

DETECTION OF GENES RELATED TO GROWTH TRAITS IN SHEEP IN COPY NUMBER VARIATION REGIONS

IDENTIFICAÇÃO DE GENES RELACIONADOS COM CARACTERÍSTICAS DE CRESCIMENTO EM OVINOS LOCALIZADOS EM REGIÕES COM VARIAÇÕES NO NÚMERO DE CÓPIAS

Cleyce Maiara Cardoso*¹; Marina Bueno Miotto¹; Nedenia Bonvino Stafuzza¹; Ana Claudia Freitas^{1,2}; Claudia Cristina Paro de Paz¹

¹Institute of Animal Science, Advanced Beef Cattle Research Center, Sertãozinho, SP, Brazil;

²Agricultural Research Agency of the State of Minas Gerais, Pato de Minas, MG, Brazil.

*Corresponding author: cleycecardosozoo@gmail.com

Abstract

In Brazil, the Santa Ines sheep breed has aptitude for meat production, standing out for its rusticity and high productivity. Genomic studies in poorly studied breeds are of great importance, as they make it possible, through genotyping, to find sources of structural variations in the sheep genome. Copy number variations (CNV) are the major source of structural variation among individuals, which are defined as modifications in the genome where the number of copies of a genomic DNA sequence differs from a reference genome. These studies make it possible to advance knowledge of the genomes of poorly studied species such as sheep. Thus, the aim of this study was to identify genes related to growth traits in Santa Ines sheep bred in copy number variation regions (CNVR). A total of 638 animals were genotyped with the Ovine SNP50 Genotyping BeadChip (Illumina). Markers with unknown genomic position and located on sex chromosomes in the reference genome Oar_v 3.1 were excluded. The adjustment of the GC content was performed to reduce the rate of false positive CNVs, considering 500 Kb around each SNP. CNV detection was performed with the PennCNV software, where CNVs identified in only one sample, with fewer than three SNPs, BAF drift>0.01 and GC wave factor>0.09 were excluded. CNVRs were inferred by concatenating individual CNVs identified in more than one animal using the CNVRuler software, excluding genomic regions with an allelic frequency lower than 0.5% and with recurrence <0.1. A total of 742 CNVs (357 deletions and 385 duplications) were detected, which were grouped in 360 CNVRs (224 gain, 132 loss and 4 both) detected on all autosomal chromosomes. The CNVRs inferred in this study had an average size of 142,655 bp (from 12,139 to 2,643,483 bp), covered 2.03% of the autosomal genome sequence, and had frequencies ranging from 0.31% to 4.85% in this population. The OAR1 showed the largest number of CNVRs (40), while the OAR20 presented the smallest number of CNVRs (3). A total of 832 Ensembl ID from protein-coding genes were identified in these CNVRs. Subsequently, functional enrichment analysis was performed with the DAVID software and revealed 44 Gene Ontology terms and 8 enriched KEGG pathways ($P<0.05$). Among the genes identified in the CNVRs, we highlight the following genes related to muscle growth and development: *CCNT2*, *UNC45B*, *DYNC1LI2*, *KLHL40*, *MYH1*, *CORO1B* and *KMT5B*. In addition, we also identified genes related to ossification, such as *NLE1*, *W5P5S7_SHEEP*, *COL1A1*, *TCIRG1*, *GIT1*, *W5P1A2_SHEEP* and *BCL*. The results of the present study provide information to help understand the genetic architecture of growth traits in sheep, their possible impacts on productive traits, and future investigations, being especially useful for researchers interested in structural variations in the sheep genome.

Keywords

CNVR, Santa Ines, Structural variation.