

Genomic association study for adaptability traits in four Colombian cattle breeds

C. De León¹, C. Manrique¹, R. Martínez² and J.F. Rocha²

¹ Universidad Nacional de Colombia, Bogotá, Colombia

² Corporación Colombiana de Investigación Agropecuaria – Agrosavia, Mosquera, Cundinamarca, Colombia

Corresponding author: J.F. Rocha
E-mail: jfmartinez@agrosavia.co

Genet. Mol. Res. 18 (3): gmr18373

Received May 25, 2019

Accepted July 07, 2019

Published September 26, 2019

DOI <http://dx.doi.org/10.4238/gmr18373>

ABSTRACT. The aim of this study was to perform a genome wide association study for the coefficient of adaptability, the heat tolerance coefficient, age at first calving and calving interval, using 58,868 and 57,482 single nucleotide polymorphisms obtained from 1262 Blanco Orejinegro and 742 Sanmartinero individuals, implementing a single-step-genomic-BLUP (ssGBLUP) methodology. A total of 25 and 11 genomic regions in BON and SM were associated with the traits evaluated. In these regions; several genes involved in the regulation of reproduction and adaptability were identified. Some of them, such as the RPTOR, TM2D1, RAB21 and ACOT13 genes, are involved in the physiological response to heat stress. SLC6A16, TMEM50A, NLRP9, KHDRBS2 and UBE2C are possibly associated to reproductive performance. The results of this study indicate an association between some genomic regions and reproductive and adaptability-related traits in Colombian cattle breeds. Several genes contained in these regions can potentially be used in animal breeding programs to select for adaptability within cattle production systems located in a tropical environment.

Key words: Blanco Orejinegro; Sanmartinero; SNP; Heat tolerance coefficient; Coefficient of adaptability

INTRODUCTION

Genomic wide association studies (GWAS) have been broadly used as a robust methodological procedure in the identification of regions of the genome and genes that

influence the expression of characteristics of interest (Meuwissen et al., 2001; Hayes, 2011; Nicolazzi et al., 2015) in the implementation of genetic improvement programs (Costa et al., 2015). Most traits analyzed using GWAS in cattle are related to milk production, growth and reproduction (Berry et al., 2013), but few GWAS have been performed on adaptation traits in cattle. Bejarano et al. (2013) found several single nucleotide polymorphisms (SNPs) associated with the coefficient of adaptability (CA) and the heat tolerance coefficient (HTC) in four Colombian creole cattle breeds. Dikmen et al. (2013) reported genomic regions and genes associated with the control of the rectal temperature in dairy cattle exposed to a stressful environment. These authors observed a high genetic variance (0.44%) for rectal temperature explained by SNPs located on BTA24. On the other hand, Hawken et al. (2012) reported significant association of SNPs with age at puberty, postcalving interval and first postcalving ovulation, located on chromosomes 5, 14 and 16.

The Colombian Creole cattle have been raised in a tropical environment for centuries, and they are part of conservation and genetic improvement programs, implementing pure and cross-breeding mating systems in order to maintain the genetic diversity of *in vivo* and *in vitro* germplasm banks, but also to improve the productivity of cattle populations in the country. Several Colombian cattle breeds show a better reproductive performance measured by age at first calving (AFC) and calving interval (CI) performance (Rocha et al., 2012), disease resistance and high heat tolerance (Rocha et al., 2019), attributes that make these animals an attractive genetic resource (Correal and Henao, 2000). The objective of this study was to perform a GWAS for traits that indicate adaptability such as CA, HTC, AFC and CI in the Colombian cattle breeds BON and SM, identifying genomic regions and genes associated with the phenotypic variance of these traits.

MATERIAL AND METHODS

Study location, animals and phenotypic data

A total of 1262 Blanco Orejinegro (BON) and 742 Sanmartinero (SM) animals from the *in vivo* germplasm bank of the Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA) were used for this study. BON animals were maintained in the research center El Nus in Antioquia, in a mountainous area in the Andean region with altitudes ranging from 800 to 1800 m. Topography is undulating with annual average temperatures between 18 and 24°C and a bimodal rainfall regime. SM animals were located in the La Libertad research center in Villavicencio, Meta, in Colombian Orinoquia on flat lands with heights of 336 m above sea level in a humid tropical forest zone. There is a rainy season from April to November and a dry season from December to March, with annual average temperatures that range between 27 and 41°C and relative humidity from 36 to 87%. Both BON and SM populations were maintained under semi-extensive grazing conditions.

Metadata

To estimate CA and HTC, rectal temperature (RT) and respiratory rate (RR) data was obtained. RT was measured using a veterinary digital thermometer and the RR by

direct observation of respiratory movements. These variables were taken between 12:00 P.M. and 3:30 P.M.

To estimate the HTC, the following equation was used:

$$HTC = 100 - [18 \times (RT - 38.3)] \quad (\text{Eq. 1})$$

where: 100 is the maximum efficiency to maintain the temperature between normal ranges, 18 is a constant, RT is the average rectal temperature recorded during the current study, and 38.3 is the normal rectal temperature of cattle (Bianca, 1963).

The coefficient of adaptability (CA) was estimated employing the following equation:

$$CA = (RT/38.33) + (RR/23.0) \quad (\text{Eq. 2})$$

where: RT is the average rectal temperature recorded during the current study, 38.3 is the normal bovine temperature, RR is the respiratory rate recorded during the study, and 23.0 is the normal RR of animals (Benezra, 1954).

The temperature-humidity index (THI) was estimated from data on environmental temperature (ET) and relative humidity (RH) recorded during the same years and time that physiological constants RT and RR were recorded for both breeds. The records of these environmental variables were taken from meteorological databases of each research center. The ET and RH were recorded using ultralight and small button-sized sensor devices (ibuttons), which were positioned directly in the paddock where animals grazed, and were programmed to measure ET and RH every 15 minutes from 12:00 to 3:30 P.M.

The THIs was estimated using the following equation:

$$THI = 1.8xET - (1 - RH/100)x(ET - 14.3) + 32 \quad (\text{Eq. 3})$$

where: ET is the environmental temperature (°C) and RH is the relative humidity (%) (Valtorta et al., 2000).

Genealogical and reproductive records were collected between 1980 and 2017 for BON and SM. A total of 1927 AFC and 5150 CI records were used. Genealogical information included between four and six generations in these populations.

Records from animals of doubtful or with repeated identifications and with extreme or negative numbers were eliminated. In the BON breed, CA values of 1.6 - 4.8, HTC values of 27.0 - 183.0, AFC values of 775 - 1306 days and CI values of 313 - 1004 days, were accepted. For SM, CA values of 2.0 - 4.54, HTC values of 70 - 116, AFC of 667 - 1501 days and CI values of 310 - 919 days, were accepted.

Genotyping and quality control of genotypes

In total, 58,868 SNPs for BON and 57,482 SNPs for SM, were included in the analysis. These SNPs were obtained from 1262 BON and 742 SM genotyped animals (Table 1).

From the group of BON, 129 animals were genotyped using the IlluminaSNP7K low density chip, 490 animals with the 20KGeneSeek chip and 643 animals with the BovineSNP50K medium density chip. In the case of SM, 126 were genotyped with low density chip IlluminaSNP7K, 418 animals with 20KGeneSeek chip and 198 with the medium density chip BovineSNP50K. SNP and their genotypes from the 7K and 20K chips

were imputed to the 50K chip using the Fimpute program (Sargolzaei et al., 2014). After imputation, quality control (Quality Control-QC) was performed on 58,868 available SNPs for the BON breed and 57,482 for the SM breed, using the options for this purpose included in the family of BLUPF90 programs. SNPs were excluded when the call rate was <90%, if they had an extreme deviation from Hardy-Weinberg equilibrium ($P < 0.01$) (for example, SNPs in autosomal chromosomes presenting both homozygous genotypes, but without heterozygotes), if the SNP had an unknown genomic position, if they were located in sex chromosomes, if they were monomorphic or if the minor allele frequency (MAF) was <0.03. Likewise, animals that had a call rate <90% or with pedigree errors or Mendelian conflicts between relatives were eliminated. Finally, 51,283 and 49,009 effective polymorphic SNPs located in autosomal chromosomes were obtained in BON and SM, respectively. The genomic kinship matrix (G), the numerator matrix or pedigree of genotyped animals (A22) and the H matrix and its inverse were constructed with PREGF90. Matrix H, combines genomic and pedigree information. These matrices, (matrix G or its inverse, matrix A or its inverse) were used to perform quality control process of the SNP. To develop this quality control, we used the OPTIONS FreqFilev <file> and OPTION chrinfo < file> to build chromosome mapping files (Misztal et al., 2015).

Table 1. Number of Blanco orejinegro (BON) and Sanmartinero (SM) cattle analyzed according to the genotyping density.

Breed	20KGeneSeek	BovineSNP50K	IlluminaSNP7K	Total
BON	490	643	129	1262
SM	418	198	126	742

Statistical analysis

Association analysis was carried out following the methodology of Single-step genomic Association Study (ssGWAS), approach proposed by Wang et al. (2012), based on the single-step genomic-BLUP method (ssGBLUP) (Aguilar et al., 2010; Misztal et al., 2013). The H matrix inverse was constructed by combining genomic matrix inverse (G^{-1}) derived from SNPs (G) and numerator matrix inverse (A^{-1}) derived from, as follows:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \quad (\text{Eq. 4})$$

where: A_{22}^{-1} is the inverse numerator matrix for genotyped animals and G is the genomic relationships matrix. Matrix G is constructed by weighting the contribution from expected variance of each marker, according to the methodology described by VanRaden (2008), with:

$$G = ZDZ' \quad (\text{Eq. 5})$$

where: Z is the incidence matrix of markers containing recoded genotypes (0, 1 or 2), corrected by the allele's frequency, D is a diagonal inverse matrix containing the expected variance of markers:

$$D_{ii} = \frac{1}{m[2p_i(1-p_i)]}, \quad (\text{Eq. 6})$$

were: m is the number of markers, p_i is the frequency of the second allele. Matrix H^{-1} was replaced in mixed model equation.

The mixed animal model implemented is described below:

$$y = X\beta + Za + e \quad (\text{Eq. 7})$$

were: y = vector $n \times 1$ of trait observations; β = vector $p \times 1$ of solutions for fixed family effects, month in which physiological constants were recorded, THI and animal age as covariates for HTC and CA; solutions vector $p \times 1$ to fixed effects of family, year and calving month to AFC; solutions vector $p \times 1$ to fixed family effects, calving age and delivery order as covariates for CI; X = known incidence matrix $n \times p$ that associates β with y ; a = solutions vector $n \times 1$ to genetic values; Z = random effects matrix that associates a with y ; e = vector $n \times 1$ of peculiar random residual effect of observations not explained by other parts of the model.

Mixed animal model equations for CA, HTC, AFC and CI are described below:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \alpha H^{-1} \end{bmatrix} \begin{bmatrix} \beta \\ a \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (\text{Eq. 8})$$

where: H^{-1} is a matrix that integrates A^{-1} and G^{-1} matrices. Matrix A is the numerator or pedigree matrix and G genomic kinship matrix constructed from markers frequencies (Legarra et al., 2009); X and Z are incidence matrices mentioned in the model. Term α is a scalar that relates residual variance and the genetic variance: $\alpha = \sigma^2R/\sigma^2A$ (Mrode and Thompson, 2014).

To calculate variance and covariance components, a uni-character analysis was performed with AIREMLF90 program in the BLUPF90 family programs (Misztal, 2010) using the restricted maximum likelihood (REML) algorithm (Thompson and Mäntysaari, 1999) and applying a mixed model for each trait.

Then, the solutions of fixed and random effects (GEBV) were estimated using a single-step genomic BLUP (ssGBLUP) included in the software BLUPF90 (Misztal et al., 2009; Aguilar et al., 2010). In this study, PREGSF90 and POSTGSF90 were used to perform ssGWAS analyzes for each characteristic, following the methodology described by Wang et al. (2012). Estimated breeding values from genotyped animals (GEBV) were converted in allelic substitution effect of each SNP (Wang et al., 2012). The equation to predict SNPs effects using weighted genomic relations matrix was:

$$\hat{u} = DZ' [ZDZ']^{-1} \hat{a}g \quad (\text{Eq. 9})$$

where: \hat{u} is a vector of estimated SNP markers effects, D is a diagonal significance matrix for SNP effect variances, Z is a matrix relating genotypes of each locus for each individual and ag is additive genetic effect for animals with genotype. Individual variance for effect of each SNP (as in D) was estimated as:

$$\sigma^2 u_i = u_i^2 2p_i(1-p_i) \quad (\text{Eq. 10})$$

where: u_i^2 is square effect of the i -th SNP, p_i is allelic frequency observed from second allele of the i -th marker in current population. When windows of n adjacent SNP are used, like in this study, variance attributed to these Windows, is calculated by adding variance of n adjacent SNPs, for each SNP.

Genetic variance explained

The percentage of genetic variance explained by the i -th region was calculated as follows:

$$\frac{\text{Var}(a_i)}{\sigma_a^2} \times 100\% = \frac{\text{Var}(\sum_{j=1}^n Z_j \hat{u}_j)}{\sigma_a^2} \times 100\% \quad (\text{Eq. 11})$$

where: a_i is the genetic value of the i -th region consisting from n contiguous SNP, a_2a is the total genetic variance, Z_j is a vector of gene content of the j -th SNP for all individuals, and \hat{u}_j is marker effect of the j -th SNP within the i -th region (Wang et al., 2012). Solution files for SNPs indicating allele substitution effect and variance explained by each window of four contiguous SNPs, were used to build Manhattan Plot graphs using POSTGSF90.

Gene mapping and functional analysis

Regions surrounding significant SNP, were inspected for identification of possible QTLs and candidate genes located within each region at a distance of 1 Mb, using public domain databases such as Genome National Center for Biotechnology Information (Ncbi, 2019) and Ensembl Genome Browser (Ensembl, 2019). Functional analyses of mapped genes were carried out through the website UniProt (2019) and GeneCards® (2019), which were used to verify the functional information of the genes.

RESULTS AND DISCUSSION

Descriptive statistics for CA, HTC, AFC and CI in both BON and SM are summarized in Table 2. Average values found showed a high phenotypic variation.

Table 2. Mean, standard deviation (SD), minimum (min) and maximum (max) values for adaptability traits in Blanco Orejinegro (BON) and Sanmartinero (SM) cattle.

Indices	Breed	Mean	SD	Min	Max	Total
Coefficient of variability	BON	2.7	0.43	1.6	4.8	1,742
	SM	3.5	0.46	2.0	4.5	1,075
Heat tolerance coefficient	BON	95.7	5.8	27.0	183.0	1,742
	SM	90.7	13.0	70.0	1,16.2	1,072
Age at first calving (days)	BON	1,086	54.5	775	1,301	958
	SM	1,172	157.7	667	1,501	970
Calving interval (days)	BON	509.0	159.5	313	1,004	2,293
	SM	467.0	135.2	310	919	2,857

For the identification of genomic regions associated with the traits evaluated, a threshold was defined according to the percentage of genomic variability (gVar) explained by four adjacent SNP windows, accepting as significant regions those with a gVar (%) ≥ 0.13 . The identified significant regions for the four traits are shown in Table 3.

Table 3. Number of genomic regions with a significant effect on four adaptability traits in from GWAS adaptability traits in Blanco Orejinegro (BON) and Sanmartinero (SM) cattle.

Trait	Breed	Genomic regions ¹	%gVar ²
Coefficient of adaptability	BON	6	1.17
	SM	2	0.27
Heat tolerance coefficient	BON	6	0.95
	SM	0.0	0.00
Age at first calving	BON	6	0.96
	SM	6	0.87
Calving interval	BON	7	1.12
	SM	3	0.44

¹Number of significant regions that explain a proportion of gVar ≥ 0.13 for each trait. ²Proportion of gVar explained by the genomic regions associated with a particular trait.

In BON, six regions explained 1.17% of the gVar for CA (Table 3). These regions were identified on BTA3, BTA5, BTA7, BTA8, BTA9 and BTA27 (Figure 1). In SM, two regions on BTA3 and BTA19 explained 0.277% of the gVar for this trait.

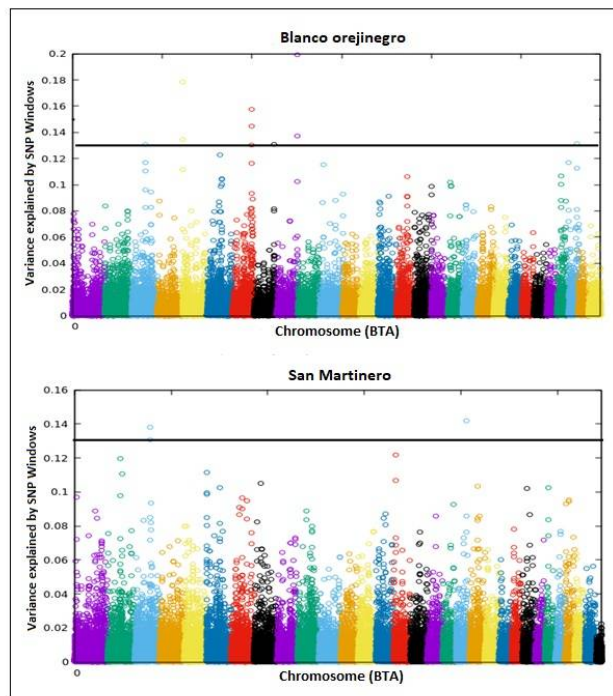


Figure 1. Manhattan plot for coefficient of adaptability in Blanco orejinegro and Sanmartinero cattle breeds.

Table 4. Genomic regions, allelic substitution effect and possible candidate genes associated to the coefficient of adaptability (CA), the heat tolerance coefficient (HTC), age at first calving (AFC) and calving interval (CI) in Blanco Orejinegro (BON) and Sanmartinero (SM) cattle.

Trait	Breed	BTA	Window QTL (bp)	%gVar	Significant SNP	Allelic substitution effect	Possible Candidate Genes (PCG)
CA	BON	3	66.486.992-66.554.349	0.13	BTA-110572-no-rs	3.7	<i>GIPC2</i>
		5	1.339.951-1.419.707	0.17	BovineHD0500000347	1.2	<i>RAB21</i>
		7	98.485.273-98.534.197	0.15	UFL-rs137601357	0.4	<i>Non- annotated genes</i>
		8	97.968.235-98.056.575	0.13	BTA-102583-no-rs	3.0	<i>Non- annotated genes</i>
		9	103.612.265-103.700.559	0.19	ARS-BFGL-NGS-112262	2.1	<i>Non- annotated genes</i>
	SM	27	40.870.573-40.951.386	0.13	BovineHD2700011878	3.2	<i>Sin genes anotados</i>
		3	83.833.592-83.952.776	0.13	ARS-BFGL-NGS-65863	2.9	<i>PATJ</i> <i>TM2D1</i>
		19	51.767.412 – 51.842-200	0.14	UA-IFASA-8764	2.8	<i>RPTOR</i>
		1	56.648.072-56.778.172	0.15	BovineHD0100016005	1.6	<i>ABHD10</i> <i>TAGNLN3</i>
		2	122.461.245-122.551.247	0.18	BTA-107470-no-rs	0.6	<i>PUM1</i>
HTC	BON	12	2.947.397-2.991.581	0.13	BovineHD1200000945	3.7	<i>ENSBTAG00000048324</i>
		12	41.635.295-41.872.77	0.16	BTA-100318-no-r	0.8	<i>Sin genes anotados</i>
		23	14.170.863-14.246.801	0.20	ARS-BFGL-NGS-34854	2.5	<i>Sin genes anotados</i>
		23	32.998.188-33.085.360	0.13	Hapmap57845-rs29014813	3.4	<i>ARMH2</i> <i>GMNN</i> <i>C23H6ORF62</i> <i>ACOT13</i> <i>TDP2</i>
		2	127.626.236-127.645.905	0.13	BovineHD0200037053	2.9	<i>ENSBTAG00000001297</i>
	SM	2	127.646.623-127.653.392	0.23	BovineHD0200037058	1.5	<i>RF00026</i> <i>ENSBTAG00000001297</i>
		2	127.653.946-127.664.126	0.17	BovineHD020003706	1.2	<i>TMEM50A</i> <i>ENSBTAG00000001297</i>
		8	93.404.723-93.506.193	0.16	BTB-01364301	1.0	<i>Sin genes anotados</i>
		10	102.935.702-103.048.273	0.13	Hapmap49784-BTA-87663	2.8	<i>CTDSPL2</i> <i>EIF3J</i> <i>SPG11</i> <i>RF0026</i> <i>RF0020</i>
		18	55.779.877-55.831.611	0.14	BovineHD1800016339	4.3	<i>SLC6A16</i> <i>TEAD2</i> <i>CD37</i>
AFC	BON	8	97.989.840-98.097.344	0.14	ARS-BFGL-NGS-101099	3.3	<i>ENSBTAG00000048201.</i>
		14	41.696.728-41.809.704	0.15	BTB-01670622	3.2	<i>PKIA.</i>
		15	10.517.607-10.700.852	0.15	BTB-00580688	3.0	<i>Sin genes anotados</i>
		23	402.797-764.842	0.16	BTA-55427-no-rs	2.8	<i>KHDRBS2</i>
	SM	25	30.643.276-30.736.444	0.14	ARS-BFGL-NGS-106963	2.8	<i>Sin genes anotados</i>
		29	34.676.788-34.746.856	0.13	ARS-BFGL-NGS-102334	2.9	<i>NTM.</i>

Trait	Breed	BTA	Window QTL (bp)	%gVar	Significant SNP	Allelic substitution effect	Possible Candidate Genes (PCG)
IEP	BON	2	1.164.256–1.216.101	0.16	BFGL-NGS-6050	3.4	<i>CYFIP1</i>
		5	93.016.099–93.093.486	0.15	BovineHD0500026412	4.8	<i>Sin genes anotados</i>
		8	79.660.034–79.817.203	0.22	BTA-90142-no-rs	4.2	<i>Sin genes anotados</i>
		11	38.942.740–39.104.584	0.17	BovineHD1100011547	4.6	<i>CCDC85A</i>
		11	73.198.161–73.317.267	0.13	Hapmap58572-ss46526046	2.8	<i>ADGRF3 HADHB HADHA GAREM2 ENSBTAG00000054613 NLRP9 VSTM1 ENSBTAG00000051657 ENSBTAG00000052278 ENSBTAG00000053115 ENSBTAG0000004963</i>
		18	61.701.988–61.854.328	0.16	BovineHD1800017801	3.6	<i>ENSBTAG00000052278 ENSBTAG00000053115 ENSBTAG0000004963</i>
		20	22.205.878–22.207.811	0.13	ARS-BFGL-NGS-65373	2.9	<i>Sin genes anotados</i>
	SM	5	5.550.983–5.644.397	0.14	BTB-00215816	3.3	<i>PHLDA1.</i>
		13	74.579.328–74.66.7264	0.14	ARS-BFGL-NGS-111155	3.6	<i>DNTTIP1, UBE2C, TNNC2, SNX21, ACOT8, ZSWIM3, ZSWIM1, NEURL2, SPATA25, CTSB, PLTP.</i>
		23	41.064.620–41.173.272	0.16	Hapmap60333-rs29018032	3.4	<i>Non- annotated genes</i>

There were several polymorphisms associated with CI. In BON, seven regions located on BTA2, BTA5, BTA8, BTA11, BTA18 and BTA20 were highly associated and explained 1.12% of the gVar for CI. In the case of SM, three genomic regions on BTA5, BTA13 and BTA23 explained 0.44% of the gVar of this trait (Table 3).

DISCUSSION

There were 14 genomic regions significantly associated with the coefficients of adaptability CA and HTC located on chromosomes BTA1, BTA2, BTA3, BTA5, BTA7, BTA8, BTA9, BTA12, BTA19, BTA23 and BTA27 (Table 4). The TM2 domain-containing protein 1 gene (TM2D1) present on BTA3 has a G protein-coupled receptor activity in the cell signaling pathway and in apoptosis. Salunga et al. (2007) found a relation between this gene and the expression of heat shock proteins in humans. Moreover, the associated SNP BovineHD0500000347 found on BTA5 (Table 4) is close to the Ras-related protein Rab-21 gene (RAB21) that codes for G-RAS (GTPase) proteins in the signaling cell pathway acting on immune system and inhibiting bacterial endotoxin receptors (Pingli et al., 2019). Genes intervening in microorganisms or tick's immune responses in cattle have been associated with adaptive mechanisms. The expression of these genes might explain the higher tick burden in cattle with lower adaptability and higher susceptibility to experience a thermal discomfort (CA >2), as it has been observed in Colombian cattle breeds (Rocha et al., 2019). Another important associated region on BTA19 contains the regulatory associated protein of MTOR complex 1 gene (RPTOR). This gene affects environmental stress processes in chickens (Chen et al., 2010), mediating the transcription of the HSP70, which

codes for a heat shock protein. Likewise, Sun et al. (2010) reported an allelic variant of RPTOR associated with the response to environmental stress in humans. On BTA23, the SNP Hapmap57845-rs29014813 (Table 4), close to the Acyl-CoA thioesterase 13 gene (ACOT13), has a high expression level during cellular oxidative processes, regulating hepatic glucose and lipid metabolism, and recently associated with thermogenesis mechanisms in mice by adaptive responses at high temperatures (Won et al., 2013), which might explain why these genomic regions were significantly associated with HTC and CA. Several DNA regions associated with CA and HTC were reported by Bejarano et al. (2013) in Colombian creole cattle breeds. These were present on BTA1, BTA3, BTA5, BTA9, BTA11, BTA12, BTA15, BTA19, BTA20, BTA21, BTA24 and BTA25. Some of these chromosomes correspond to those where significant associations were identified in the present study

A GWAS performed in dairy cattle by Macciotta et al. (2017) showed a SNP (ARS-BFGL-NGS-29678 on BTA26) associated to heat tolerance, and Dikmen et al. (2013) identified a SNP also located on BTA26 that was associated with sweating rate in Holstein cows during heat stress. Another aspect to be taken into account in the GWAS for heat tolerance traits is the associated phenotype. Dikmen et al. (2015), identified 1 of 4 SNPs previously associated to milk production and then was associated to physiological response to heat stress.

A total of 22 regions located on BTA2, BTA5, BTA8, BTA10, BTA11, BTA13, BTA14, BTA15, BTA18, BTA20, BTA23, BTA25 and BTA29 were associated with AFC and CI (Table 4). The SNP ARS-BFGL-NGS-111155 on BTA13 is located close to the Ubiquitin conjugating enzyme E2C gene (UBE2C), that is highly expressed in the testis (Uniprot, 2019) and it was reported by Fujioka et al. (2018) intervening in the meiosis progression of porcine oocytes. The SNP Bovine HD020003706 identified on BTA2 is likely associated with the transmembrane protein 50A gene (TMEM50A). Similarly, Reverter et al. (2016) found an association between the SNP ARS-BFGL-NGS-116714 in the coding region of the TMEM27 gene and some fertility indicators in Brahman cows. In humans, the TMEM50A gene was reported like a tumor suppressor (Schmit and Michiels, 2018). Nevertheless, the TMEM18 (BTA2) and TMEM38B (BTA9) genes were associated with the first menstruation of women (Chizu et al. 2013), indicating the strong association between transmembrane genes and reproductive traits. On BTA23, the SNP BTA-55427-no-rs is located close to the KH RNA Binding Domain Containing, Signal Transduction Associated 2 gene (KHDRBS2) that has been shown to influence the fertility of goats (Guo et al., 2018). A SNP (BovineHD230000143) detected within the coding region of KHDRBS2 was found significantly associated with fertility traits in Brahman cows (Reverter et al., 2016).

On BTA18, the regions associated with AFC contain the solute carrier family 6 member 16 (SLC6A16), the TEA domain transcription factor 2 (TEAD2) and the NLR Family Pyrin Domain Containing 9 (NLRP9) genes. A GWAS performed in Holstein found an association between a gene that also belongs to this solute carrier family (SLC40A1) and the interval between the luteal activity and the first insemination (Tenghe et al., 2016). Similarly, Yodklaew et al. (2017) found a significant association between AFC and the SNP ARS -BFGL-NGS-44288 that is close to the SLC7A2 gene on BTA27 in crossbred cows from Thailand. Furthermore, other studies carried out in Brahman (Reverter et al., 2016) and Holstein cows (Dikmen et al., 2013) have shown an association between adaptability

traits and genes such as *SLC7A5* and *SLCO1C1*. The former is highly expressed in the endometrium of pregnant cows, while the latter is associated with the response to environmental stress. In the case of *TEAD2*, it has been associated with embryonic development during the early organ's and muscles formation, possibly due to its intervention on the embryonic heart morphogenesis, development of mesoderm, vasculogenesis and embryonic cell differentiation (Wang et al., 2018). The other gene of interest found on BTA18, located close to a SNP (Bovine HD1800017801) highly associated with fertility traits in the present study, was *NLRP9*. This is highly expressed in ovaries, testis, oocytes and participates during embryonic preimplantation, gametogenesis, folliculogenesis and early embryonic development (Ponsuksili et al., 2006).

Several studies have found DNA regions highly associated with fertility traits in *Bos taurus* and *Bos indicus* cattle. A GWAS carried out in crossbred cows in Thailand found 24 SNP on BTA1, BTA3, BTA9, BTA10, BTA11, BTA19, BTA25, BTA27 and BTA29 with an effect on the fertility of these animals (Yodklaew et al., 2017). Costa et al. (2015) found 42 regions associated with AFC in Nellore cows, and located on BTA2, BTA3, BTA4, BTA7, BTA8, BTA9, BTA10, BTA11, BTA13, BTA14, BTA16, BTA18, BTA20, BTA21, BTA23, BTA25 and BTA27. Hawken et al. (2012) reported 165 SNP associated with age at puberty in pure and crossbred Brahman cows, from which 65 SNP were located on BTA14 and BTA15. Similarly, Fortes et al. (2012) showed 134 SNP with a significant effect on age at puberty, and 32 of them were found on BTA14. All these studies show multiple regions where polymorphisms with a marked effect on adaptability traits in cattle were found. Several of these chromosomes coincide with those found with significant regions in the present study. However, differences among the specific regions associated with adaptability traits in these studies might be attributed to variations in the number of animals genotyped, the density of genotyping, differences in the genetic structure of the populations evaluated (linkage disequilibrium, allelic frequencies, etc), the phenotype selected and the amount and quality of records collected.

CONCLUSIONS

A total of 14 regions and 56 SNPs associated with CA and HTC, and 22 regions and 88 SNP associated with AFC and CI was identified. Several SNPs were located near genes with a role in biological mechanisms and pathways that might influence the adaptability of cattle to a tropical environment. Genes such as *RPTOR*, *TM2D1*, *RAB21* and *ACOT13* are possibly associated with the physiological response to environmental stress, while *SLC6A16*, *TEAD2*, *TMEM50A*, *NLRP9*, *KHDRBS2* and *UBE2C* could affect fertility by intervening in reproductive processes. The associated genotypes that were identified can potentially be used in animal breeding programs to select for adaptability within cattle production systems in a tropical environment.

REFERENCES

- Aguilar I, Misztal I, Johnson DL, Legarra A, et al. (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: 743-752. doi: 10.3168/jds.2009-2730.
- Bejarano GDH, Martínez SRA, Cardozo CJ, Lucero EC, et al. (2013). Caracterización fisiológica y genética para atributos de adaptación a estrés ambiental en cuatro razas criollas colombianas. rccp: *Revista Nacional de Ciencias*

- Pecuarias*. 26: 2256-2958, suplemento. Memorias XII Encuentro Nacional y V Internacional de los Investigadores de las Ciencias Pecuarias ENCIP.
- Benezra MV (1954). A new index for measuring the adaptability of cattle to tropical conditions. *J. Animal Sci.* 13: 1015.
- Berry DP, Kearney JF, Twomey K, Evans RD (2013). Genetics of reproductive performance in seasonal calving dairy cattle production systems. *Irish J. Agr. Food Res.* 52: 1-16.
- Bianca W (1963). Rectal temperature and respiratory rate as indicators of heat tolerance in cattle. *J. Agr. Sci.* 60: 113-120.
- Chen ZY, Zhang WW, Gan JK, Kong LN, et al. (2010). Genetic effect of an A/G polymorphism in the *HSP70* gene on thermotolerance in chicken. *Genet. Mol. Res.* 15(2): gmr8271.
- Chizu T, Yukinori O, Atsushi T, Katsutoshi O, et al. (2013). Genome Wide Association Study of Age at Menarche in the Japanese Population. *PLoS one.* 8: 5, e63821.
- Correal GM and Henao FG (2000). El ganado criollo Sanmartinero (SM) y su potencial productivo. *Anim. Genet. Resour. Inf.* 28: 7-17.
- Costa RB, Camargo GMF, Diaz IDPS, Irano NMM, et al. (2015). Genome-wide association study of reproductive traits in Nellore heifers using Bayesian inference. *Genet. Sel. Evol.* 47(1): 67. doi: 10.1186/s12711-015-0146-0.
- Dikmen S, Cole JB, Null DJ and Hansen PJ (2013). Genome-wide association mapping for identification of quantitative trait loci for rectal temperature during heat stress in Holstein cattle. *PLoS one.* 8: e69202. doi: 10.1371/journal.pone.0069202.
- Dikmen S, Wang X, Ortega MS, Cole JB, et al. (2015). Single nucleotide polymorphisms associated with thermoregulation in lactating dairy cows exposed to heat stress. *J. Anim. Breed Genet.* 132: 409-419. <https://doi.org/10.1111/jbg.12176>.
- Ensembl (2019). Ensembl Genome Browser. <http://www.ensembl.org/index.html>. Accessed 17 January 2019.
- Fortes MRS, Li Y, Collis E, Zhang Y, et al. (2012). The IGF1 pathway genes and their association with age of puberty in cattle. *Anim. Genet.* 44: 91-95. <https://doi.org/10.1111/j.1365-2052.2012.02367.x>
- Fujioka YA, Onuma A, Fujii W, Sugiura K, et al. (2018). Contributions of UBE2C and UBE2S to meiotic progression of porcine oocytes. *J. Reprod. Devel.* 64: 253-259. doi: 10.1262/jrd.2018-006. Epub 2018 Mar 23.
- GeneCards (2019). GeneCards Human Gene Database. In: Weizmann Inst. Sci. <http://www.genecards.org/>. Accessed 17 March 2019.
- Guo J, Tao H, Li PLL, et al. (2018). Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Sci. Rep.* 8: 10405.
- Hawken RJ, Zhang YD, Fortes MR, Collis E, et al. (2012). Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J. Anim. Sci.* 90: 1398-1410.
- Hayes BJ (2011). Ben Hayes course notes. Toulouse.
- Legarra A, Aguilar I, Misztal I (2009). A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92: 4656-4663. doi: 10.3168/jds.2009-2061.
- Macciotta NPP, Biffani S, Bernabucci U, Lacetera N, et al. (2017). Derivation and genome-wide association study of a principal component-based measure of heat tolerance in dairy cattle. *J. Dairy Sci.* 100: 4683-4697. <https://doi.org/10.3168/jds.2016-12249>.
- Meuwissen THE, Hayes BJ and Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* 157: 1819-1829. doi: 11290733.
- Misztal I, Legarra A and Aguilar I (2009). Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.* 92: 2009-2064. doi: 10.3168/jds.2009-2064.
- Misztal I (2010). BLUPF90. A flexible mixed model program in Fortran 90. Animal and Dairy Science, University of Georgia.
- Misztal I, Aggrey SE and Muir WM. (2013). Experiences with a single-step genome evaluation. *Poult. Sci.* 92: 2530-2534. doi: 10.3382/ps.2012-02739.
- Misztal I, Tsuruta S, Aguilar I, Legarra A, et al. (2015). Manual for BLUPF90 family of programs.
- Mrode RA and Thompson R (2014). Linear Model for the Prediction of Animal Breeding. *Cabi* 3rd Edition. Editorial Gutenberg, Malta.
- Ncbi (2019). National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/snp>. Accessed 17 January 2019.
- Nicolazzi EL, Biffani S, Biscarini F, Orozco TWP, et al. (2015). Software solutions for the livestock genomics SNP array revolution. *Anim. Genet.* 46: 343-353. doi: 10.1111/age.12295.
- Pingli Yong-Hongwu, Yan-Tingzhu, Man-Xiangli, et al. (2019). Requirement of Rab21 in LPS-induced TLR4 signaling and pro-inflammatory responses in macrophages and monocytes. *Biochem. Biophys. Res. Commun.* 508: 169-176.
- Reverter A, Porto-Neto LR, Fortes MRS, McCulloch R, et al. (2016). Genomic analyses of tropical beef cattle fertility based on genotyping pools of Brahman cows with unknown pedigree. *J. Animal Sci.* 94: 4096-4108. doi:10.2527/jas2016-0675.
- Rocha JF, Gallego JL, Vásquez RF, Pedraza JA, et al. (2012). Estimation of genetic parameters for age at first calving and calving interval in Blanco Orejinegro (BON) breed cattle populations in Colombia. *Rev. Colomb. Cienc. Pec.* 25: 220-228.

- Rocha JF, Martinez R, Morris ST and Lopez-Villalobos N (2019). Tick burden in *Bos taurus* cattle and its relationship with heat stress in three agroecological zones in the tropics of Colombia. *Parasite Vector*. 12: 73.
- Salunga THL, Tabuchi Y, Takasaki I, Feril LB JR, et al. (2007). Identification of genes responsive to paeoniflorin, a heat shock protein-inducing compound, in human leukemia U937 cells. *International Journal of Hyperthermia*. *Int. J. Hyperth.* 23: 529-537.
- Sargolzaei M, Chesnais JP and Schenkel FS (2014). A new approach for efficient genotype imputation using information from relatives. *BMC Genomics*. 15: 478 doi: 10.1186/1471-2164-15-478.
- Schmit K and Michiels C (2018). *TMEM* Proteins in Cancer: A Review. *Front Pharmacol*. 9: 1345.
- Sun CH, Southard C, Witonsky DB, Kittler R, et al. (2010). Allele-Specific Down-Regulation of *RPTOR* Expression Induced by Retinoids Contributes to Climate Adaptations. *PLoS one Genetics*. 6: e1001178.
- Tenghe AMM, Bouwman AC, Berglund B, Strandberg DE, et al. (2016). Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. *J. Dairy Sci.* 99: 5470-5485 <http://dx.doi.org/10.3168/jds.2015-10533>.
- Thompson R and Mäntysaari E (1999). Prospects for Statistical Methods in Dairy Cattle Breeding. *Interbull Bulletin*. 20: 71-78.
- UniProt (2019). UniProt. <http://www.uniprot.org/>. Accessed 17 January 2019.
- Valtorta SE, Scarpati E, Leva R and Gallardo MR (2000). Summer Environmental effects on milk production and composition in an argentine grazing System. *Biometeorology and urban climatology at the turn of the millennium*. In: Dear RJ de, Kalma JD, Oke TR, Auliciems A (eds) WCASP-50, WMO/TD 1026. *Biometeorology and urban climatology at the turn of the millennium. Selected papers from the Conference ICB-ICUC'99*. World Meteorological Organization, Geneva, pp. 347-352
- VanRaden PM (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91: 4414-4423.
- Wang H, Misztal I, Aguilar I, Legarra A, et al. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res. (Camb)* 94: 73-83. doi: 10.1017/S0016672312000274.
- Wang J, Zhang F, Yang H, Wu H, et al. (2018). Effect of TEAD4 on multilineage differentiation of muscle-derived stem cells. *Am. J. Transl. Res.* 10: 3 998-1011.
- Won KH, Ozdemir C, Kawano Y, Leclair KB, et al. (2013). Thioesterase Superfamily Member 2/Acy1-CoA Thioesterase 13 (Them2/Acot13) Regulates Adaptive Thermogenesis in Mice. *J. Biol. Chem.* 288: 33376-33386.
- Yodklaew P, Skorn Koonawootrittriron S, Elzo MA, Suwanasopee T, et al. (2017). Genome-wide association study for lactation characteristics, milk yield and age at first calving in a Thai multibreed dairy cattle population. *ANRES*. 51: 223-230. doi 0.1016/j.anres.2017.04.002.