Genome wide association study (GWAS) for body weight traits in Santa Inês sheep breed

S.T. Amorim¹, J.P. Eler³, E. Mattos³, L. Grigoleto³, M.V.A Lemos², F. Baldi² & J.B. Ferraz³

¹Universidade Federal de Santa Catarina, Departamento de Zootecnia e Desenvolvimento Rural, Rodovia Admar Gonzaga, 1346, 88034-001, Florianópolis, Brazil
Sabrina0207amorim@gmail.com (Corresponding Author)
²Faculdade de Ciências Agrárias e Veterinárias, UNESP, Jaboticabal, 14884-000, Brazil
³Research Center on Animal Breeding, Biotechnology and Transgenesis, College of Animal Science and Food Engineering, University of Sao Paulo, 13635-970, Pirassununga, SP, Brazil

Summary

Sheep farming in tropical semi-arid regions like Northeast of Brazil, is characterized by hair breeds like Santa Inês. Santa Inês is a locally adapted sheep breed and have distinctive characteristics that could be of value for specialized production systems. Therefore, the aim of this study was to estimate variance components and identify genomic regions and pathways associated with adult weight (AW), metabolic adult weight (MAW), and corporal condition (CC) in animals raised in Southeast and Northeast of Brazil. In our study, heritability values were moderate for AW (0.32) and MAW(0.33) and low for CC (0.04). Our study also showed candidates genes related to growth regulation, skeletal muscle tissue development, growth factor activity, regulation of muscle contraction, muscle cell differentiation, energy production and conversion, lipid binding, lipid catabolic process, and fat cell differentiation. These genes may help in the selection of animals with a higher growth rate and a better body condition.

Keywords: GWAS, Ovis aries, tropical climate, genomic selection, Santa Inês

Introduction

Brazilian sheep are descendants of animals imported during the colonization period, which led to the development of local breeds adapted to Brazilian environmental conditions (Mariante et al., 2009). Santa Inês hair sheep breed is originally from the north-eastern region of Brazil, which is a tropical region known by water and food scarcity throughout the year. Santa Inês animals represents an unique genetic resource to study biological mechanisms of adaptation in extreme environments. Traits like adult weight are directly related to nutritional requirements. Several studies of GWAS identified genetic variants for productive traits in...
different sheep breeds (Zhang et al., 2013; Al-Mamum et al., 2015), however, these studies were performed in wooly breeds. Berton et al. (2017), identified genomic regions and pathways associated with gastrointestinal parasites resistance in Santa Inês breed. The aim of this study was to estimate variance components and identify genomic regions and pathways associated with adult weight (AW), metabolic adult weight (MAW), and corporal condition (CC) within an adapted Santa Inês population.

Material and methods

The phenotypic records were collected from 1,333 females of Santa Inês breed belonging to four flocks located in southeast and northeast states of Brazil. A total of 576 animals were genotyped with the Ovine SNP12k BeadChip (Illumina, Inc.), which contained 12,785 biallelic SNP markers. Relationship matrix was composed by 32,292 animals. Adult weight (AW) admitted as the weight at 4 years old, and metabolic adult weight (MAW) as MAW= AW$^{0.75}$. To evaluate CC, Suiter (1994) methodology was used.

The variance components were estimated using a single animal trait model by single step genomic BLUP (ssGBLUP) procedure. The ssGBLUP methodology allows the combination of genomic relationship matrix ($G$) and relationship matrix ($A$) in genetic evaluations, through $H$ matrix (Aguilar et al., 2010). The linear model was used to analyse AW and MAW:

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

where $y$ is the observations vector; $\beta$ is the vector of fixed effects (contemporary group and covariate); $a$ is the additive direct vector; $X$ is the incidence matrix; $Z$ is the incidence genetic random effects additive direct matrix; $e$ is the residual effect vector. The quantitative genetic analyses were performed using the AIREMLF90 programs (Misztal et al., 2002; Aguilar et al., 2010). The GWAS analyses were performed using the single-step GWAS methodology (Wang et al., 2012) considering the same linear animal model described before to estimate the variance components afore mentioned. The segments that explained values equal to or greater than 1% of the additive genetic variance were selected to determine the possible QTL regions. The iteration process was used to increase the weights of SNP with large effects and decrease those with small effects. The selected segments were identified and located in the ovine genome by surveying the database available in the "National Center for Biotechnology Information" (NCBI - http://www.ncbi.nlm.nih.gov) and Ensembl Genome Browser. In these databanks, it was possible to identify segments located within or close to genes that could explain the variability in the expression of such traits.

Results and Discussion

For AW, MAW and CC the mean values were 50.30 (±9.76), 19.16 (±2.33) and 2.76 (±0.72), respectively. Heritability estimates showed moderate results for AW (0.32) and MAW (0.33) and low for CC (0.04). These values are similar to the values presented in Safari and Fogarty (2005) review. Corporal condition had low heritability, which can be related to the nutritional state of the dam. The estimated heritabilities in this study indicates that the genetic gain for these traits are possible, and all traits presented genetic variability, so it's
important to investigate the presence of genomic regions or genes affecting AW, MAW and CC.

The SNPs windows regions which accounted for more than 1% of the genetic additive variance were used to search for candidate genes (CG). A total of 23 windows were found for AW, and the window that is responsible for the most part of additive genetic variance was located in chromosome 12, where two candidate genes called CACNA1S and TNNT2 were related to this trait. Our study showed 16 candidates genes related to growth regulation, skeletal muscle tissue development, growth factor activity, and muscle cell differentiation for AW. CACNA1S mutations have been associated with hypokalemic periodic paralysis and malignant hyperthermia (MHS). Several polymorphisms in the CNTFR gene like De Mars et al. (2007), that reported in relation to skeletal muscle strength in humans. TNNT2 is essential for the contraction of striated muscles because it has a central role in the calcium regulation of the actin thin filament function.

For MAW, we found 25 windows that were capable to explain until 14.32% of the additive genetic variance. We found many regions in common with AW, but MAW have shown more candidates genes than AW. MAW have several genes associated with energy production and conversion like LDHA, known as the key regulator of glycolysis. Related to growth and development, MDFIC was also reported with the improvement of piglet birth weight (Zhang et al., 2014). MSTN was described as the gene responsible for double muscling in cattle breeds like Belgian Blue and Piedmontese was also reported in this study (McPherron & Lee, 1997; Kambadur et al., 1997).

For CC, 14 windows were found. In these windows genes linked to lipid binding, lipid catabolic process, fat cell differentiation, generation of precursor metabolites and energy were associated to this trait. Corporal condition is related to prolificity and fertility, females that present highest scores present lowest anestrus postpartum. Related to insulin, PAK1 is responsible for cellular response to insulin stimulus and it was reported in Injuin & Takenawa (2012) research. PAK1 has also been implicated as a positive effector of mechanisms in skeletal myotubes that is crucial to maintaining glucose homeostasis in vivo. LTBP1 is an insulin-like growth factor binding protein, and known to be calcium binding, this interaction is thought to induce structural changes and protect the protein from proteolysis. THRSP is expressed in mammary gland, liver, white and brown adipose tissue of rats and is involved in lipogenic processes. This gene is also similar to the gene product of S14, a rat gene whose expression is limited to liver and adipose tissue and is controlled by nutritional and hormonal factors.

The large number of genes identified associated with AW, MAW and CC in this study should help to understand the genetic mechanism underlying corporal traits in Santa Inês breed. Strategies such as genomic selection can help to select animals adapted to tropical conditions that can improve sheep farming.

**Conclusion**

Several candidate regions related to skeletal muscle, growth factor, energy and response to insulin were found in this study. These candidate regions and candidate genes may help in the selection of animals with better corporal condition and muscle development, leading to reduce the production cycle, therefore it may help to reduce the production cycle and improve the environmental adaptability.
List of References


Berton, MP; de Oliveira Silva, RM; Peripolli, E; Stafuzza, NB; Martin, JF; Álvarez, MS; Gavinã, BV; Toro, MA; Banchero, G; Oliveira, PS; Eler, JP; Baldi, F; Ferraz, JBS. Genomic regions and pathways associated with gastrointestinal parasites resistance in Santa Inês breed adapted to tropical climate. J Anim Sci Biotechnol. 2017 Sep 4;8:73.


Zhang, L; Liu, J; Zhao, F; Ren, H; Xu, L; Lu, J; Zhang, S; Zhang, X; Wei, C; Lu, G; Zheng, Y; Du, L. Genome-wide association studies for growth and meat production traits in sheep. PLoS One. 2013 Jun 25;8(6):e66569.

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