Copy number variation analysis of the East Adriatic sheep breeds

M. Ferenčaković[1], M. Spehar[2], V. Brajkovic[1], I. Drzaic[1], I. Curik[1] and V. Cubric-Curik[1]
[1]Faculty of Agriculture, University of Zagreb, Department of Animal Science, Svetošimunska cesta 25, 10000 Zagreb, Croatia, [2]Ministry of Agriculture, Vukovarska 78, 10000 Zagreb, Croatia; mferencakovic@agr.hr

Copy number variations (CNV) are structural variations in the genome of an individual in the form of losses or gains of DNA fragments greater than 1 kb in size, or in recent times, greater than 50 bp. CNVs are a significant source of genetic and phenotypic variation. More overlapping CNVs detected in two different samples are copy number variation regions (CNVRs). The existence of CNV in the protein coding region alters the protein function, whereas in the regulatory region alters the gene expression level. So far, studies of diversity of CNV distribution in worldwide sheep populations revealed differentiation in CNV between geographical area as well as between diverse groups. Gene ontology (GO) analysis of CNVRs genes gave association with genes related to environmental response and biological functions. Along the East Adriatic, we have analysed 200 individuals belonging to eight sheep breeds (Istria sheep, Krk Island sheep, Cres Island sheep, Rab Island sheep, Lika sheep, Pag Island sheep, Dalmatian Pramenka sheep, Dubrovnik Ruda sheep, 25 individuals from each breed). Animals were genotyped with Ovine Infinium® HD SNP BeadChip (606,006 SNPs). Quality control of SNP data was done using Golden Helix SVS v8.8.3 software with the following parameters: call rate higher than 0.9, a departure from Hardy-Weinberg equilibrium at the 0.001 level, missing genotype rate less than 0.05 and missing data rate less than 0.1. Only markers with GC scores higher than 0.7 were used. Wave detection and correction on autosomes were based on the Oar_v4.0 with a minimum training marker distance of 1 Mb and a recommended wave factor threshold of 0.05. Detected CNVs and CNVRs were used to explore lineage-specific CNVs as well as for GO enrichment analysis. To our knowledge, CNV analysis of any of breeds from Balkan was never performed and our study is a first attempt to perform a comprehensive study of the CNV population genetic properties of the East Adriatic sheep breeds. Our findings contribute to published sheep CNVs and will be helpful for future studies of the CNVRs associated with traits, history, migrations and genetic diversity in sheep.