Genome-wide association studies for growth traits in Colombian Creole cattle using a single-step genomic best linear unbiased prediction (gBLUP)

R. Martínez *, D. Bejarano*, Y Gómez*, F. Martínez*, J. Ramírez†, J Gallego†, Y. Abuabara‡.
*Colombian Corporation of Agricultural Research, Biotechnology and Bioindustry Center[1], Bogotá, Colombia,
†El Nus Research Center, San Roque, Antioquia, Colombia, ‡Turipaná Research Center, Cereté, Cordoba, Colombia.

ABSTRACT: We evaluated 3747 animals in a genome-wide association study using the Illumina 50K for growth traits including: birth weight (BWT), weaning weight (WWT) and 16 month weight (16MW) using a single-step gBLUP. We found statistically significant genes associated with BWT, WWT, and 16MW in two regions on the chromosome BTA14 and one region on chromosome BTA6. The genomic regions on chromosome 14 were consistent between traits, and SNP´s were located on coding region of PKIA and TOX gene and near to HAS2 and PLAG1 genes, this last related to body gain weight and growth in *Bos taurus*, indicating their potential use in selection to obtain more efficiency in meat production with this breed.
Key words: genomic; birth weight; weaning
SNP

Introduction

In Colombia the Creole cattle breeds are adapted to tropical areas and are used extensively to produce meat, because they have important adaptation traits as high fertility, low mortality, natural resistance to some diseases and heat tolerance (Martínez et al. (2008)). However, these Creole breeds have not been used for improvement programs to increase the productive traits as growth, feed efficiency and conformation. The growth traits are an economically important topic in beef cattle, birth weight is associated with mature size and carcass weight (Meyer K.(1995)), thus being a valuable selection criterion to improve the meat production system. Today with the abundance of SNPs in the bovine genome they have become effective for genomic prediction and genome-wide association studies (Georges (2007); Goddard and Hayes (2009)).. Currently, genomic evaluations use multiple-step procedures, which are prone to biases and errors. A single-step procedure may be applicable when genomic predictions can be obtained by modifying the numerator relationship matrix A (Aguilar et al. (2009)). Misztal et al. (2009) proposed a single-step BLUP (ssGBLUP) for genomic evaluation where an additive relationship matrix is modified to incorporate the genomic information. The ssGBLUP was successfully applied for scoring over 6 million Holsteins with accuracy superior to a multi-step procedure (Aguilar et al. (2010)). The aim of this paper was to estimate effects of individual SNP on the growth traits like birth weight (BWT), weaning weight (WWT) and 16 month weight (16MW) in Colombian creole cattle breed, Blanco Orejinegro, using a single step GBLUP, in a univariate model.

Materials and Methods: The Animals were raised in the Research Center El Nus, located in northwest region of Antioquia department. The following information was queried from the historical databases: genealogical information, date and year of birth, group, pedigree, age of dam, number of calving, sex, and productive traits like: birth, weaning, and 16 month weight, in a population of 3637 animals with productive information.

DNA and Genotypes: A single 5-mL blood sample from each animal was collected with vacutainer tubes coated with EDTA and shipped to our laboratory in CORPOICA, Colombia. The tubes were homogenized and the DNA extraction was done using a commercial kit (Mobio MR) and the samples were quantified and diluted to 100 ng/ul, to use in genotyping.

The SNP marker data used in this study were obtained from the Bovine SNP50K BeadChip (Illumina, San Diego, CA), using a Hi Scan equipment in the Animal molecular Genetic Laboratory, CORPOICA, Colombia. The procedures used are as described by Matukumalli et al. (2009). Genotype call rates averaged 99.4 ± 0.06% for 54,609 SNP. The genotypes were recorded in the Illumina A/B allele format, after were transformed to compute a value at each locus coded as 0, 1, or 2, representing the number of B alleles, this was done using a program in Fortran language: Illumina2preGS.

The single-trait model was $y = Xb + Zu + Zm + e$, where y is the vector of observations for BW, WW and 16MW; b is the vector of fixed effects including contemporary group (year, month, sex); u is the vector of random additive genetic effects, combining polygenic (breeding values based on pedigree) and genomic (breeding values based on genotypes) breeding values, and m is the vector of maternal genetic effects; X and Z are incidence matrices; e is the vector of random residuals. For genomic analysis were used a ssGBLUP, where the A matrix was replaced by the H matrix (Aguilar et al. (2010))

Results and discussion

Genotype and quality control: From the initial set of 54609 SNPs, a total of 4993 (9.14%) markers were excluded due to minor allele frequency (MAF) < 0.05, and 239 monomorphic SNP were removed. The identity by state (IBS) check revealed 17 no unexpected sample duplicates. The number of SNPs excluded due to SNP

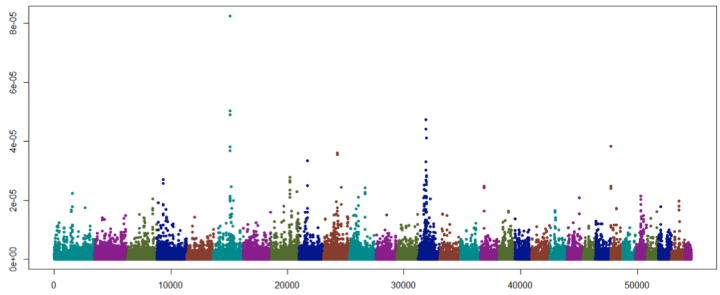

Figure 1. Manhatan plot for GWAS to birth weight in Colombian creole cattle

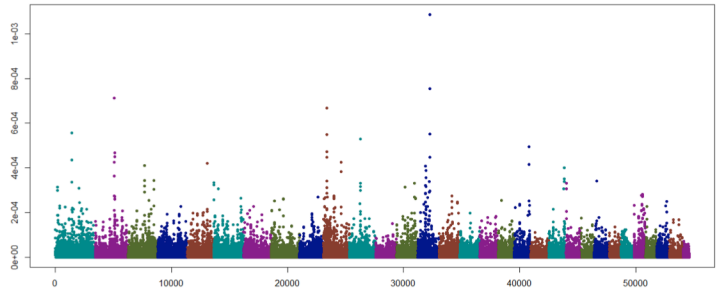

Figure 2. Manhatan plot for GWAS to weaning weight in Colombian creole cattle

Call Rate (CRSNP) < 0.90 was 370 SNPs for a total number of effective SNPs of 49297 and due to Fisher's exact test P-value for Hardy-Weinberg Equilibrium (HWE) < 1 × 10-5 were eliminated 13523 SNPs (24.66%) and 10 individuals were removed due to low Call Rate (CRIND). The final dataset included data for 419 individuals and a total of 35,774 SNPs with a mean allele frequency of 0,54.

Association analysis: were found two regions on BTA14 with SNPs associated to birth weight, this regions are located between 17,6 to 21,3 MB and comprising 10 SNPs which were highly linked (p<1.50e-5) and the second region between 24,7 to 27,7 MB, comprising 20 SNPs which were highly linked (p<2.0e-5) (figure 1 and figure 4). The most significant SNP in the first region (BTB00555019), was located at BTA14:19349948, near to hyaluronan synthase 2 gene (HAS2) (0,4MB), for the second region, the SNP most significant (Hapmap28828-BTC-011250), was located at BTA14: 26713734, within of coding region of thymocyte selection-associated high mobility group box gene (TOX), which was identified as a key transcription factor (TF) responsible for the regulation of puberty in two tropical breeds (Fortes et al. (2011)) and this gene is located near the PLAG1 gene, which have been found associated with important effects on body weight, regulation of stature and weight in *Bos Taurus*. Littlejohn et al. (2011) suggests robust association of PLAG1 genetic variation in determining bovine growth rates and animal size and further implicates PLAG1 as a major regulator of mammalian growth across species and Utsunomiya et al. (2013) also have found Five SNPs on BTA14 associated with BW in Nellore cattle, whose surrounding region has been shown to contain many QTLs and genes as PLAG1 with significant effect on stature-related traits in cattle. This region on BTA14 has also been shown to be associated with reproductive traits. Cole et al. (2011) reported a QTL on BTA14 associated with stillbirth, which also has been associated with body size in dairy cattle.

We also found a significant region on chromosome BTA6, located within coding region of gamma-aminobutyric acid (GABA) A receptor, alpha 4 (GABRA4) gene and less significant SNP´s were found on chromosomes BTA8, BTA9, BTA10 and BTA25. For weaning weight (Figure 2) similarly were found two regions highly significant associated on BTA14: the first located between 26,16 to 27,57 MB, with the same highly linked SNP´s as the one found for birth weight (Hapmap28828-BTC-011250). Additionally, we found other 10 SNPs with p<2.0e-4. A second smaller region was located between 43,64 to 43,92 MB, in this region the most highly linked SNP (ARS-BFGL-BAC-21447) was located at BTA14:43899216, within coding region of the protein kinase (cAMP-dependent, catalytic) inhibitor alpha (PKIA), which may be an important factor in the control of weaning weight. Moreover, BTA14:43899216 was also found highly related to 16 MW (Figure 3). Additionally, we found a region containing the SNP (Hapmap31252-BTC-010762) located within the coding region of carbonic anhydrase VIII (CA8) gene and another SNP on the coding region of the PKIA gene. All these evidences show a consistent relationship of the genetic control between growth traits in Colombian creole cattle. Another chromosome also showing important effects for

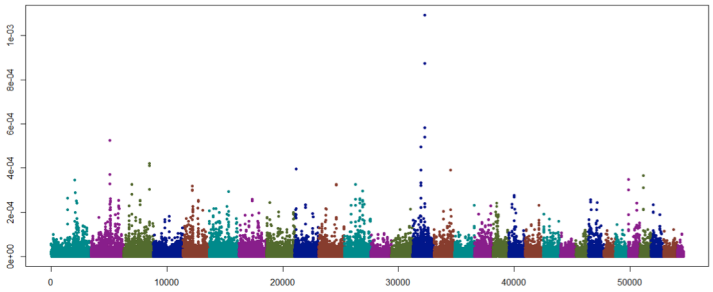

Figure 3. Manhatan plot for GWAS to 16 month weight in Colombian creole cattle

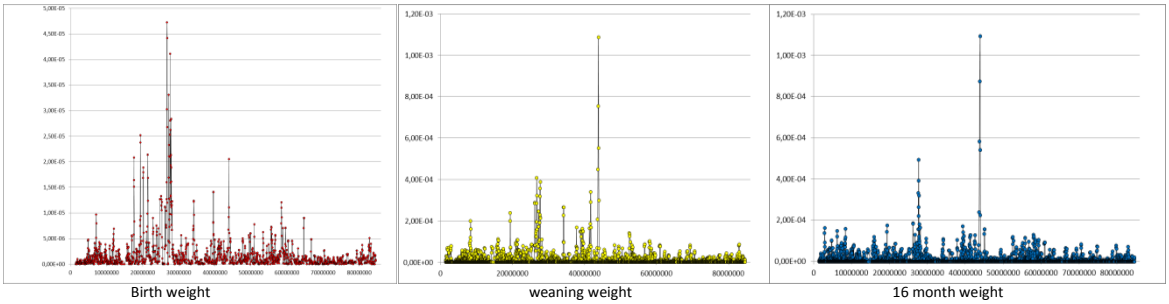

Figure 4. Distribution of SNPs on Chromosome 14 related to growth traits, birth weight, weaning weight and 16 month weight.

two traits: weaning weight, and 16 MW with significant regions on chromosomes BTA1, BTA2 and BTA3, which not were seen in birth weight.

Conclusion

Based on the results we can conclude that similar regions have been found controlling all phases of growth in Bos taurus and desirable genotypes. These SNPs may be useful in breeding decisions on Colombian creole cattle. The most significant SNPs are found located near or within genes like HAS2, PLAG1, TOX GABRA4, PKIA. The genomic regions with highly significant associations displayed in this study are coincident with the same region on chr14, as previously described in several QTL´s for growth and correlated with traits in *Bos Taurus* and *Bos indicus* breeds confirming that the important effect of genes on BTA14 related to growth also may be applied to Colombian creole cattle.

Literature Cited

Aguilar I., A. Legarra, I. Misztal. (2009). *J Dairy Sci*. 92:94648–4655

Aguilar, I., I. Misztal, D. L. Johnson, A. et al. (2010). *J. Dairy Sci.* 93:743–752.

Chen, C. Y., I. Misztal, I. Aguilar, S. et al. (2011*). J. Anim. Sci.* 89:23–28

Cole JB, Wiggans GR,Ma L, Sonstegard TS, et al. (2011). *BMC Genomics*, 12:408.

Fortes M., Reverter A., Nagaraj SH., et al. (2011). *J Anim Sci*. 89(6):1669-83

Georges, M. (2007). *Annu. Rev. Genomics Hum. Genetic*. 8:131–62.

Goddard, M. E., and B. J. Hayes (2009). Nat. Rev. Genet. 10:381–391.

Littlejohn M, Grala T, Sanders K, Walker C, et al. (2012). *Anim Genet* 43(5):591–594.

Martinez R., García D., Gallego J.,et al. (2008*) J. Anim Sci*. 86:545-552

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. et al. (2009). *PLoS One* 4:e5350.

Meyer K. (1995). *Livest Prod Sci* 1995, 44:125–137. Utsunomiya Y., do Carmo A., Carvalheiro R., et. al. (2013). *BMC Genetics* 14:52