STRUCTURAL AND MOLECULAR-GENETIC CHARACTERISTICS OF THE *CSN2* GENE IN CATTLE BREEDS: A REVIEW AND FUTURE PERSPECTIVES

Javohir D. Ramazonov[1], Zarina R. Askarova[1], Bobur Q. Abdukarimov[1]

[1]Unique Object Laboratory of Phytopathogenic and Other Microorganisms, Institute of Genetics and Plant Experimental Biology, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan

Abstract

This study investigates the structural and molecular-genetic characteristics of the *CSN2* gene, which encodes β-casein-one of the key milk protein components in cattle. Located on chromosome 6 of the bovine genome, the *CSN2* gene is a protein-coding sequence comprising nine exons, according to contemporary genomic annotation data. Nucleotide substitutions within the *CSN2* gene drive the formation of diverse allelic variants that directly alter the amino acid composition of the β-casein protein. Notably, the A1 and A2 β-casein variants are distinguished by a single amino acid substitution at position 67. This genetic variability is of paramount importance for characterizing the genetic profiles of cattle breeds, evaluating milk protein composition, and identifying promising molecular markers for marker-assisted selection. Elucidating the structure and polymorphisms of the *CSN2* gene represents a critical research direction, particularly for molecular-level characterization and conservation of the gene pool in local and genetically valuable cattle breeds.

Keywords: cattle, *CSN2*, β-casein, gene structure, genetic polymorphism, A1, A2, molecular genetics, milk protein, genotyping.

Introduction.

In dairy cattle breeding, the genetic evaluation of milk protein composition and

technological properties, alongside milk yield, represents a crucial objective of modern selection programs. Caseins constitute the principal fraction of milk proteins, encompassing αS1-casein, αS2-casein, β-casein, and κ-casein. These proteins are encoded by the *CSN1S1*, *CSN1S2*, *CSN2*, and *CSN3* genes, respectively. Genetic variants of caseins are recognized as significant sources of variation in milk protein composition and processing traits. The *CSN2* gene encodes the β-casein polypeptide chain. According to NCBI database records for *Bos taurus*, the *CSN2* gene is localized on chromosome 6, functions as a protein-coding sequence, and is annotated with nine exons within the current ARS-UCD2.0 genome assembly. Its reference genome coordinates span from position 85,449,178 to 85,457,699 on the NC_037333.1 sequence in the complementary orientation. Investigating the molecular architecture and genetic variants of the *CSN2* gene enables the elucidation of inter-breed genetic differentiation and the evaluation of milk protein-associated traits using molecular markers.

Structural Characteristics of the *CSN2* Gene

Within the bovine genome, the *CSN2* gene acts as a protein-coding sequence responsible for β-casein synthesis in mammary gland tissues. The genomic organization of the gene comprises alternating exon and intron regions, whose transcription yields messenger RNA (mRNA) encoding the precursor form of the β-casein protein. According to NCBI data, the current genome annotation of the *Bos taurus CSN2* gene contains nine exons. As one of the principal caseins in milk, β-casein plays an essential role in the formation of casein micelles. Genetic variants of milk proteins can influence both molecular characteristics and the physicochemical and technological properties of milk. Consequently, deciphering the structural organization and allelic variants of the *CSN2* gene is of paramount importance for livestock genetics.

Molecular-Genetic Variability and Inter-Breed Distribution

A prominent characteristic of the *CSN2* gene is its high degree of polymorphism. The distinct genetic variants of β-casein are primarily associated with nucleotide substitutions within the *CSN2* sequence. Among these, the A1 and A2 variants have been extensively investigated, differing by a single amino acid substitution at position 67 of the β-casein polypeptide chain: histidine is present in the A1 variant, whereas proline occupies this position in the A2 variant.

The distribution of *CSN2* polymorphisms varies significantly across cattle breeds. Breed origin, historical selection pressures, and population genetic structure act as primary determinants shaping allele frequencies. Analyzing polymorphisms in milk protein genes facilitates the characterization of genetic structures and inter-breed differentiation within bovine populations. Furthermore, previous studies have demonstrated significant associations between *CSN2* genotypes and the relative composition of milk proteins. Research on Dutch Holstein-Friesian cattle revealed that the β-casein genotype is significantly correlated with the relative concentrations of β-casein and other specific casein fractions, as well as overall protein yield. These findings highlight the potential of utilizing the *CSN2* gene as a promising molecular marker for assessing milk protein composition from a molecular-genetic perspective.

Methodological Framework for *CSN2* Analysis *(Added Scientific Context)*

Molecular-genetic analysis of the *CSN2* gene is typically conducted using methods such as genomic DNA extraction, PCR amplification, PCR-RFLP, allele-specific PCR, SNP genotyping, or direct sequencing. These methodological approaches enable the targeted amplification of specific gene regions to identify nucleotide substitutions and allelic variants. The identification and polymorphism analysis of the *CSN2* gene are equally critical for indigenous and regionally adapted cattle breeds. Specifically, performing molecular-genetic identification of the *CSN2* gene in Bushuyev cattle establishes foundational molecular resources for investigating the genetic diversity linked to milk proteins in this indigenous breed. Future

methodological perspectives involve sequencing the identified *CSN2* fragments to align nucleotide sequences against the reference genome, detecting novel single nucleotide polymorphisms (SNPs), and calculating corresponding allele and genotype frequencies to evaluate robust genotype-phenotype associations with milk quality metrics.

Conclusion.

The *CSN2* gene is a vital protein-coding sequence responsible for β-casein synthesis in cattle, mapping to chromosome 6 of the *Bos taurus* genome and comprising nine exons according to contemporary genomic annotations. Molecular variability within this gene drives the formation of distinct β-casein genetic variants, notably the A1 and A2 variants, which are differentiated by a single amino acid substitution at position 67 of the polypeptide chain. Elucidating the structural and polymorphic properties of the *CSN2* gene establishes a solid scientific foundation for molecular-level characterization of bovine genetic resources, identifying genetic variations in milk protein composition, and developing future marker-assisted selection strategies. Consequently, continuing the investigation of the *CSN2* gene in genetically valuable cattle breeds-such as the Bushuyev breed-specifically through identifying its polymorphisms and evaluating their associations with productivity traits, represents a highly promising research avenue.

References.

1. Farrell, H. M., Jr., Jiménez-Flores, R., Bleck, G. T., et al. (2004). Nomenclature of the proteins of cows' milk-sixth revision. *Journal of Dairy Science*, 87(6), 1641-1674. https://doi.org/10.3168/jds.S0022-0302(04)73319-6

2. Caroli, A. M., Chessa, S., & Erhardt, G. J. (2009). Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of Dairy Science*, 92(11), 5335-5352. https://doi.org/10.3168/jds.2009-2461

3. Heck, J. M. L., Schennink, A., van Valenberg, H. J. F., et al. (2009). Effects of milk protein variants on the protein composition of bovine milk. *Journal of Dairy Science*, 92(3), 1192-1202. https://doi.org/10.3168/jds.2008-1208

4. Jiang, Y., et al. (2024). Recent advances in the bovine β-casein gene mutants on functional characteristics and nutritional health of dairy products: Status, challenges, and prospects.

5. National Center for Biotechnology Information (NCBI). Gene ID: 281099. *CSN2* casein beta [*Bos taurus*]. Reference genome: ARS-UCD2.0, Chromosome