melanosomes transport, which are structures that produce the melanin pigment responsible for pigmentation of skin, hair and eyes (Pastural et al., 2000; Hume and Seabra, 2011). Melanocytes play an important role in the innate immune response, controlling microbial infection and inflammation since these cells contribute to phagocytosis of invading pathogens. While presenting antigen to the competent immune cells, melanocytes produce key cytokines, such as Interleukin-1 Beta (**IL-1β**), IL6 and Tumor Necrosis Factor Alpha (**TNF-α**), as well as chemokines, which alert macrophages, neutrophils and T cells. Melanocytes can further express pattern

recognition receptors (**PRRs**), which detect specific pathogenic structures and induce the production of Type I Interferon and other immune system proteins (Gasque and Jaffar-Bandjee, 2015).

The TREML1 gene (window 6, Table 1) showed connection to TAL1::TCF3, REL, PAX5 and EGR1 TF. TREM-like transcript-1 (TLT-1/TREML1) belongs to the TREM family and is known to be selectively expressed in activated platelets and to facilitate platelet aggregation by binding to fibrinogen, acting as a chemotactic mediator in the attachment of thrombi to endothelium vascular (Esponda et al., 2015). Its soluble form (sTLT-1) participates in the regulation of inflammation during sepsis, attenuating leukocyte activity and modulating the interference between platelets and neutrophils (Morales et al., 2010; Derive et al., 2012).

The PRSS16 gene (window 8, Table 1), which did not share any ontology in the biological processes network (Figure 3), was related to the most enriched TF (TAL1::TCF3 and PAX5) and shows evidence to be associated with the immune system (Brisson et al., 2015). The importance of this gene is paramount in the positive regulation of $CD4^+$ T lymphocytes since it encodes the Thymus Specific Serine Protein (**TSSP**), which is predominantly expressed by the thymic epithelial cells in the cortex (**cTECs**) and is involved in the generation of peptides presented by class II MHC molecules in the thymus (Viret et al., 2011b; Fornari et al., 2012). Thymus Specific Serine Protein is required for the intracellular development of cells expressing T cell receptors (**TCR**), contributing to the diversification of T cell receptors repertoire of endogenous functional CD4 T cells in the thymus (Viret et al., 2011a). These genes are also involved in the organism defense against pathogens and injuries and thus may be involved in bovine resistance against ticks.

Among all genes of the immune system associated with the host resistance to pathogens, the major histocompatibility complex genes (MHC or bovine lymphocyte antigen - **BoLA**) are the most studied. Located on chromosome 23, these genes encode cell surface glycoproteins which work as receptors on antigen presenting cells, connecting and presenting antigenic peptides to T lymphocytes, thus triggering the onset of immune response in the host (Tizard, 2008). In the same position on chromosome 23, Machado et al. (2010) found highly significant QTLs associated with tick counts using the same experimental population of this current work. The association between MHC class II alleles and tick resistance was also reported by Martinez et al. (2006), which found an association between low number of ticks and gene-BoLA-DRB 3.2, alleles * 18, * 20 and * 27, suggesting that BoLA-DRB 3.2 alleles could be used to aid in the selection of tick-resistant animals. Other studies also reported a significant relationship between MHC genes and tick counts (Acosta-Rodríguez et al., 2005; Untalan et al., 2007). Most of the genes identified in this study (e.g. TREM1, TREM2, CD83 and PRSS16) show important roles associated with the activities of MHC class I and II, suggesting that these genes could be involved in the host immune response to tick infestation.

Based on the networks, we selected genes to assess the origin of alleles associated with bovine tick resistance. Considering the LD between SNPs and genes, the physical position of TREM1, TREM2, CD83, MYO5A, TREML1 and PRSS16 were annotated for adjacent SNPs selection. We selected a total of 11 SNPs: three SNPs in window three, four in window six, two in window eight and two in window 10 (Table 1). The animals were grouped as resistant or susceptible animals based on their genomic breeding values. Thus, the alleles of the selected SNPs were classified and quantified according to their breed of origin and paternity: HH for both paternal and maternal alleles originated from Holstein

breed, GG for both paternal and maternal alleles originated from Gir breed, HG for paternal allele originated from Holstein breed and maternal allele from Gir breed, and GH for paternal allele originated from Gir breed and maternal allele from Holstein breed. According to De Roos et al. (2009), the effects of SNPs on the phenotype depend on the LD between marker and QTL, but this LD can differ between Bos taurus and Bos indicus since they are distinct sub-species, which makes the search for the origin of alleles extremely important.

Based on Chi-square test applied to resistant and susceptible animal genotypes, we observed three significant SNPs ($P < 0.05$) in resistant animals (SNPs 8, 10 and 11) and nine in susceptible animals (SNPs 1 to 9) (Table 3), which showed a deviation of the expected gametic segregation of 25% for each genotypic class (1HH: 1GG: 1HG: 1GH). In resistant animals, the average proportion of animals for GG significant SNPs was 44.8% whereas animals classified as HH were represented by only 11.3%. On the other hand, in susceptible animals, HH represented 44.6% on average, whereas GG animals represented 9.3% of the animals in this group. Based on these findings, it was observed that most resistant animals inherited both Gir alleles whereas a small part received both Holstein alleles. Machado et al. (2010), evaluating the same experimental population of this study, identified a QTL with a negative additive effect on chromosome 5 from Holstein breed, showing that alleles from this susceptible breed also contributed to the increased resistance of the evaluated animals for tick resistance.

In the two evaluated groups, animals classified as HG and GH represented 43.9 and 46% of resistant and susceptible animals, respectively. Since we were not able to identify influence of the origin and paternity breed of the alleles in the SNPs evaluated in these two classes, the phenotypes could be influenced by other factors not detected by this study such

as the presence of regulatory genes. As tick resistance is a polygenic trait, other genes with lower and epistatic effects, not detected in this GWAS analysis, may also have influenced this phenotype.

Table 3. Percentage of individuals classified by breed of origin and paternity of each SNP allele analyzed in tick resistant and susceptible animals in an experimental Holstein x Gir F2 population

Alleles	Window 3			Window 6			Window 8		Window 10	
	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	SNP 6	SNP 7	SNP 8	SNP 9	SNP 10	SNP 11
Resistant Animals											
HH[1]	11.8	12.1	12.5	18.2	18.2	20.0	20.7	16.7*	17.1	8.6*	8.8*
GG[2]	38.2	39.4	37.5	39.4	39.4	33.3	34.5	44.4*	42.9	45.7*	44.1*
HG[3]	29.4	27.3	28.1	27.3	27.3	30.0	27.6	25.0*	25.7	31.4*	32.4*
GH[4]	20.6	21.2	21.9	15.2	15.2	16.7	17.2	13.9*	14.3	14.3*	14.7*
Susceptible Animals											
HH	42.1*	40.5*	42.1*	41.2*	41.9*	41.9*	43.8*	54.1*	54.1*	43.3	45.2
GG	10.5*	10.8*	10.5*	5.9*	6.5*	6.5*	6.3*	13.5*	13.5*	23.3	22.6
HG	31.6*	32.4*	31.6*	32.4*	29.0*	29.0*	28.1*	16.2*	16.2*	13.3	12.9
GH	15.8*	16.2*	15.8*	20.6*	22.6*	22.6*	21.9*	16.2*	16.2*	20.0	19.4

[1]Paternal and maternal alleles originated from Holstein breed.

[2]Paternal and maternal alleles originated from Gir breed.

[3]Paternal allele originated from Holstein breed and maternal from Gir breed.

[4]Paternal allele originated from Gir breed and maternal from Holstein breed.

* Significant SNPs for allele origin in the chi-square test ($P < 0.05$).

In summary, GWAS allowed us to identify candidate genes for tick resistance in an experimental F2 (Gir x Holstein) population. The network analysis highlighted gene interactions that were consistent with the biology of the immune system and provided known TF that allowed the identification of candidate genes for cattle tick resistance in both rainy and dry seasons (e.g. TREM1, TREM2, TREML1, CD83, MYO5A and PRSS16).

According to origin of allele analysis, we observed that most of the animals classified as resistant showed two alleles from Gir breed, while the susceptible animals showed alleles from Holstein breed. Based on these results, the identified genes may be deeply investigated in additional experiments aiming to validate their effects on tick resistance phenotype in cattle.

ACKNOWLEDGEMENTS

We thank EMBRAPA, FAPEMIG, CAPES, CNPq and MCTI/CNPq/INCT-CA for the financial support, Embrapa Dairy Cattle Research Center for providing the data needed to carry out this study and Mario P. L. Calus for his assistance in BOA methodology. The authors wish to thank and acknowledge all the team effort and the visionary work done by the late Dr. Mário Luiz Martinez, former leader of this project at Embrapa Dairy Cattle Research Center.

REFERENCES

Acosta-Rodríguez, R., R. Alonso-Morales, S. Balladares, H. Flores-Aguilar, Z. García-Vazquez, and C. Gorodezky. 2005. Analysis of BoLA class II microsatellites in cattle infested with Boophilus microplus ticks: Class II is probably associated with susceptibility. Vet. Parasitol. 127:313–321. doi:10.1016/j.vetpar.2004.10.007.

Aerts-Toegaert, C., C. Heirman, S. Tuyaerts, J. Corthals, J.L. Aerts, A. Bonehill, K. Thielemans, K. Breckpot, and K. Breckpot. 2007. CD83 expression on dendritic cells and T cells: Correlation with effective immune responses. Eur. J. Immunol. 37:686–695. doi:10.1002/eji.200636535.

Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta, and T.J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic

information for genetic evaluation of Holstein final score.. J. Dairy Sci. 93:743–52. doi:10.3168/jds.2009-2730.

Aguilar, I., I. Misztal, A. Legarra, and S. Tsuruta. 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. J. Anim. Breed. Genet. 128:422–428. doi:10.1111/j.1439-0388.2010.00912.x.

Aspland, S.E., H.H. Bendall, and C. Murre. 2001. The role of E2A-PBX1 in leukemogenesis.. Oncogene 20:5708–17. doi:10.1038/sj.onc.1204592.

Ayres, D.R., R.J. Pereira, A.A. Boligon, F. Baldi, V.M. Roso, and L.G. Albuquerque. 2015. Genetic parameters and investigation of genotype × environment interactions in Nellore × Hereford crossbred for resistance to cattle ticks in different regions of Brazil. J. Appl. Genet. 56:107–113. doi:10.1007/s13353-014-0238-5.

Benyoucef, A., J. Calvo, L. Renou, M.-L. Arcangeli, A. Van Den Heuvel, S. Amsellem, M. Mehrpour, J. Larghero, E. Soler, I. Naguibneva, and F. Pflumio. 2015. The SCL / TAL1 Transcription Factor Represses the Stress Protein DDiT4 / REDD1 in Human Hematopoietic Stem / Progenitor Cells. Stem Cells 2268–2279.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi:10.1093/bioinformatics/btp101.

Bolormaa, S., J.E. Pryce, K.E. Kemper, B.J. Hayes, Y. Zhang, B. Tier, W. Barendse, A. Reverter, and M.E. Goddard. 2013. Detection of quantitative trait loci in Bos indicus and Bos taurus cattle using genome-wide association studies.. Genet. Sel. Evol. 45:43. doi:10.1186/1297-9686-45-43.

Bosco, M.C., D. Pierobon, F. Blengio, F. Raggi, C. Vanni, M. Gattorno, A. Eva, F. Novelli, P. Cappello, M. Giovarelli, and L. Varesio. 2010. Hypoxia modulates the gene expression profile of immunoregulatory receptors in human mature dendritic cells: Identification of TREM-1 as a novel hypoxic marker in vitro and in vivo. Blood 117:2625–2639. doi:10.1182/blood-2010-06-292136.

Breloer, M., and B. Fleischer. 2008. CD83 regulates lymphocyte maturation, activation and homeostasis. Trends Immunol. 29:186–194. doi:10.1016/j.it.2008.01.009.

Cao, W., S.H. Lee, and J. Lu. 2005. CD83 is preformed inside monocytes, macrophages and dendritic cells, but it is only stably expressed on activated dendritic cells.. Biochem. J. 385:85–93. doi:10.1042/BJ20040741.

Capron, C., Y. Lécluse, A.L. Kaushik, A. Foudi, C. Lacout, D. Sekkai, I. Godin, O. Albagli, I. Poullion, F. Svinartchouk, E. Schanze, W. Vainchenker, F. Sablitzky, A. Bennaceur-Griscelli, and D. Duménil. 2006. The SCL relative LYL-1 is required for fetal and adult hematopoietic stem cell function and B-cell differentiation. Blood 107:4678–4686. doi:10.1182/blood-2005-08-3145.

Cardoso, C.P., B.F. Silva, D.S. Gonçalves, N.J. Tagliari, M.E. Saito, and A.F.T. Amarante. 2014. Resistência contra ectoparasitas em bovinos da raça crioula lageana e meio-sangue angus avaliada em condições naturais. Pesqui. Vet. Bras. 34:141–146. doi:10.1590/S0100-736X2014000200008.

Carter, J.H., J.M. Lefebvre, D.L. Wiest, and W.G. Tourtellotte. 2007. Redundant role for early growth response transcriptional regulators in thymocyte differentiation and survival.. J. Immunol. 178:6796–805.

Carvalho, W.A., R. Domingues, M.C. de Azevedo Prata, M.V.G.B. da Silva, G.C. de Oliveira, S.E.F. Guimarães, and M.A. Machado. 2014. Microarray analysis of tick-

infested skin in resistant and susceptible cattle confirms the role of inflammatory pathways in immune activation and larval rejection. Vet. Parasitol. 205:307–317. doi:10.1016/j.vetpar.2014.07.018.

Chen, L.C., J.D. Laskin, M.K. Gordon, and D.L. Laskin. 2008. Regulation of TREM Expression in Hepatic Macrophages and Endothelial Cells during Acute Endotoxemia. Exp. Mol. Pathol. 84:145–155. doi:10.1016/j.yexmp.2007.11.004.Regulation.

Derive, M., Y. Bouazza, N. Sennoun, S. Marchionni, L. Quigley, V. Washington, F. Massin, J.-P. Max, J. Ford, C. Alauzet, B. Levy, D.W. McVicar, and S. Gibot. 2012. Soluble Trem-like Transcript-1 Regulates Leukocyte Activation and Controls Microbial Sepsis. J. Immunol. 188:5585–5592. doi:10.4049/jimmunol.1102674.

Esponda, O.L., R. Hunter, J.R.R. Del Río, and A.V. Washington. 2015. Levels of Soluble TREM-Like Transcript 1 in Patients Presenting to the Emergency Department With Chest Pain. Clin Appl Thromb Hemost. 21:30–34. doi:10.1177/1076029614547298.Levels.

Ferro, F.R. de A., C.C.C. Neto, M. da R.T. Filho, S.T.S. Ferri, and Y.C. Montaldo. 2010. Efeito Do Estresse Calórico No Desempenho Reprodutivo Effect of Caloric Stress on Reproductive Performance of. Rev. Verde 5:1–25.

Fornari, T.A., M.M. Marques, C. Nguyen, A. Carrier, and G.A. Passos. 2012. Transcription profiling of Prss16 (Tssp) can be used to find additional peptidase genes that are candidates for self-peptide generation in the thymus. Mol. Biol. Rep. 39:4051–4058. doi:10.1007/s11033-011-1186-3.

Fufa, T.D., M.L. Harris, D.E. Watkins-Chow, D. Levy, D.U. Gorkin, D.E. Gildea, L. Song, A. Safi, G.E. Crawford, E. V. Sviderskaya, D.C. Bennett, A.S. Mccallion, S.K.

Loftus, and W.J. Pavan. 2015. Genomic analysis reveals distinct mechanisms and functional classes of SOX10-regulated genes in melanocytes. Hum. Mol. Genet. 24:5433–5450. doi:10.1093/hmg/ddv267.

Gao, S., L. Yuan, Y. Wang, and C. Hua. 2017. Enhanced expression of TREM-1 in splenic cDCs in lupus prone mice and it was modulated by miRNA-150. Mol. Immunol. 81:127–134. doi:10.1016/j.molimm.2016.12.006.

Gasparin, G., M. Miyata, L.L. Coutinho, M.L. Martinez, R.L. Teodoro, J. Furlong, M.A. Machado, M.V.G.B. Silva, T.S. Sonstegard, and L.C.A. Regitano. 2007. Mapping of quantitative trait loci controlling tick [Riphicephalus (Boophilus) microplus] resistance on bovine chromosomes 5, 7 and 14. Anim. Genet. 38:453–459. doi:10.1111/j.1365-2052.2007.01634.x.

Gasque, P., and M.C. Jaffar-Bandjee. 2015. The immunology and inflammatory responses of human melanocytes in infectious diseases. J. Infect. 71:413–421. doi:10.1016/j.jinf.2015.06.006.

Genua, M., S. Rutella, C. Correale, and S. Danese. 2014. The triggering receptor expressed on myeloid cells (TREM) in inflammatory bowel disease pathogenesis. J Transl. Med. 12:293. doi:10.1186/s12967-014-0293-z.

Hayden, M.S., and S. Ghosh. 2011. NF-κB in immunobiology. Cell Res. 21:223–244. doi:10.1038/cr.2011.13.

Herrington, F., and R. Nibbs. 2016. Regulation of the Adaptive Immune Response by the IκB Family Protein Bcl-3. Cells 5:14. doi:10.3390/cells5020014.

Hickey, J.M., B.P. Kinghorn, B. Tier, J.F. Wilson, N. Dunstan, and J.H.J. van der Werf. 2011. A combined long-range phasing and long haplotype imputation method to

impute phase for SNP genotypes.. Genet. Sel. Evol. 43:12. doi:10.1186/1297-9686-43-12.

Hume, A.N., and M.C. Seabra. 2011. Melanosomes on the move: a model to understand organelle dynamics. Biochem. Soc. Trans. 39:1191–1196. doi:10.1042/BST0391191.

Ibelli, A.M.G., A.R.B. Ribeiro, R. Giglioti, L.C.A. Regitano, M.M. Alencar, A.C.S. Chagas, A.L. Paço, H.N. Oliveira, J.M.S. Duarte, and M.C.S. Oliveira. 2012. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. Vet. Parasitol. 186:425–430. doi:10.1016/j.vetpar.2011.11.019.

Ju, X., P.A. Silveira, W.-H. Hsu, Z. Elgundi, R. Alingcastre, N.D. Verma, P.D. Fromm, J.L. Hsu, C. Bryant, Z. Li, F. Kupresanin, T.-H. Lo, C. Clarke, K. Lee, H. McGuire, B. Fazekas de St. Groth, S.R. Larsen, J. Gibson, K.F. Bradstock, G.J. Clark, and D.N.J. Hart. 2016. The Analysis of CD83 Expression on Human Immune Cells Identifies a Unique CD83 + -Activated T Cell Population. J. Immunol. 1600339. doi:10.4049/jimmunol.1600339.

Kretschmer, B., K. Lüthje, A.H. Guse, S. Ehrlich, F. Koch-Nolte, F. Haag, B. Fleischer, and M. Breloer. 2007. CD83 modulates B cell function in vitro. Increased IL-10 and reduced Ig secretion by CD83Tg B cells. PLoS One 2. doi:10.1371/journal.pone.0000755.

Léger, E., G. Vourc'h, L. Vial, C. Chevillon, and K.D. McCoy. 2013. Changing distributions of ticks: Causes and consequences. Exp. Appl. Acarol. 59:219–244. doi:10.1007/s10493-012-9615-0.

Lüthje, K., B. Kretschmer, B. Fleischer, and M. Breloer. 2008. CD83 regulates splenic B cell maturation and peripheral B cell homeostasis. Int. Immunol. 20:949–960. doi:10.1093/intimm/dxn054.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

Maere, S., K. Heymans, and M. Kuiper. 2005. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21:3448–3449. doi:10.1093/bioinformatics/bti551.

Mailliard, R.B., S.M. Alber, H. Shen, and S.C. Watkins. 2005. IL-18–induced CD83+CCR7+ NK helper cells. J. Exp. Med. 202:941–953. doi:10.1084/jem.20050128.

Mapholi, N.O., A. Maiwashe, O. Matika, V. Riggio, S.C. Bishop, M.D. MacNeil, C. Banga, J.F. Taylor, and K. Dzama. 2016. Genome-wide association study of tick resistance in South African Nguni cattle. Ticks Tick. Borne. Dis. 7:487–497. doi:10.1016/j.ttbdis.2016.02.005.

Martinez, M.L., M.A. Machado, C.S. Nascimento, M.V.G.B. Silva, R.L. Teodoro, J. Furlong, M.C.A. Prata, A.L. Campos, M.F.M. Guimarães, A.L.S. Azevedo, M.F.A. Pires, and R.S. Verneque. 2006. Association of BoLA-DRB3.2 alleles with tick (Boophilus microplus) resistance in cattle. Genet. Mol. Res. 5:513–524. doi:0211 [pii].

Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D.H. Lee. 2002. BLUPF90 and related programs (BGF90). Proc. 7th World Congr. Genet. Appl. to Livest. Prod. 28:21–22. doi:9782738010520.

Molento, M.B., C.J. Veríssimo, A.T. Amarante, J.A. Van Wyk, A.C.S. Chagas, J. V De Araújo, and F.A. Borges. 2013. Alternativas Para O Controle De Nematoides Gastrintestinais De Pequenos Ruminantes. Arq. Inst. Biol 80:253–263.

Morales, J., K. Villa, J. Gattis, W. Castro, K. Colon, J. Lubkowski, P. Sanabria, R. Hunter, and A.V. Washington. 2010. Soluble TLT-1 modulates platelet-endothelial cell interactions and actin polymerization.. Blood Coagul. fibrinolysis 21:229–236. doi:10.1097/MBC.0b013e3283358116.Soluble.

Pastural, E., F. Ersoy, N. Yalman, N. Wulffraat, E. Grillo, F. Ozkinay, I. Tezcan, G. Gediköglu, N. Philippe, A. Fischer, and G. de Saint Basile. 2000. Two genes are responsible for Griscelli syndrome at the same 15q21 locus.. Genomics 63:299–306. doi:10.1006/geno.1999.6081.

Porto Neto, L.R., N.N. Jonsson, M.J. D'Occhio, and W. Barendse. 2011. Molecular genetic approaches for identifying the basis of variation in resistance to tick infestation in cattle. Vet. Parasitol. 180:165–172. doi:10.1016/j.vetpar.2011.05.048.

Prayaga, K.C., N.J. Corbet, D.J. Johnston, M.L. Wolcott, G. Fordyce, and H.M. Burrow. 2009. Genetics of adaptive traits in heifers and their relationship to growth, pubertal and carcass traits in two tropical beef cattle genotypes. Anim. Prod. Sci. 49:413–425. doi:10.1071/EA08247.

De Roos, A.P.W., B.J. Hayes, and M.E. Goddard. 2009. Reliability of genomic predictions across multiple populations. Genetics 183:1545–1553. doi:10.1534/genetics.109.104935.

Sandelin, A., W. Alkema, P. Engström, W.W. Wasserman, and B. Lenhard. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles.. Nucleic Acids Res. 32:D91-4. doi:10.1093/nar/gkh012.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2011. FImpute – An efficient imputation algorithm for dairy cattle populations. J. Anim. Sci. 89:333.

Schenk, M., A. Bouchon, F. Seibold, and C. Mueller. 2007. TREM-1–expressing intestinal macrophages crucially amplify chronic inflammation in experimental colitis and inflammatory bowel diseases. J. Clin. Invest. 117. doi:10.1172/JCI30602DS1.

Seillet, C., L.C. Rankin, J.R. Groom, L.A. Mielke, J. Tellier, M. Chopin, N.D. Huntington, G.T. Belz, and S. Carotta. 2014. Nfil3 is required for the development of all innate lymphoid cell subsets. J Exp Med 211:1733–1740. doi:10.1084/jem.20140145.

Seno, H., H. Miyoshi, S.L. Brown, M.J. Geske, M. Colonna, and T.S. Stappenbeck. 2009. Efficient colonic mucosal wound repair requires Trem2 signaling.. Proc. Natl. Acad. Sci. U. S. A. 106:256–61. doi:10.1073/pnas.0803343106.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, and M.P.L. Calus. 2016. Empirical determination of breed-of-origin of alleles in three-breed cross pigs. Genet. Sel. Evol. 48:55. doi:10.1186/s12711-016-0234-9.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13:2498–2504. doi:10.1101/gr.1239303.metabolite.

Sharif, O., and S. Knapp. 2008. From expression to signaling: Roles of TREM-1 and TREM-2 in innate immunity and bacterial infection. Immunobiology 213:701–713. doi:10.1016/j.imbio.2008.07.008.

Shyma, K.P., J.P. Gupta, and V. Singh. 2015. Breeding strategies for tick resistance in tropical cattle : a sustainable approach for tick control. J. Parasit. Dis. 39:1–6. doi:10.1007/s12639-013-0294-5.

Souroullas, G.P., J.M. Salmon, F. Sablitzky, D.J. Curtis, and M.A. Goodell. 2009. Adult Hematopoietic Stem and Progenitor Cells Require Either Lyl1 or Scl for Survival. Cell Stem Cell 4:180–186. doi:10.1016/j.stem.2009.01.001.

Tizard, I.R. 2008. Imunologia Veterinária: Uma Introdução. 9th ed. Elsevier, ed. Rio de Janeiro.

Touzet, H., and J.-S. Varré. 2007. Efficient and accurate P-value computation for Position Weight Matrices.. Algorithms Mol. Biol. 2:15. doi:10.1186/1748-7188-2-15.

Turnbull, I.R., S. Gilfillan, M. Cella, T. Aoshi, M. Miller, L. Piccio, M. Hernandez, and M. Colonna. 2006. Cutting edge: TREM-2 attenuates macrophage activation.. J. Immunol. 177:3520–3524. doi:10.4049/jimmunol.177.6.3520.

Turner, L.B., B.E. Harrison, R.J. Bunch, L.R.P. Neto, Y. Li, and W. Barendse. 2010. A genome-wide association study of tick burden and milk composition in cattle. Anim. Prod. Sci. 50:235–245. doi:10.1071/AN09135.

Tze, L.E., K. Horikawa, H. Domaschenz, D.R. Howard, C.M. Roots, R.J. Rigby, D. a Way, M. Ohmura-Hoshino, S. Ishido, C.E. Andoniou, M. a Degli-Esposti, and C.C. Goodnow. 2011. CD83 increases MHC II and CD86 on dendritic cells by opposing IL-10-driven MARCH1-mediated ubiquitination and degradation.. J. Exp. Med. 208:149–165. doi:10.1084/jem.20092203.

Untalan, P.M., J.H. Pruett, and C.D. Steelman. 2007. Association of the bovine leukocyte antigen major histocompatibility complex class II DRB3*4401 allele with host

resistance to the Lone Star tick, Amblyomma americanum. Vet. Parasitol. 145:190–195. doi:10.1016/j.vetpar.2006.12.003.

Vandenplas1, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 2016 481 48:401–413. doi:10.1186/S12711-016-0240-Y.

Veer, A. van der, V. van der Velden, M. Willemse, P. Hoogeveen, E. Petricoin, H. Beverloo, G. Escherich, M. Horstmann, R. Pieters, and M. den Boer. 2014. Interference with pre-B-cell receptor signaling offers a therapeutic option for TCF3-rearranged childhood acute lymphoblastic leukemia.. Blood Cancer J. 4:e181. doi:10.1038/bcj.2014.5.

Verardo, L.L., F.F. Silva, M.S. Lopes, O. Madsen, J.W.M. Bastiaansen, E.F. Knol, M. Kelly, L. Varona, P.S. Lopes, and S.E.F. Guimarães. 2016. Revealing new candidate genes for reproductive traits in pigs : combining Bayesian GWAS and functional pathways. Genet. Sel. Evol. 48:1–13. doi:10.1186/s12711-016-0189-x.

Viret, C., C. Lamare, M. Guiraud, N. Fazilleau, A. Bour, B. Malissen, A. Carrier, and S. Guerder. 2011a. Thymus-specific serine protease contributes to the diversification of the functional endogenous CD4 T cell receptor repertoire.. J. Exp. Med. 208:3–11. doi:10.1084/jem.20100027.

Viret, C., S. Leung-Theung-Long, and L. Serre. 2011b. Thymus-specific serine protease controls autoreactive CD4 T cell development and autoimmune diabetes in mice. J. Clin. 121. doi:10.1172/JCI43314DS1.

Walker, A.R. 2011. Eradication and control of livestock ticks: biological, economic and social perspectives.. Parasitology 138:945–959. doi:10.1017/S0031182011000709.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, R.L. Fernando, Z. Vitezica, R. Okimoto, T. Wing, R. Hawken, and W.M. Muir. 2014. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front. Genet. 5:1–10. doi:10.3389/fgene.2014.00134.

Yamashiro, S., J.M. Wang, D. Yang, W.H. Gong, H. Kamohara, and T. Yoshimura. 2000. Expression of CCR6 and CD83 by cytokine-activated human neutrophils.. Blood 96:3958–3963.

CHAPTER 3

Genome wide association studies for heat stress response in Bos taurus x Bos indicus crossbred cattle

Pamela I. Otto[*], Simone E. F. Guimarães[*], Lucas L. Verardo[*], Ana Luísa S. Azevedo[†], Jeremie Vandenplas[‡], Claudia A. Sevillano[‡§], Daniele B. D. Marques[*], Maria de Fatima A. Pires[†], Célio de Freitas[†], Rui S. Verneque[†], Marta Fonseca Martins[†], João Cláudio C. Panetto[†], Wanessa A. Carvalho[†], Diego O. R. Gobo[*], Marcos Vinícius G. B. da Silva[†], Marco A. Machado[†]

[*]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [†]EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, 36038-330, Brazil; [‡]Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands; [§]Topigs Norsvin Research Center, 6640 AA Beuningen, the Netherlands

Journal of Dairy Science (accepted)

ABSTRACT

Heat stress is an important issue in global dairy industry. In tropical areas an alternative to overcome heat stress is the use of crossbred animals or synthetic breeds, such as the Girolando breed. In this study, we performed genome wide association study (GWAS) and post-GWAS analyses for heat stress in an experimental Gir x Holstein F2 population. Rectal temperature (RT) was measured in heat stressed F2 animals and the variation between two consecutive RT measurements (ΔRT) was used as dependent variable. Illumina BovineSNP50v1 BeadChip and single-SNP approach were used for GWAS. Post-GWAS analyses were performed by gene ontology terms enrichment and gene-transcription factors (TF) networks, generated from enriched TF. The breed origin of marker alleles in the F2 population was assigned using the breed-of-origin of alleles (BOA) approach. Heritability and repeatability estimates for ΔRT were 0.13 ± 0.08 and 0.29 ± 0.06, respectively. After the association analysis, six SNPs were significantly associated with ΔRT. Genes involved with biological processes in response to heat stress effects (LIF, OSM, TXNRD2 and DGCR8) were identified as putative candidate genes. After performing the BOA approach, 10% of F2 animals with the lowest breeding values for ΔRT were classified as Low-ΔRT, while the 10% with the highest breeding values for ΔRT were classified as High-ΔRT. On average, 49.4% of Low-ΔRT animals had two alleles from Holstein breed (HH), whereas 39% had both alleles from Gir breed (GG). On the other hand, in the High-ΔRT animals, the average proportion of animals for HH and GG were 1.4 and 50.2%, respectively. This study allowed the identification of candidate genes for ΔRT in Gir x Holstein crossbred animals. According to BOA approach, Holstein breed alleles could be associated to better response to heat stress effects, which could be explained by the fact that Holstein animals are more affected by heat stress than Gir

animals and so they require a genetic architecture to defend the body of the deleterious effects of heat stress. Future studies can provide further knowledge to uncover the genetic architecture underlying heat stress in crossbred cattle.

Key Words: crossbred cattle, gene network, heat stress, post-GWAS analyses

INTRODUCTION

Heat stress is an important factor in global dairy industry since it is responsible for large losses in milk production, growth, and reproductive performance, and negatively affects animal health and welfare (Nardone et al., 2010). Climate changes may induce increase in the average temperature and decrease in rainfall, leading to worsened environmental conditions to livestock production in tropical areas. Harsh environmental conditions are common in Brazil, where two-thirds of its territory is located in tropical regions, with prevalence of high temperatures due to solar radiation and grazing systems dependent on the rainy season (Cruz et al., 2011).

Basically, when animals cannot dissipate sufficient body heat to prevent a rise in their body temperature, the heat stress is installed (West, 2003). Modern high-producing dairy cows are more susceptible to the deleterious effects of heat stress (Renaudeau et al., 2012) due to their very high metabolic rate associated with milk production. Bos indicus animals have a greater ability to regulate their body temperature since they have undergone natural selection for thousands of years under high temperatures in India. Part of this ability is the result of lower metabolic rates and higher capacity of heat loss through thermoregulatory mechanisms, which helps in heat stress prevention (Hansen, 2004). In tropical areas an alternative to overcome heat stress in dairy production systems is the use of crossing or synthetic breeds containing some portion of Bos indicus genetics, such as the Girolando breed which is derived from crossing Gir and Holstein animals.

Respiration rate (**RR**) acts as a powerful thermoregulatory mechanism which helps maintaining the body temperature through evaporative cooling. This physiological parameter is important in the prediction of heat stress in dairy cattle (Mcdowell et al., 1975). The rectal temperature (**RT**) is an indicator of thermal equilibrium in which an

increase in this parameter expresses failure or exhaustion of the thermoregulatory mechanisms, resulting in heat stress (Ferrazza et al., 2017).

Molecular markers can help understanding heat stress adaptive parameters of dairy cattle which are relevant to production systems in tropical regions. This approach allows the identification of genomic regions and genes associated to heat stress response, becoming an additional attribute in the process of genetic selection of heat tolerant animals (Dikmen et al., 2015). Dikmen et al. (2013) found quantitative trait loci (**QTL**) for RT close to genes involved in cell protection against heat stress in Holstein and concluded that some SNPs may be useful in genetic selection and identification of genes controlling heat tolerance. In a subsequent study, Dikmen et al. (2015) identified specific genetic markers responsible for genetic variation of thermoregulation, showing the importance of genome wide association studies (**GWAS**) in finding markers and genes involved in physiological response to heat stress. However, studies about genes underlying RT, via gene-transcription factors (**TF**) networks and gene ontology enrichment analyses, in Bos indicus x Bos taurus animals have not been reported yet.

Recently, Vandenplas et al. (2016) developed an approach that enables to assign breed-of-origin of alleles (**BOA**) in crossbred animals. Results from this approach can be used to estimate SNP effects depending on the breed-of-origin of alleles and to improve genomic prediction or GWAS in crossbred animals. According to Sevillano et al. (2018), estimated effects and explained variance of SNPs strongly associated with crossbred performance are different depending upon from which parental breed they were inherited. Therefore, including breed-specific SNP effect in the genomic evaluation model allowed, in some cases, a better prediction for crossbred performance (Sevillano et al., 2017). The BOA approach was used by Otto et al. (2018) to evaluate the origin of marker alleles of the

candidate genes identified for tick resistance, in the same population used in the current study. Those authors observed that most of animals classified as resistant for tick infestation showed two alleles from the Gir breed, whereas the susceptible animals presented two alleles from the Holstein breed.

In this study, we performed a GWAS in a Gir x Holstein F2 experimental population to identify SNP associated with heat stress response and gene networks were used to identify the most likely candidate genes. The origin of alleles of selected genes was assigned to help understanding the gene mechanisms involved in heat stress response.

MATERIALS AND METHODS

Data

The data was obtained from a previous Embrapa research, when an experimental F2 population was produced by crossing four Holstein bulls with 27 Gir cows to generate 150 F1 (1/2 Gir : 1/2 Holstein) animals. Of this population, 65 F1 females were mated to four F1 males to generate 376 F2 animals and the pedigree file contains a total of 476 animals. All F2 animals were raised together in the Embrapa Dairy Cattle experimental station, located in the southeast of Brazil (Machado et al., 2010).

A total of 341 Gir x Holstein F2 animals were submitted to a heat chamber at 42 °C and 60% relative humidity (**RH**), after an adaptation period of 12 hours at 22 °C and 50% RH. Each evaluation included a total of six animals and during the whole period that the animals were housed into the heat chamber, water and food were not offered aiming at reducing their effects on the animals' responses. Parental Gir, Holstein and F1 animals were not submitted to heat stress.

The RT and RR of the F2 animals were evaluated in two replicates. First evaluation was done after the adaptation period, and the second evaluation was done six hours after the heat chamber reached 42 °C. To measure RT, a digital clinical thermometer was inserted approximately 7.5 cm into each animal's rectum. For RR measurement, flank movements were counted for 30 seconds. The difference between both rectal temperature measurements were calculated to estimate the variation of rectal temperature (**ΔRT**) in each animal by:

$$\Delta RT = RTA - RTB$$

where RTA is the rectal temperature after the heat stress and RTB is the rectal temperature before the heat stress (after the adaptation period). The variation of respiration rate (**ΔRR**) was defined similarly to ΔRT. Each animal was submitted twice to heat chamber to perform the measurement of ΔRT and ΔRR.

Animals were evaluated in contemporary groups aged from 10 and 14 months. Additional traits that might affect heat stress evaluations were assessed during this period: coat color, coat thickness and length, density and type of hair (Machado et al., 2010). Ancestry was estimated using Admixture software (Alexander et al., 2009) with a reduced panel (7,425 SNPs), pruned by linkage disequilibrium (**LD**) between subsequent markers.

All animals of the experimental population were genotyped with the Illumina BovineSNP50 v1 BeadChip (Illumina, San Diego, California, USA), including parental Gir, Holstein, F1 and F2 population. The genotype quality control was implemented using the SNPStats package in R software (Solé et al., 2006). Samples with a call rate < 0.90 and heterozygosity of 3.0 standard deviations above or below the observed mean were removed. For the quality control of mapped SNPs, only autosomal SNPs with call rate > 0.90 and

minor allele frequency (**MAF**) > 0.03 were considered. After quality control, the genotype file contained 40,283 SNPs markers.

Genome wide association analyses

A single-SNP GWAS was performed with the ASReml software (Gilmour et al., 2009) applying the following model:

$$\boldsymbol{y_{ijklmnop}} = \boldsymbol{\mu} + \boldsymbol{CG_i} + \boldsymbol{CC_j} + \boldsymbol{HT_k} + \boldsymbol{SNP_l} + \boldsymbol{S_m} + \boldsymbol{\beta_1}\left(\boldsymbol{CL_{ijklmnop}} - \overline{\boldsymbol{CL}}\right) + \boldsymbol{\beta_2}\left(\boldsymbol{CT_{ijklmnop}} - \overline{\boldsymbol{CT}}\right) + \boldsymbol{\beta_3}\left(\boldsymbol{HD_{ijklmnop}} - \overline{\boldsymbol{HD}}\right) + \boldsymbol{\beta_4}\left(\boldsymbol{VRR_{ijklmnop}} - \overline{\boldsymbol{VRR}}\right) + \boldsymbol{\beta_5}\left(\boldsymbol{BC_{ijklmnop}} - \overline{\boldsymbol{BC}}\right) + \boldsymbol{a_n} + \boldsymbol{pe_o} + \boldsymbol{e_{ijklmnop}}$$

where $\boldsymbol{y_{ijklmnop}}$ is the variation of rectal temperature (ΔRT); μ is the general mean; $\boldsymbol{CG_i}$ is the effect accounting for contemporary groups; $\boldsymbol{CC_j}$ is the effect accounting for coat color (j ranges from 1 to 4); $\boldsymbol{HT_k}$ is the effect accounting for hair type (**k** ranges from 1 to 4); $\boldsymbol{SNP_l}$ is the fixed effect of SNP genotype, where l was coded as 0, 1 or 2 copies of one of the alleles; $\boldsymbol{S_m}$ is the effect accounting for season (dry or rainy season); $\boldsymbol{CL_{ijklmnop}}$ is a covariate for the effect of coat length; $\boldsymbol{CT_{ijklmnop}}$ is a covariate for the effect of coat thickness; $\boldsymbol{HD_{ijklmnop}}$ is a covariate for the effect of hair density; $\boldsymbol{VRR_{ijklmnop}}$ is a covariate for the effect of ΔRR, calculated by the difference between both measurements; $\boldsymbol{BC_{ijklmnop}}$ is a covariate for the estimated indicine percent of the individual; $\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}, \boldsymbol{\beta_4}$ and $\boldsymbol{\beta_5}$ are the regression coefficients a_n is a random additive genetic effect of animal n, assuming $a \sim N(0, \mathbf{A}\sigma_a^2)$, with additive pedigree relationship matrix **A** and the additive genetic variance σ_a^2; pe_o is the random permanent environmental effects, assuming $pe \sim N\left(0, \mathbf{I}\sigma_{pe}^2\right)$, with identity matrix **I** and permanent variance σ_{pe}^2; $\boldsymbol{e_{ijklmno}}$ is the random

residual effect, assuming $e \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_e^2 is the residual variance. The variance components were estimated based on same model without the SNP effect.

Although the additional evaluated traits show correlation with RT control, in this work we selected ΔRT as the main trait. For this reason, we included these additional traits only as covariates in the model in order to look for significant SNP that could be only linked to heat stress response disregarding the effects of these additional traits on ΔRT. The descriptive statistics of phenotype and covariates are shown in Table 1.

Table 1. Descriptive statistics of phenotype and covariates included in the model to evaluate the variation of rectal temperature (ΔRT) in an experimental Gir x Holstein F2 population evaluated for heat stress response.

Variables	Parameters						
	Season[1]	Number	Mean	SD[2]	CV[3]	Minimum	Maximum
ΔRT	1	314	2.65	0.66	24.78	1.00	4.45
	2	339	2.01	0.57	28.22	0.55	3.40
Coat length	1	320	1.04	0.22	21.49	0.51	1.89
	2	334	1.81	0.38	21.07	0.79	3.28
Coat thickness	1	322	3.63	0.73	20.07	1.48	6.00
	2	341	4.68	1.15	24.58	1.73	9.89
Hair density	1	320	184.91	55.70	30.12	82.00	412.00
	2	332	178.35	56.64	31.76	67.67	375.00
ΔRR	1	316	104.88	19.93	19.01	58.00	154.00
	2	343	88.34	28.48	32.23	12.00	160.00
BC	-	369	0.48	0.05	10.26	0.31	0.66

[1]Season 1 = rainy, 2 = dry.

[2]Standard deviation.

[3]Coefficient of variation (%).

[4]Variation of respiration rate.

[5]Breed composition.

After the association analysis, the genome-wide false discovery rate (**FDR**) was applied to avoid false positives due to multiple testing. The R package qvalue (Dabney et al., 2014) was used to provide the p-values corrected for FDR (q-values) for the SNP association tests. Associations with a q-value ≤ 0.05 were considered significant.

A single-SNP GWAS for ΔRR trait was performed with the same model used to evaluate the ΔRT trait, in which, the ΔRT trait was included as covariate. Nevertheless, no significant SNP was found to be associated with ΔRR.

QTL regions

QTL regions were defined based on the location of the significant SNPs and the average linkage disequilibrium (r^2) in the evaluated population. The Haploview software (Barrett et al., 2005) was used to calculate the pairwise LD between markers. Since the significant SNPs identified in the GWAS analysis were located on the chromosome 17, we based on the LD of this chromosome to fix the QTL size. In addition, the LD between significant SNPs identified in the GWAS (Additional file 1 and 2) also was used to define the QTL regions. An average $r^2 \geq 0.15$ was found for SNPs within a distance of 500 Kb. Thus, all significant SNPs located within a region of 500 Kb were considered as belonging to the same QTL region.

Gene Search and Generation of Gene Networks

Putative candidate genes within the QTL regions were identified based on the UMD 3.1 assembly of the bovine genome (Zimin et al., 2009), using the National Center for Biotechnology Information (**NCBI**) (http://www.ncbi.nlm.nih.gov).

The gene set was used for gene ontology enrichment analysis via the ClueGO application in Cytoscape (Bindea et al., 2009) to generate the Gene Ontology (**GO**) network. This analysis was performed based on a unilateral hypergeometric test with Bonferroni correction and aims to analyze the gene set on the search for a functional gene ontology term or pathway that might show correlation with selected genes.

The same gene set was used to perform the analysis of regulatory sequences. For that, promoter sequences (FASTA format) were obtained by considering 3000 base pairs upstream and 300 base pairs downstream from genes transcription start site in the bovine genome (Soares et al., 2017). These data were screened for the identification of enriched TF using the TFM-explorer software (http://bioinfo.lifl.fr/TFM/TFME/), which uses weighting matrices from JASPAR database (http://jaspar.binf.ku.dk/) (Sandelin et al., 2004) to detect potential transcription factor binding sites (**TFBS**) of a set of gene sequences and search for over-represented TFBS. In addition, this software extracts significant groups - TFBS regions of the selected gene sequences associated with a factor - by calculating a score function threshold which was chosen to generate a p-value $\leq 10^{-3}$ for each position in each sequence, as described in Touzet and Varré (2007).

The TF list was then analyzed in Cytoscape software (Shannon et al., 2003) using the Biological Networks Gene Ontology tool (**BiNGO**) (Maere et al., 2005) to determine significantly overrepresented functional gene ontology terms. Based on the over-represented biological processes in Biological Networks Gene Ontology tool and evidences from literature data related to the investigated traits, we were able to identify the heat stress most related TF allowing to generate gene-TF interaction networks.

Gene-TF network analysis was performed in Cytoscape software (Shannon et al., 2003) based on the number of TFBS (connections with the selected TF) in each gene to

determine the associations with ΔRT that aided in the identification of the most likely candidate genes.

Assignment of Breed of Origin of Alleles

In order to assign the breed origin of marker alleles in the F2 population, the breed-of- origin of alleles approach (Vandenplas et al., 2016) was used with the parameter settings recommended by Sevillano et al. (2016). The BOA approach consists of three steps: (1) phasing the purebred and crossbred animals genotypes with AlphaPhase 1.1 software (Hickey et al., 2011). Haplotype phasing was performed using pedigree and nine combinations haplotypes length, in which each combination was run both considering "Offset" and "NotOffset" modes, allowing each allele to be considered 18 times through different haplotypes of variable length. "Offset" analyses shifts the beginning of each core to halfway along the first core, creating 50% overlaps between cores (Vandenplas et al., 2016); (2) determining the unique haplotypes for each pure breed. In this step, it is necessary that 80% of a haplotype's copies were required to be observed in a specific breed to assigning its breed of origin; and (3) assigning breed origin to each allele observed in F2 animals based on the determined unique haplotypes.

After performing the BOA approach, 10% of F2 animals with the lowest breeding values for ΔRT were designated as Low-ΔRT, while the 10% with the highest breeding values for ΔRT were designated as High-ΔRT. Considering the LD between SNPs and genes, the physical positions of candidate genes were annotated and used as coordinates for selection of adjacent SNPs, which were classified as HH (i.e., both alleles originated from Holstein breed), GG (i.e., both alleles originated from Gir breed), and HG/GH (i.e., one allele originated from Holstein and one from Gir breed), based on the results of the BOA

approach. Finally, the alleles of each selected and classified SNP were quantified to evaluate the gametic segregation and check if there is a prevalence of alleles from a certain breed when animals are classified as Low-ΔRT or High-ΔRT. For that, a Chi-square test was used to evaluates the expected gametic segregation of 25% HH, 50% HG/GH and 25% GG in the both groups. SNPs with a difference in prevalence of alleles from a certain breed, which did not show the expected gametic segregation, was considered significant ($P \leq 0.05$).

RESULTS AND DISCUSSION

Genetic parameters

We found heritability estimates of 0.13 ± 0.08 and repeatability estimates of 0.29 ± 0.06 for ΔRT trait. During the heat stress evaluation period, the RT was evaluated at 42 °C and 60% RH, in which the Temperature Humidity Index (**THI**) was 97. Dikmen et al. (2012) estimated heritability of 0.17 ± 0.13 for RT recorded between 1,500 and 1,700 h and different THI in Holstein cattle during summer in north Florida. In beef cattle, heritability estimates of 0.22 for RT, recorded only during summer months when the ambient temperature was > 30 ºC, were found in Brahman animals and estimates of 0.14 in crossbred animals showing different tropical breeds genetic background (Porto-Neto et al., 2014)

Most studies use the Temperature Humidity Index (**THI**) to estimate the presence or absence of heat stress phenotype in a given environment (Biffani et al., 2016; Hagiya et al., 2017). However, evaluation of the response of each animal individually during the heat stress is relevant, as animals respond in different ways to environmental stimuli, which is

even more accentuated in the F2 crossbred animals used in this current study, since each animal possesses a unique genetic composition.

QTL regions

In the current study, the GWAS for ΔRT identified six significant SNPs ($q \leq 0.05$) in the F2 population (Figure 1, Table 2). These SNPs were distributed over three QTL regions located on Bos taurus chromosome 17 (BTA17). These QTL regions were defined based on the average LD between markers at the BTA17 and the LD between significant SNPs.

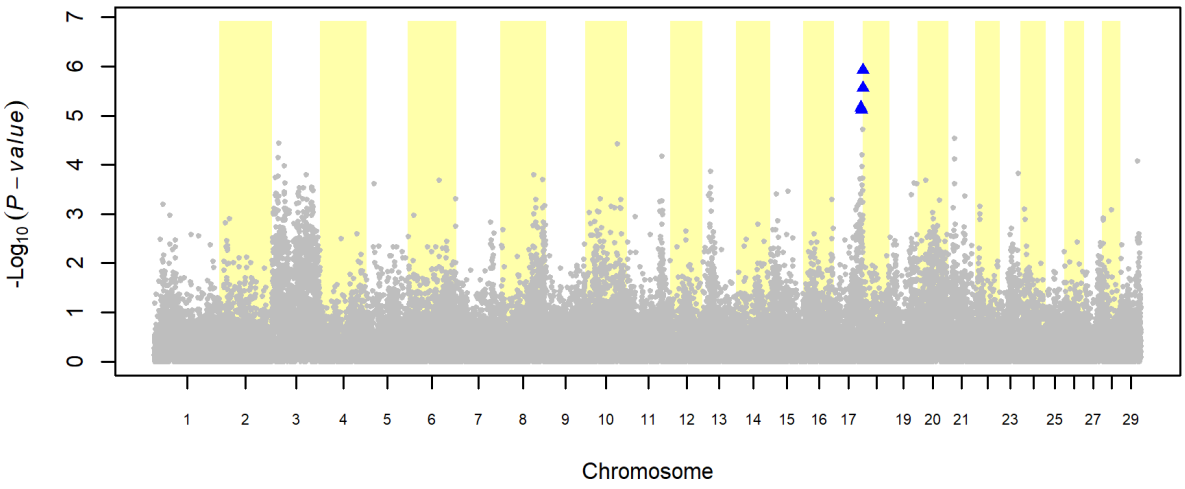


Figure 1. Manhattan plot from the association analysis of variation of rectal temperature (ΔRT) in the experimental Gir x Holstein F2 population. The significant SNPs (q-value ≤ 0.05) are shown as blue triangles and the gray dots denote SNPs which were not significant in GWAS.

Several QTLs for RT single records were also reported on other chromosomes. In purebreds cattle (Angus, Simmental, and Piedmontese), QTLs were identified on BTA1, 8, 10, 11, 12, 20, 22, 23, 25, 26 and 27 (Howard et al., 2013). However, the same authors annotated genes related to heat stress on BTA10, 12, 22, 23 and 25 only. Porto-Neto et al.

(2014) identified significant SNPs associated with RT on BTA6 and 17, in a Brahman and tropical composite population. In a GWAS for RT single records evaluated at different THI in heat stressed Holstein animals, Dikmen et al. (2013) identified QTLs on BTA4, 5, 16, 24 and 26, using the single-step genomic BLUP approach. In a subsequent study, based on previously selected SNPs related to thermotolerance, Dikmen et al. (2015) identified significant SNPs associated with RT single records on BTA4, 6 and 24. Based on selected SNPs located in coding regions of genes previously associated with reproduction, production or health traits, the same authors identified significant SNPs associated with RT single records on BTA6, 7, 9, 10, 11, 15, 17, 18, 22 and 25. Since the regions on BTA17 were shown to be associated with RT in this current study as well as in previous studies, and also because these regions contain candidate genes that have been shown to be involved in heat stress response (Porto-Neto et al., 2014; Dikmen et al., 2015), it is possible that these BTA17 regions may have a role on cattle heat stress.

Table 2. Significant SNP located on the chromosome 17 associated with variation of rectal temperature (ΔRT) in an experimental Gir x Holstein F2 population evaluated for heat stress response.

SNP	QTL[1]	Pos[2]	Effect[3]	−log10 (p-value)	q-value[4]	MAF[5]
BTA.41874.no.rs	1	71165342	0.15	5.45	0.05	0.33
ARS.BFGL.NGS.78729	1	71231159	0.15	5.43	0.05	0.34
ARS.BFGL.NGS.10248	1	71260851	0.15	5.43	0.05	0.34
ARS.BFGL.NGS.58770	2	74292319	0.18	5.75	0.01	0.24
ARS.BFGL.NGS.24012	3	74948921	0.17	5.44	0.03	0.27
ARS.BFGL.NGS.18349	3	74998349	0.17	5.44	0.03	0.27

[1]QTL regions where the significant SNP is located.

[2]Position on the chromosome (in megabase pairs).

[3]Allele substitution effect.

[4]FDR-based q-value.

[5]Minor allele frequency.

Candidate genes

In this current study, 54 genes encoding proteins with different functions were annotated for ΔRT. The gene sets were grouped according to their gene ontology terms enrichment in a GO network to better understand their functions (Figures 2). In this network (Figure 2), we identified genes involved in janus kinase/signal transducers and activators of transcription (**JAK/STAT**) pathway, the interleukin 6 family cytokine (**LIF**), oncostatin M (**OSM**) and DGCR8, microprocessor complex subunit (**DGCR8**), and a gene linked to response to oxygen radical biological process, the thioredoxin reductase 2 (**TXNRD2**). The OSM and LIF genes are cytokines that belong to the Interleukin 6 (**IL6**) family, and they are closely related to their structures and functions (Tanaka and Miyahima, 2003). These IL6-type cytokines, upon binding to cell surface receptors, activate members of the JAK tyrosine kinase family, resulting in the activation of the STAT TF family members, which consist in the JAK/STAT pathway (Hong et al., 2013). During the heat shock, this signaling pathway can be quickly activated by OSM and LIF cytokines, aiding in the reduction of the effects stimulated by heat shock as well as in the regulation of the expression of Hsp70 and Hsp90 genes (Chatterjee et al., 2007; Allegra et al., 2011; Stephanou and Latchman, 2011). STAT3 also shows important biological functions associated to apoptosis control, cell survival, as well as activation and inhibition of the immune and inflammatory responses (Loor, 2010).

The DGCR8 gene, together with Nuclear RNase III enzyme Drosha, comprise the Microprocessor complex, which is essential in the processing of microRNA in animals

(Triboulet et al., 2009; Faller et al., 2010). According to Knuckles et al. (2017), the DGCR8, together with the N6-adenosine-methyltransferase Mettl3, are reallocated to heat-shock genes under acute temperature stress. They work in consonance to mark Hsp70 mRNAs allowing subsequent RNA degradation to control the timing and magnitude of the heat-shock response. Previous studies described a significant increase in the protein expression levels of heat shock proteins (**HSPs**) in cattle submitted to heat stress (Belhadj Slimen et al., 2016; Hu et al., 2016). The clearing of this HSPs after the stress is very important to avoid its accumulation and presumably worse recovery after stress (Knuckles et al., 2017).

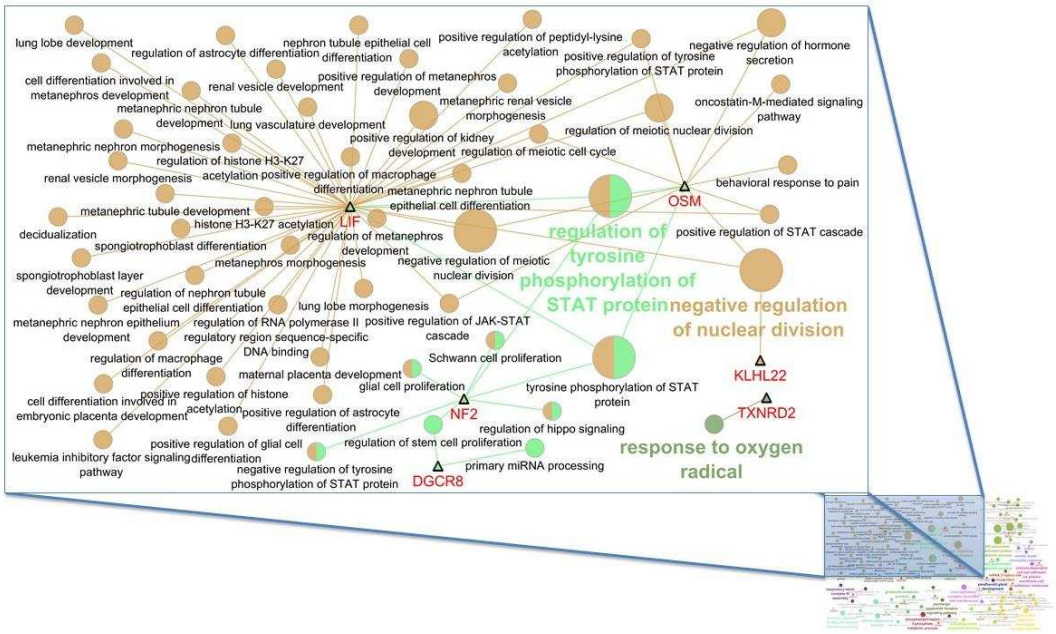


Figure 2. Functional networks showing gene interactions (triangles nodes) related to variation of rectal temperature (ΔRT) and the relationship across genes and their sub-networks related to positive regulation of tyrosine phosphorylation of STAT3 protein in the experimental Gir x Holstein F2 population.

The TXNRD2 gene, together with thioredoxin 2 (**TRX2**) and peroxiredoxin 3 (**PRX3**) genes, comprise the mitochondrial thioredoxin system, which provide a primary line of redox regulation through the high control over reactive oxygen species (**ROS**) emission

from mitochondria (Stanley et al., 2011; Yoshioka, 2015), which is a central source of H_2O_2. Animals exposed to heat stress, show an increase in the respiration rate and consequently H_2O_2 production exceeds its scavenging, leading to oxidative stress (Bernabucci et al., 2002; Srikandakumar and Johnson, 2004). TRXR2 supplies the NADPH electrons to PRX3, which controls H_2O_2 levels and establishes mitochondrial redox homeostasis (Stanley et al., 2011; Aon et al., 2012). Inhibition of TRXR2 leads to impaired redox homeostasis and results in increased levels of H_2O_2, endangering cell functions (Prasad et al., 2014).

The same gene set was used to perform the analysis of regulatory sequences, in which, based on the biological process and literature review, we were able to identify three heat stress associated TF: nuclear factor kappa B (**NF-kB**), aryl hydrocarbon receptor nuclear translocator (**ARNT**) and STAT3 (Table 3). A gene-TF network was generated with the three selected TF and genes which show potential binding sites for these TF (Figure 3).

Table 3. Enriched transcription factors (TF) associated with genes identified for cattle variation of rectal temperature (ΔRT), based on their gene ontology biological process and literature review. Data obtained from an experimental Gir x Holstein F2 population evaluated for heat stress response.

TF[1]	Biological Process	Literature review[3]
ARNT	Regulation of homeostatic process	Involved in the negative regulation of necrosis and can be regulated by heat shock proteins before and after hypoxia (Belenichev et al., 2011)
NF-kB	Anti-apoptosis	The heat stress can inhibit the activity of NF-kB and induce massive apoptosis (Belardo et al. 2010)
STAT3	Homeostatic process	Cell survival regulator (Terui et al., 2004)

	which participates in intracellular ROS homeostasis (He et al., 2010; Li et al., 2010)

[1]Transcription factor. ARNT = aryl hydrocarbon receptor nuclear translocator; NF-kB = Nuclear Factor Kappa B; STAT3 = signal transducer and activator of transcription family 3.
[2]The cited literature studies are just a sample of the vast available literature.

The NF-kB was the most enriched TF, linked to a higher number of genes. The response to heat shock has been implicated in the negative regulation of NF-kB signaling pathway (Yenari et al., 2005). According to Belardo et al. (2010), the effects of hyperthermic treatment are related to the increase in temperature above physiological conditions and the duration of exposure. These authors observed that heat stress induced an inhibition of constitutive NF-kB activity in HS-Sultan cells, a B-cell lymphoma directly related to temperature increase, besides inhibiting the activity of NF-kB and inducing massive apoptosis in other types of aggressive B-cell neoplasms. In a study regarding crosstalk among cytokines and induced hyperthermia using genomic approaches, Janus et al. (2015) found a substantial overlap among the set of genes potentially regulated by heat shock protein / heat shock transcription factor 1 (**HSF1**) and the set of genes potentially regulated by tumor necrosis factor alfa (**TNFα**) / NF-kB. This allowed the identification of predictable inhibitory effects of heat shock on the expression of classical NF-kB target genes and new patterns of activation (or coactivation) related to responses to different types of stress and hormonal stimuli. In addition, it helped to identify repression (or co-repression) in subsets of genes associated with processes mediated by NF-kB signaling and humoral immune responses, as well as chemokine and cytokine responses. In this way, pleiotropic effects of heat stress on the regulation of NF-kB-dependent genes should be expected.

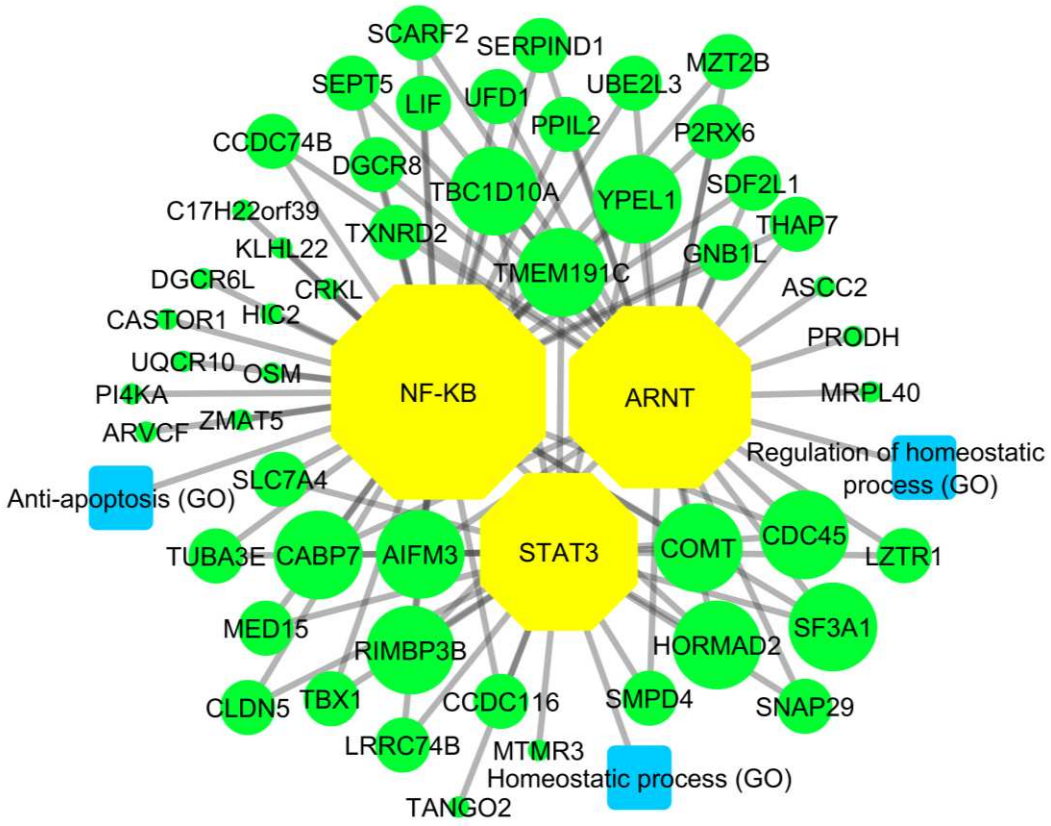


Figure 3. Gene-transcription factor (TF) network: genes located in the QTL regions associated with variation of rectal temperature (ΔRT) (green circle nodes) and their associated TF (yellow octagon nodes). The node size corresponds to the network analyses (Cytoscape, Shannon et al., 2003) in which the larger nodes denotes a higher edge density associated with the number of TF binding sites. Blue square nodes show the gene ontology biological processes related to TF. Data obtained in an experimental Gir x Holstein F2 population evaluated for heat stress response.

The ARNT TF, also known as hypoxia inducible factor 1 beta subunit (**HIF1B**), is a subunit of transcription factor hypoxia-inducible factor-1 (**HIF-1**) (Ahn et al., 2013), a key-regulator of the developmental and physiological networks required for the maintenance of O_2 homeostasis. HIF-1 plays an important role in preventing the overproduction of mitochondrial ROS in hypoxia conditions (Semenza, 2011). After preconditioning at intermediate temperatures, HSF-1 promotes survival under extreme thermal stress (Kourtis

et al., 2012) and, according to Horowitz and Assadi (2010), its expression increased after heat acclimation-mediated cross-tolerance. ARNT is involved in the negative regulation of necrosis and can be stabilized and have its expression regulated as well as its lifetime increased by heat shock proteins before and after hypoxia (Belenichev et al., 2011).

STAT3, the less enriched TF in the gene-TF network (Figure 3), regulates the expression of Hsp70 and Hsp90 genes, as previously described, it is also an important a regulator of cellular survival after apoptotic stimuli (Terui et al., 2004) and directly protects cells from oxidative stress (Li et al., 2015). This TF is sensitive to intracellular oxidants and can be activated in response to ROS accumulation. Furthermore, it participates in intracellular ROS homeostasis and shows an inverse relationship with NF-kB (He et al., 2010; Li et al., 2010).

From the gene-TF network, candidate genes for ΔRT were highlighted based on TFBS. The LIF, DGCR8 and TXNRD2 genes, which also shares ontology in the gene ontology enrichment analysis, are part of the most enriched genes group in this network (Figure 3).

Breed of Origin of Alleles for candidate genes

Based on the GO and gene-TF networks, we were able to select four candidate genes to assess the origin of alleles associated to ΔRT in cattle. Considering the LD between SNPs and genes, the physical positions of LIF, OSM, TXNRD2 and DGCR8 genes were annotated and used as coordinates for selection of adjacent SNPs. We selected a total of 7 SNPs: two SNPs associated with LIF and OSM genes, three SNPs associated with TXNRD2 gene and two SNPs associated with DGCR8 gene. The animals were grouped as Low-ΔRT or High-ΔRT based on their genomic breeding values for ΔRT, and the alleles of the

selected SNPs were classified and quantified according to their breed-of-origin (HH, GG or HG/GH).

Based on Chi-square test results obtained from Low-ΔRT and High-ΔRT F2 animals, we observed that all selected SNPs for allele origin were significant ($P < 0.05$) in both evaluated groups (Table 4) and showed a deviation of the expected gametic segregation for each genotypic class (1HH: 2HG/GH: 1GG). These results suggest a prevalence of alleles from a certain breed in the Low-ΔRT and High-ΔRT groups. On average, 53% of Low-ΔRT animals had two alleles from Holstein breed, whereas 6.4% had two alleles from Gir breed. On the other hand, in the High-ΔRT animals, the average proportion of animals for HH and GG were 2.9% and 49.9%, respectively. Based on theses results, we can observe that Holstein alleles are more frequent in Low-ΔRT animals and Gir alleles are more abundant in High-ΔRT animals.

It is known that Bos indicus animals have a greater ability to regulate their body temperature, which helps in the heat stress prevention (Hansen, 2004; Cattelam and Martinez, 2013), in comparison to Bos taurus animals which are affected by heat stress (Renaudeau et al., 2012). In the heat chamber, the F2 animals were submitted to high temperatures and relative humidity, which may have prejudiced the evaporation mechanisms, responsible for 80% of body heat loss to the environment (Shearer and Beede, 1990; West, 2003), and favored the installation of heat stress. Given the fact that the genes identified in our study play important functions to control the effects resulting from the current heat stress, and not the heat stress prevention per se, it is reasonable to assume that the significant SNP alleles were derived from the Holstein breed, which requires a more efficient genetic architecture to defend the body of deleterious effects of heat stress in comparison to the Gir animals.

Table 4. Percentage of individuals classified by breed-of-origin of each SNP allele located in the candidate genes (LIF, OSM, TXNRD2 and DGCR8) associated with variation of rectal temperature (ΔRT) in Low-ΔRT and High-ΔRT animals from an experimental Gir x Holstein F2 population evaluated for heat stress response.

	LIF/OSM		TXNRD2			DGCR8	
Alleles	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7
Low-ΔRT Animals							
HH[1]	55.3*	55.3*	51.4*	51.4*	51.4*	54.5*	51.4*
GG[2]	7.9*	7.9*	5.7*	5.7*	5.7*	6.1*	5.7*
HG/GH[3]	36.8*	36.8*	42.9*	42.9*	42.9*	39.4*	42.9*
High-ΔRT Animals							
HH	10.0*	10.0*	0.0*	0.0*	0.0*	0.0*	0.0*
GG	52.5*	52.5*	48.1*	48.1*	50.0*	48.1*	50.0*
HG/GH	37.5*	37.5*	51.9*	51.9*	50.0*	51.9*	50.0*

[1]Both alleles originated from Holstein breed.

[2]Both alleles originated from Gir breed.

[3]One allele originated from Holstein and one from Gir breed.

*Significant SNPs for allele origin in the chi-square test ($P < 0.05$)

In summary, the ΔRT trait is a relevant phenotype to individual evaluation of heat stress response. GWAS allowed the identification of significant SNPs associated with ΔRT in the experimental F2 Gir x Holstein population. Gene-TF networks and gene ontology enrichment analyses allowed the identification of the candidate genes (LIF, OSM, TXNRD2 and DGCR8) that are biologically related to heat stress response. According to BOA analysis, Holstein breed alleles could be associated to a more complex response to heat stress effects, which can be explained by the fact that Holstein animals are more affected by heat stress than Gir animals and so they require an intricate genetic architecture to defend the body of the deleterious effects of heat stress. Future studies are needed to provide

further knowledge to try understanding the genetic architecture underlying the heat stress response in crossbred cattle.

ACKNOWLEDGEMENTS

We thank EMBRAPA - Dairy Cattle Research Center, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Ministério da Ciência Tecnologia e Inovação (MCTI)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia - Ciência Animal (INCT-CA) for the financial support, Embrapa Dairy Cattle Research Center for providing the data needed to carry out this study and Mario P. L. Calus (Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands) for his assistance in BOA methodology. The authors wish to thank and acknowledge all the team effort and the visionary work done by the late Dr. Mário Luiz Martinez, former leader of this project at Embrapa Dairy Cattle Research Center.

REFERENCES

Ahn, J.E., X. Zhou, S.E. Dowd, R.S. Chapkin, and K. Zhu-Salzman. 2013. Insight into Hypoxia Tolerance in Cowpea Bruchid: Metabolic Repression and Heat Shock Protein Regulation via Hypoxia-Inducible Factor 1. PLoS One 8. doi:10.1371/journal.pone.0057267.

Alexander, D.H., J. Novembre, and K. Lange. 2009. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. Genome Res. 19:1655–1664. doi:10.1101/gr.094052.109.vidual.

Allegra, A., E. Sant'Antonio, G. Penna, A. Alonci, A. D'Angelo, S. Russo, A. Cannavò, D. Gerace, and C. Musolino. 2011. Novel therapeutic strategies in multiple myeloma: Role of the heat shock protein inhibitors. Eur. J. Haematol. 86:93–110. doi:10.1111/j.1600-0609.2010.01558.x.

Aon, M.A., B.A. Stanley, V. Sivakumaran, J.M. Kembro, B. O'Rourke, N. Paolocci, and S. Cortassa. 2012. Glutathione/thioredoxin systems modulate mitochondrial H 2 O 2 emission: An experimental-computational study. J. Gen. Physiol. 139:479–491. doi:10.1085/jgp.201210772.

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. doi:10.1093/bioinformatics/bth457.

Belardo, G., R. Piva, and M.G. Santoro. 2010. Heat stress triggers apoptosis by impairing NF-kappaB survival signaling in malignant B cells. Leukemia 24:187–96. doi:10.1038/leu.2009.227.

Belenichev, I.F., Y.M. Kolesnik, S. V. Pavlov, E.P. Sokolik, and N. V. Bukhtiyarova. 2011. Disturbance of HSP70 chaperone activity is a possible mechanism of mitochondrial dysfunction. Neurochem. J. 5:251–256. doi:10.1134/S1819712411040040.

Belhadj Slimen, I., T. Najar, A. Ghram, and M. Abdrrabba. 2016. Heat stress effects on livestock: Molecular, cellular and metabolic aspects, a review. J. Anim. Physiol. Anim. Nutr. (Berl). 100:401–412. doi:10.1111/jpn.12379.

Bernabucci, U., B. Ronchi, N. Lacetera, and A. Nardone. 2002. Markers of Oxidative Status in Plasma and Erythrocytes of Transition Dairy Cows During Hot Season. J. Dairy Sci. 85:2173–2179. doi:10.3168/jds.S0022-0302(02)74296-3.

Biffani, S., U. Bernabucci, A. Vitali, N. Lacetera, and A. Nardone. 2016. Short communication : Effect of heat stress on nonreturn rate of Italian Holstein cows. J. Dairy Sci. 1–7. doi:10.3168/jds.2015-10491.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi:10.1093/bioinformatics/btp101.

Cattelam, J., and M. Martinez. 2013. Estresse térmico em bovinos. Rev. Port. Ciências Veterinárias 108:96–102.

Chatterjee, M., S. Jain, T. Stühmer, M. Andrulis, U. Ungethüm, R.J. Kuban, H. Lorentz, K. Bommert, M. Topp, D. Krämer, H.K. Müller-Hermelink, H. Einsele, A. Greiner, and R.C. Bargou. 2007. STAT3 and MAPK signaling maintain overexpression of heat shock proteins 90α and β in multiple myeloma cells, which critically contribute to tumor-cell survival. Blood 109:720–728. doi:10.1182/blood-2006-05-024372.

Cruz, L.V. da, D. de S.R. Angrimani, B.R. Rui, and M.A. da Silva. 2011. Efeitos do estresse térmico na produção leiteira: revisão de literatura. Rev. científica eletrônica Med. veterinária. doi:16797353.

Dabney, A., Storey, J.D., Warnes, G.R., 2014. qvalue: Q-value estimation forfalse discovery rate control. R package version 1.36.0.

Dikmen, S., J.B. Cole, D.J. Null, and P.J. Hansen. 2012. Heritability of rectal temperature and genetic correlations with production and reproduction traits in dairy cattle. J. Dairy Sci. 95:3401–3405. doi:10.3168/jds.2011-4306.

Dikmen, S., J.B. Cole, D.J. Null, and P.J. Hansen. 2013. Genome-Wide Association Mapping for Identification of Quantitative Trait Loci for Rectal Temperature during Heat Stress in Holstein Cattle. PLoS One 8:1–7. doi:10.1371/journal.pone.0069202.

Dikmen, S., X. -z. Wang, M.S. Ortega, J.B. Cole, D.J. Null, and P.J. Hansen. 2015. Single nucleotide polymorphisms associated with thermoregulation in lactating dairy cows exposed to heat stress. J. Anim. Breed. Genet. 132:n/a-n/a. doi:10.1111/jbg.12176.

Faller, M., D. Toso, M. Matsunaga, I. Atanasov, R. Senturia, Y. Chen, Z.H. Zhou, and G. Feng. 2010. DGCR8 recognizes primary transcripts of microRNAs through highly cooperative binding and formation of higher-order structures. Rna 16:1570–1583. doi:10.1261/rna.2111310.

Ferrazza, R. de A., H.D.M. Garcia, V.H.V. Aristizábal, C. de S. Nogueira, C.J. Veríssimo, J.R. Sartori, R. Sartori, and J.C.P. Ferreira. 2017. Thermoregulatory responses of Holstein cows exposed to experimentally induced heat stress. J. Therm. Biol. 66:68–80. doi:10.1016/j.jtherbio.2017.03.014.

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

Hagiya, K., K. Hayasaka, T. Yamazaki, T. Shirai, T. Osawa, Y. Terawaki, Y. Nagamine, Y. Masuda, and M. Suzuki. 2017. Effects of heat stress on production , somatic cell score and conception rate in Holsteins. Anim. Sci. J. 88:3–10. doi:10.1111/asj.12617.

Hansen, P.J. 2004. Physiological and cellular adaptations of zebu cattle to thermal stress. Anim. Reprod. Sci. 82–83:349–360. doi:10.1016/j.anireprosci.2004.04.011.

He, G., G.Y. Yu, V. Temkin, H. Ogata, C. Kuntzen, T. Sakurai, W. Sieghart, M. Peck-Radosavljevic, H.L. Leffert, and M. Karin. 2010. Hepatocyte IKKb/NF-kB Inhibits Tumor Promotion and Progression by Preventing Oxidative Stress-Driven STAT3 Activation. Cancer Cell 17:286–297. doi:10.1016/j.ccr.2009.12.048.

Hickey, J.M., B.P. Kinghorn, B. Tier, J.F. Wilson, N. Dunstan, and J.H.J. van der Werf. 2011. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes.. Genet. Sel. Evol. 43:12. doi:10.1186/1297-9686-43-12.

Hong, D.S., U. Banerji, B. Tavana, G.C. George, J. Aaron, and R. Kurzrock. 2013. Targeting the molecular chaperone heat shock protein 90 (HSP90): Lessons learned and future directions. Cancer Treat. Rev. 39:375–387. doi:10.1016/j.ctrv.2012.10.001.

Horowitz, M., and H. Assadi. 2010. Heat acclimation-mediated cross-tolerance in cardioprotection: Do HSP70 and HIF-1?? play a role?. Ann. N. Y. Acad. Sci. 1188:199–206. doi:10.1111/j.1749-6632.2009.05101.x.

Howard, J.T., S.D. Kachman, W.M. Snelling, E.J. Pollak, D.C. Ciobanu, L.A. Kuehn, and M.L. Spangler. 2013. Beef cattle body temperature during climatic stress: A genome-wide association study. Int. J. Biometeorol. 58:1665–1672. doi:10.1007/s00484-013-0773-5.

Hu, H., Y. Zhang, N. Zheng, J. Cheng, and J. Wang. 2016. The effect of heat stress on gene expression and synthesis of heat-shock and milk proteins in bovine mammary epithelial cells. Anim. Sci. J. 87:84–91. doi:10.1111/asj.12375.

Janus, P., T. Stokowy, R. Jaksik, K. Szoltysek, L. Handschuh, J. Podkowinski, W. Widlak, M. Kimmel, and P. Widlak. 2015. Cross talk between cytokine and hyperthermia-

induced pathways: identification of different subsets of NF-κB-dependent genes regulated by TNFα and heat shock. Mol. Genet. Genomics 290:1979–1990. doi:10.1007/s00438-015-1055-1.

Knuckles, P., S.H. Carl, M. Musheev, C. Niehrs, A. Wenger, and M. Bühler. 2017. RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat. Struct. Mol. Biol. 24:561–569. doi:10.1038/nsmb.3419.

Kourtis, N., V. Nikoletopoulou, and N. Tavernarakis. 2012. Small heat-shock proteins protect from heat-stroke-associated neurodegeneration. Nature 490:213–218. doi:10.1038/nature11417.

Li, L., S. -h. Cheung, E.L. Evans, and P.E. Shaw. 2010. Modulation of Gene Expression and Tumor Cell Growth by Redox Modification of STAT3. Cancer Res. 70:8222–8232. doi:10.1158/0008-5472.CAN-10-0894.

Li, L., W. Wei, Y. Zhang, G. Tu, Y. Zhang, J. Yang, and Y. Xing. 2015. SirT1 and STAT3 protect retinal pigmented epithelium cells against oxidative stress. Mol. Med. Rep. 12:2231–2238. doi:10.3892/mmr.2015.3570.

Loor, J.J. 2010. Genomics of metabolic adaptations in the peripartal cow. Animal 4:1110–1139. doi:doi:10.1017/S1751731110000960.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

Maere, S., K. Heymans, and M. Kuiper. 2005. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21:3448–3449. doi:10.1093/bioinformatics/bti551.

Mcdowell, R.E., N.W. Hooven, and J.K. Camoens. 1975. Effect of Climate on Performance of Holsteins in First Lactation. J. Dairy Sci. 59:965–971. doi:10.3168/jds.S0022-0302(76)84305-6.

Nardone, A., B. Ronchi, N. Lacetera, M.S. Ranieri, and U. Bernabucci. 2010. Effects of climate changes on animal production and sustainability of livestock systems. Livest. Sci. 130:57–69. doi:10.1016/j.livsci.2010.02.011.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Porto-Neto, L.R., A. Reverter, K.C. Prayaga, E.K.F. Chan, D.J. Johnston, R.J. Hawken, G. Fordyce, J.F. Garcia, T.S. Sonstegard, S. Bolormaa, M.E. Goddard, H.M. Burrow, J.M. Henshall, S.A. Lehnert, and W. Barendse. 2014. The genetic architecture of climatic adaptation of tropical cattle. PLoS One 9:1–22. doi:10.1371/journal.pone.0113284.

Prasad, R., L.F. Chan, C.R. Hughes, J.P. Kaski, J.C. Kowalczyk, M.O. Savage, C.J. Peters, N. Nathwani, A.J.L. Clark, H.L. Storr, and L.A. Metherell. 2014. Thioredoxin reductase 2 (TXNRD2) mutation associated with familial glucocorticoid deficiency (FGD). J. Clin. Endocrinol. Metab. 99:1556–1563. doi:10.1210/jc.2013-3844.

Renaudeau, D., A. Collin, S. Yahav, V. de Basilio, J.L. Gourdine, and R.J. Collier. 2012. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. Animal 6:707–728. doi:10.1017/S1751731111002448.

Sandelin, A., W. Alkema, P. Engström, W.W. Wasserman, and B. Lenhard. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles.. Nucleic Acids Res. 32:D91-4. doi:10.1093/nar/gkh012.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2011. FImpute – An efficient imputation algorithm for dairy cattle populations. J. Anim. Sci. 89:333.

Semenza, G.L. 2011. Hypoxia-inducible factor 1: Regulator of mitochondrial metabolism and mediator of ischemic preconditioning. Biochim. Biophys. Acta - Mol. Cell Res. 1813:1263–1268. doi:10.1016/j.bbamcr.2010.08.006.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, and M.P.L. Calus. 2016. Empirical determination of breed-of-origin of alleles in three-breed cross pigs. Genet. Sel. Evol. 48:55. doi:10.1186/s12711-016-0234-9.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13:2498–2504. doi:10.1101/gr.1239303.metabolite.

Shearer, J. K.; Beede, D. K. Thermoregulation and physiological responses of dairy cattle in hot weather. Agri-Practice, Santa Bárbara, v. 11, n. 4, p. 5-8, 13, 16-17, July/Aug. 1990.

Soares, A.C.C., S.E.F. Guimarães, M.J. Kelly, M.R.S. Fortes, F.F. e Silva, L.L. Verardo, R.R. Mota, and S. Moore. 2017. Multiple-trait genomewide mapping and gene network analysis for scrotal circumference growth curves in Brahman cattle. J. Anim. Sci. 95:3331–3345. doi:10.2527/jas.2017.1409.

Solé, X., E. Guinó, J. Valls, R. Iniesta, and V. Moreno. 2006. SNPStats: A web tool for the analysis of association studies. Bioinformatics 22:1928–1929. doi:10.1093/bioinformatics/btl268.

Srikandakumar, A., and E.H. Johnson. 2004. Effect of heat stress on milk production, rectal temperature, respiratory rate and blood chemistry in Holstein, Jersey and Australian milking Zebu cows. Trop. Anim. Health Prod. 36:685–692. doi:10.1023/B:TROP.0000042868.76914.a9.

Stanley, B.A., V. Sivakumaran, S. Shi, I. McDonald, D. Lloyd, W.H. Watson, M.A. Aon, and N. Paolocci. 2011. Thioredoxin Reductase-2 Is Essential for Keeping Low Levels of H2O2 Emission from Isolated Heart Mitochondria. J. Biol. Chem. 286:33669–33677. doi:10.1074/jbc.M111.284612.

Stephanou, A., and D.S. Latchman. 2011. Transcriptional modulation of heat-shock protein gene expression. Biochem. Res. Int. 2011. doi:10.1155/2011/238601.

Tanaka, M., and A. Miyahima. 2003. Onconstatin M, a multifunctional cytokine. Rev. Physiol. Biochem. Pharmacol. 39–52. doi:10.1007/s10254-003-0013-1.

Terui, K., S. Enosawa, S. Haga, H.Q. Zhang, H. Kuroda, K. Kouchi, T. Matsunaga, H. Yoshida, J.F. Engelhardt, K. Irani, N. Ohnuma, and M. Ozaki. 2004. Stat3 confers

resistance against hypoxia/reoxygenation-induced oxidative injury in hepatocytes through upregulation of Mn-SOD. J. Hepatol. 41:957–965. doi:10.1016/j.jhep.2004.08.019.

Touzet, H., and J.-S. Varré. 2007. Efficient and accurate P-value computation for Position Weight Matrices.. Algorithms Mol. Biol. 2:15. doi:10.1186/1748-7188-2-15.

Triboulet, R., H.M. Chang, R.J. Lapierre, and R.I. Gregory. 2009. Post-transcriptional control of DGCR8 expression by the Microprocessor. Rna 15:1005–1011. doi:10.1261/rna.1591709.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Vandenplas1, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 2016 481 48:401–413. doi:10.1186/S12711-016-0240-Y.

West, J.W. 2003. Effects of Heat-Stress on Production in Dairy Cattle. J. Dairy Sci. 86:2131–2144. doi:10.3168/jds.S0022-0302(03)73803-X.

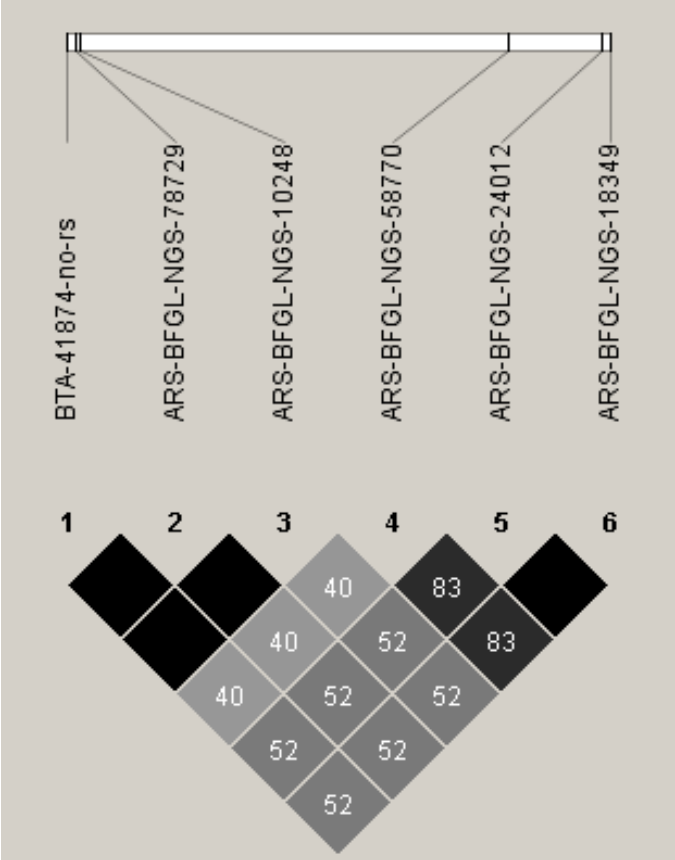
Yenari, M.A., J. Liu, Z. Zheng, Z.S. Vexler, J.E. Lee, and R.G. Giffard. 2005. Antiapoptotic and anti-inflammatory mechanisms of heat-shock protein protection. Ann. N. Y. Acad. Sci. 1053:74–83. doi:10.1196/annals.1344.007.

Yoshioka, J. 2015. Thioredoxin Reductase 2 (Txnrd2) Regulates Mitochondrial Integrity in the Progression of Age-Related Heart Failure. J. Am. Heart Assoc. 4:1–3. doi:10.1161/JAHA.115.002278.

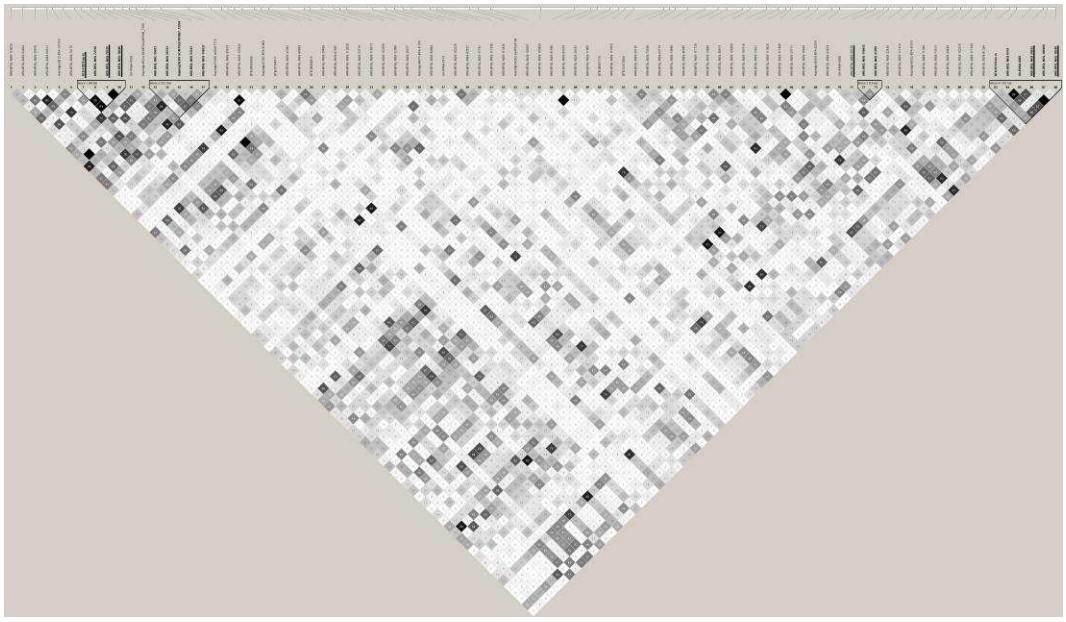
Zimin, A. V., A.L. Delcher, L. Florea, D.R. Kelley, M.C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C.P. Van Tassell, T.S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian,

J.A. Yorke, and S.L. Salzberg. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 10. doi:10.1186/gb-2009-10-4-r42.

Supplementary data



Additional file 1. LD plot (r2) for the six significant SNPs for variation of rectal temperature (ΔRT).



Additional file 2. LD plot (r2) for SNPs located within and between selected three QTL regions. The underlined SNPs represent the significant SNPs for variation of rectal temperature (ΔRT).

CHAPTER 4

Genome Wide Association Studies for gastrointestinal nematodes trait in Bos taurus x Bos indicus crossbred cattle

Pamela I. Otto*, Simone E. F. Guimarães*, Lucas L. Verardo*, Ana Luísa S. Azevedo†, Márcia Cristina A. Prata†, John Furlong†, Célio de Freitas†, Rui S. Verneque†, Marta Fonseca Martins†, João Cláudio C. Panetto†, Wanessa A. Carvalho†, Marcos Vinícius G. B. da Silva†, Marco A. Machado†

*Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; †EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, 36038-330, Brazil; ‡Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands; §Topigs Norsvin Research Center, 6640 AA Beuningen, the Netherlands

Article to be submitted to Journal of Dairy Science

ABSTRACT

Gastrointestinal nematodes (GIN) parasitism is a very important problem on livestock production, especially in ruminants, where the infections are usually subclinical making difficult the diagnosis and aggravating the losses caused by these parasites. This study aimed to identify genomic regions associated with nematode fecal egg count (FEC) in a Gir x Holstein F2 experimental population and to evaluate the breed of origin of SNPs adjacent and within the sequence of candidate genes located in significant genomic regions. The FEC phenotypes were evaluated in F2 animals during dry and rainy seasons. Illumina BovineSNP50 BeadChip was used for single-SNP genome wide association studies. The breed-of-origin of alleles (BOA) approach was used to infer the origin of the alleles in the F2 population. Heritability estimates of 0.51 ± 0.08 and repeatability of 0.55 ± 0.06 were computed for FEC. Four significant SNPs on chromosome 15 (33.64 and 37.89 M) were found to be associated with FEC. The myogenic differentiation 1 (MYOD1) gene, which plays an important role in the myogenesis and muscle repair, and the calcitonin related polypeptide beta (CALCB) gene, responsible to modulate immune cell lines that play an important role in the host response against gastrointestinal nematodes, were identified as putative candidate genes. After performing the BOA approach, we observed that 41% of animals in the FEC resistant group showed two alleles inherited from the Holstein breed, whereas 12.8% had two alleles from Gir breed. Thus, Holstein breed alleles could be associated to FEC resistance. These findings provided useful information for GIN parasitism resistance prediction in a Holstein x Gir population and suggest new candidate genes, which can be used for a deeper study of the mechanisms involved in the regulation of the assessed trait.

Key Words: crossbred cattle, helminth infections, post-GWAS analyses, resistance trait

INTRODUCTION

Gastrointestinal nematodes (**GIN**) parasitism is one of the most important problems affecting livestock production (Stear et al., 2007). In ruminants, GIN infections are usually subclinical, making difficult the diagnosis and aggravating the losses caused by these parasites (Charlier et al., 2009). According to Grisi et al. (2014), GIN infections in Brazil are responsible for the main economic loss in ruminant production - approximately $ 7.11 billion - in comparison to tick losses ($ 3,24 billion) and other parasites.

GIN infections and immune response of the host are the result of a prolonged co-evolutionary relationship between them (Anthony et al., 2007). A long period of exposure to parasites is necessary for the animals to develop an effective level of immunity protection (Gasbarre et al., 2001). This ability is influenced by the genetic composition of the host (Gasbarre et al., 1990). Genetic variability for GIN resistance in bovine has been reported, which makes the use of genetic selection promising for the improvement of this trait (Coppieters et al., 2009; Kim et al., 2013).

Previous studies have found that Bos taurus animals showed significant highest fecal egg count (**FEC**) than Bos indicus and crossbred animals (Silva et al., 2012a; Neves et al., 2014). The nematode species is also important to define resistance or susceptibility of a host-specific breed or crossbred cattle (Achi et al., 2003; Oliveira et al., 2009, 2013). Nevertheless, experimental FEC trait evaluation require time and large investments, as well as exposure of animals to biological parasites. Thus, the use of markers associated with quantitative trait loci (**QTL**) as a prior indicator of FEC resistance or susceptibility of an animal could be very useful to breeding programs. Genome wide association studies (**GWAS**) have found genomic regions associated with FEC and allowed the identification

of important genes related to immune system in cattle (Gasbarre et al., 2001; Xu et al., 2014; Kim et al., 2015).

Recently, Vandenplas et al. (2016) developed a procedure that enables breed-of-origin assignment (BOA) of alleles in crossbred animals. Results from this approach can be used in genomic evaluation models - that take breed-specific SNP effects into account - to better predict crossbred performance. Additionally, it can also be used in genome wide association studies (GWAS) to identify the parental breed of the (associated) SNP alleles and haplotypes (Sevillano et al., 2017, 2018). Otto et al. (2018) have used the BOA approach to evaluate the origin of marker alleles of candidate genes identified for tick resistance. Most of tick resistant animals had two alleles inherited from the Gir breed, whereas the susceptible animals had two alleles from the Holstein breed.

This work aimed to perform a GWAS to identify genomic regions associated to FEC in a Gir x Holstein F2 experimental population and to investigate the breed of origin of SNPs adjacent and within the sequence of candidate genes found to be associated to this trait.

MATERIALS AND METHODS

Data

Data was obtained at Embrapa Dairy Cattle Research Center, located in Southeastern Brazil. An experimental population was generated as described by Machado et al. (2010) and evaluated for GIN.

A total of 266 Gir x Holstein F2 animals were continually exposed to a mixed GIN infection by grazing for 28 weeks. Fecal samples were weekly collected in two seasons (dry and rainy). Fecal egg counts per gram of feces were determined by centrifugal-flotation

technique using a sugar solution (Sloss and R.L. Kemp, 1978). According to Gasbarre et al. (2001), animals develop immune responses that limit parasite egg output and can be classified in: 1) animals that never show high FEC values (Type I); 2) animals that show an increase in FEC values in the first two months on pasture followed by a decrease and steady levels similar to Type I values (Type II); and 3) animals that maintain high FEC levels throughout the evaluation period (Type III). In order to avoid a misclassification of the resistant and susceptible animals according to the FEC, the average of five measurements taken one week apart at the end of the grazing evaluation period was determined for FEC (Silva et al., 2012b). Animals were evaluated in contemporary groups from 10 to 14 months of age.

All animals of the experimental population were genotyped with the Illumina BovineSNP50 v2 BeadChip (Illumina, San Diego, California, USA). Genotype quality control was applied on the same dataset used in a previous study (Otto et al., 2018), in which more details are provided. After quality control, the genotype file contained 40,283 SNP markers and 476 samples, including 4 Holstein bulls, 23 Gir cows, 65 F1, and 384 F2 animals. Ancestry was estimated using Admixture software (Alexander et al., 2009) on a LD filtered dataset (7,425 SNP).

Genome wide association analyses

FEC trait was transformed prior to GWAS to account for its non-normal distribution, using the natural log-transformation: ln (FEC + 1). A single-SNP GWAS was performed with ASReml software (Gilmour et al., 2009) applying the following model:

$$\mathbf{y = xb + X\beta + Za + Wp + e}$$

in which **y** is the vector of observations of ln transformed FEC; **b** is the fixed effect of the SNP tested for association; **x** is a vector containing the genotype scores; **β** is the vector of fixed effects (i.e., contemporary group, season and sex) and the covariates (i.e., age and breed composition); **a** is the vector of random additive genetic effects; **p** is the vector of random permanent environmental effects; and **e** is the vector of residual effects; **X, Z** and **W** are the incidence matrices of fixed, additive genetic and permanent environmental effects, respectively.

It was assumed that $\mathbf{a} \sim N(0, \mathbf{G}\sigma_a^2)$, $\mathbf{p} \sim N(0, \mathbf{I}\sigma_p^2)$ and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_a^2, σ_p^2 and σ_e^2 are the additive genetic, permanent environmental and residual variances, respectively; **G** is the genomic relationship matrix and **I** is the identity matrix.

The **G** matrix was built as described by VanRaden (2008):

$$G = \frac{ZZ'}{2\sum_{i=1}^{n} p_i(1-p_i)}$$

where **Z** is a matrix that contains all SNP genotypes of all animals; **n** is the total number of SNPs present in **Z** and $\boldsymbol{p_i}$ is the frequency of the allele B of SNP i. The SNP genotypes were coded as 0, 1 and 2, being 0 = AA, 1=AB and 2 =BB.

To account for multiple testing, a false discovery rate (FDR) implemented in the R package qvalue (Dabney et al., 2014) was applied to provide the p-values corrected for FDR (q-values) for the SNP association tests. Associations with a q-value ≤ 0.05 were considered significant. The variance components were estimated based on same model without a SNP effect.

QTL regions and Candidate gene

QTL regions were defined based on the location of the significant SNPs and the average linkage disequilibrium (r^2) among the SNP markers in each chromosome. The

pairwise LD between markers was estimated using the Haploview software (Barrett et al., 2005). Since, all the significant SNPs identified with the GWAS were located on the chromosome 15, the QTL size was based on the average LD of the chromosome 15. An average $r^2 > 0.20$ was found for SNPs within a distance of 500 Kb. Thus, all SNPs located within 500 Kb, considering 250 Kb upstream and 250 Kb downstream from significant SNPs, were considered to belong to the same QTL region.

Putative candidate genes within the QTL regions were identified based on the UMD 3.1 assembly of the bovine genome (Zimin et al., 2009), using the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov).

Assignment of Breed of Origin of Alleles

To infer the breed-of-origin of the alleles on the chromosome 15 in the F2 population, the breed-of-origin of alleles (**BOA**) approach (Vandenplas et al., 2016) was used. It consists of three steps: (1) phasing the haplotypes of purebred and crossbred animals with AlphaPhase 1.1 software (Hickey et al., 2011); (2) determining the unique haplotypes for each pure breed; and (3) assigning the breed-of-origin for each allele detected in the F2 animals based on the unique haplotypes. This approach was applied to the same dataset in a previous study (Otto et al., 2018), in which more details are provided. In total, 90.6% of the alleles of chromosome 15 of the F2 animals were assigned a breed-of-origin.

After performing the BOA approach, 10% of the F2 animals showing the lowest genomic breeding values for FEC were designated as resistant, while the 10% showing the highest breeding values for FEC were designated as susceptible. For the resistant and susceptible animals, SNPs adjacent and within the sequence of candidate genes were selected and classified as HH (i.e., both alleles originated from Holstein breed), GG (i.e.,

both alleles originated from Gir breed), and HG/GH (i.e., one allele originated from Holstein and one from Gir breed), based on the results of the BOA approach. Finally, the alleles of each SNP were quantified to check if there is a prevalence of alleles from a certain breed in the resistant or susceptible groups. For that, a Chi-square test was used to evaluates the expected gametic segregation of 25% HH, 50% HG/GH and 25% GG in the both groups. SNPs with a difference in prevalence of alleles from a specific breed were considered significant ($P \leq 0.05$).

RESULTS AND DISCUSSION

Genetic parameters

In this study, the estimated heritability was equal to 0.51 ± 0.08 and the estimated repeatability was 0.55 ± 0.08 for FEC trait. In the literature, heritability estimates reported for GIN in cattle ranged from 0.05 to nearly 0.8 (Kloosterman et al., 1992; Coppieters et al., 2009; May et al., 2017). In small ruminants heritability range from 0.01 to 0.65 (Lôbo et al., 2009; Pollott et al., 2018). The FEC heritability estimates can be influenced by population sampled, statistical model, nature of infection (i.e., natural or artificial infections), time and replication of fecal sampling and animal age (Burrow, 2001; Amarante et al., 2009; Passafaro et al., 2015). Our results indicate the presence of genetic variability for bovine FEC, indicating a potential to promote genetic and phenotypic changes, and could be used as a complementary tool in parasite control programs.

QTL regions

A total of three significant SNPs ($q \leq 0.05$) associated with FEC were detected (Figure 1 and Table 2). These SNPs are located on the Bos taurus chromosome 15

(BTA15) between 33.64 and 37.89 Mb and distributed over three QTL regions. The most significant SNP (ARS.BFGL.NGS.99491) showed a -log10 (p-value) of 5.91 (q-value = 0.02) and allele substitution effect of 1.62 (Table 1). The estimated genomic inflation factor was 1.04 (Figure 2).

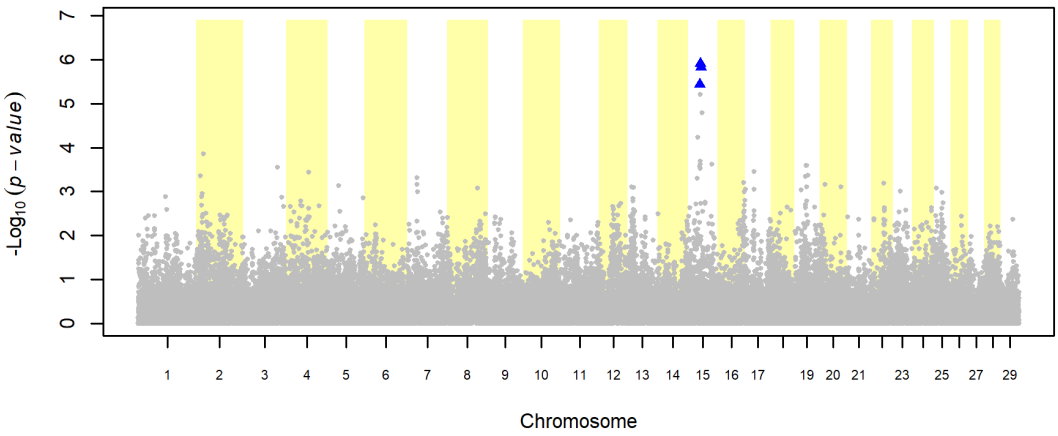


Figure 1. Manhattan plot showing the association analysis of average fecal egg count (FEC) in a Gir x Holstein F2 population evaluated for gastrointestinal nematode infection. The significant SNPs (q-*value* ≤ *0.05*) are shown as blue triangles and the gray dots denote SNPs which were not significant in GWAS.

Table 1. Significant SNP located on the chromosome 15 associated to average fecal egg count (FEC) in an experimental Gir x Holstein F2 population evaluated for gastrointestinal nematode infection.

SNP	QTL[1]	Pos[2]	Effect[3]	−log10(p-value)	q-value[4]
Hapmap39348-BTA-63139	1	33648898	1.61	5.44	0.04
ARS-BFGL-NGS-99491	2	35333265	1.63	5.91	0.02
Hapmap47787-BTA-95921	3	37893014	1.54	5.83	0.02

[1]QTL regions where the significant SNP is located.

[2]Position on the chromosome (in megabase pairs).

[3]Allele substitution effect.

[4]FDR-based q-value

Although several SNPs on different chromosomes were reported as associated to FEC (Coppieters et al., 2009; Kim et al., 2015; Benavides et al., 2016), none of the SNPs located on the QTL regions identified in this study have been previously reported.

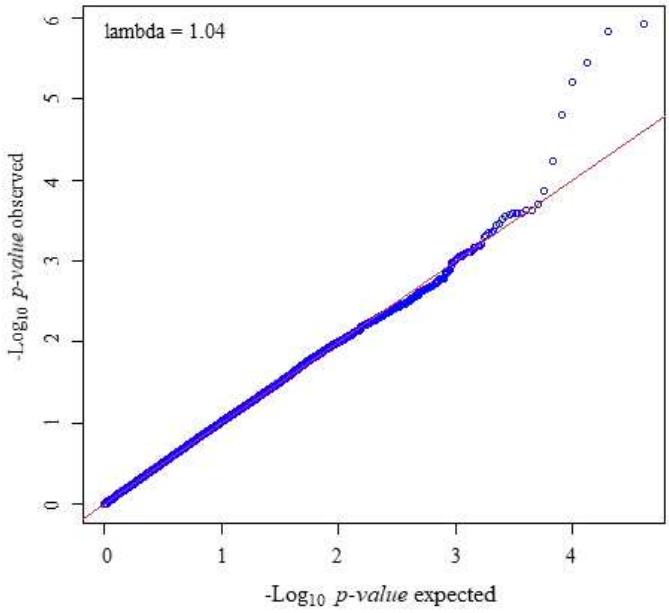


Figure 2. Quantile–Quantile plot with lambda value. On the y-axis, observed −log10 (p-values) (blue dots) are plotted against expected −log10 (p-values) under the null hypothesis of no association between SNP markers and average fecal egg count (FEC) in a Gir x Holstein F2 population evaluated for gastrointestinal nematodes infections (red line).

Candidate genes

Based on the three detected QTL regions, we identified eight genes encoding proteins with various functions. The candidate gene myogenic differentiation 1 (**MYOD1**), also known as MYOD, is located on the second identified FEC QTL region. This gene is expressed at different stages of cell differentiation by satellite cells found in injured muscle (Cooper et al., 1999; Odelberg et al., 2000; Jung and Ko, 2010), and plays an important role in the myogenesis and muscle repair (Berkes and Tapscott, 2005). Satellite cells are responsible to post-natal growth and repair in adult skeletal muscle, and their activation is an important step in the muscle regeneration ability (Wu et al., 2001). Studies have

describing that Trichinella spp. infections in mice, an intracellular parasitic nematode, causes severe damage in muscle cells, which results in activation and proliferation of satellite cells (Boonmars et al., 2004; Wu et al., 2005). According to Wu et al. (2001), the MYOD gene was over-expressed in infected muscle tissue during Trichinella spp. infections, and the MYOD expression was detected exclusively in satellite cells adjacent to infected muscle cells.

The third QTL region contains another candidate gene, the calcitonin related polypeptide beta (**CALCB**) gene, also known as CGRPB. The CGRPB gene is one of the isoforms of the CGRP gene, expressed primarily in the enteric sensory system, by enteric neurons, gut, and inner organs. The CGRPB and CGRPA genes, the another CGRP's isoform, were generated from an ancestor gene and have the same biological activities in the vasculature (Mulderry et al., 1988; Wimalawansa, 1996; Rezaeian et al., 2008). Holzmann (2013) described the CGRP as a potent neuropeptide, which downregulates the immune response and influences the key processes in autoimmunity through its local and systemic levels rapidly increased during inflammatory response. This gene may modulate the adaptive immune responses of CD4+ T helper cells and amplify the IL-10 production (Holzmann, 2013), which play an important role in the host response against GIN (Claerebout and Vercruysse, 2000). According to Jonge et al. (2003), the intestinal schistosomiasis caused by a Trematoda parasite, which also results in host tissue damage, is associated with an upregulation of the number of CGRP-immunoreactive (ir) nerve fibers, which is involved in the recruitment of mucosal mast cells during infection. Thus, the CGRP and mucosal mast cells constitute a major part of the defense reaction against injury and noxious stimuli during this intestinal parasitosis (Jonge et al., 2003). Later studies described the importance of mucosal mast cells in the mechanism of parasite expulsion

from the intestine in nematode infections (McDermott et al., 2003; Pennock and Grencis, 2006). No evidence was found for additional candidate genes in the QTL regions identified for FEC.

Breed of Origin of Alleles for candidate genes

By means of the GWAS results and literature data, we were able to select the MYOD1 and CALCB genes as candidate genes for FEC. To access the origin of alleles of these genes, we took into account the LD between SNPs and genes to annotate the physical positions of the genes and select the adjacent SNPs. We selected three SNPs associated with MYOD1 gene and two associated with CALCB gene. The genomic breeding value for FEC trait was used to classify the animals as resistant or susceptible and the alleles of the selected SNPs were classified and quantified according to their breed-of-origin (HH, GG or HG/GH).

Chi-square results allowed us to identify a significant SNP for allele origin ($P < 0.05$) in the resistant animals (Table 2). Regarding this SNP, we observed that 41% of animals had two alleles inherited from the Holstein breed, whereas 12.8% had two alleles from the Gir breed. Although the proportion of animals that had one allele inherited from each breed (Holstein and Gir) was higher than the proportion of animals showing both alleles from Holstein breed (47.6 vs 41.0%), the deviation of expected gametic segregation (25%) was higher in the second case. Based on these findings, we suggest that Holstein breed alleles could be associated to GIN resistance in this QTL region. Silva et al. (2012a) observed that purebred Holstein cows showed highest significant FEC in comparison to purebred Gir and crossbred animals (1/2 Holstein x ½ Gir and 3/8 Holstein x 5/8 Gir), corroborating the results by Neves et al. (2014), that indicated highest FEC in the purebred Holsteins and

Simental animals when compared to crossbred animals (Simental x Guzerá and Holstein x Gir). No statistically significant differences in the average FEC in the genetic groups evaluated were found by Oliveira et al. (2009, 2013), however, these authors and Achi et al. (2003) identified significant susceptibility to infection of some different nematodes species in B. taurus and B. indicus. Oliveira et al. (2009, 2013) reported significantly larger numbers of Cooperia spp in the Nellore breed compared to crossbred cattle, proposing the susceptibility of this parasite to the breed. According to Achi et al. (2003), the Haemonchus placei was highly dominant in B. taurus, whereas the highest percentage of H. similis was identified in B. indicus. In this way, it can be suggested that the resistance of bovines to GIN can be influenced by nematode species.

Table 2. Percentage of individuals classified by breed-of-origin of each SNP allele located in the candidate genes (MYOD1 and CALCB) associated with average fecal egg count (FEC) in resistant and susceptible animals from an experimental Gir x Holstein F2 population evaluated for gastrointestinal nematode infection.

	MYOD1			CALCB	
Alleles	SNP3	SNP4	SNP5	SNP6	SNP7
Resistant animals					
HH[1]	37.8	39.5	39.5	41.0*	37.8
GG[2]	13.5	13.2	13.2	12.8*	13.5
HG/GH[3]	48.6	47.4	47.4	46.2*	48.6
Susceptible animals					
HH	35.1	34.2	34.2	32.4	31.4
GG	32.4	31.6	31.6	32.4	34.3
HG/GH	32.4	34.2	34.2	35.1	34.3

[1]Both alleles originated from Holstein breed.

[2]Both alleles originated from Gir breed.

[3]One allele originated from Holstein and one from Gir breed.

*Significant SNPs for allele origin in the chi-square test ($P < 0.05$).

In the current study, the Holstein alleles associated with the resistance could behave as cryptic alleles, which play an important role in the hybrid vigor of crossbred populations. According to Abasht and Lamont (2007), these alleles are important in population heterosis since they influence the expression of phenotypes in these populations and can show opposite effect or no significant effect on one of the parental populations. This cryptic effect can be detected as a result of allele selection, genetic drift, epistasis and pleiotropic effects of the QTL on other traits that are under selection (Abasht and Lamont, 2007).

In summary, we found three SNPs associated with fecal egg count of gastrointestinal nematodes in a Holstein x Gir F2 experimental population. Two candidate genes (MYOD1 and CALCB) with important roles linked to gastrointestinal infections were identified. The results of BOA analysis indicate that Holstein alleles could be associated to FEC resistance. These findings provide additional information regarding the complex mechanisms involved in the parasite resistance in cattle and could be useful to help predicting the FEC resistance in Holstein x Gir crossbreds that is being mostly and increasingly used for milk production in Brazil and other tropical countries in the world.

ACKNOWLEDGEMENTS

We thank Embrapa Dairy Cattle Research Center (EMBRAPA), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Ministério da Ciência Tecnologia e Inovação (MCTI)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia - Ciência Animal (INCT-CA) for the

financial support, Embrapa Dairy Cattle Research Center for providing the data needed to carry out this study, Mario P. L. Calus and Jeremie Vandenplas (Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands) for their assistance in the BOA methodology. The authors wish to thank and acknowledge all the team effort and the visionary work done by the late Dr. Mário Luiz Martinez, former leader of this project at Embrapa Dairy Cattle Research Center.

REFERENCES

Abasht, B., and S.J. Lamont. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. Anim. Genet. 38:491–498. doi:10.1111/j.1365-2052.2007.01642.x.

Achi, Y.L., J. Zinsstag, K. Yao, N. Yeo, P. Dorchies, and P. Jacquiet. 2003. Host specificity of Haemonchus spp. for domestic ruminants in the savanna in northern Ivory Coast. Vet. Parasitol. 116:151–158. doi:10.1016/S0304-4017(03)00258-9.

Alexander, D.H., J. Novembre, and K. Lange. 2009. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. Genome Res. 19:1655–1664. doi:10.1101/gr.094052.109.vidual.

Amarante, A.F.T., I. Susin, R.A. Rocha, M.B. Silva, C.Q. Mendes, and A. V Pires. 2009. Veterinary Parasitology Resistance of Santa Ines and crossbred ewes to naturally acquired gastrointestinal nematode infections. Vet. Parasitol. 165:273–280. doi:10.1016/j.vetpar.2009.07.009.

Anthony, R.M., L.I. Rutitzky, J.F. Urban, M.J. Stadecker, and W.C. Gause. 2007. Protective immune mechanisms in helminth infection. Nat. Rev. Immunol. 7:975–87. doi:10.1038/nri2199.

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. doi:10.1093/bioinformatics/bth457.

Benavides, M.V., T.S. Sonstegard, and C. Van Tassell. 2016. Genomic Regions Associated with Sheep Resistance to Gastrointestinal Nematodes. Trends Parasitol. 32:470–480. doi:10.1016/j.pt.2016.03.007.

Berkes, C.A., and S.J. Tapscott. 2005. MyoD and the transcriptional control of myogenesis. Semin. Cell Dev. Biol. 16:585–595. doi:10.1016/j.semcdb.2005.07.006.

Boonmars, T., Z. Wu, I. Nagano, T. Nakada, and Y. Takahashi. 2004. Differences and similarities of nurse cells in cysts of Trichinella spiralis and T . pseudospiralis. J. Helminthol. 78:7–16. doi:10.1079/JOH2003203.

Burrow, H.M. 2001. Variances and covariances between productive and adaptive traits and temperament in a composite breed of tropical beef cattle. Livest. Prod. Sci. 70:213–233. doi:10.1016/S0301-6226(01)00178-6.

Charlier, J., J. Höglund, G. von Samson-Himmelstjerna, P. Dorny, and J. Vercruysse. 2009. Gastrointestinal nematode infections in adult dairy cattle: Impact on production, diagnosis and control. Vet. Parasitol. 164:70–79. doi:10.1016/j.vetpar.2009.04.012.

Claerebout, E., and J. Vercruysse. 2000. The immune response and the evaluation of acquired immunity against gastrointestinal nematodes in cattle : a review. Parasitology 120:25–42.

Cooper, R.N., S. Tajbakhsh, V. Mouly, G. Cossu, M. Buckingham, and G.S. Butler-Browne. 1999. In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle. J. od Cell Sci. 2901:2895–2901.

Coppieters, W., T.H.M. Mes, T. Druet, F. Farnir, N. Tamma, C. Schrooten, A.W.C. a Cornelissen, M. Georges, and H.W. Ploeger. 2009. Mapping QTL influencing gastrointestinal nematode burden in Dutch Holstein-Friesian dairy cattle.. BMC Genomics 10:96. doi:10.1186/1471-2164-10-96.

Dabney, A., Storey, J.D., Warnes, G.R., 2014. qvalue: Q-value estimation forfalse discovery rate control. R package version 1.36.0.

Gasbarre, L.C., E.A. Leighton, and C.J. Davies. 1990. Genetic control of immunity to gastrointestinal nematodes of cattle. Vet. Parasitol. 37:257–272. doi:10.1016/0304-4017(90)90009-Z.

Gasbarre, L.C., E.A. Leighton, and T. Sonstegard. 2001. Role of the bovine immune system and genome in resistance to gastrointestinal nematodes. Vet. Parasitol. 98:51–64. doi:10.1016/S0304-4017(01)00423-X.

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

Grisi, L., R.C. Leite, J.R.D.S. Martins, A.T.M. De Barros, R. Andreotti, P.H.D. Cançado, A.A.P. De León, J.B. Pereira, and H.S. Villela. 2014. Reassessment of the potential economic impact of cattle parasites in Brazil. Brazilian J. Vet. Parasitol. 23:150–156. doi:10.1590/S1984-29612014042.

Hickey, J.M., B.P. Kinghorn, B. Tier, J.F. Wilson, N. Dunstan, and J.H.J. van der Werf. 2011. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes. Genet. Sel. Evol. 43:12. doi:10.1186/1297-9686-43-12.

Holzmann, B. 2013. Modulation of immune responses by the neuropeptide CGRP. Amino Acids 45:1–7. doi:10.1007/s00726-011-1161-2.

Jonge, F. De, L. Van Nassauw, D. Adriaensen, F. Van Meir, H.R.P. Miller, E. Van Marck, and J.-P. Timmermans. 2003. Effect of intestinal inflammation on capsaicin-sensitive afferents in the ileum of Schistosoma mansoni-infected mice. Histochem Cell Biol. 119:477–484. doi:10.1007/s00418-003-0532-5.

Jung, S., and Y. Ko. 2010. TRIM72, a novel negative feedback regulator of myogenesis, is transcriptionally activated by the synergism of MyoD (or myogenin) and MEF2. Biochem. Biophys. Res. Commun. 396:238–245. doi:10.1016/j.bbrc.2010.04.072.

Kim, E.S., T. S.Sonstegard, M.V.G.B. Da Silva, L.C. Gasbarre, and C.P. VanTassell. 2015. Genome-Wide scan of gastrointestinal nematode resistance in closed angus population selected for minimized influence of MHC. PLoS One 10:1–18. doi:10.1371/journal.pone.0119380.

Kim, E.S., T.S. Sonstegard, M.V.G.B. Silva, L.C. Gasbarre, and C.P. Van Tassell. 2013. Identification of quantitative trait loci affecting gastrointestinal parasite resistance in an experimental Angus population. Anim. Genet. 45:117–121. doi:10.1111/age.12101.

Kloosterman, A., H.K. Parmentier, and H.W. Ploeger. 1992. Breeding Cattle and Sheep for Resistance to Gastrointestinal Nematodes. Parasitol. Today 8:330–335.

Lôbo, R.N.B., L.S. Vieira, A.A. De Oliveira, E.N. Muniz, and J.M. Silva. 2009. Genetic parameters for faecal egg count , packed-cell volume and body-weight in Santa Inês lambs. Genet. Mol. Biol. 294:288–294.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan

for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

May, K., K. Brügemann, T. Yin, C. Scheper, C. Strube, and S. König. 2017. Genetic line comparisons and genetic parameters for endoparasite infections and test-day milk production traits. J. Dairy Sci. 100:1–15. doi:10.3168/jds.2017-12901.

McDermott, J.R., R.E. Bartram, P.A. Knigh, H.R.P. Miller, D.R. Garrod, and R.K. Grencis. 2003. Mast cells disrupt epithelial barrier function during enteric nematode infection. PNAS 100:7761–7766. doi:https://doi.org/10.1073/pnas.1231488100.

Mulderry, P.K., M.A. Ghatei, R.A. Spokes, P.M. Jones, A.M. Pierson, and Q.A. Hamid. 1988. Differential expression of alpha-CGRP and beta-CGRP by primary sensory neurons and enteric autonomic neurons of the rat. Neuroscience 25:195–205.

Neves, J.H. das, N. Carvalho, L. Rinaldi, G. Cringoli, and A.F.T. Amarante. 2014. Diagnosis of anthelmintic resistance in cattle in Brazil: A comparison of different methodologies. Vet. Parasitol. 206:216–226. doi:10.1016/j.vetpar.2014.10.015.

Odelberg, S.J., A. Kollhoff, and M.T. Keating. 2000. Dedifferentiation of Mammalian Myotubes Induced by msx1. Cell 103:1099–1109.

Oliveira, M.C.S., M.M. Alencar, A.C.S. Chagas, R. Giglioti, and H.N. Oliveira. 2009. Gastrointestinal nematode infection in beef cattle of different genetic groups in Brazil. Vet. Parasitol. 166:249–254. doi:10.1016/j.vetpar.2009.09.006.

Oliveira, M.C.S., M.M. Alencar, R. Giglioti, M.C.D. Beraldo, F.F. Aníbal, R.O. Correia, L. Boschini, A.C.S. Chagas, T.B. Bilhassi, and H.N. Oliveira. 2013. Resistance of beef cattle of two genetic groups to ectoparasites and gastrointestinal nematodes in the state of São Paulo, Brazil. Vet. Parasitol. 197:168–175. doi:10.1016/j.vetpar.2013.06.021.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Passafaro, T.L., J.P.B. Carrera, L.L. dos Santos, F.S.S. Raidan, D.C.C. dos Santos, E.P. Cardoso, R.C. Leite, and F.L.B. Toral. 2015. Genetic analysis of resistance to ticks, gastrointestinal nematodes and Eimeria spp. in Nellore cattle. Vet. Parasitol. 210:224–234. doi:10.1016/j.vetpar.2015.03.017.

Pennock, J.L., and R.K. Grencis. 2006. The Mast Cell and Gut Nematodes: Damage and Defence. Chem. Immunol. Allergy 90:128–140. doi:10.1159/000088885.

Pollott, G.E., L.J.E. Karlsson, S. Eady, and J.C. Greeff. 2018. Genetic parameters for indicators of host resistance to parasites from weaning to hogget age in Merino sheep. J. Anim. Sci. 82:2852–2864.

Rezaeian, A.H., T. Katafuchi, M. Yoshizawa, N. Hiraiwae, T. Saito, M. Nishibori, K. Hamano, N. Minamino, and H. Yasue. 2008. Genomic organization, expression and evolution of porcine CRSP1, 2, and 3. Cytogenet Genome Res. 121:41–49. doi:10.1159/000124380.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Silva, J.B. Da, C.P. Rangel, B.D.A. Baêta, and A.H. Da Fonseca. 2012a. Risk factors relating to helminth infections in cows during the peripartum. Rev. Bras. Parasitol. Veterinária 21:92–96. doi:10.1590/S1984-29612012000200004.

Silva, M.V.B., T.S. Sonstegard, O. Hanotte, J.M. Mugambi, J.F. Garcia, S. Nagda, J.P. Gibson, F.A. Iraqi, A.E. McClintock, S.J. Kemp, P.J. Boettcher, M. Malek, C.P. Van Tassell, and R.L. Baker. 2012b. Identification of quantitative trait loci affecting resistance to gastrointestinal parasites in a double backcross population of Red Maasai and Dorper sheep. Anim. Genet. 43:63–71. doi:10.1111/j.1365-2052.2011.02202.x.

Sloss, M.W., and R.L. Kemp. 1978. Veterinary Clinical Parasitology. 5th ed. Ames, ed. Iowa University Press.

Stear, M.J., M. Doligalska, and K. Donskow-Schmelter. 2007. Alternatives to anthelmintics for the control of nematodes in livestock. Parasitology 134:139–151. doi:10.1017/S0031182006001557.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980.

Wimalawansa, S.J. 1996. Calcitonin Gene-Related Peptide and Its Receptors: Molecular Genetics, Physiology, Pathophysiology, and Therapeutic Potentials. Endocr. Rev. 17:533–585. doi:10.1210/edrv-17-5-533.

Wu, Z., A. Matsuo, T. Nakada, I. Nagano, and Y. Takahashi. 2001. Different response of satellite cells in the kinetics of myogenic regulatory factors and ultrastructural pathology after Trichinella spiralis and T . pseudospiralis infection. Parasitology 123:85–94.

Wu, Z., I. Nagano, T. Boonmars, and Y. Takahashi. 2005. A spectrum of functional genes mobilized after Trichinella spiralis infection in skeletal muscle. Parasitology 130:561–573. doi:10.1017/S0031182004006912.

Xu, L., Y. Hou, D.M. Bickhart, J. Song, C.P. Van Tassell, T.S. Sonstegard, and G.E. Liu. 2014. A genome-wide survey reveals a deletion polymorphism associated with resistance to gastrointestinal nematodes in Angus cattle. Funct. Integr. Genomics 14:333–9. doi:10.1007/s10142-014-0371-6.

Zimin, A. V., A.L. Delcher, L. Florea, D.R. Kelley, M.C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C.P. Van Tassell, T.S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J.A. Yorke, and S.L. Salzberg. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 10. doi:10.1186/gb-2009-10-4-r42.

CHAPTER 5

Single-step GWAS and Post-GWAS analyses identify loci and candidate genes for milk yield in Girolando cattle

Pamela I. Otto[*], Mario P. L. Calus[‡], Simone E. F. Guimarães[*], Jeremie Vandenplas[‡], Marta Fonseca Martins[†], Marco A. Machado[†], João Cláudio C. Panetto[†], Marcos Vinícius G. B. da Silva[†]

[*]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil; [†]EMBRAPA - Dairy Cattle Research Center, Juiz de Fora, MG, 36038-330, Brazil; [‡]Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands; [§]Topigs Norsvin Research Center, 6640 AA Beuningen, the Netherlands

Article to be submitted to Journal of Dairy Science

ABSTRACT

Milk production is an economically important sector of Brazilian agribusiness and the majority of the milk produced comes from Girolando cows. This study aimed to identify QTL and candidate genes associated with 305-day milk yield (305MY) in the Girolando population. In addition, we investigated the SNP variances specific for both breeds of origin. A single-step GBLUP procedure was used to identify QTL associated with 305MY and the most likely candidate genes were identified through follow-up analyses. Genomic breeding values specific for Holstein and Gir were estimated for Girolando animals using a model that uses breed-specific partial relationship matrices and those were converted to SNP effects by breed-of-origin. Differences between breeds-of-origin were evaluated by comparing estimated SNP variances between breeds. From ten regions explaining most additive genetic variance for 305MY in Girolando cattle, seven candidate genes were identified on the chromosomes 1, 4, 6 and 26. For these seven candidate genes, the SNP variances were different depending upon the breed-of-origin, while the alleles coming from the Gir breed-of-origin showed the highest genetic variance. The results reported in the current study are valuable to QTL mapping in the Girolando breed and combined with the candidate genes can be implemented in breeding programs by genomic selection.

Key Words: gene network, BOA model, tropical dairy cattle, breed origin effect

INTRODUCTION

Milk production is an economically important sector of Brazilian agribusiness. Most of the milk in Brazil is produced in systems relying on grazing in challenging environments, and these circumstances should be considered in the breeding goal (Ferraz and Felício, 2010). To meet these requirements, the Girolando breed was created, which aims to produce milk sustainably in tropical and subtropical regions. Animals crossed between Holstein and Gir cattle are present in the majority of Brazilian dairy farms and are responsible for 80% of the milk produced in the country (Silva et al., 2018b).

The first crosses between Holstein (**H**) and Gir (**G**) in Brazil emerged in the 1940's, in an unintentional manner in a Brazilian farm. Years later, based on the increased importance of milk production, a program was implemented to create the Girolando breed, in order to direct the crosses to generate offspring with high genetic capacity for productivity traits, high fertility indexes and heterosis. The Girolando's genetic composition ranges from 1/4 H + 3/4 G to 7/8 H + 1/8 G, however, the crossing schemes have been directed to focus the breed's genetic composition into 5/8 H + 3/8 G. The Girolando progeny testing scheme started in 1997, and allowed dissemination in Brazil of Girolando semen based on reliable breeding values, which contributed greatly to genetic improvement in this breed (Silva et al., 2018b).

The Girolando is the breed with the fastest growth in semen production in Brazil, while the average 305-day milk yield increased by 51.29% (3,599 to 5,445 kg) from 2000 to 2016. The genetic evaluations of Girolando animals published yearly, and the recent inclusion of genomic information in the model, allow a more efficient selection of bulls and cows, which may contribute to further advancements in the Girolando breed genetic improvement (Silva et al., 2017, 2018a).

In Brazilian tropical dairy cattle, researchers have been studying milk production and milk composition traits (Canaza-cayo et al., 2016) and associations of these traits to variants of specific gene polymorphisms (Cardoso et al., 2011; Bervian et al., 2019), in addition to reproduction and resistance traits (Canaza-Cayo et al., 2017; Bervian et al., 2018; Otto et al., 2018). However, genome wide association studies (**GWAS**) for milk yield in this breed have not been reported yet. GWAS have been used to identify single nucleotide polymorphisms (**SNPs**) that are associated with QTL with major effect in dairy cattle (Silva et al., 2011; Meredith et al., 2012). Moreover, Post-GWAS analyses, including gene networks and gene ontology enrichment analyses, have provided a better understanding of the molecular mechanisms involved in the evaluated trait and allow the identification of the most likely candidate genes (Soares et al., 2017; Marques et al., 2018).

Breed origin of SNP alleles (BOA) can be assigned in crossbred animals using the so-called BOA approach (Vandenplas et al., 2016). Results from this approach have been used in models that take breed-specific SNP effects into account, to improve genomic prediction, or GWAS (Sevillano et al., 2017, 2018). According to Sevillano et al. (2018), the parental breed origin affects estimated SNP effects and explained variance of SNPs associated with crossbred pigs performance.

The objective of this study, therefore, was to perform a GWAS to identify QTL associated with 305-day milk yield (**305MY**) in the Girolando population and identify the most likely candidate genes through Post-GWAS studies, including gene ontology terms enrichment and gene-transcription factors networks. In addition, we investigated whether estimated SNP-specific variance is affected by the breed of origin.

MATERIALS AND METHODS

Phenotypic data

The phenotypic data consisted of 305MY from Girolando cattle, obtained from herds supervised by the Dairy Control Service or that collaborate with the Program for Genetic Improvement of the Girolando Breed's (PMGG) Sires Progeny Test. The database that was used in this study contained 166,628 lactations, from 94,124 Girolando cows, edited for the age at calving (547 to 9,095 days), calving season (dry or rainy), breed composition, which was determined by the proportions of the Holstein breed (1/4, 3/8, 1/2, 5/8, 3/4, 7/8) and contemporary group (herd-year of calving). Data from cows of age at calving over 9,095 days and lactations suspended by abnormal causes were eliminated; while size of herd and contemporary group of birth herd-year were considered for data inclusion. Records of at least three lactations and the use of at least two bulls for herd-year were considered as pre-requisites for data inclusion. In addition, contemporary groups of cows that participated in dairy tournament in one or more controls of lactation were formed. The pedigree file included 188,615 animals, being 8,725 sires, 70,691 dams, including Holstein and Gir purebred (**PB**), besides of Girolando cattle (**CB**).

Genotyping and quality control

In the genotype quality control analysis, SNP with minor allele frequency ≤ 0.02, maximum difference between observed and expected allele frequency for Hardy Weinberg Equilibrium of 0.15, GenCall score ≤ 0.70 and call rate ≤ 0.98 were excluded from the data set. Samples with a call rate < 0.90 also were excluded. After edits 397,325 SNPs and 6,760 samples remained for the further analyses.

A total of 6,760 animals were genotyped using the different SNP chip density: Illumina BovineSNP50 BeadChip v2 (50K), Illumina BovineHD BeadChip (HD), Zoetis Custom SNP chip ZL2 (20K), and Zoetis Custom SNP chip ZM2 (70K). Cattle genotyped with the 20 K, 50 K or 70K chips were imputed to the HD .v2 panel using FImpute software (Sargolzaei et al., 2014) with default parameter settings and using pedigree information.

Genome-wide association analysis

The single-step GWAS analysis was conducted using the BLUPF90 software family (Misztal et al., 2002). First, the variance components were estimated and GEBV were predicted with AIREMLF90 by applying the following single-trait animal model:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e}$$

in which **y** is the vector of phenotypic observations (305MY); **β** is the vector of fixed effects (i.e., contemporary group, calving season and breed composition) and the covariates (i.e., age at calving, with the linear and quadratic components); **a** is the vector of random additive genetic effects; **p** is the vector of random permanent environmental effects; **e** is the vector of residual effects; and **X, Z** and **W** are the incidence matrices of **β**, **a** and **p**, respectively. It was assumed that $\mathbf{a} \sim N(0, \mathbf{H}\sigma_a^2)$, $\mathbf{p} \sim N(0, \mathbf{I}\sigma_p^2)$ and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_a^2, σ_p^2 and σ_e^2 are the additive genetic, permanent environmental and residual variances, respectively; **H** is a matrix combining genomic and pedigree information, as proposed by Aguilar et al. (2010), and **I** is the identity matrix.

The SNP effects were calculated using postGSf90 software and the GWAS results were reported as the percentage of genetic variance explained by overlapping windows of 0.5-megabase, in which the windows size was determined by a previous linkage

disequilibrium study in a F2 Girolando population (unpublished data). The percentage of additive genetic variance explained by the windows was calculated as described by Wang et al. (2014). A manhattan plot showing the non-overlapping windows which exhibit the highest additive genetic variance was created using the mhtplot function of R software (R version 3.3.1).

Gene Search and Generation of Gene Networks

Within the top 10 windows explaining the highest percentage of additive genetic variance, putative candidate genes were identified based on the UMD 3.1.1 assembly of the bovine genome (Zimin et al., 2009), using the National Center for Biotechnology Information (**NCBI**) Map Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/?org=bos-taurus).

First, the gene set was used to perform the biological processes enrichment analysis to the identification of most likely candidate genes for 305MY. For that, Gene Ontology (**GO**) network highlighting biological roles and gene interactions for the 305MY trait was constructed using the ClueGO application in Cytoscape (Bindea et al., 2009), based on a unilateral hypergeometric test and Bonferroni correction.

Second, the analyses of promoter sequences of candidate genes were performed. For that, promoter sequences (FASTA format) were collected a flanking sequence of 3000 base pairs upstream and 300 base pairs downstream from genes transcription start site in the bovine genome, as used by Soares et al. (2017). These data were submitted to the TFM-explorer software (http://bioinfo.lifl.fr/TFM/TFME/), which uses weighting matrices from JASPAR database (http://jaspar.binf.ku.dk/) (Sandelin et al., 2004) to detect potential transcription factor binding sites (TFBS), as described in Touzet and Varré (2007).

The TF list was then analyzed in Cytoscape software (Shannon et al., 2003) using the Biological Networks Gene Ontology tool (BiNGO) (Maere et al., 2005) to determine significantly overrepresented functional gene ontology terms. Based on the over-represented biological processes identified and literature review, we were able to select the TF most related to 305MY.

The most likely candidate genes were then identified through the gene–TF network, which was built using the Network Analyzer tool in the Cytoscape software, to provide an overview of candidate genes and their predicted TF across the 305MY trait.

Estimation of SNP-allele effects

To estimate the SNP-allele substitution effects the back-solve was performed based on the GEBV of breeds Holstein and Gir for Girolando animals, which were estimated using the GBLUP model with breed-specific partial relationship matrices. SNP-allele effects were derived using the following steps: (1) Determine breed-of-origin of alleles. (2) Build the breed-specific partial relationship matrices, $\mathbf{G^H}$ and $\mathbf{G^G}$. and calculate GEBVs for Girolando using the BOA model. (3) Back-solve SNP-allele effects for Girolando from GEBVs. (4) Calculate SNP-genetic variance.

Assignment of Breed of Origin of Alleles. To infer the breed-of-origin of alleles in the Girolando cattle population, the BOA approach (Vandenplas et al., 2016) was used, assuming the relaxation factor parameter recommended by Sevillano et al. (2016). This analysis consists of three steps: (1) phasing the purebred and crossbred cattle genotypes with AlphaPhase 1.1 software (Hickey et al., 2011). Haplotype phasing was performed using pedigree, the nine combinations haplotypes length used by Sevillano et al. (2016) and Vandenplas1 et al. (2016), and two more combinations of core and tail lengths (500:150,

600:150). All phasing analyses were run considering "Offset" and "NotOffset" modes, allowed each allele to be considered 22 times. The "Offset" mode moves the start of each core to halfway along the first core, creating 50% overlap between cores. (2) determining the unique haplotypes for each pure breed. A haplotype was declared specific for a certain breed if 80% of its copies were observed in this breed. (3) assigning breed origin for each allele observed in Girolando animals based on the determined haplotypes. After performing the BOA approach, SNPs adjacent and located in the candidate genes for 305MY trait.

GBLUP model with breed-specific partial relationship matrices (BOA model). To account for the breed-specific effect of alleles, a model with two breed-specific partial relationship matrices ($\mathbf{G^H}$ and $\mathbf{G^G}$) was fitted (BOA model) (Sevillano et al., 2017):

$$\mathbf{y = X\beta + Zg^H + Zg^G + Wp + e}$$

in which **y** is the vector of phenotypic observations (305MY); **β** is the vector of fixed effects (the same effects used in the single-step model) and **X** is the incidence matrix of fixed effects; $\mathbf{g^H}$ and $\mathbf{g^G}$ are the vectors of the additive genetic effect of PB gametes in CB animals, and **Z** is the incidence matrix; **p** is vector of random permanent environmental effects, and **W** is the incidence matrix; **e** is the vector of residual effects. The variance of the permanent environmental effect and residual effect were:

$$\boldsymbol{Var_p = \sigma^2_p \otimes I},\ \text{and}$$

$$\boldsymbol{Var_e = \sigma^2_e \otimes I}.$$

The variance of additive genetic effect for breed G origin was:

$$\boldsymbol{Var_{g^H} = \sigma^2_{gH} \otimes G^H},$$

where $\mathbf{g^H}$ is a random vector with additive effects from the breed H gametes, and $\mathbf{G^H}$ is a breed-specific partial relationships matrix for breed H. The variance structure for the origin of breed G is defined similarly.

Both breed-specific partial relationship matrices, $\mathbf{G^H}$ and $\mathbf{G^G}$, were built using the breed-of-origin of selected alleles located in the candidate genes for 305MY in Girolando, using the first method from (VanRaden, 2008). The breed-specific partial relationship submatrices were defined, considering e.g. the breed H origin, as:

$$\mathbf{G^H} = (\mathbf{M^H} - \mathbf{C^H D^H})(\mathbf{M^H} - \mathbf{C^H D^H})'\boldsymbol{F^{-1}},$$

where $\mathbf{M^H}$ is a matrix containing Holstein allele content for the Girolando (coded as 0, 1, or 2), if an individual has on a specific locus no Holstein alleles or none of its alleles on this locus were assigned a breed-of-origin, then this genotype was set to zero in the centred matrix represented by $(\mathbf{M^H} - \mathbf{C^H D^H})$; $\boldsymbol{C^H}$ is a matrix containing counts of the number of H alleles per animal and locus, $\mathbf{D^H}$ is a diagonal matrix with breed H specific frequencies of the counted allele (p_j^H) on the diagonals, where p_j^H was calculated across Girolando animals by counting the occurrences of alleles originating from the H breed and coded as 1, divided by the total number of H alleles in the Girolando on locus j. The scaling factor was defined as $F = \sum_j 2p_j^H\left(1 - p_j^H\right)$. The breed-specific partial relationship submatrix $\mathbf{G^G}$ is defined similarly. The ASReml software (Gilmour et al., 2015) was used to estimate the variance components considering only the animals with the genotype and phenotype information. Thus, we applied the BOA model to a Girolando population with 4,310 animals.

Back-solve SNP-allele effects from GEBV. SNP-allele effects ($\hat{\alpha}_H$) were estimated based on the GEBV of breed H for Girolando ($\hat{\mathbf{a}}_\mathbf{H}$) performance similar to Wang et al. (2014), using

$$\hat{\mathbf{a}}_\mathbf{H} = \mathbf{W^H}\hat{\boldsymbol{\alpha}}_\mathbf{H}$$

where W^H contains centered genotypes, which can be obtained respectively by:

$\mathbf{W^H} = (\mathbf{M^H} - \mathbf{C^H D^H})$, and

$$\hat{\boldsymbol{\alpha}}_{\mathbf{H}} = \mathbf{W}^{\mathbf{H}'}\left(\mathbf{W}^{\mathbf{H}}\mathbf{W}^{\mathbf{H}'}\right)^{-\mathbf{1}}\hat{\mathbf{a}}_{\mathbf{H}} = F^{-1}\mathbf{W}^{\mathbf{H}'}\mathbf{G}^{(\mathbf{H})-\mathbf{1}}\hat{\mathbf{a}}_{\mathbf{H}}$$

The SNP-allele effects of breed G were calculated similarly to breed H.

SNP-genetic variance. The genetic variance by each SNP ($\sigma^2_{\mathrm{SNP_H}}$) was estimated based on the SNP-allele effects and allele frequencies of breed H for Girolando as follows:

$$\boldsymbol{\sigma^2_{\mathrm{SNP_H}}} = \mathbf{2p_H q_H}(\hat{\boldsymbol{\alpha}}_{\mathbf{H}})^{\mathbf{2}},$$

where $\mathbf{p_H}$ and $\mathbf{q_H}$ are the allele frequencies and the $\hat{\boldsymbol{\alpha}}_{\mathbf{H}}$ is the allele substitution effect. The SNP-genetic variances of breed G ($\boldsymbol{\sigma^2_{\mathrm{SNP_G}}}$) were calculated similarly to breed H.

RESULTS AND DISCUSSION

Genetics parameter

In the current study, the heritability estimate for 305MY was 0.11 ± 0.005 while the repeatability estimate was 0.33 ± 0.003. In the previous studies with Girolando cattle, the heritability and repeatability estimates were higher than the estimates found in this study, with reported heritability values ranging from 0.15 to 0.35 (Ribeiro et al., 2017; Canaza-Cayo et al., 2017) and repeatability values ranging from 0.37 to 0.47 (Facó et al., 2008, 2009). According to Canaza-Cayo et al. (2017), different factors can influence in the genetic parameter estimates such as population size, analysis model, the evaluated traits and others. In the current study, the main factor that effected the estimations was the population size.

QTL regions

Using the single step GWAS approach, the top 10 identified genome regions associated with 305MY in terms of genetic variance explained, were located on Bos taurus chromosomes (BTA) 1, 2, 4, 5, 6, 7, 16 and 26 (Figure 1), and jointly explained 4.75% of

the additive genetic variance. In a previous study with Italian Holstein cattle performed by Fontanesi et al. (2014), a significant SNP associated with milk yield, was found in the same region on BTA4 identified in the current study. Moreover, in a Holstein-Friesian cattle population, three significant SNP associated with the same trait were identified close to QTL on the BTA6 and 26 of this study (Meredith et al., 2012). No evidence was found for additional significant SNPs or regions in the QTL regions identified for 305MY in the current study.

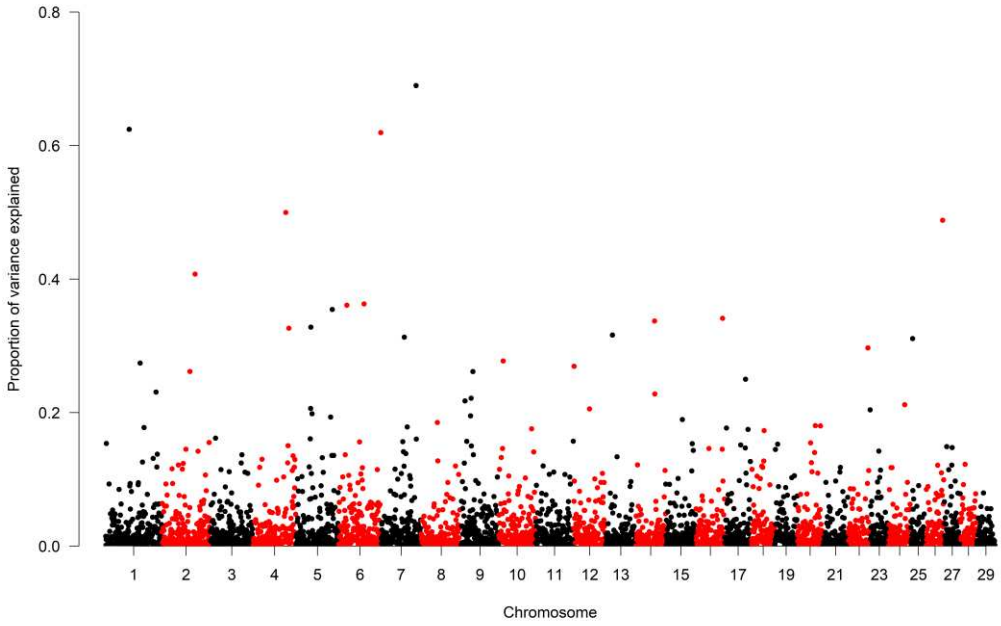


Figure 1. Manhattan Plot of additive genetic variance explained by non-overlapping windows in the Girolando population evaluated for milk yield in 305 days (305MY). Each dot denotes a window.

Candidate genes

Putative candidate genes within the top 10 windows selected were identified based on the UMD 3.1 assembly of the bovine genome. To better understand their functions, we

grouped the genes according to their GO terms enrichment (Figures 2). In this GO network, we identified genes involved in regulation of hormone secretion, the LEP and CLOCK gene, a gene associated with regulation of ATPase activity, the CASR, and a gene involved in negative regulation of Janus kinase/signal transducers and activators of transcription (JAK/STAT) cascade, the LRRC4.

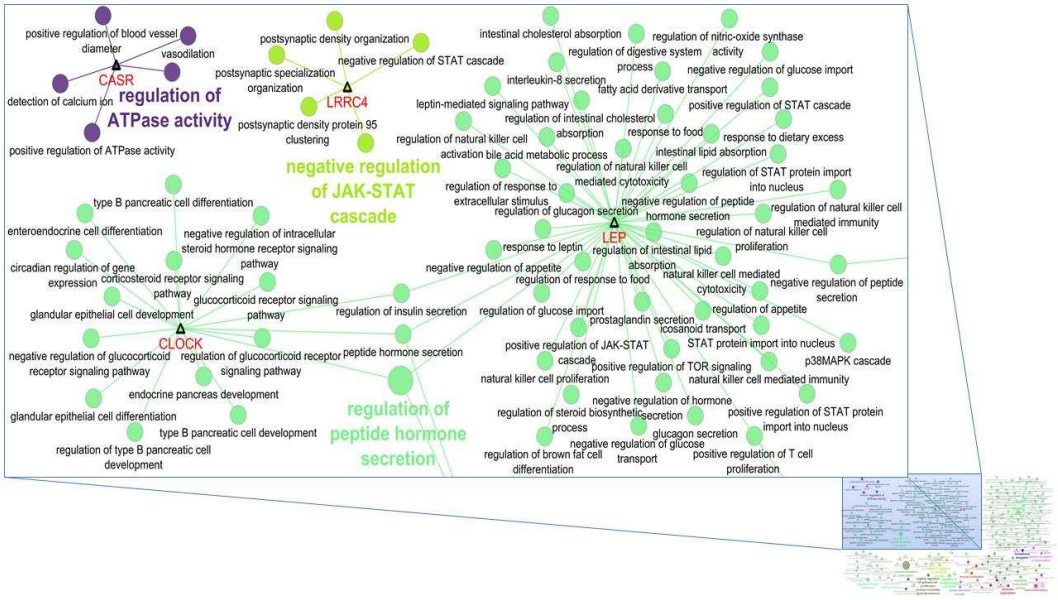


Figure 2. Functional networks showing gene interactions (triangle nodes) related to milk yield in 305 days (305MY) and the relationship across gene and their sub-networks related to regulation of hormone secretion, of ATPase activity and JAK/STAT cascade in the Girolando population. The node size represents the term enrichment significance from the ClueGO Cytoscape plug-in. The most enriched terms per group are shown in bold and in black letters the sub biological process terms.

The Leptin gene (LEP), encodes leptin, a hormone secreted by adipocytes which plays an essential role in the regulation of energy balance, fertility and milk production (Silva et al., 2002; Liefers et al., 2005). In dairy cattle, leptin concentrations increase throughout pregnancy and are affected by polymorphisms in the leptin receptors gene (LEPR) (Ingvartsen and Boisclair, 2001; Liefers et al., 2004). In previous studies, several SNP in the LEP gene had significant association with 305-day milk yield (Banos et al.,

2008; Clempson et al., 2011; Kiyici et al., 2018). Evidence of dominance effects on 305-day milk yield was found for a SNP in the LEP gene by Kiyici et al. (2018). Also, significant dominance effects on fat and protein content were described for a SNP in the LEP gene by Suchocki et al. (2010). The LEP also modulates some DGAT1 functions and suppresses the expression of DGAT2 (Chen et al., 2002; Suzuki et al., 2005). Thus, the LEP gene is considered to be an important candidate gene for milk yield of dairy cattle.

The Clock Circadian Regulator (CLOCK) is one of the core circadian clock genes, which is regulated in the mammary, liver and adipose tissue, where during the transition from pregnancy to lactation, its expression is increased (Casey et al., 2009; Casey and Plaut, 2012). The CLOCK gene products make up core clock elements that generate circadian rhythms that drive rhythmic expression of numerous genes (Casey et al., 2009). During lactation the cow's circadian system is responsible to initiate and sustain lactation, through the regulating of the hormone physiology changes and of the increased metabolic demand needed to support milk synthesis (Casey and Plaut, 2012; Casey et al., 2014b). The decrease or disrupting of CLOCK levels can result in loss of milk production (Casey et al., 2016) and decrease milk-fat yield in dairy cattle (Casey et al., 2014a). This gene also plays an important role in the regulation of epithelial cell growth and differentiation in the mammary gland (Casey et al., 2016).

The calcium sensing receptor (CASR) is a guanine nucleotide-binding protein-coupled, cell surface receptor (GPCR) (Brown et al., 1993) and is highly expressed in mice breast epithelial cells during lactation (Vanhouten et al., 2004). The CASR gene promotes calcium transport into milk and coordinates the secretion of the parathyroid hormone-related protein (PTHrP) into milk and into the blood circulation, which is responsible for activation of bone resorption of maternal skeletal calcium stores (Wysolmerski, 2012;

Mamillapalli et al., 2013; VanHouten and Wysolmerski, 2013). Furthermore, it has important functions in the maintenance of blood circulating calcium and in the bones, adjusting the milk calcium transport according to the systemic availability of calcium and stimulating the PTHrP secretion when necessary. Thus, the CASR ensures a steady supply of calcium for milk production and protects the body against hypocalcemia, which may decrease overall milk production (Vanhouten et al., 2004; VanHouten, 2005; Kim and Wysolmerski, 2016).

The Leucine Rich Repeat Containing 4 (LRRC4) is involved in the negative regulation of JAK/STAT cascade (Wu et al., 2006). The STAT family members play important functions in lactogenesis, mammary epithelial cells proliferation, differentiation and apoptosis, in different mammary developmental stages (Haricharan and Li, 2014). In the mice mammary gland, the JAK/STAT signaling pathway is the main regulator of milk protein expression (Yang et al., 2000), however, according to Bionaz and Loor (2011), the JAK2/STAT5 signaling plays a minor role in milk protein synthesis in dairy cattle. The STAT5a is responsible for development of the mammary gland, during pregnancy and lactation, and alveolar progenitor cell survival (Yamaji et al., 2009; Dong et al., 2010). In dairy cattle, SNPs in the STAT5a were associated with fertilization success, embryonic survival, and milk production traits, being potential DNA markers for improving milk yield at 305 days and milk composition (Khatib et al., 2008; He et al., 2012; Raven et al., 2014). In the mammary involution stage, the STAT5 is inactivated and STAT3 is activated. The STAT3 is responsible for involution of this gland through inducing apoptosis and later mammary epithelial cell clearing (Kritikou et al., 2003; Hughes et al., 2012). However, in bovine, the mammary involution stage is an important period to both cell death and epithelial cell renewal, therefore most part of alveolar structure is maintained. The

apoptosis accompanied by an increase in cell proliferation, allows the cell turnover to next lactation (Capuco and Akers, 1999; Rowson et al., 2012).

The gene set also was used to perform an analysis of regulatory sequences and after identification of the TF most related to 305MY (Table 1), we generated a gene-TF interaction network, which allowed us to evaluate the genes linked to them (Figure 3). The E74-like factor 5 (ELF5) and STAT1 were the TF most enriched in the gene-TF network (Figure 3). The ELF5 is a member of the ETS family of transcription factors and is found in the epithelial layer of different tissues (Zhou et al., 1998). This TF is expressed by mammary luminal progenitor cells, where during the pregnancy and lactation, its expression is increased (Harris et al., 2006; Oakes et al., 2008; Bionaz and Loor, 2011). The ELF5 TF is a key regulator in the secretory alveolar epithelium differentiation during mammary morphogenesis and in lobuloalveolar development in the mammary gland (Harris et al., 2006). The ELF5's expression is induced by progesterone and working together, both have a bigger impact on the differentiation of progenitor cells when compared to their separate impacts (Lee et al., 2013). This TF also is a key member of the set of prolactin-regulated genes, able to restore lobuloalveolar development and milk production in prolactin-deficient mammary epithelium in vivo (Harris et al., 2006). In mice's virgin mammary epithelium, forced expression of ELF5 was responsible for alveolar differentiation of mammary epithelium and precocious milk secretion during pregnancy, and its absence in the body resulted in failed alveolar morphogenesis and lactation (Oakes et al., 2008). In cattle and murine mammary glands, the insulin plays a major role on milk protein synthesis via induction of ELF5 (Menzies et al., 2010; Bionaz and Loor, 2011; Sigl et al., 2014). As such, the ELF5 TF has an import role on milk and milk protein synthesis during lactation and mammary epithelium differentiation.

Table 1. Enriched transcription factors (TF) associated with genes identified for milk yield in 305 days (305MY), based on their gene ontology biological process and literature review.

TF[1]	Biological Process	Literature review[2]
ELF5	Mammary gland epithelial cell differentiation	Key regulator in the secretory alveolar epithelium differentiation and in lobuloalveolar development in the mammary gland (Harris et al., 2006)
STAT1	JAK-STAT cascade	SNPs in the STAT1 were associated with the milk production traits (Rychtářová et al., 2014; Ozden Cobanoglu et al., 2016).
STAT3	JAK-STAT cascade	Responsible for involution of mammary gland through inducing apoptosis and later mammary epithelial cell clearing (Kritikou et al., 2003; Hughes et al., 2012).

[1]ELF5 = E74-like factor 5; STAT1 = signal transducer and activator of transcription family 1; STAT3 = signal transducer and activator of transcription family 3.

[2]The cited literature studies are just a sample of the vast available literature.

The STAT1 is involved in all stages of mammary gland development and it is expressed in the mature virgin mammary gland before pregnancy and after involution, during remodeling of the gland (Hughes and Watson, 2012; Haricharan and Li, 2014). The STAT1 and STAT3 expression can be regulated by growth hormone, cytokines, progesterone and prolactin (Boutinaud and Jammes, 2004; Haricharan and Li, 2014; Leehy et al., 2018). The STAT1 also is involved in milk production traits. Previous studies described significant association between SNPs in the STAT1 with milk production, milk fat and protein yields and percentages in different dairy cattle breeds (Cobanoglu et al., 2006; Rychtářová et al., 2014; Ozden Cobanoglu et al., 2016). A study with buffaloes also found associations between the STAT1 and protein percentage and 305-d milk yield,

providing evidence that STAT1 is important to milk production traits in different species (Deng et al., 2016).

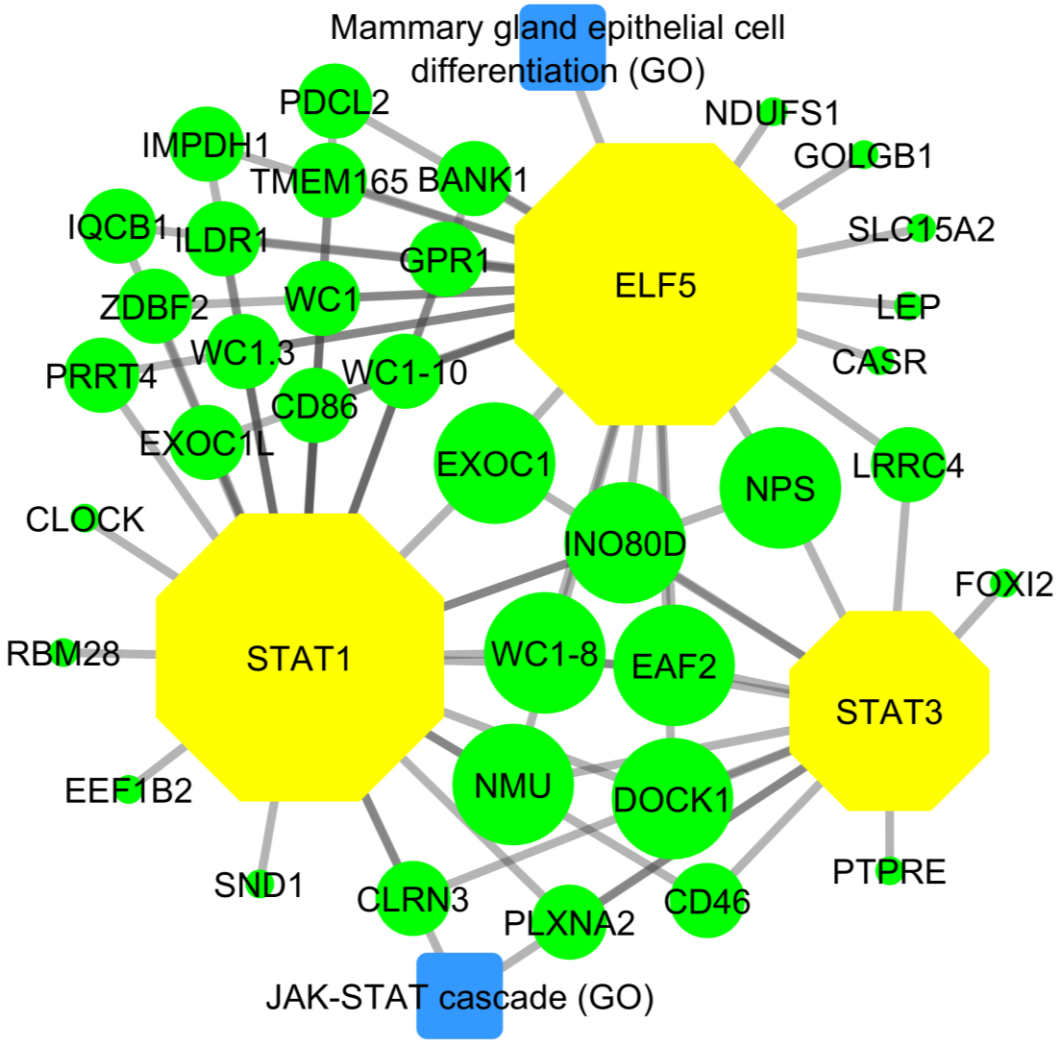


Figure 3. Gene-transcription factor (TF) network: genes located in the top 10 windows for milk yield in 305 days (305MY) (green circle nodes) and their associated TF (yellow octagon nodes). The node size corresponds to the network analyses (Cytoscape, Shannon et al., 2003) in which the larger nodes denotes a higher edge density associated with the number of TF binding sites. Blue square nodes show the gene ontology biological processes related to TF.

The STAT3 TF, showed a low number of connections in the gene-TF network (Figure 3). It is a member of the STAT family and it is responsible for mammary gland involution after the lactation period (Chapman et al., 1999; Sargeant et al., 2014). In the first phase involution, this TF is a key transcriptional regulator of genes associated with innate immunity and in the second phase, it plays a role in wound healing (Hughes et al., 2012). The STAT3 expression was observed during gestation, lactation and involution (Philp et al., 1996) and in low levels in response to altered milking frequency in dairy cattle (Murney et al., 2015). During bovine mammary involution, the balance between STAT5 and STAT3 signaling plays a role in the regulation of the transition from lactation to involution. At the onset of involution of the mammary gland, the STAT5 is downregulated at the same time the STAT3 is upregulated (Singh et al., 2016, 2017). The negative relationship between STAT5 and STAT3 activation can to plays an important role in the preventing further STAT5 activation when the milk production is shuts down (Granillo et al., 2007). According recent studies, when involution is experimentally induced by abrupt cessation of milking, increased expression of STAT3 can to detected 72 h postmilking, being able it occurs at different rates between cows, including between quarters within mammary gland (Singh et al., 2016, 2017).

From the gene-TF network, candidate genes for 305MY trait were highlighted based on TFBS. The dedicator of cytokinesis 1 (DOCK1), also known as DOCK180, was one of gene most enriched in this network. This gene plays an important role in the initiation of mammary gland involution and epithelial phagocytes to mediate dead cell clearance in vitro and in vivo (Bagci et al., 2014). Previous studies described an impaired STAT3 activation at the onset of mammary involution in mutant mice with DOCK1 inactivated, which resulted in delay in the initiation and progression of morphological changes associated with

involution and impairing of phagocytosis of apoptotic cells (Akakura et al., 2004; Bagci et al., 2014). Mammary gland involution is essential to ensure that the tissue conserves its ability to initiate another lactation cycle.

We also identified the SLC15A2 and SND1 genes in the gene-TF network, although not enriched in the network, they present important roles for milk production. The Solute Carrier Family 15 Member 2 (SLC15A2), also known as oligopeptide transporter 2 (PEPT2), is one of the main transporters for peptides (Kim et al., 1972). The PEPT2 is responsible for independent transport of several different short-chain dipeptides and tripeptides and a variety of pharmacologically important compounds (Daniel and Kottra, 2004). According to Groneberg et al. (2002), the PEPT2 mRNA and protein are expressed at high levels in lactating rat mammary gland epithelial cells and in human epithelial cells collected from breast milk. The luminal expression of PEPT2 provides an efficient pathway for reuptake of short-chain peptides derived from hydrolysis of milk proteins secreted into the lumen, contributing to a reduction in drug disposition in the milk (Groneberg et al., 2002). Later studies demonstrated that the PEPT2 mRNA and protein also are expressed in lactating bovine mammary glands and cultured mammary epithelial cells, and that in the presence of lactogenic hormones (prolactin, hydrocortisone, and insulin) the PEPT2 gene expression was significantly enhanced (Zhou et al., 2011; Cui et al., 2017). In the bovine mammary gland, the reabsorption of peptides is very important for milk protein synthesis, and the PEPT2 is essential to maintain the peptide homoeostasis and reducing drug contamination in milk, especially during lactation (Zhou et al., 2011). Thus, the PEPT2 gene can influence milk production in cattle (Yang et al., 2015; Cui et al., 2017).

The staphylococcal nuclease domain containing 1 (SND1), also known as Tudor staphylococcal nuclease (Tudor-SN) or p100, encodes the SND1 protein, a multifunctional

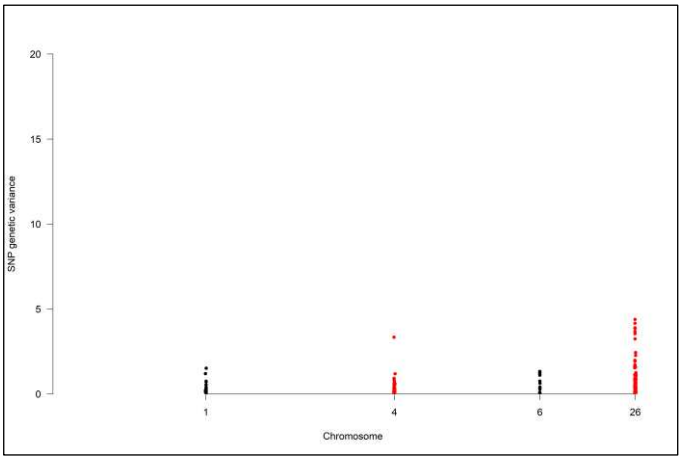
highly conserved protein, with homologues found in the genomes of 21 diverse eukaryotes (Broadhurst et al., 2005). This protein was initially identified in HeLa cells and its conservation throughout evolution suggests SND1 protein is an essential transcriptional coactivator and it has important physiological functions (Tong et al., 1995). The SND1 gene expression was observed in mice and bovine in different tissues, and it can be upregulated in methionine-promoted milk synthesis and cell proliferation, and in response to lactogenic hormones during lactation (Broadhurst et al., 2005; Lu et al., 2012; Fashe et al., 2013). The SND1 is responsible to positively regulate STAT5, mTOR, SREBP-1, and Cyclin D1 signaling pathways, which play an important role in milk synthesis and proliferation of bovine mammary epithelial cells (Ao et al., 2015), in the b-casein expression (Broadhurst and Wheeler, 2001) and the loop in prolactin-induced transcription, cooperating with STAT5 in transcriptional activation (Paukku et al., 2003). As such, the SDN1 is a very important gene for milk synthesis and proliferation of bovine mammary epithelial cells (Lu et al., 2012; Ao et al., 2015).

SNP-genetic variance

Post-GWAS analyses allowed the identification of seven candidate genes associated with 305MY, which were located on BTA1, 4, 6 and 26. After performing the BOA approach for these chromosomes, the SNPs located in these genes were selected and their alleles used to estimate breed-specific effect of alleles. Based on this, we observed that the SNP-specific variance varied across breeds and genes. Within the seven candidate genes, the alleles coming from the Gir breed-of-origin showed larger genetic variance than alleles coming from the Holstein breed-of-origin (Figure 4). Alleles frequency, SNP-specific effects and SNP-specific variances by breed are detailed in Additional files 1. Large

variations were observed in SNP-specific variances for both breeds, ranging from 0.01 to 18.85 and from 0.01 to 4.39 when alleles came from the Gir and Holstein breed, respectively. These results suggest that the SNP-specific variances in the candidate genes for 305MY in Girolando depends upon the breed-of-origin.

According to Sevillano et al. (2018), estimated effects and explained variance of SNPs associated with crossbred pigs performance depends on the breed of origin of the alleles. In the current study, the main difference across breeds are the SNP-effects which is affected to the additive genetic variance (Additional files 1). A larger additive genetic variance for 305MY was estimated for the Gir breed. The difference in SNP-specific between alleles coming from the Gir and Holstein breed can be the result of genetic selection. The selection for one specific allele or trait in the subsequent generations, can result in fixation of alleles and substantial losses in genetic diversity, which can be intensified by the use of genomic selection (Bulmer, 1971; Grevenhof et al., 2012; Engelsma et al., 2014). Given the fact that the Gir breed didn't underwent as intense a selection process for milk production as the Holstein breed, the genetic variability for this trait is highest in Gir, which can explain the higher SNP-specific variance of evaluated alleles coming from the Gir breed.



a

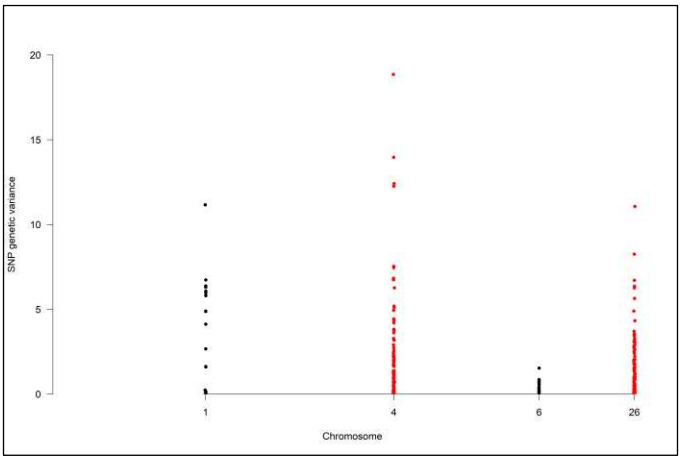
b 

Figure 4. SNP-genetic variance for milk yield in 305 days (305MY) when alleles originate from Holstein (a) or Gir (b) breed. Each dot denotes a SNP.

This was the first study to evaluate 305MY in a Girolando population using GWAS and post-GWAS analyses. From ten regions explaining most additive genetic variance for 305MY in Girolando cattle, candidate genes were identified on the chromosomes 1, 4, 6 and 26. A total of seven most likely candidate genes (LEP, CLOCK, CASR, LRRC4, DOCK1, SLC15A2 and SND1), which play important roles in different stages of milk synthesis, were identified based on the gene-TF networks and gene ontology enrichment analyses. Moreover, we observed that alleles coming from the Gir breed-of-origin showed larger genetic variance than alleles coming from the Holstein breed-of-origin. These results suggest that the SNP-genetic variance of alleles on the candidate genes for 305MY in Girolando animals depends upon from which parental breed they were inherited. The results reported in the current study are valuable to QTL mapping in the Girolando breed and combined with the candidate genes can be implemented in breeding programs by genomic selection.

Additional file 1: of Effects of SNP alleles located in the candidate genes (LEP, CLOCK, CASR, LRRC4, DOCK1, SLC15A2 and SND1) associated with 305MY in Girolando cattle estimated for genomic prediction depend on their breed-of-origin.

SNP	Effect H[1]	Effect G[2]	2pq H	2pq G	Variance H	Variance G
SLC15A2						
1	0.88	0.97	0.23	0.24	0.21	0.23
2	3.34	0.01	0.36	0.33	1.19	0.00
3	0.21	23.19	0.48	0.48	0.10	11.16
CASR						
1	0.50	0.32	0.64	15.30	0.32	4.89
2	0.49	0.50	0.19	12.71	0.10	6.30
3	0.50	0.49	0.37	12.91	0.18	6.37
4	0.06	0.02	0.04	0.08	0.00	0.00
5	0.50	0.31	1.07	5.26	0.54	1.62
6	0.50	0.32	0.60	15.24	0.30	4.88
7	0.50	0.47	0.47	5.65	0.23	2.66
8	0.50	0.49	0.30	12.30	0.15	6.07
9	0.50	0.49	0.37	13.65	0.19	6.73
10	0.50	0.49	0.26	12.11	0.13	5.97
11	0.50	0.49	0.28	11.78	0.14	5.80
12	0.36	0.50	0.37	8.32	0.13	4.12
13	0.36	0.34	0.43	0.43	0.15	0.14
14	0.50	0.30	1.44	0.23	0.71	0.07
15	0.26	0.02	1.03	0.03	0.27	0.00
16	0.49	0.34	1.51	0.25	0.75	0.09
17	0.43	0.47	0.90	3.37	0.39	1.60
18	0.24	0.08	6.37	1.73	1.51	0.13
19	0.09	0.01	0.13	0.00	0.01	0.00
20	0.09	0.02	0.20	0.00	0.02	0.00
SND1 and LRRC4						
1	0.48	0.10	0.48	0.10	0.09	0.11
2	0.48	0.11	0.48	0.11	0.06	0.16

3	0.49	0.25	0.49	0.25	0.00	0.07
4	0.48	0.28	0.48	0.28	0.09	0.08
5	0.47	0.10	0.47	0.10	0.14	0.17
6	0.47	0.10	0.47	0.10	0.09	0.22
7	0.47	0.10	0.47	0.10	0.09	0.22
8	0.48	0.10	0.48	0.10	0.08	0.22
9	0.47	0.10	0.47	0.10	0.16	0.17
10	0.49	0.18	0.49	0.18	0.03	0.48
11	0.49	0.25	0.49	0.25	0.02	0.03
12	0.49	0.32	0.49	0.32	0.00	0.19
13	0.49	0.32	0.49	0.32	0.00	0.19
14	0.49	0.32	0.49	0.32	0.00	0.19
15	0.49	0.50	0.49	0.50	0.07	18.85
16	0.48	0.13	0.48	0.13	0.02	0.22
17	0.48	0.36	0.48	0.36	0.08	1.70
18	0.49	0.33	0.49	0.33	0.02	0.72
19	0.19	0.01	0.19	0.01	0.20	0.00
20	0.19	0.38	0.19	0.38	0.30	0.93
21	0.49	0.40	0.49	0.40	0.37	1.34
22	0.42	0.20	0.42	0.20	0.00	0.24
23	0.23	0.29	0.23	0.29	0.03	2.20
24	0.48	0.46	0.48	0.46	0.92	0.05
25	0.50	0.22	0.50	0.22	3.33	0.52
26	0.44	0.00	0.44	0.00	0.04	0.00
27	0.44	0.00	0.44	0.00	0.08	0.00
28	0.48	0.29	0.48	0.29	0.49	2.34
29	0.48	0.40	0.48	0.40	0.62	1.26
30	0.44	0.42	0.44	0.42	0.00	2.90
31	0.48	0.29	0.48	0.29	0.52	2.46
32	0.44	0.39	0.44	0.39	0.00	6.74
33	0.44	0.39	0.44	0.39	0.00	6.83
34	0.48	0.50	0.48	0.50	0.83	0.13

35	0.42	0.00	0.42	0.00	0.11	0.00
36	0.05	0.40	0.05	0.40	0.00	1.20
37	0.47	0.45	0.47	0.45	0.30	0.00
38	0.23	0.21	0.23	0.21	0.07	0.68
39	0.46	0.21	0.46	0.21	0.00	1.00
40	0.50	0.30	0.50	0.30	0.00	0.97
41	0.48	0.15	0.48	0.15	0.61	0.32
42	0.45	0.15	0.45	0.15	0.00	0.26
43	0.48	0.15	0.48	0.15	0.80	0.28
44	0.45	0.15	0.45	0.15	0.00	0.26
45	0.45	0.15	0.45	0.15	0.00	0.27
46	0.45	0.15	0.45	0.15	0.01	0.27
47	0.26	0.49	0.26	0.49	0.01	3.28
48	0.45	0.15	0.45	0.15	0.00	0.28
49	0.36	0.43	0.36	0.43	0.02	4.42
50	0.45	0.15	0.45	0.15	0.00	0.28
51	0.36	0.44	0.36	0.44	0.02	4.94
52	0.46	0.42	0.46	0.42	0.12	13.96
53	0.36	0.19	0.36	0.19	0.09	0.24
54	0.49	0.46	0.49	0.46	0.15	0.14
55	0.50	0.32	0.50	0.32	0.26	0.01
56	0.49	0.46	0.49	0.46	0.07	1.65
57	0.49	0.47	0.49	0.47	0.03	1.95
58	0.49	0.46	0.49	0.46	0.05	1.65
59	0.49	0.46	0.49	0.46	0.13	1.77
60	0.33	0.33	0.33	0.33	0.10	0.69
61	0.49	0.30	0.49	0.30	0.00	0.76
62	0.50	0.36	0.50	0.36	0.72	2.01
63	0.49	0.45	0.49	0.45	0.00	2.17
64	0.49	0.45	0.49	0.45	0.00	1.65
65	0.50	0.42	0.50	0.42	0.24	3.61
66	0.49	0.41	0.49	0.41	0.04	7.45

67	0.49	0.30	0.49	0.30	0.00	0.18
68	0.49	0.30	0.49	0.30	0.02	0.10
69	0.48	0.30	0.48	0.30	0.00	0.09
70	0.48	0.30	0.48	0.30	0.01	0.11
71	0.48	0.40	0.48	0.40	0.00	0.07
72	0.48	0.49	0.48	0.49	0.04	2.01
73	0.48	0.47	0.48	0.47	0.02	7.53
74	0.48	0.30	0.48	0.30	0.00	0.17
75	0.48	0.41	0.48	0.41	0.01	1.09
76	0.49	0.36	0.49	0.36	0.12	3.79
77	0.49	0.27	0.49	0.27	0.25	0.74
78	0.48	0.45	0.48	0.45	0.00	0.44
79	0.48	0.45	0.48	0.45	0.03	0.44
80	0.50	0.28	0.50	0.28	0.17	0.00
81	0.48	0.45	0.48	0.45	0.00	0.46
82	0.49	0.30	0.49	0.30	0.12	1.90
83	0.49	0.45	0.49	0.45	0.01	0.05
84	0.50	0.50	0.50	0.50	0.01	2.72
85	0.50	0.49	0.50	0.49	0.01	0.00
86	0.50	0.49	0.50	0.49	0.01	0.00
87	0.50	0.45	0.50	0.45	0.06	0.00
88	0.49	0.19	0.49	0.19	0.51	0.08
89	0.46	0.42	0.46	0.42	0.08	0.83
90	0.46	0.42	0.46	0.42	0.13	0.80
91	0.46	0.42	0.46	0.42	0.13	0.80
92	0.46	0.42	0.46	0.42	0.11	0.80
93	0.44	0.49	0.44	0.49	0.24	0.81
94	0.46	0.42	0.46	0.42	0.11	0.80
95	0.42	0.44	0.42	0.44	0.07	0.31
96	0.42	0.44	0.42	0.44	0.09	0.34
97	0.46	0.33	0.46	0.33	0.25	0.41
98	0.38	0.34	0.38	0.34	0.00	0.29

99	0.46	0.44	0.46	0.44	0.18	0.37
100	0.46	0.44	0.46	0.44	0.17	0.37
101	0.41	0.20	0.41	0.20	0.01	0.41
102	0.45	0.20	0.45	0.20	0.28	0.25
103	0.46	0.39	0.46	0.39	0.11	0.75
104	0.41	0.18	0.41	0.18	0.05	0.62
105	0.46	0.39	0.46	0.39	0.10	0.94
106	0.41	0.33	0.41	0.33	0.04	0.64
107	0.41	0.48	0.41	0.48	0.14	0.00
108	0.46	0.48	0.46	0.48	0.18	12.25
109	0.46	0.47	0.46	0.47	0.08	0.22
110	0.41	0.29	0.41	0.29	0.01	4.20
111	0.45	0.29	0.45	0.29	0.15	3.82
112	0.45	0.29	0.45	0.29	0.15	3.82
113	0.45	0.37	0.45	0.37	0.10	0.19
114	0.44	0.39	0.44	0.39	0.10	0.41
115	0.44	0.38	0.44	0.38	0.10	0.26
116	0.47	0.39	0.47	0.39	0.48	0.35
117	0.48	0.38	0.48	0.38	0.59	0.22
118	0.47	0.37	0.47	0.37	0.00	0.34
119	0.47	0.37	0.47	0.37	0.00	0.24
120	0.38	0.28	0.38	0.28	0.08	1.63
121	0.50	0.27	0.50	0.27	0.11	2.55
122	0.46	0.20	0.46	0.20	0.07	1.38
123	0.50	0.36	0.50	0.36	0.20	2.23
124	0.50	0.36	0.50	0.36	0.16	2.23
125	0.50	0.32	0.50	0.32	0.10	1.25
126	0.50	0.36	0.50	0.36	0.11	1.28
127	0.50	0.36	0.50	0.36	0.10	1.28
128	0.50	0.43	0.50	0.43	0.21	4.35
129	0.50	0.47	0.50	0.47	0.18	1.14
130	0.50	0.30	0.50	0.30	0.08	0.90

131	0.50	0.30	0.50	0.30	0.08	0.90
132	0.49	0.41	0.49	0.41	0.34	1.34
133	0.50	0.36	0.50	0.36	0.10	1.71
134	0.50	0.36	0.50	0.36	0.10	1.71
135	0.50	0.30	0.50	0.30	0.07	1.04
136	0.46	0.21	0.46	0.21	0.29	1.95
137	0.46	0.21	0.46	0.21	0.24	1.95
138	0.46	0.45	0.46	0.45	0.17	0.04
139	0.46	0.45	0.46	0.45	0.18	0.04
140	0.46	0.21	0.46	0.21	0.23	2.10
141	0.46	0.21	0.46	0.21	0.23	2.10
142	0.46	0.21	0.46	0.21	0.23	2.10
143	0.50	0.34	0.50	0.34	0.06	0.05
144	0.46	0.21	0.46	0.21	0.26	2.10
145	0.46	0.21	0.46	0.21	0.26	2.10
146	0.47	0.44	0.47	0.44	0.17	0.01
147	0.47	0.44	0.47	0.44	0.16	0.01
148	0.47	0.44	0.47	0.44	0.17	0.01
149	0.47	0.47	0.47	0.47	0.08	5.18
150	0.47	0.49	0.47	0.49	0.08	3.74
151	0.47	0.49	0.47	0.49	0.08	3.17
152	0.46	0.06	0.46	0.06	0.16	0.00
153	0.47	0.47	0.47	0.47	0.25	12.41
154	0.46	0.50	0.46	0.50	0.30	0.14
155	0.50	0.46	0.50	0.46	0.08	5.11
LEP						
1	0.47	0.46	1.27	13.76	0.60	6.26
2	0.34	0.40	3.48	1.81	1.18	0.71
CLOCK						
1	0.45	0.31	0.04	1.05	0.02	0.33
2	0.45	0.48	0.21	0.00	0.09	0.00
3	0.11	0.47	0.00	1.55	0.00	0.73

4	0.11	0.47	0.00	1.55	0.00	0.73
5	0.48	0.48	0.07	0.23	0.04	0.11
6	0.48	0.48	0.05	0.47	0.02	0.22
7	0.47	0.50	2.64	0.00	1.24	0.00
8	0.47	0.50	2.78	0.06	1.31	0.03
9	0.47	0.50	2.73	0.02	1.29	0.01
10	0.47	0.50	2.69	0.02	1.26	0.01
11	0.47	0.50	2.72	0.17	1.28	0.08
12	0.47	0.50	2.79	0.02	1.31	0.01
13	0.49	0.48	1.53	0.01	0.75	0.00
14	0.47	0.49	2.44	0.43	1.15	0.21
15	0.11	0.47	0.00	1.42	0.00	0.67
16	0.47	0.50	2.82	0.19	1.32	0.10
17	0.47	0.50	2.82	0.19	1.32	0.10
18	0.47	0.50	2.77	0.38	1.30	0.19
19	0.47	0.50	2.77	0.38	1.30	0.19
20	0.47	0.49	2.35	0.83	1.10	0.40
21	0.49	0.50	0.67	0.23	0.33	0.11
22	0.49	0.27	0.57	3.11	0.28	0.85
23	0.49	0.48	0.67	0.04	0.33	0.02
24	0.49	0.50	0.65	1.35	0.32	0.67
25	0.49	0.50	0.68	1.11	0.34	0.55
26	0.49	0.50	0.82	1.56	0.40	0.78
27	0.49	0.48	0.58	0.02	0.29	0.01
28	0.49	0.48	0.70	0.41	0.34	0.20
29	0.47	0.50	0.06	0.01	0.03	0.00
30	0.47	0.44	0.02	0.05	0.01	0.02
31	0.47	0.44	0.03	0.36	0.01	0.16
32	0.47	0.44	0.03	0.49	0.01	0.22
33	0.47	0.44	0.04	0.53	0.02	0.24
34	0.47	0.44	0.04	0.46	0.02	0.20
35	0.47	0.44	0.04	0.53	0.02	0.24

36	0.37	0.48	1.68	3.20	0.63	1.53
DOCK1						
1	0.32	0.01	4.76	0.05	1.54	0.00
2	0.48	0.50	3.31	0.22	1.59	0.11
3	0.47	0.31	2.41	0.41	1.13	0.13
4	0.46	0.49	1.62	4.05	0.74	2.00
5	0.00	0.00	0.00	0.00	0.00	0.00
6	0.23	0.01	0.49	0.06	0.11	0.00
7	0.23	0.01	0.71	0.06	0.16	0.00
8	0.39	0.10	1.14	2.70	0.44	0.27
9	0.28	0.12	0.02	0.00	0.01	0.00
10	0.37	0.10	1.25	2.55	0.46	0.26
11	0.22	0.11	0.08	0.08	0.02	0.01
12	0.37	0.10	1.26	2.81	0.47	0.29
13	0.34	0.33	2.66	0.32	0.91	0.11
14	0.34	0.33	2.53	0.31	0.87	0.10
15	0.50	0.25	0.11	19.40	0.05	4.90
16	0.18	0.00	2.16	0.01	0.40	0.00
17	0.00	0.00	0.02	0.01	0.00	0.00
18	0.50	0.26	0.35	7.39	0.18	1.93
19	0.44	0.50	0.73	0.00	0.33	0.00
20	0.50	0.26	0.32	7.67	0.16	2.01
21	0.50	0.43	0.35	6.37	0.17	2.76
22	0.44	0.32	1.19	0.20	0.52	0.06
23	0.49	0.43	1.18	1.11	0.58	0.47
24	0.44	0.27	0.02	2.66	0.01	0.72
25	0.29	0.25	0.01	2.21	0.00	0.55
26	0.46	0.25	0.86	3.43	0.39	0.87
27	0.00	0.00	0.01	0.02	0.00	0.00
28	0.36	0.27	0.15	0.23	0.06	0.06
29	0.28	0.25	0.07	2.80	0.02	0.69
30	0.41	0.01	2.62	0.03	1.08	0.00

31	0.48	0.37	1.26	0.29	0.60	0.11
32	0.48	0.50	0.03	0.04	0.01	0.02
33	0.49	0.50	1.14	7.39	0.56	3.70
34	0.50	0.29	3.85	3.44	1.92	0.99
35	0.50	0.47	3.45	0.03	1.71	0.02
36	0.41	0.00	2.77	0.01	1.14	0.00
37	0.47	0.44	4.22	6.25	1.98	2.73
38	0.47	0.50	3.56	6.17	1.66	3.08
39	0.48	0.50	8.08	3.62	3.87	1.81
40	0.48	0.46	7.42	6.14	3.54	2.85
41	0.27	0.29	0.06	0.00	0.02	0.00
42	0.48	0.50	7.68	3.53	3.68	1.76
43	0.48	0.50	8.05	2.95	3.85	1.47
44	0.48	0.50	8.12	2.70	3.89	1.35
45	0.48	0.50	8.68	2.89	4.16	1.44
46	0.48	0.46	6.78	5.48	3.24	2.54
47	0.46	0.29	3.43	0.03	1.59	0.01
48	0.48	0.13	9.23	0.76	4.39	0.10
49	0.16	0.01	0.55	0.01	0.09	0.00
50	0.16	0.00	0.46	0.03	0.08	0.00
51	0.16	0.12	0.52	0.59	0.08	0.07
52	0.16	0.12	0.48	0.59	0.08	0.07
53	0.18	0.50	0.72	3.05	0.13	1.52
54	0.16	0.00	0.53	0.00	0.09	0.00
55	0.16	0.00	0.60	0.02	0.10	0.00
56	0.16	0.00	0.53	0.02	0.09	0.00
57	0.16	0.00	0.53	0.02	0.09	0.00
58	0.16	0.00	0.53	0.02	0.09	0.00
59	0.16	0.00	0.53	0.02	0.09	0.00
60	0.16	0.00	0.53	0.02	0.09	0.00
61	0.16	0.00	0.53	0.02	0.09	0.00
62	0.32	0.44	0.77	7.93	0.25	3.46

63	0.25	0.01	0.54	0.01	0.14	0.00
64	0.32	0.44	0.77	7.99	0.25	3.52
65	0.16	0.00	0.53	0.00	0.09	0.00
66	0.25	0.01	0.54	0.01	0.14	0.00
67	0.16	0.28	0.72	6.72	0.12	1.87
68	0.16	0.00	0.52	0.02	0.08	0.00
69	0.16	0.00	0.52	0.02	0.08	0.00
70	0.16	0.00	0.48	0.02	0.08	0.00
71	0.16	0.00	0.52	0.00	0.08	0.00
72	0.16	0.00	0.48	0.02	0.08	0.00
73	0.16	0.00	0.52	0.02	0.08	0.00
74	0.44	0.40	0.18	6.43	0.08	2.55
75	0.22	0.15	0.04	0.76	0.01	0.11
76	0.22	0.15	0.06	0.93	0.01	0.14
77	0.18	0.01	0.05	0.03	0.01	0.00
78	0.31	0.43	0.14	4.82	0.04	2.05
79	0.27	0.15	0.07	1.19	0.02	0.18
80	0.27	0.15	0.17	1.19	0.05	0.18
81	0.36	0.42	0.02	3.66	0.01	1.55
82	0.28	0.42	0.10	0.56	0.03	0.24
83	0.27	0.15	0.15	0.99	0.04	0.15
84	0.43	0.48	0.13	17.08	0.05	8.25
85	0.43	0.48	0.09	5.01	0.04	2.40
86	0.32	0.13	0.00	1.40	0.00	0.18
87	0.17	0.01	0.09	0.06	0.02	0.00
88	0.38	0.46	0.04	14.73	0.02	6.71
89	0.38	0.46	0.01	13.75	0.00	6.26
90	0.18	0.34	0.10	1.49	0.02	0.51
91	0.18	0.34	0.01	1.14	0.00	0.38
92	0.18	0.34	0.01	1.40	0.00	0.47
93	0.18	0.34	0.01	1.75	0.00	0.59
94	0.32	0.49	0.00	0.03	0.00	0.01

95	0.32	0.49	0.00	0.04	0.00	0.02
96	0.07	0.00	0.00	0.00	0.00	0.00
97	0.19	0.49	0.01	0.49	0.00	0.24
98	0.18	0.34	0.03	1.15	0.00	0.39
99	0.18	0.34	0.00	1.15	0.00	0.39
100	0.41	0.47	0.11	0.10	0.04	0.05
101	0.36	0.47	0.00	0.15	0.00	0.07
102	0.36	0.47	0.00	0.01	0.00	0.00
103	0.26	0.47	0.00	0.14	0.00	0.07
104	0.26	0.47	0.00	0.14	0.00	0.07
105	0.30	0.50	0.21	3.17	0.06	1.59
106	0.30	0.50	0.21	3.17	0.06	1.59
107	0.07	0.00	0.02	0.00	0.00	0.00
108	0.09	0.22	0.01	1.67	0.00	0.37
109	0.30	0.50	0.10	6.94	0.03	3.46
110	0.24	0.29	0.00	0.00	0.00	0.00
111	0.09	0.22	0.12	1.83	0.01	0.40
112	0.29	0.49	0.21	7.26	0.06	3.54
113	0.37	0.46	0.14	0.58	0.05	0.27
114	0.37	0.46	0.12	0.62	0.04	0.29
115	0.09	0.22	0.01	1.83	0.00	0.40
116	0.09	0.22	0.01	1.83	0.00	0.40
117	0.08	0.22	0.08	1.94	0.01	0.42
118	0.09	0.22	0.01	1.71	0.00	0.38
119	0.07	0.00	0.04	0.00	0.00	0.00
120	0.07	0.00	0.04	0.00	0.00	0.00
121	0.14	0.03	0.00	0.24	0.00	0.01
122	0.14	0.11	0.00	3.52	0.00	0.40
123	0.27	0.50	2.03	1.27	0.56	0.63
124	0.34	0.26	1.38	0.05	0.48	0.01
125	0.21	0.47	0.94	0.95	0.20	0.44
126	0.00	0.00	0.00	0.04	0.00	0.00

127	0.43	0.50	0.68	0.14	0.29	0.07
128	0.47	0.50	1.94	2.23	0.91	1.11
129	0.49	0.49	0.25	2.20	0.12	1.08
130	0.31	0.50	1.59	11.31	0.49	5.64
131	0.49	0.49	0.29	3.20	0.14	1.57
132	0.42	0.50	1.53	12.74	0.65	6.36
133	0.18	0.48	0.63	6.63	0.12	3.19
134	0.35	0.48	1.57	6.80	0.55	3.28
135	0.00	0.01	0.01	0.19	0.00	0.00
136	0.35	0.50	4.44	2.84	1.57	1.42
137	0.32	0.47	0.13	0.06	0.04	0.03
138	0.32	0.47	0.15	0.07	0.05	0.03
139	0.35	0.24	3.26	3.48	1.13	0.84
140	0.38	0.47	0.22	0.06	0.08	0.03
141	0.38	0.50	0.20	0.08	0.08	0.04
142	0.37	0.46	0.22	0.04	0.08	0.02
143	0.38	0.47	0.21	0.04	0.08	0.02
144	0.01	0.34	0.01	0.11	0.00	0.04
145	0.40	0.36	0.07	0.07	0.03	0.02
146	0.38	0.47	0.23	0.14	0.09	0.06
147	0.00	0.00	0.00	0.00	0.00	0.00
148	0.38	0.47	0.25	0.04	0.09	0.02
149	0.32	0.50	7.50	4.04	2.43	2.02
150	0.46	0.45	0.10	0.00	0.05	0.00
151	0.46	0.47	0.10	0.71	0.05	0.33
152	0.49	0.50	4.57	5.78	2.26	2.89
153	0.25	0.28	3.66	3.36	0.93	0.93
154	0.43	0.45	0.00	0.08	0.00	0.04
155	0.47	0.45	0.96	0.02	0.45	0.01
156	0.48	0.43	1.24	0.53	0.59	0.23
157	0.47	0.45	0.93	0.02	0.43	0.01
158	0.47	0.45	0.93	0.02	0.43	0.01

159	0.46	0.45	0.10	0.00	0.05	0.00
160	0.47	0.49	0.88	2.32	0.41	1.13
161	0.44	0.45	0.49	0.00	0.21	0.00
162	0.42	0.43	0.21	0.44	0.09	0.19
163	0.40	0.45	0.12	0.02	0.05	0.01
164	0.46	0.50	0.58	3.50	0.26	1.74
165	0.46	0.50	0.58	1.67	0.26	0.84
166	0.46	0.36	0.62	6.14	0.28	2.23
167	0.49	0.45	0.81	5.91	0.40	2.63
168	0.49	0.45	0.82	5.91	0.40	2.63
169	0.49	0.35	0.25	0.33	0.12	0.11
170	0.49	0.35	0.26	0.35	0.13	0.12
171	0.49	0.25	0.27	2.16	0.13	0.53
172	0.50	0.25	0.37	2.26	0.18	0.57
173	0.34	0.49	1.87	0.04	0.63	0.02
174	0.43	0.48	0.44	6.41	0.19	3.07
175	0.50	0.50	0.50	0.53	0.25	0.26
176	0.50	0.50	0.29	0.00	0.14	0.00
177	0.50	0.45	0.18	1.01	0.09	0.46
178	0.50	0.20	0.74	6.21	0.37	1.24
179	0.44	0.48	1.50	0.42	0.65	0.20
180	0.41	0.48	0.79	0.75	0.33	0.36
181	0.31	0.12	2.19	0.09	0.67	0.01
182	0.50	0.40	0.78	0.02	0.39	0.01
183	0.50	0.20	0.76	6.21	0.38	1.24
184	0.50	0.20	0.57	7.92	0.28	1.58
185	0.50	0.50	0.55	8.67	0.27	4.32
186	0.21	0.00	0.14	0.00	0.03	0.00
187	0.22	0.21	0.12	1.31	0.03	0.27
188	0.21	0.00	0.12	0.00	0.03	0.00
189	0.39	0.18	2.67	6.28	1.03	1.15
190	0.41	0.50	2.06	22.13	0.84	11.06

191	0.40	0.27	2.52	7.45	1.02	2.05
192	0.40	0.19	3.05	4.78	1.22	0.92
193	0.45	0.50	0.21	0.05	0.09	0.02
194	0.41	0.50	2.79	0.00	1.14	0.00
195	0.40	0.46	3.04	0.01	1.23	0.00
196	0.44	0.44	2.53	7.01	1.12	3.05
197	0.40	0.19	3.02	4.62	1.21	0.88
198	0.41	0.48	2.93	6.36	1.19	3.03
199	0.40	0.19	3.11	4.62	1.24	0.88
200	0.41	0.44	3.05	0.01	1.24	0.01
201	0.41	0.43	2.62	0.21	1.07	0.09
202	0.41	0.40	2.03	2.96	0.82	1.18
203	0.41	0.40	2.66	2.96	1.08	1.18
204	0.41	0.40	2.68	2.96	1.09	1.18
205	0.41	0.40	2.55	2.96	1.03	1.18
206	0.41	0.40	2.18	2.55	0.88	1.02

[1]H = Holstein breed.

[2]G = Gi breed.

ACKNOWLEDGEMENTS

We thank EMBRAPA - Dairy Cattle Research Center, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Ministério da Ciência Tecnologia e Inovação (MCTI)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)/Instituto Nacional de Ciência e Tecnologia - Ciência Animal (INCT-CA) for the financial support, Embrapa Dairy Cattle Research Center for providing the data needed to carry out this study and Mario P. L. Calus (Wageningen University & Research Animal Breeding and Genomics,

6700 AH Wageningen, the Netherlands) for his assistance in BOA and CALC_GRM methodologies.

REFERENCES

Aguilar, I., I. Misztal, D.L. Johnson, A. Legarra, S. Tsuruta, and T.J. Lawlor. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score.. J. Dairy Sci. 93:743–52. doi:10.3168/jds.2009-2730.

Akakura, S., S. Singh, M. Spataro, R. Akakura, J. Il Kim, M.L. Albert, and R.B. Birge. 2004. The opsonin MFG-E8 is a ligand for the αvβ5 integrin and triggers DOCK180-dependent Rac1 activation for the phagocytosis of apoptotic cells. Exp. Cell Res. 292:403–416. doi:10.1016/j.yexcr.2003.09.011.

Ao, J., C. Wei, Y. Si, C. Luo, W. Lv, Y. Lin, Y. Cui, and X. Gao. 2015. Tudor-SN regulates milk synthesis and proliferation of bovine mammary epithelial cells. Int. J. Mol. Sci. 16:29936–29947. doi:10.3390/ijms161226212.

Bagci, H., M. Laurin, J. Huber, W.J. Muller, and J.F. Côté. 2014. Impaired cell death and mammary gland involution in the absence of Dock1 and Rac1 signaling. Cell Death Dis. 5:1–11. doi:10.1038/cddis.2014.338.

Banos, G., J.A. Woolliams, B.W. Woodward, A.B. Forbes, and M.P. Coffey. 2008. Impact of Single Nucleotide Polymorphisms in Leptin , Leptin Receptor , Growth Hormone Receptor , and Diacylglycerol Acyltransferase (DGAT1) Gene Loci on Milk Production , Feed , and Body Energy Traits of UK Dairy Cows. J. Dairy Sci. 91:3190–3200. doi:10.3168/jds.2007-0930.

Bervian, G., F.Z.B. A, S. Ribeiro, E. Kenji, N. Arashiro, J. Gabriel, V. De Grazia, F.C.V. D, P.I.O. E, and B.C.D.C. C. 2018. Metabolic , follicular and embryo production responses of postpartum crossbred Holstein · Gir dairy cows fed diets with different energy levels. Anim. Prod. Sci.. doi:https://doi.org/10.1071/AN18090 Metabolic,.

Bervian, G., F.Z. Brandão, S. Ribeiro, A.L. Ferreira, M.M. Campos, F.S. Machado, and B.C. De Carvalho. 2019. Milk production and composition , food consumption , and energy balance of postpartum crossbred Holstein-Gir dairy cows fed two diets of different energy levels. Trop. Anim. Health Prod. 51:65–71. doi:https://doi.org/10.1007/s11250-018-1660-1.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi:10.1093/bioinformatics/btp101.

Bionaz, M., and J.J. Loor. 2011. Gene networks driving bovine mammary protein synthesis during the lactation cycle. Bioinform. Biol. Insights 5:83–98. doi:10.4137/BBI.S7003.

Boutinaud, M., and H. Jammes. 2004. Growth hormone increases Stat5 and Stat1 expression in lactating goat mammary gland: A specific effect compared to milking frequency. Domest. Anim. Endocrinol. 27:363–378. doi:10.1016/j.domaniend.2004.04.002.

Broadhurst, M.K., R.S.F. Lee, S. Hawkins, and T.T. Wheeler. 2005. The p100 EBNA-2 coactivator: A highly conserved protein found in a range of exocrine and endocrine cells and tissues in cattle. Biochim. Biophys. Acta - Gene Struct. Expr. 1681:126–133. doi:10.1016/j.bbaexp.2004.10.009.

Broadhurst, M.K., and T.T. Wheeler. 2001. The p100 coactivator is present in the nuclei of mammary epithelial cells and its abudance is increased in response to prolactin in culture and in mammary tissue during lactation. J. Endocrinol. 171:329–337. doi:10.1677/joe.0.1710329.

Brown, E.M., G. Gamba, D. Riccardl, M. Lombardi, R. Butters, O. Kifor, A. Sun, M.A. Hediger, J. Lytton, and S.C. Hebert. 1993. Cloning and characterization of an extracellular Ca(2+)-sensing receptor from bovine parathyroid. Nature 366:575–580.

Bulmer, M.G. 1971. The Effect of Selection on Genetic Variability. Am. Nat. 105:201–211.

Canaza-cayo, A.W., J. Araújo, P. Sávio, R. De Almeida, M. Fonseca, S. Daltro, M. Vinicius, and G. Barbosa. 2016. Genetic trend estimates for milk yield production and fertility traits of the Girolando cattle in Brazil. Livest. Sci. 190:113–122. doi:10.1016/j.livsci.2016.06.009.

Canaza-Cayo, A.W., P.S. Lopes, J.A. Cobuci, M.F. Martins, and M.V.G.B. da Silva. 2017. Genetic parameters of milk production and reproduction traits of Girolando cattle in Brazil. Ital. J. Anim. Sci. 17:22–30. doi:10.1080/1828051X.2017.1335180.

Capuco, A. V., and R.M. Akers. 1999. Mammary Involution in Dairy Animals. J. Mammary Gland Biol. Neoplasia 4:137–144. doi:10.1023/A:1018769022990.

Cardoso, S.R., L.B. Queiroz, V.A. Goulart, G.B. Mourão, E. Benedetti, and L.R. Goulart. 2011. Productive performance of the dairy cattle Girolando breed mediated by the fat-related genes DGAT1 and LEP and their polymorphisms. Res. Vet. Sci. 91:e107–e112. doi:10.1016/j.rvsc.2011.02.006.

Casey, T., J. Crodian, S.S. Donkin, and K. Plaut. 2014a. Continuously Changing Light-Dark Phase Decreases Milk Yield , Fat , Protein and Lactose in Dairy Cows. Adv. Dairy Res. 2. doi:10.4172/2329-888X.1000119.

Casey, T., J. Crodian, A. Suárez-Trujillo, E. Erickson, B. Weldon, K. Crow, S. Cummings, Y. Chen, A. Shamay, S.J. Mabjeesh, and K. Plaut. 2016. CLOCK regulates mammary epithelial cell growth and differentiation. Am. J. Physiol. - Regul. Integr. Comp. Physiol. 311:R1125–R1134. doi:10.1152/ajpregu.00032.2016.

Casey, T., O. Patel, K. Dykema, H. Dover, K. Furge, and K. Plaut. 2009. Molecular signatures reveal circadian clocks may orchestrate the homeorhetic response to lactation. PLoS One 4. doi:10.1371/journal.pone.0007395.

Casey, T.M., J. Crodian, E. Erickson, K.K. Kuropatwinski, A.S. Gleiberman, and M.P. Antoch. 2014b. Tissue-Specific Changes in Molecular Clocks During the Transition from Pregnancy to Lactation in Mice1. Biol. Reprod. 90:1–15. doi:10.1095/biolreprod.113.116137.

Casey, T.M., and K. Plaut. 2012. Lactation biology symposium: Circadian clocks as mediators of the homeorhetic response to lactation. J. Anim. Sci. 90:744–754. doi:10.2527/jas.2011-4590.

Chapman, R.S., P.C. Lourenco, E. Tonner, D.J. Flint, S. Selbert, K. Takeda, S. Akira, A.R. Clarke, and C.J. Watson. 1999. Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3.. Genes Dev. 13:2604–16.

Chen, H.C., P.M. Elias, R.V.F. Jr, H.C. Chen, S.J. Smith, B. Tow, P.M. Elias, and R. V Farese. 2002. Leptin modulates the effects of acyl CoA : diacylglycerol acyltransferase deficiency on murine fur and sebaceous glands Find the latest version : Leptin modulates the effects of acyl CoA : diacylglycerol acyltransferase deficiency on murine fur and sebace. J. Clin. Investig. 109:175–181. doi:10.1172/JCI200213880.Introduction.

Clempson, A.M., G.E. Pollott, J.S. Brickell, N.E. Bourne, N. Munce, and D.C. Wathes. 2011. Evidence that leptin genotype is associated with fertility , growth , and milk production in Holstein cows. J. Dairy Sci. 94:3618–3628. doi:10.3168/jds.2010-3626.

Cobanoglu, O., I. Zaitoun, Y.M. Chang, G.E. Shook, and H. Khatib. 2006. Effects of the signal transducer and activator of transcription 1 (STAT1) gene on milk production traits in Holstein dairy cattle.. J. Dairy Sci. 89:4433–7. doi:10.3168/jds.S0022-0302(06)72491-2.

Cui, Y., X. Zhang, C. Guo, R. Du, G. Ailun, C. Ao, and M. Gao. 2017. Identification and characterization of bovine mammary peptide transporters in response to tripeptide and lactogenic hormone treatment. Czech J. Anim. Sci. 62:296–305. doi:10.17221/93/2015-CJAS.

Daniel, H., and G. Kottra. 2004. The proton oligopeptide cotransporter family SLC15 in physiology and pharmacology. Pflugers Arch. Eur. J. Physiol. 447:610–618. doi:10.1007/s00424-003-1101-4.

Deng, T.X., C.Y. Pang, X.R. Lu, P. Zhu, A.Q. Duan, and X.W. Liang. 2016. Associations between polymorphisms of the STAT1 gene and milk production traits in water buffaloes. J. Anim. Sci. 94:927–935. doi:10.2527/jas.2015-0139.

Dong, J., T. Tong, A.M. Reynado, J.M. Rosen, S. Huang, and Y. Li. 2010. Genetic manipulation of individual somatic mammary cells in vivo reveals a master role of STAT5a in inducing alveolar fate commitment and lactogenesis even in the absence of ovarian hormones. Dev. Biol. 346:196–203. doi:10.1016/j.ydbio.2010.07.027.

Engelsma, K.A., R.F. Veerkamp, M.P.L. Calus, and J.J. Windig. 2014. Consequences for diversity when animals are prioritized for conservation of the whole genome or of one specific allele. J. Anim. Breed. Genet. 131:61–70. doi:10.1111/jbg.12052.

Facó, O., R.M. Filho, R. Nonato, B. Lobo, D. Maria, and M. Ribeiro. 2009. Efeito da redução da variação da duração de lactação na avaliação genética de bovinos leiteiros mestiços. Rev. Ciência Agronômica 40:287–292.

Facó, O., R.N.B. Lôbo, R.M. Filho, G.A. Martins, S.M.P. de Oliveira, and D.M.M.R. Azevêdo. 2008. Efeitos genéticos aditivos e não-aditivos para características produtivas e reprodutivas em vacas mestiças Holandês × Gir. Rev. Bras. Zootec. 37:48–53.

Fashe, T., J. Saarikettu, P. Isomäki, J. Yang, and O. Silvennoinen. 2013. Expression analysis of Tudor-SN protein in mouse tissues. Tissue Cell 45:21–31. doi:10.1016/j.tice.2012.09.001.

Ferraz, J.B.S., and P.E. de Felício. 2010. Production systems – An example from Brazil. Meat Sci. 84:238–243. doi:10.1016/j.meatsci.2009.06.006.

Fontanesi, L., D.G. Calò, G. Galimberti, R. Negrini, R. Marino, A. Nardone, P. Ajmone-Marsan, and V. Russo. 2014. A candidate gene association study for nine economically important traits in Italian Holstein cattle. Anim. Genet. 45:576–580. doi:10.1111/age.12164.

Granillo, A.R., J.C. Boffi, L. Barañao, E. Kordon, A. Pecci, and A. Guberman. 2007. STAT5 transcriptional activity is impaired by LIF in a mammary epithelial cell line. Biochem. Biophys. Res. Commun. 356:727–732. doi:10.1016/j.bbrc.2007.03.040.

Grevenhof, E.M. Van, J.A.M. Van Arendonk, and P. Bijma. 2012. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet. Sel. Evol. 44:1–10.

Groneberg, D.A., F. Döring, S. Theis, M. Nickolaus, A. Fischer, and H. Daniel. 2002. Peptide transport in the mammary gland: expression and distribution of PEPT2

mRNA and protein. Am. J. Physiol. - Endocrinol. Metab. 282:E1172–E1179. doi:10.1152/ajpendo.00381.2001.

Haricharan, S., and Y. Li. 2014. STAT signaling in mammary gland differentiation, cell survival and tumorigenesis. Mol. Cell. Endocrinol. 382:560–569. doi:10.1016/j.mce.2013.03.014.

Harris, J., P.M. Stanford, K. Sutherland, S.R. Oakes, M.J. Naylor, F.G. Robertson, K.D. Blazek, M. Kazlauskas, H.N. Hilton, S. Wittlin, W.S. Alexander, G.J. Lindeman, J.E. Visvader, and C.J. Ormandy. 2006. Socs2 and Elf5 Mediate Prolactin-Induced Mammary Gland Development. Mol. Endocrinol. 20:1177–1187. doi:10.1210/me.2005-0473.

He, X., M.X. Chu, L. Qiao, J.N. He, P.Q. Wang, T. Feng, R. Di, G.L. Cao, L. Fang, and Y.F. An. 2012. Polymorphisms of STAT5A gene and their association with milk production traits in Holstein cows. Mol. Biol. Rep. 39:2901–2907. doi:10.1007/s11033-011-1051-4.

Hickey, J.M., B.P. Kinghorn, B. Tier, J.F. Wilson, N. Dunstan, and J.H.J. van der Werf. 2011. A combined long-range phasing and long haplotype imputation method to impute phase for SNP genotypes.. Genet. Sel. Evol. 43:12. doi:10.1186/1297-9686-43-12.

Hughes, K., and C.J. Watson. 2012. The spectrum of STAT functions in mammary gland development. Landes Biosci. 1:151–158. doi:10.4161/jkst.19691.

Hughes, K., J.A. Wickenden, J.E. Allen, and C.J. Watson. 2012. Conditional deletion of Stat3 in mammary epithelium impairs the acute phase response and modulates immune cell numbers during post-lactational regression. J. Pathol. 227:106–117. doi:10.1002/path.3961.

Ingvartsen, K.L., and Y.R. Boisclair. 2001. Leptin and the regulation of food intake , energy homeostasis and immunity with special focus on periparturient ruminants. Domest. Anim. Endocrinol. 21:215–250.

Khatib, H., R.L. Monson, V. Schutzkus, D.M. Kohl, G.J.M. Rosa, and J.J. Rutledge. 2008. Mutations in the STAT5A Gene Are Associated with Embryonic Survival and Milk Composition in Cattle. J. Dairy Sci. 91:784–793. doi:10.3168/jds.2007-0669.

Kim, W., and J.J. Wysolmerski. 2016. Calcium-sensing receptor in breast physiology and cancer. Front. Physiol. 7:1–11. doi:10.3389/fphys.2016.00440.

Kim, Y.S., W. Birtwhistle, and Y.W. Kim. 1972. Peptide hydrolases in the bruch border and soluble fractions of small intestinal mucosa of rat and man.. J. Clin. Invest. 51:1419–1430. doi:10.1172/JCI106938.

Kiyici, J.M., K. Arslan, and B. Akyuz. 2018. Relationships between polymorphisms of growth hormone , leptin and myogenic factor 5 genes with some milk yield traits in Holstein dairy cows. Int. J. Dairy Technol. 70:1–7. doi:10.1111/1471-0307.12539.

Kritikou, E.A., A. Sharkey, K. Abell, P.J. Came, E. Anderson, R.W.E. Clarkson, and C.J. Watson. 2003. A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. Development 130:3459–3468. doi:10.1242/dev.00578.

Lee, H.J., D. Gallego-Ortega, A. Ledger, D. Schramek, P. Joshi, M.M. Szwarc, C. Cho, J.P. Lydon, R. Khokha, J.M. Penninger, and C.J. Ormandy. 2013. Progesterone drives mammary secretory differentiation via RankL-mediated induction of Elf5 in luminal progenitor cells. Development 140:1397–1401. doi:10.1242/dev.088948.

Leehy, K.A., T.H. Truong, L.J. Mauro, and C.A. Lange. 2018. Progesterone receptors (PR) mediate STAT actions: PR and prolactin receptor signaling crosstalk in breast cancer models. J. Steroid Biochem. Mol. Biol. 176:88–93. doi:10.1016/j.jsbmb.2017.04.011.

Liefers, S.C., R.F. Veerkamp, M.F.W. Pas, C. Delavaud, Y. Chilliard, and T. Van Der Lende. 2004. A missense mutation in the bovine leptin receptor gene is associated with leptin concentrations during late pregnancy. Anim. Genet. 35:138–141. doi:10.1111/j.1365-2052.2004.01115.x.

Liefers, S.C., R.F. Veerkamp, M.F.W. Te Pas, Y. Chilliard, and T. Van Der Lende. 2005. Genetics and physiology of leptin in periparturient dairy cows. Domest. Anim. Endocrinol. 29:227–238. doi:10.1016/j.domaniend.2005.02.009.

Lu, L., X. Gao, Q. Li, J. Huang, R. Liu, and H. Li. 2012. Comparative phosphoproteomics analysis of the effects of L-methionine on dairy cow mammary epithelial cells. Can. J. Anim. Sci. 92:433–442. doi:10.4141/cjas2012-063.

Maere, S., K. Heymans, and M. Kuiper. 2005. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21:3448–3449. doi:10.1093/bioinformatics/bti551.

Mamillapalli, R., J. Van Houten, P. Dann, D. Bikle, W. Chang, E. Brown, and J. Wysolmerski. 2013. Mammary-specific ablation of the calcium-sensing receptor during lactation alters maternal calcium metabolism, milk calcium transport, and neonatal calcium accrual. Endocrinology 154:3031–3042. doi:10.1210/en.2012-2195.

Marques, D.B.D., J.W.M. Bastiaansen, M.L.W.J. Broekhuijse, M.S. Lopes, E.F. Knol, B. Harlizius, S.E.F. Guimarães, F.F. Silva, and P.S. Lopes. 2018. Weighted single - step GWAS and gene network analysis reveal new candidate genes for semen traits in pigs. Genet. Sel. Evol. 50:1–14. doi:10.1186/s12711-018-0412-z.

Menzies, K.K., H.J. Lee, C. Lefèvre, C.J. Ormandy, K.L. Macmillan, and K.R. Nicholas. 2010. Insulin, a key regulator of hormone responsive milk protein synthesis during lactogenesis in murine mammary explants. Funct. Integr. Genomics 10:87–95. doi:10.1007/s10142-009-0140-0.

Meredith, B.K., F.J. Kearney, E.K. Finlay, D.G. Bradley, A.G. Fahey, D.P. Berry, and D.J. Lynn. 2012. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. BMC Genet. 13:1–11.

Murney, R., K. Stelwagen, T.T. Wheeler, J.K. Margerison, and K. Singh. 2015. Activation of signal transducer and activator of transcription 5 (STAT5) is linked to β1-integrin protein abundance in unilaterally milked bovine mammary glands. J. Dairy Sci. 98:3133–3142. doi:10.3168/jds.2014-9003.

Oakes, S.R., M.J. Naylor, M.-L. Asselin-Labat, K.D. Blazek, M. Gardiner-Garden, H.N. Hilton, M. Kazlauskas, M.A. Pritchard, L.A. Chodosh, P.L. Pfeffer, G.J. Lindeman, J.E. Visvader, and C.J. Ormandy. 2008. The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev. 22:581–586. doi:10.1101/gad.1614608.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Ozden Cobanoglu, Eser Kemal Gurcan, Soner Cankaya, Ertugrul Kul, and Hasan Samet Abaci. 2016. The Detection of STAT1 Gene Influencing Milk Related Traits in

Turkish Holstein and Jersey Cows. J. Agric. Sci. Technol. A 6. doi:10.17265/2161-6256/2016.04.006.

Paukku, K., J. Yang, and O. Silvennoinen. 2003. Tudor and Nuclease-Like Domains Containing Protein p100 Function as Coactivators for Signal Transducer and Activator of Transcription 5. Mol. Endocrinol. 17:1805–1814. doi:10.1210/me.2002-0256.

Philp, J.A., T.G. Burdon, and C.J. Watson. 1996. Differential activation of STATs 3 and 5 during mammary gland development. FEBS Lett 396:77–80.

Raven, L.A., B.G. Cocks, M.E. Goddard, J.E. Pryce, and B.J. Hayes. 2014. Genetic variants in mammary development, prolactin signalling and involution pathways explain considerable variation in bovine milk production and milk composition. Genet. Sel. Evol. 46:1–13. doi:10.1186/1297-9686-46-29.

Ribeiro, V.M.P., F.A. Merlo, G.C. Gouveia, L.K. Winkelstroter, L.R.A. Abreu, M.V.G.B. da Silva, J.C. do C. Panetto, L. de C. Paiva, M. de A.R. Cembranelli, and F.L.B. Toral. 2017. Genetic analysis of productive and reproductive traits in multiple-breed dairy cattle populations. Pesqui. Agropecu. Bras. 52:1109–1117. doi:10.1590/S0100-204X2017001100017.

Rowson, A.R., K.M. Daniels, S.E. Ellis, and R.C. Hovey. 2012. Growth and development of the mammary glands of livestock: A veritable barnyard of opportunities. Semin. Cell Dev. Biol. 23:557–566. doi:10.1016/j.semcdb.2012.03.018.

Rychtářová, J., Z. Sztankóová, J. Kyselová, V. Zink, and M. Štípková. 2014. Effect of DGAT1, BTN1A1, OLR1, and STAT1 genes on milk production and reproduction traits in the Czech Fleckvieh breed. Czech J. Anim. Sci. 2014:45–53.

Sandelin, A., W. Alkema, P. Engström, W.W. Wasserman, and B. Lenhard. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles.. Nucleic Acids Res. 32:D91-4. doi:10.1093/nar/gkh012.

Sargeant, T.J., B. Lloyd-Lewis, H.K. Resemann, A. Ramos-Montoya, J. Skepper, and C.J. Watson. 2014. Stat3 controls cell death during mammary gland involution by regulating uptake of milk fat globules and lysosomal membrane permeabilization. Nat. Cell Biol. 16:1057–1068. doi:10.1038/ncb3043.

Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:1–12.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, and M.P.L. Calus. 2016. Empirical determination of breed-of-origin of alleles in three-breed cross pigs. Genet. Sel. Evol. 48:55. doi:10.1186/s12711-016-0234-9.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13:2498–2504. doi:10.1101/gr.1239303.metabolite.

Sigl, T., H.H.D. Meyer, and S. Wiedemann. 2014. Gene expression analysis of protein synthesis pathways in bovine mammary epithelial cells purified from milk during

lactation and short-term restricted feeding. J. Anim. Physiol. Anim. Nutr. (Berl). 98:84–95. doi:10.1111/jpn.12039.

Silva, A.A., A.L.S. Azevedo, R.S. Verneque, K. Gasparini, M.G.C.D. Peixoto, M.V.G.B. da Silva, P.S. Lopes, S.E.F. Guimarães, and M.A. Machado. 2011. Quantitative trait loci affecting milk production traits on bovine chromosome 6 in zebuine Gyr breed. J. Dairy Sci. 94:971–980. doi:10.3168/jds.2009-2970.

Silva, L.F.P., M.J. VandeHaar, M.S. Weber Nielsen, and G.W. Smith. 2002. Evidence for a local effect of leptin in bovine mammary gland.. J. Dairy Sci. 85:3277–3286. doi:10.3168/jds.S0022-0302(02)74416-0.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, and B.R. Coelho. 2017. Girolando Breed Genetic Improvement Program – Genetic Evaluation of Cows – June/2017. Juiz de Fora: Embrapa Gado de Leite.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, and D.R. de L.R. Faza. 2018a. Programa de Melhoramento Genético Da Raça Girolando – Avaliação Genéteica de Vacas – Junho 2018. E.G. de Leite, ed. Juiz de Fora : Embrapa Gado de Leite.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018b. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Singh, K., C.V.C. Phyn, M. Reinsch, J.M. Dobson, K. Oden, S.R. Davis, K. Stelwagen, H.V. Henderson, and A.J. Molenaar. 2017. Temporal and spatial heterogeneity in milk and

immune-related gene expression during mammary gland involution in dairy cows. J. Dairy Sci. 100:7669–7685. doi:10.3168/jds.2017-12572.

Singh, K., I. Vetharaniam, J.M. Dobson, M. Prewitz, K. Oden, R. Murney, K.M. Swanson, R. McDonald, H.V. Henderson, and K. Stelwagen. 2016. Cell survival signaling in the bovine mammary gland during the transition from lactation to involution. J. Dairy Sci. 99:7523–7543. doi:10.3168/jds.2015-10515.

Soares, A.C.C., S.E.F. Guimarães, M.J. Kelly, M.R.S. Fortes, F.F. e Silva, L.L. Verardo, R.R. Mota, and S. Moore. 2017. Multiple-trait genomewide mapping and gene network analysis for scrotal circumference growth curves in Brahman cattle. J. Anim. Sci. 95:3331–3345. doi:10.2527/jas.2017.1409.

Suchocki, T., J. Komisarek, and J. Szyda. 2010. Testing candidate gene effects on milk production traits in dairy cattle under various parameterizations and modes of inheritance. J. Dairy Sci. 93:2703–2717. doi:10.3168/jds.2009-2550.

Suzuki, R., K. Tobe, M. Aoyama, K. Sakamoto, M. Ohsugi, N. Kamei, S. Nemoto, A. Inoue, Y. Ito, S. Uchida, K. Hara, T. Yamauchi, N. Kubota, Y. Terauchi, and T. Kadowaki. 2005. Expression of DGAT2 in White Adipose Tissue Is Regulated by Central Leptin Action *. J. Biol. Chem. 280:3331–3337. doi:10.1074/jbc.M410955200.

Tong, X., R. Drapkin, R. Yalamanchili, G. Mosialos, and E. Kieff. 1995. The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE.. Mol. Cell. Biol. 15:4735–4744. doi:10.1128/MCB.15.9.4735.

Touzet, H., and J.-S. Varré. 2007. Efficient and accurate P-value computation for Position Weight Matrices.. Algorithms Mol. Biol. 2:15. doi:10.1186/1748-7188-2-15.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Vanhouten, J., P. Dann, G. Mcgeoch, E.M. Brown, K. Krapcho, M. Neville, and J.J. Wysolmerski. 2004. The calcium-sensing receptor regulates mammary gland parathyroid hormone – related protein production and calcium transport. J. Clin. Invest. 113:598–608. doi:10.1172/JCI200418776.Introduction.

VanHouten, J.N. 2005. Calcium sensing by the mammary gland. J. Mammary Gland Biol. Neoplasia 10:129–139. doi:10.1007/s10911-005-5396-y.

VanHouten, J.N., and J.J. Wysolmerski. 2013. The Calcium-Sensing Receptor in the Breast Joshua. Best Pract. Res. Clin. Endocrinol. Metab. 6:247–253. doi:10.1111/j.1743-6109.2008.01122.x.Endothelial.

VanRaden, P.M. 2008. Efficient Methods to Compute Genomic Predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980.

Wang, H., I. Misztal, I. Aguilar, A. Legarra, R.L. Fernando, Z. Vitezica, R. Okimoto, T. Wing, R. Hawken, and W.M. Muir. 2014. Genome-wide association mapping including phenotypes from relatives without genotypes in a single-step (ssGWAS) for 6-week body weight in broiler chickens. Front. Genet. 5:1–10. doi:10.3389/fgene.2014.00134.

Wu, M., C. Huang, K. Gan, H. Huang, Q. Chen, J. Ouyang, Y. Tang, X. Li, Y. Yang, H. Zhou, Y. Zhou, Z. Zeng, L. Xiao, D. Li, K. Tang, S. Shen, and G. Li. 2006. LRRC4 , a Putative Tumor Suppressor Gene, Requires a Functional Leucine-rich Repeat Cassette Domain to Inhibit Proliferation of Glioma Cells In Vitro by Modulating the

Extracellular Signal-regulated Kinase/Protein Kinase B/Nuclear Factor-κB Pathway. Mol. Biol. Cell 17:3534–3542. doi:10.1091/mbc.e05-11-1082.

Wysolmerski, J.J. 2012. Parathyroid Hormone-Related Protein: An Update. J. Clin. Endocrinol. Metab. 97:2947–2956. doi:10.1210/jc.2012-2142.

Yamaji, D., R. Na, Y. Feuermann, S. Pechhold, W. Chen, G.W. Robinson, and L. Hennighausen. 2009. Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. Genes Dev. 23:2382–2387. doi:10.1101/gad.1840109.

Yang, J.-X., C.-H. Wang, Q.-B. Xu, F.-Q. Zhao, J.-X. Liu, and H.-Y. Liu. 2015. Methionyl-Methionine Promotes a-s1 Casein Synthesis in Bovine Mammary Gland Explants by Enhancing Intracellular Substrate Availability and Activating JAK2-STAT5 and mTOR-Mediated Signaling Pathways. J. Nutr. 145:1748–1753. doi:10.3945/jn.114.208330.

Yang, J., J.J. Kennelly, and V.E. Baracos. 2000. The activity of transcription factor Stat5 responds to prolactin , growth hormone , and IGF-I in rat and bovine mammary explant culture. J. Anim. Sci. 3114–3125.

Zhou, J., A.Y.N. Ng, M.J. Tymms, L.S. Jermiin, A.K. Seth, R.S. Thomas, and I. Kola. 1998. A novel transcription factor, ELF5, belongs to the ELF subfamily of ETS genes and maps to human chromosome 11p13-15, a region subject to LOH and rearrangement in human carcinoma cell lines. Oncogene 17:2719–2732. doi:10.1038/sj.onc.1202198.

Zhou, M.M., Y.M. Wu, H.Y. Liu, K. Zhao, and J.X. Liu. 2011. Effects of tripeptides and lactogenic hormones on oligopeptide transporter 2 in bovine mammary gland. J.

Anim. Physiol. Anim. Nutr. (Berl). 95:781–789. doi:10.1111/j.1439-0396.2010.01110.x.

Zimin, A. V., A.L. Delcher, L. Florea, D.R. Kelley, M.C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C.P. Van Tassell, T.S. Sonstegard, G. Marçais, M. Roberts, P. Subramanian, J.A. Yorke, and S.L. Salzberg. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 10. doi:10.1186/gb-2009-10-4-r42.

CHAPTER 6

GENERAL DISCUSSION

Introduction

In the genomic era, the number of studies using association analysis to identify QTL related with traits of interest is increasing. However, GWAS is only the first step in the process to find out causal variants, and further studies are needed to validate these results. In addition, the BOA approach recently created for 2- or 3-way crosses in pigs and poultry (Vandenplas et al., 2016) and adapted to crossbred cattle of the current study, has provided additional tools for genomic studies, contributing to better biological understanding of the evaluated traits. Despite of its relevance in dairy production, Girolando Breed still lack information about its genomic architecture and genomic data regarding such important traits for production and survival. In this chapter, I discuss the importance of GWAS and post-GWAS analyses, the benefit and applicability of its results, and the impact of BOA results in adaptive and milk production traits in Girolando cattle.

The importance of genomic studies for adaptive traits in the crossbred cattle

The GWAS described at the present study for tick, heat stress and nematodes resistance traits in a Gir x Holstein experimental population enabled us to identify different sets of SNPs in QTL regions. Most QTL detection studies have been done in experimental F2 populations, created by crossing 2 genetically different breeds (Machado et al., 2010; Soma et al., 2014). Studies in F2 populations can be efficient in the genomic architecture dissection, in the identification of genomic regions bearing candidate genes and can also be used in GWAS using SNP markers (Ledur et al., 2010; Stratz et al., 2014).

Moreover, the Post-GWAS analyses provided useful information to identify candidate genes for the evaluated traits and increase their biological understanding. The identified genes may be deeply investigated in additional experiments aiming to validate their effects on these phenotypes in cattle. Subsequently, the combination of validated candidate genes and information on the QTL identified in this study might provide useful information that can be used for marker-assisted selection (MAS) or genomic selection applications in breeding programs for dairy cattle.

In livestock, one of the main scientific, economic and social concerns is the control of parasites exclusively by use of drugs. The intensive use of anthelmintic and acaricides, the underdose and the lack of chemicals turnover have favored the selection and proliferation of resistant parasites populations (Souza et al., 2008), making the conduction of researchs for new chemicals necessary (Biegelmeyer et al., 2012). Moreover, the misuse of drugs leaves chemical residues in meat, milk and in environment. In these circumstances, the evaluation of adaptive traits is of great importance to livestock production systems, in terms that it allows the identification of the most tolerant animals to biotic and abiotic stress conditions. Genetic variability for the three resistance traits was observed in the current research and it has been reported in previous studies (Ravagnolo and Misztal, 2000; Kim et al., 2013; Cardoso et al., 2014), which makes the use of genetic selection promising for the improvement of this traits. Thus, the genetic selection can be an effective alternative to decrease the use of drugs in parasites control in Tropical regions, contributing to control parasites populations resistant to drugs and to decrease the contamination of pastures and products of animal origin (Peña et al., 2000; Oliveira et al., 2009).

Given the fact that most of the milk in Brazil is produced in systems relying on grazing in challenging environments and the negative influence and impact of heat stress on

milk production, the use of marker assisted selection or genomic selection for a better response to heat stress effects can also contributes for success of milk production.

The BOA approach for adaptive traits in the crossbred cattle

In this thesis, we used the BOA approach results to evaluate the gametic segregation and check if there is a prevalence of alleles from a certain breed in animals classified based on their breeding values for each adaptive evaluated trait. On average, 92% of the alleles of F2 animals were assigned a breed-of-origin. In Chapter 2, as expected, we found that most of resistant animals for tick infestation showed two alleles from Gir breed, while the susceptible ones showed alleles from Holstein.

In Chapter 3, we observed that Holstein breed alleles could be associated to better response to heat stress effects. Several studies have described the great ability of Bos indicus animals to regulate their body temperature, due the lower metabolic rates and higher capacity of heat loss through thermoregulatory mechanisms (Hansen, 2004; Cattelam and Vale, 2013). Thus, it is possible that these abilities prevent the installation of heat stress in the body, and consequently its negative effects. On the other hand, we highlighted that due to the large amount of metabolic heat produced by high milk yield and limited ability to dissipate body heat, Bos taurus animals often are affected by heat stress effects (Renaudeau et al., 2012). So, we assumed that the Holstein breed might have a more efficient genetic architecture to defend their body from deleterious effects of heat stress in comparison to Gir animals.

In Chapter 4, the results showed that Holstein breed alleles could be associated to AVFEC resistance. Although there were no statistically significant differences in the average FEC in the genetic groups evaluated, Achi et al. (2003) and Oliveira et al. (2009,

2013) described significant susceptibility to infection of some nematodes species in B. taurus and B. indicus, suggesting that bovine resistance for gastrointestinal nematodes can be influenced by nematode species. The alleles associated with the resistance from Holstein breed can behave as cryptic alleles, which play an important role in crossbred populations. According to Abasht and Lamont, (2007), these alleles are important in population heterosis since they influence the expression of phenotypes in these populations and can show opposite effect or no significant effect on one of the parental populations. This cryptic effect can be detected as a result of allele selection, genetic drift, epistasis and pleiotropic effects of the QTL on other traits that are under selection (Abasht and Lamont, 2007).

These results demonstrate the need for a deeper biological understanding of genetic architecture underlying response in adaptive traits in purebred and crossbred animals. This enlightenment is meaningful to establish crossbreeds taking into account the complementarity between breeds.

Genomic studies for 305-day milk yield in Girolando cattle

In the previous chapters we have seen the importance of Girolando animals in the milk production in Brazil, due its notable excellent productivity, high fertility indexes and good vigor in tropical and subtropical regions. However, we highlighted in Chapter 5 that GWAS for milk yield in this breed have not been reported yet. So, this study is an important step in the QTL mapping and identification of candidate genes associated with this trait in the Girolando breed.

As previously described, the identified QTL and candidate genes may be implemented in Girolando breed breeding program by marker-assisted selection. In dairy cattle, young candidate bulls can be selected by genomic selection before progeny testing,

which may result in increase of selection differentials, shortening the generation interval, and increasing genetic gain (García-Ruiz et al., 2016; Weller et al., 2017).

Within the seven candidate genes for 305MY, the SNP variances were different depending upon the breed-of-origin, while the alleles coming from the Gir breed-of-origin showed the highest genetic variance. These results suggest that the SNP-specific variances in the candidate genes for 305MY in Girolando depends upon the breed-of-origin. In the current study, the main difference across breeds was the SNP-effects which is affected to the additive genetic variance. A highest additive genetic variance for 305MY was estimated for the Gir breed. The difference in SNP-specific between alleles coming from the Gir and Holstein breed can be the result of genetic selection. The selection for one specific allele or trait in the subsequent generations, can result in substantial losses in genetic diversity, which can be intensified by the use of genomic selection (Bulmer, 1971; Grevenhof et al., 2012; Engelsma et al., 2014). Given the fact that the Gir breed didn't underwent as intense a selection process for milk production as the Holstein breed, the genetic variability for this trait is highest in Gir, which can explain the higher SNP-specific variance of evaluated alleles coming from the Gir breed. Future studies are needed to provide further knowledge to better understand the genetic architecture underlying the influence of breed-of-origin of alleles in the milk yield in Girolando cattle.

Practical application of GWAS results in breeding programs

The GWAS results might provide useful information that can be used for MAS or genomic selection applications in breeding programs for dairy cattle. In MAS one or a few markers linked to QTL are used as a selection tool. This strategy has been used by Girolando breed genetic improvement for the identification of animals with genotypes

related to genetic diseases, i.e. BLAD, DUMPS and CVM, and higher potential for milk production, fat and protein milk content, i.e. Kappa-casein, Beta-casein, Beta-lactoglobulin (β-LGB) and DGAT1 (Silva et al., 2018). These information can be used to direct mating and choose semen, contributing for Girolando breeding. The inclusion of molecular markers information linked to adaptive traits and 305MY studied in the current research can provide more genomic information.

An alternative approach is to use GWAS results in genetic prediction to improve the prediction accuracy. According to Zhang et al. (2010), GWAS results can be used to generate a trait-specific G matrix that gives different weights to each SNP. Thus, SNPs that contribute more to the genetic variance of the evaluated trait are favored. This approach allowed the estimation of breeding values more accurate compared than model that applies the traditional genomic relationship matrix (G matrix). In dairy cattle, the combination of commonly-used 54K SNP chip with custom low-density SNP array resulted in higher reliability of the breeding values compared to traditional GBLUP. The low-density SNP chip was built with significant markers from GWAS using whole-genome sequence data (Brøndum et al., 2015). According to Lopes et al. (2017), the use of GWAS findings in marker-assisted models (MA-BLUP and MA-GBLUP) also resulted in increase of prediction accuracies for the number of teats trait in pig populations. For that, the most significant SNP from GWAS were included as fixed effect in the prediction models (MA-BLUP and MA-GBLUP), in this way, a specific set of SNPs will be used per trait, which gives a higher weight to each marker with large effect, resulting in highest prediction accuracies. These researches demonstrate that marker breeding values can be an important tool for selection of animals and to contribute to further advancements in breeding programs.

BOA approach in breeding programs

The availability of dense SNPs panels (Van Tassell et al., 2008; Ramos et al., 2009) made possible the incorporation of genomic data in genetic evaluations of breeding programs. Genomic selection compute estimated breeding values using a large number of markers (typically SNPs) spread across the whole genome (Meuwissen et al., 2001). Several factors impact SNP effects. Individuals originated from crossings of different populations, as Girolando cattle, present allele frequencies and divergent linkage disequilibrium across the genome in different genomic content (Gautier et al., 2007; O'Brien et al., 2014). Thus, the effect of the same allele, may differ because of different levels of LD between SNPs and a QTL, depending from which parental breed the QTL was inherited (Sevillano et al., 2018). Different MAF can be observed for functional variation underling the inherited QTL, depending on the breed of origin (Villa-Angulo et al., 2009; Wientjes et al., 2015). The dominance and epistatic interactions can have different effects in purebred and crossbred animals. Therefore, SNP effects may be breed-specific, this has to be considered a relevant factor for Girolando breeding program.

The BOA approach recently created for 2- or 3-way crosses in pigs and poultry (Vandenplas et al., 2016) and adapted to crossbred cattle in the current study (Otto et al., 2018), allows empirical testing of the model that accounts for breed-specific SNP effects. Results from this approach have been applied in models that take the breed-specific SNP effects, to improve genomic prediction, or GWAS (Sevillano et al., 2017, 2018). According to Sevillano et al. (2017), the inclusion of breed-specific SNP effect in the genomic evaluation model allowed, in some cases, a better prediction for crossbred performance.

In the current research, on average, 94% of the alleles of Girolando cattle were assigned a breed-of-origin. These results suggest that the breed origin of alleles for

Girolando cattle population was accurately assigned, allowing its use in genomic prediction and GWAS in this breed. Our team believes that once Holstein and Gir breeds present large genetic differences, so the inclusion of breed-specific SNP effect in the evaluation model will improve genomic prediction accuracy in Girolando cattle.

Conclusion

This thesis provides valuable results to QTL mapping in Girolando cattle and candidate genes identified using genome wide association and post genomic approaches for adaptative traits and milk yield. We have shown that selection for improved adaptive traits, besides contributing for success of milk production, might also benefit consumers, milk industry and the environment. The BOA approach provided important insights into the genetic architecture of the evaluated traits that no tool had previously made possible, becoming a powerful auxiliary tool in breeding programs. The present study has shown a rich information resource to add more pieces to the intricate nature of Girolando cattle genome, providing better biological comprehension of its nature and physiology of production under tropical conditions.

REFERENCES

Abasht, B., and S.J. Lamont. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. Anim. Genet. 38:491–498. doi:10.1111/j.1365-2052.2007.01642.x.

Achi, Y.L., J. Zinsstag, K. Yao, N. Yeo, P. Dorchies, and P. Jacquiet. 2003. Host specificity of Haemonchus spp. for domestic ruminants in the savanna in northern Ivory Coast. Vet. Parasitol. 116:151–158. doi:10.1016/S0304-4017(03)00258-9.

Biegelmeyer, P., L.Q. Nizoli, F.F. Cardoso, and E. Dionello. 2012. Bovine Resistance To

Rhipicephalus (Boophilus) Microplus Ticks. Arch. Zootec 61:1–11.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 8:4107–4116. doi:10.3168/jds.2014-9005.

Bulmer, M.G. 1971. The Effect of Selection on Genetic Variability. Am. Nat. 105:201–211.

Cardoso, C.P., B.F. Silva, D.S. Gonçalves, N.J. Tagliari, M.E. Saito, and A.F.T. Amarante. 2014. Resistência contra ectoparasitas em bovinos da raça crioula lageana e meio-sangue angus avaliada em condições naturais. Pesqui. Vet. Bras. 34:141–146. doi:10.1590/S0100-736X2014000200008.

Cattelam, J., and M.M. do Vale. 2013. Estresse térmico em bovinos. Rev. Port. Ciências Veterinárias 108:96–102.

Engelsma, K.A., R.F. Veerkamp, M.P.L. Calus, and J.J. Windig. 2014. Consequences for diversity when animals are prioritized for conservation of the whole genome or of one specific allele. J. Anim. Breed. Genet. 131:61–70. doi:10.1111/jbg.12052.

García-Ruiz, A., J.B. Cole, P.M. VanRaden, G.R. Wiggans, F.J. Ruiz-Lópeza, and C.P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. PNAS E3995–E4004. doi:10.1073/pnas.1519061113.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A. Boland, J.G. Garnier, D. Boichard, G.M. Lathrop, I.G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics 177:1059–1070. doi:10.1534/genetics.107.075804.

Grevenhof, E.M. Van, J.A.M. Van Arendonk, and P. Bijma. 2012. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet. Sel. Evol. 44:1–10.

Hansen, P.J. 2004. Physiological and cellular adaptations of zebu cattle to thermal stress. Anim. Reprod. Sci. 82–83:349–360. doi:10.1016/j.anireprosci.2004.04.011.

Kim, E.S., T.S. Sonstegard, M.V.G.B. Silva, L.C. Gasbarre, and C.P. Van Tassell. 2013. Identification of quantitative trait loci affecting gastrointestinal parasite resistance in an experimental Angus population. Anim. Genet. 45:117–121. doi:10.1111/age.12101.

Ledur, M., N. Navarro, and M. Pérez-Enciso. 2010. Large-scale SNP genotyping in crosses between outbred lines: How useful is it?. Heredity (Edinb). 105:173–182. doi:10.1038/hdy.2009.149.

Lopes, M.S., H. Bovenhuis, M. van Son, Ø. Nordbø, E.H. Grindflek, E.F. Knol, and J.W.M. Bastiaansen. 2017. Using markers with large effect in genetic and genomic predictions. J. Anim. Sci. 59–71. doi:10.2527/jas2016.0754.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829. doi:11290733.

Oliveira, M.C.S., M.M. Alencar, A.C.S. Chagas, R. Giglioti, and H.N. Oliveira. 2009.

Gastrointestinal nematode infection in beef cattle of different genetic groups in Brazil. Vet. Parasitol. 166:249–254. doi:10.1016/j.vetpar.2009.09.006.

O'Brien, A.M.P., G. Mészáros, Y.T. Utsunomiya, T.S. Sonstegard, J.F. Garcia, C.P. Van Tassell, R. Carvalheiro, M.V.B. da Silva, and J. SooLkner. 2014. Linkage disequilibrium levels in Bos indicus and Bos taurus cattle using medium and high density SNP chip data and different minor allele frequency distributions. Livest. Sci. 166:121–132. doi:10.1016/j.livsci.2014.05.007.

Oliveira, M.C.S., M.M. Alencar, R. Giglioti, M.C.D. Beraldo, F.F. Aníbal, R.O. Correia, L. Boschini, A.C.S. Chagas, T.B. Bilhassi, and H.N. Oliveira. 2013. Resistance of beef cattle of two genetic groups to ectoparasites and gastrointestinal nematodes in the state of São Paulo, Brazil. Vet. Parasitol. 197:168–175. doi:10.1016/j.vetpar.2013.06.021.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Peña, M.T., J.E. Miller, W. Wyatt, and M.T. Kearney. 2000. Differences in susceptibility to gastrointestinal nematode infection between Angus and Brangus cattle in south Louisiana. Vet. Parasitol. 89:51–61. doi:10.1016/S0304-4017(99)00229-0.

Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.-L. Hu, H.H. Kerstens, A.S. Law, H.-J. Megens, D. Milan, D.J. Nonneman, G.A.

Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook, and M.A.M. Groenen. 2009. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PLoS One 4. doi:10.1371/journal.pone.0006524.

Ravagnolo, O., and I. Misztal. 2000. Genetic Component of Heat Stress in Dairy Cattle , Parameter Estimation. J. Dairy Sci. 83:2126–2130. doi:10.3168/jds.S0022-0302(00)75095-8.

Renaudeau, D., A. Collin, S. Yahav, V. de Basilio, J.L. Gourdine, and R.J. Collier. 2012. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. Animal 6:707–728. doi:10.1017/S1751731111002448.

Soma, Y., Y. Uemoto, S. Sato, T. Shibata, H. Kadowaki, E. Kobayashi, and K. Suzuki. 2014. Genome-wide mapping and identification of new quantitative trait loci affecting meat production, meat quality, and carcass traits within a Duroc purebred population. J. Anim. Sci. 89:601–608. doi:10.2527/jas.2010-3119.

Souza, A.P. De, C.I. Ramos, V. Bellato, A.A. Sartor, and C.A. Schelbauer. 2008. Resistência de helmintos gastrintestinais de bovinos a anti-helmínticos no Planalto Catarinense. Ciência Rural 38:1363–1367. doi:10.1590/S0103-84782008000500026.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C.

Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Stratz, P., R. Wellmann, S. Preuss, K. Wimmers, and J. Bennewitz. 2014. Genome-wide association analysis for growth , muscularity and meat quality in Pi etrain pigs. Anim. Genet. 45:350–356. doi:10.1111/age.12133.

Van Tassell, C.P., T.P.L. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, W.C. Warren, and T.S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5:247–252. doi:10.1038/NMETH.1185.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Villa-Angulo, R., L.K. Matukumalli, C.A. Gill, J. Choi, C.P. Van Tassell, and J.J. Grefenstette. 2009. High-resolution haplotype block structure in the cattle genome. BMC Genet. 10:19. doi:10.1186/1471-2156-10-19.

Wientjes, Y.C.J., R.F. Veerkamp, P. Bijma, H. Bovenhuis, C. Schrooten, and M.P.L. Calus. 2015. Empirical and deterministic accuracies of across-population genomic prediction. Genet. Sel. Evol. 47:1–14. doi:10.1186/s12711-014-0086-0.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 8:4107–4116. doi:10.3168/jds.2014-9005.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A.

Boland, J.G. Garnier, D. Boichard, G.M. Lathrop, I.G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics 177:1059–1070. doi:10.1534/genetics.107.075804.

Lopes, M.S., H. Bovenhuis, M. van Son, Ø. Nordbø, E.H. Grindflek, E.F. Knol, and J.W.M. Bastiaansen. 2017. Using markers with large effect in genetic and genomic predictions. J. Anim. Sci. 59–71. doi:10.2527/jas2016.0754.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829. doi:11290733.

O'Brien, A.M.P., G. Mészáros, Y.T. Utsunomiya, T.S. Sonstegard, J.F. Garcia, C.P. Van Tassell, R. Carvalheiro, M.V.B. da Silva, and J. SooLkner. 2014. Linkage disequilibrium levels in Bos indicus and Bos taurus cattle using medium and high density SNP chip data and different minor allele frequency distributions. Livest. Sci. 166:121–132. doi:10.1016/j.livsci.2014.05.007.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.-L. Hu, H.H. Kerstens, A.S. Law, H.-J. Megens, D. Milan, D.J. Nonneman, G.A. Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook, and M.A.M. Groenen. 2009. Design of a High Density

SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PLoS One 4. doi:10.1371/journal.pone.0006524.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Van Tassell, C.P., T.P.L. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, W.C. Warren, and T.S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5:247–252. doi:10.1038/NMETH.1185.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Villa-Angulo, R., L.K. Matukumalli, C.A. Gill, J. Choi, C.P. Van Tassell, and J.J. Grefenstette. 2009. High-resolution haplotype block structure in the cattle genome. BMC Genet. 10:19. doi:10.1186/1471-2156-10-19.

Abasht, B., and S.J. Lamont. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. Anim.

Genet. 38:491–498. doi:10.1111/j.1365-2052.2007.01642.x.

Achi, Y.L., J. Zinsstag, K. Yao, N. Yeo, P. Dorchies, and P. Jacquiet. 2003. Host specificity of Haemonchus spp. for domestic ruminants in the savanna in northern Ivory Coast. Vet. Parasitol. 116:151–158. doi:10.1016/S0304-4017(03)00258-9.

Biegelmeyer, P., L.Q. Nizoli, F.F. Cardoso, and E. Dionello. 2012. Bovine Resistance To Rhipicephalus (Boophilus) Microplus Ticks. Arch. Zootec 61:1–11.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 8:4107–4116. doi:10.3168/jds.2014-9005.

Bulmer, M.G. 1971. The Effect of Selection on Genetic Variability. Am. Nat. 105:201–211.

Cardoso, C.P., B.F. Silva, D.S. Gonçalves, N.J. Tagliari, M.E. Saito, and A.F.T. Amarante. 2014. Resistência contra ectoparasitas em bovinos da raça crioula lageana e meio-sangue angus avaliada em condições naturais. Pesqui. Vet. Bras. 34:141–146. doi:10.1590/S0100-736X2014000200008.

Cattelam, J., and M.M. do Vale. 2013. Estresse térmico em bovinos. Rev. Port. Ciências Veterinárias 108:96–102.

Engelsma, K.A., R.F. Veerkamp, M.P.L. Calus, and J.J. Windig. 2014. Consequences for diversity when animals are prioritized for conservation of the whole genome or of one specific allele. J. Anim. Breed. Genet. 131:61–70. doi:10.1111/jbg.12052.

García-Ruiz, A., J.B. Cole, P.M. VanRaden, G.R. Wiggans, F.J. Ruiz-Lópeza, and C.P. Van Tassell. 2016. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. PNAS E3995–E4004.

doi:10.1073/pnas.1519061113.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A. Boland, J.G. Garnier, D. Boichard, G.M. Lathrop, I.G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics 177:1059–1070. doi:10.1534/genetics.107.075804.

Grevenhof, E.M. Van, J.A.M. Van Arendonk, and P. Bijma. 2012. Response to genomic selection: The Bulmer effect and the potential of genomic selection when the number of phenotypic records is limiting. Genet. Sel. Evol. 44:1–10.

Hansen, P.J. 2004. Physiological and cellular adaptations of zebu cattle to thermal stress. Anim. Reprod. Sci. 82–83:349–360. doi:10.1016/j.anireprosci.2004.04.011.

Kim, E.S., T.S. Sonstegard, M.V.G.B. Silva, L.C. Gasbarre, and C.P. Van Tassell. 2013. Identification of quantitative trait loci affecting gastrointestinal parasite resistance in an experimental Angus population. Anim. Genet. 45:117–121. doi:10.1111/age.12101.

Ledur, M., N. Navarro, and M. Pérez-Enciso. 2010. Large-scale SNP genotyping in crosses between outbred lines: How useful is it?. Heredity (Edinb). 105:173–182. doi:10.1038/hdy.2009.149.

Lopes, M.S., H. Bovenhuis, M. van Son, Ø. Nordbø, E.H. Grindflek, E.F. Knol, and J.W.M. Bastiaansen. 2017. Using markers with large effect in genetic and genomic predictions. J. Anim. Sci. 59–71. doi:10.2527/jas2016.0754.

Machado, M.A., A.L.S. Azevedo, R.L. Teodoro, M.A. Pires, M.G.C.D. Peixoto, C. de Freitas, M.C.A. Prata, J. Furlong, M.V.G.B. da Silva, S.E.F. Guimarães, L.C.A. Regitano, L.L. Coutinho, G. Gasparin, and R.S. Verneque. 2010. Genome wide scan

for quantitative trait loci affecting tick resistance in cattle (Bos taurus x Bos indicus).. BMC Genomics 11:280. doi:10.1186/1471-2164-11-280.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829. doi:11290733.

Oliveira, M.C.S., M.M. Alencar, A.C.S. Chagas, R. Giglioti, and H.N. Oliveira. 2009. Gastrointestinal nematode infection in beef cattle of different genetic groups in Brazil. Vet. Parasitol. 166:249–254. doi:10.1016/j.vetpar.2009.09.006.

O'Brien, A.M.P., G. Mészáros, Y.T. Utsunomiya, T.S. Sonstegard, J.F. Garcia, C.P. Van Tassell, R. Carvalheiro, M.V.B. da Silva, and J. SooLkner. 2014. Linkage disequilibrium levels in Bos indicus and Bos taurus cattle using medium and high density SNP chip data and different minor allele frequency distributions. Livest. Sci. 166:121–132. doi:10.1016/j.livsci.2014.05.007.

Oliveira, M.C.S., M.M. Alencar, R. Giglioti, M.C.D. Beraldo, F.F. Aníbal, R.O. Correia, L. Boschini, A.C.S. Chagas, T.B. Bilhassi, and H.N. Oliveira. 2013. Resistance of beef cattle of two genetic groups to ectoparasites and gastrointestinal nematodes in the state of São Paulo, Brazil. Vet. Parasitol. 197:168–175. doi:10.1016/j.vetpar.2013.06.021.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Peña, M.T., J.E. Miller, W. Wyatt, and M.T. Kearney. 2000. Differences in susceptibility to

gastrointestinal nematode infection between Angus and Brangus cattle in south Louisiana. Vet. Parasitol. 89:51–61. doi:10.1016/S0304-4017(99)00229-0.

Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.-L. Hu, H.H. Kerstens, A.S. Law, H.-J. Megens, D. Milan, D.J. Nonneman, G.A. Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook, and M.A.M. Groenen. 2009. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PLoS One 4. doi:10.1371/journal.pone.0006524.

Ravagnolo, O., and I. Misztal. 2000. Genetic Component of Heat Stress in Dairy Cattle , Parameter Estimation. J. Dairy Sci. 83:2126–2130. doi:10.3168/jds.S0022-0302(00)75095-8.

Renaudeau, D., A. Collin, S. Yahav, V. de Basilio, J.L. Gourdine, and R.J. Collier. 2012. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. Animal 6:707–728. doi:10.1017/S1751731111002448.

Soma, Y., Y. Uemoto, S. Sato, T. Shibata, H. Kadowaki, E. Kobayashi, and K. Suzuki. 2014. Genome-wide mapping and identification of new quantitative trait loci affecting meat production, meat quality, and carcass traits within a Duroc purebred population. J. Anim. Sci. 89:601–608. doi:10.2527/jas.2010-3119.

Souza, A.P. De, C.I. Ramos, V. Bellato, A.A. Sartor, and C.A. Schelbauer. 2008. Resistência de helmintos gastrintestinais de bovinos a anti-helmínticos no Planalto Catarinense. Ciência Rural 38:1363–1367. doi:10.1590/S0103-84782008000500026.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-

origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Stratz, P., R. Wellmann, S. Preuss, K. Wimmers, and J. Bennewitz. 2014. Genome-wide association analysis for growth , muscularity and meat quality in Pi etrain pigs. Anim. Genet. 45:350–356. doi:10.1111/age.12133.

Van Tassell, C.P., T.P.L. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, W.C. Warren, and T.S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5:247–252. doi:10.1038/NMETH.1185.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Villa-Angulo, R., L.K. Matukumalli, C.A. Gill, J. Choi, C.P. Van Tassell, and J.J. Grefenstette. 2009. High-resolution haplotype block structure in the cattle genome. BMC Genet. 10:19. doi:10.1186/1471-2156-10-19.

Wientjes, Y.C.J., R.F. Veerkamp, P. Bijma, H. Bovenhuis, C. Schrooten, and M.P.L. Calus. 2015. Empirical and deterministic accuracies of across-population genomic prediction. Genet. Sel. Evol. 47:1–14. doi:10.1186/s12711-014-0086-0.

Brøndum, R.F., G. Su, L. Janss, G. Sahana, B. Guldbrandtsen, D. Boichard, and M.S. Lund. 2015. Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. J. Dairy Sci. 8:4107–4116. doi:10.3168/jds.2014-9005.

Gautier, M., T. Faraut, K. Moazami-Goudarzi, V. Navratil, M. Foglio, C. Grohs, A. Boland, J.G. Garnier, D. Boichard, G.M. Lathrop, I.G. Gut, and A. Eggen. 2007. Genetic and haplotypic structure in 14 European and African cattle breeds. Genetics 177:1059–1070. doi:10.1534/genetics.107.075804.

Lopes, M.S., H. Bovenhuis, M. van Son, Ø. Nordbø, E.H. Grindflek, E.F. Knol, and J.W.M. Bastiaansen. 2017. Using markers with large effect in genetic and genomic predictions. J. Anim. Sci. 59–71. doi:10.2527/jas2016.0754.

Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829. doi:11290733.

O'Brien, A.M.P., G. Mészáros, Y.T. Utsunomiya, T.S. Sonstegard, J.F. Garcia, C.P. Van Tassell, R. Carvalheiro, M.V.B. da Silva, and J. SooLkner. 2014. Linkage disequilibrium levels in Bos indicus and Bos taurus cattle using medium and high density SNP chip data and different minor allele frequency distributions. Livest. Sci. 166:121–132. doi:10.1016/j.livsci.2014.05.007.

Otto, P.I., S.E.F. Guimarães, L.L. Verardo, A.L.S. Azevedo, J. Vandenplas, A.C.C. Soares, C.A. Sevillano, R. Veroneze, M. de Fatima A. Pires, C. de Freitas, M.C.A. Prata, J. Furlong, R.S. Verneque, M.F. Martins, J.C.C. Panetto, W.A. Carvalho, D.O.R. Gobo, M.V.G.B. da Silva, and M.A. Machado. 2018. Genome-wide association studies for tick resistance in Bos taurus × Bos indicus crossbred cattle: A deeper look into this intricate mechanism. J. Dairy Sci. 101:1–13. doi:10.3168/jds.2017-14223.

Ramos, A.M., R.P.M.A. Crooijmans, N.A. Affara, A.J. Amaral, A.L. Archibald, J.E. Beever, C. Bendixen, C. Churcher, R. Clark, P. Dehais, M.S. Hansen, J. Hedegaard, Z.-L. Hu, H.H. Kerstens, A.S. Law, H.-J. Megens, D. Milan, D.J. Nonneman, G.A. Rohrer, M.F. Rothschild, T.P.L. Smith, R.D. Schnabel, C.P. Van Tassell, J.F. Taylor, R.T. Wiedmann, L.B. Schook, and M.A.M. Groenen. 2009. Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. PLoS One 4. doi:10.1371/journal.pone.0006524.

Sevillano, C.A., J. Napel, S.E.F. Guimarães, F.F. Silva, and M.P.L. Calus. 2018. Effects of alleles in crossbred pigs estimated for genomic prediction depend on their breed-of-origin 1–15. doi:10.1186/s12864-018-5126-7.

Sevillano, C.A., J. Vandenplas, J.W.M. Bastiaansen, R. Bergsma, and M.P.L. Calus. 2017. Genomic evaluation for a three-way crossbreeding system considering breed-of-origin of alleles. Genet. Sel. Evol. 49:1–14. doi:10.1186/s12711-017-0350-1.

Silva, M.V.G.B. da, M.F. Martins, M. de A.R. Cembranelli, J.C. do C. Panetto, L. de C. Paiva, G.S. Gonçalves, M.A. Machado, D.R. de L.R. Faza, and B.R.C. Alves. 2018. Girolando Breed Genetic Improvement Program Sire - Summary - Progeny Test Results - June 2018. 1st ed. E.G. de Leite, ed. Juiz de Fora: Embrapa Gado de Leite.

Van Tassell, C.P., T.P.L. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, W.C. Warren, and T.S. Sonstegard. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. Nat. Methods 5:247–252. doi:10.1038/NMETH.1185.

Vandenplas, J., M.P.L. Calus, C.A. Sevillano, J.J. Windig, and J.W.M. Bastiaansen. 2016. Assigning breed origin to alleles in crossbred animals. Genet. Sel. Evol. 48:1–22. doi:10.1186/s12711-016-0240-y.

Villa-Angulo, R., L.K. Matukumalli, C.A. Gill, J. Choi, C.P. Van Tassell, and J.J. Grefenstette. 2009. High-resolution haplotype block structure in the cattle genome. BMC Genet. 10:19. doi:10.1186/1471-2156-10-19.

Wientjes, Y.C.J., R.F. Veerkamp, P. Bijma, H. Bovenhuis, C. Schrooten, and M.P.L. Calus. 2015. Empirical and deterministic accuracies of across-population genomic prediction. Genet. Sel. Evol. 47:1–14. doi:10.1186/s12711-014-0086-0.

Zhang, Z., J. Liu, X. Ding, P. Bijma, D. De Koning, and Q. Zhang. 2010. Best Linear Unbiased Prediction of Genomic Breeding Values Using a Trait-Specific Marker-Derived Relationship Matrix. PLoS One 5:1–8. doi:10.1371/journal.pone.0012648.

Weller, J.I., E. Ezra, and M. Ron. 2017. Invited review : A perspective on the future of genomic selection in dairy cattle. J. Dairy Sci. 100:8633–8644. doi:10.3168/jds.2017-12879.

Wientjes, Y.C.J., R.F. Veerkamp, P. Bijma, H. Bovenhuis, C. Schrooten, and M.P.L. Calus. 2015. Empirical and deterministic accuracies of across-population genomic prediction. Genet. Sel. Evol. 47:1–14. doi:10.1186/s12711-014-0086-0.

Zhang, Z., J. Liu, X. Ding, P. Bijma, D. De Koning, and Q. Zhang. 2010. Best Linear Unbiased Prediction of Genomic Breeding Values Using a Trait-Specific Marker-Derived Relationship Matrix. PLoS One 5:1–8. doi:10.1371/journal.pone.0012648.