

ABSTRACT:

Milk and dairy products derived from cow's milk are an important nutritional component for humans. Beta-casein is a major component of cow's milk which is coded by various alleles such as A¹ and A² alleles. It has been shown that beta-casein protein in tauros cow's milk is an integrated product of the A¹ and A² alleles. However, in Indicus cows, only A² allele encodes this protein. Consuming type A¹ and A² milk can cause symptoms of various diseases, such as the prevalence of type 1 diabetes, neurological disorders such as autism, schizophrenia and other diseases, including celiac disease, while consumption of type A² milk doesn't cause symptoms of mentioned disease. Therefore, the aim of this study was to investigate the polymorphism pattern, gene expression difference and caseinate gene network of milk in two populations including 1) American Holstein (*Bos taurus*) cows and 2) Pakistani cholistani cows (*Bos indicus*). In this regard, RNA-Seq data of Holstein and cholistani populations were used. Editing and quality control of RNA-Seq data was performed by FASTQC (0.11.1 version) and Trimmomatic (0.36.0 version) software. To construct the transcriptome, alignment and locating RNA-Seq reads on the reference genome, TopHat2 software (2.1.1 version) was used. SNP detection and screening analysis on the transcriptome was performed using the samtools software package. To compare the expression of casein genes in *Bos taurus* (Holstein) and *Bos indicus* (cholistani) cows and investigating the information about chromosome number, mRNA length, number of amino acids, and gene access number, sequences of mRNA was extracted from the NCBI database and compared with data from RNA-Seq data. According to the data from the output file 1) Cuffdiff, *CSN1S1*, *CSN1S2*, *CSN2*, *CSN3*, *CSNK1A1*, *CSNK1B*, *CSNK1C*, *CSNK1A2*, *CSNK1G1*, *CSNK1G2*, *CSNK1G3*, *CSNK1E*, and *CSNK1D* genes related to Casein genes were found on the transcriptome of the US Holstein and Pakistani cholistani population samples, in which all 1) genes identified by differential expression (P.value > 0.0001) did not have different expressions. The String-db site was used to investigate the association and coexistence of casein genes. The generated network showed high correlation between *CSNK1A1* and *CSNK1B* genes with 0.998 combined score and lowest correlation was between *CSNK1B* and *CSNK1G2* with a combined score of 0.429. The results of this study showed that casein genes in Holstein and cholistani cows did not have different differential expression, so the differences may be due to gene network, interaction and genes association based on combined scores, extracted articles, performed experiments, and the effect of genes co-expression on each other. The highest correlation was between *CSNK1A1* and *CSNK1B* genes and the least between *CSNK1B* and *CSNK1G2* genes.

KEYWORDS: Casein Proteins Encoding Genes, Cattle, Gene Expression, Gene network



University of Zabol
Graduate School

Faculty of Sciences
Department of Animal Science

**The Thesis Submitted for the Degree of M.Sc
(in the field of Genetic and Breeding)**

The Single Nucleotide Polymorphisms (SNPs), Allele Specific Expression (ASE), codon usage bias and gene expression profiling on the genes coding of Caseinate proteins in milk for Holstein (*Bos taurus*) and Cholistani (*Bos indicus*) cattle

Supervisors:

Dr. A. Magsudi

Dr. M. H. Banabazi

Advisors

Dr. G. Dashab

By

Masoud Ebrahimi Ghale Jogh

Date

November. ۲۰۱۹