УДК 575.2

Efficient re-sequencing approach to Toll-like receptor polymorphism screening in cattle using the PacBio platform

Alexander E. Kalashnikov, PhD, Ing, Karel Novák Csc, prom biol
Institute of Animal Science, Prague, Czech Republic
Ernst All-Russian Research Institute of Animal Breeding, Dubrovitsy, Russian Federation

The innate immunity receptors of Toll-like receptor (TLR) family are essential for the formation of immune response to different pathogens of fungal, bacterial and viral origin. The present study involved the gDNA samples of cows and bulls of selected cattle breeds, namely Kholmogory breed, Yakut, Yaroslavl breed, Simmental dairy and meat variants, Holstein, and the cattle x forest bison hybrid population.

The PCR amplicons almost completely covering ORFs of five bovine TLR genes participating in antibacterial immune response were sequenced on the PacBio RSII platform (Pacific Biosciences). The sequencing of the entire population of 270 animals in one run of the instrument provided the coverage up to 76 reads of a particular amplicon per individual. The primary processing of sequencing data was performed with the Ugene software package followed by estimation of quality with FastQS.

Reads of suitable quality have been filtered on the reference genome of Bostau6 (https://genome.ucsc.edu/) via a BWA-MEM module (Li H. and Durbin R., 2009). Subsequently, duplicate PCR was removed by Cigar MDWMC script (Picard software package). Having aligned the combined reads from the population to a reference genome, SNP calling was applied to identify variable sites. SNPs were initially identified by FreeBayes.

The SNapShot genotyping further validated the suspected SNPs and allowed to assign the genotypes to individual animals. Annotation and functional interpretation of variant alleles was performed by VeIP (www.ensembl.org). The used pipeline of targeted resequencing allows to obtain quickly and efficiently information about the total variability in the group of *TLR* genes.

As the result of the screening, more than twenty polymorphic sites in the five *TLR* genes (*TLR1-10)* have been identified within the studied combined population. In process of sequencing has been identified from 8 to 70 mutations in the gene sequences of which can be validated to 70%. The coverage was 5,2-13,4 readings per sample was determined by analyzing data. There were obtained before 4,7.10^6 readings for each gene. The data are expected to be useful in further breeding of cattle in Russia for improved disease resistance, partly to economically important bacterial diseases.

The work was supported by the "Program for the research institution development" of the Ministry of Agriculture of the CR no MZERO0714 and RNF project "The study, conservation and management of biodiversity of animals as a basis for obtaining a healthy, safe and high quality food" no 14-36-00039.