Egyptian Journal of Veterinary Sciences

https://ejvs.journals.ekb.eg/

Molecular Characterization of *CSN1S2* in Egyptian River Buffalo

Alhassan T. Shahwan[1], Nadia A. Aboul-Maaty[1], Amal A. Hassan*[1], Eman M. Saleh[2], Hanaa A. Oraby[1] and Usama A. Ibrahim[2]

[1] *Cell Biology Department, Biotechnology Research Institute, National Research Centre (NRC), Cairo, Egypt.*

[2] *Biochemistry Department, Faculty of Science, Ain Shams University.*

Abstract

IN the process of elucidating the molecular characteristics of the *CSN1S2* gene in Egyptian buffalo, various haplotypes for the complete coding sequence (CDS) of *CSN1S2* in both river and swamp buffaloes were retrieved from GenBank. Two non-synonymous single-nucleotide polymorphisms (SNPs): c.484T>A (g.12803T>A) in exon 14 and c.568G>A (g.14067G>A) in exon 16, were the most frequent among the investigated haplotypes and were found exclusively in river buffalo. Therefore, DNA samples collected from thirty-eight Egyptian river buffalo were genotyped for these SNPs. An additional SNP in the splice donor site of *CSN1S2* intron 7 (g.7197 G>C), known to cause exon 7 skipping, was also genotyped. The prevalent genotypes at g.7197 and g.12803 were GG and AA, respectively, with frequencies of 0.84. The genotypes CC and TT were not detected at these two sites (g.7197 and g.12803), respectively. The genotype frequencies at g.14067 were 0.21 for GG, 0.11 for GA, and 0.68 for AA. Results analysis revealed that Egyptian buffalo possesses a high frequency of the genotype GG at g.7197 in *CSN1S2*, suggesting a reduced likelihood of exon 7 skipping, while the prevalent AA genotype at g.14067 is more likely to be linked to increased palmitic acid in the milk, suggesting a possible health risk. To mitigate this, breeding programs utilizing marker-assisted selection should focus on enhancing the frequency of the *CSN1S2* GG genotype at g.14067. Additionally, this study examined the genetic relationships among the *CSN1S2* coding sequence haplotypes in buffalo and constructed a phylogenetic tree that differentiates between buffalo and cattle samples.

Keywords: αS2-casein, buffalo, DNA sequencing, genotyping, polymorphism.

Introduction

Milk proteins are essential for a healthy diet and offer numerous health benefits (1). Genetic variants of milk proteins often influence technological and nutritional characteristics of milk (2). The primary protein in milk is casein (CN), derived from the Latin word *Caseus*, meaning "cheese." It comprises four fractions: αS1-, β-, αS2-, and κ-CN, which together account for approximately 80% of all milk proteins (3-4). In ruminants, these four casein fractions are secreted as stable calcium–phosphate micelles. Caseins provide essential amino acids, calcium, and phosphate to suckling infants and are linked to lactation traits in dairy animals as well as the processing characteristics of milk (5-6).

The domestic water buffalo (*Bubalus bubalis*) is a vital farm animal in tropical and subtropical regions, with a global population of approximately 208 million (7-8-9). The species originated in Southeast Asia, where 97% of the world's buffalo population is still reared (10). Based on morphological and behavioral characteristics, domestic buffalo can be classified into two categories: swamp buffalo and river buffalo. The swamp type (2n = 48) is found exclusively on the native Asian continent, while the river type (2n = 50) is more widely distributed across other continents (11-12). Swamp buffalo are primarily utilized for draft purposes, while river buffalo are mainly raised for milk production (13). The water Buffalo species contributes approximately 13% of the total milk production worldwide (14).

*Corresponding authors: Amal A. Hassan, E-mail: amal.a.hassan@hotmail.com, Tel.: 01273628999
(Received 24 July 2025, accepted 22 September 2025)
DOI: 10.21608/EJVS.2025.407115.2991

Buffalo milk contains all four casein fractions αS1, β, αS2, and κ encoded by four closely linked autosomal genes: *CSN1S1*, *CSN2*, *CSN1S2*, *and CSN3*, respectively. The four genes were mapped to chromosome 7 in river buffalo (15).

The buffalo *CSN1S2* gene encodes αS2-casein (αS2-CN), the third most abundant casein fraction and the most hydrophilic of all caseins. αS2-CN accounts for approximately 14% of the total casein in buffalo milk (16). The proportion of αS2-CN in milk varies significantly between species and is absent in human and marsupial milk (17). αS2-CN, along with αS1-CN and β-CN, plays a vital role as a calcium-sensitive protein that can form stable micelles with calcium and phosphorus, which are crucial for ensuring optimal bone growth in young animals (18-19). Additionally, αS2-CN is associated with the nutritive value of milk (5).

The *CSN1S2* gene of the Mediterranean river buffalo (GenBank accession no. MW159135.1) contains 18 exons ranging in size from 21 (exon 4) to 267 bp (exon 18). Exons 1 and 18 are not coding. The whole highly conserved signal peptide is encoded by the nucleotides 13–57 of exon 2, exons 2-16 are located in the coding sequence (CDS) region, and the stop codon TAA is created by nucleotides 10–12 of exon 17. The deduced CDS length of the buffalo *CSN1S2* gene is 669 bp long and encodes a protein consisting of 222 amino acids (aas), of which the first 15 aas form an N-terminal signal peptide, and the remaining 207 aas represent the mature peptide (20).

Research on the polymorphism of the *CSN1S2* gene in buffalo is limited. A single-nucleotide polymorphism (SNP) in the splice donor site of intron 7 in the Mediterranean river buffalo *CSN1S2* leads to the skipping of exon 7, resulting in a deleted transcript (21). It is known that the splice donor site is a short nucleotide sequence (GT in DNA). It is typically found at the 5' end of the intron, indicating the start of the intron to be spliced out (22). Nucleotide polymorphisms in the coding sequence of *CSN1S2* have been detected and analyzed in Indian riverine and Chinese swamp buffaloes using direct sequencing of polymerase chain reaction (PCR) products (3). Furthermore, Cosenza and colleagues (20) explored the potential association of the SNP in the splice donor site of intron 7 and a non-synonymous SNP (g.14067G>A) in exon 16 with traits that could influence the nutritional and technological quality of buffalo milk.

The river buffalo is the most important domestic animal in Egypt. It serves as the primary source of milk production. The *CSN1S2* gene is significant in milk production but has not yet been fully characterized in the Egyptian buffalo. In the process of elucidating the molecular characteristics of the *CSN1S2* gene in Egyptian buffalo, various haplotypes for the complete coding region of *CSN1S2* in river and swamp buffalo were retrieved from GenBank databases at the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/nuccore/?term=buffalo+CSN1S2+CDS). These haplotypes were compared to identify the locations of SNPs within the coding sequences. Particular attention was given to the most frequent non-synonymous SNPs found among the different *CSN1S2* CDS haplotypes (c.484T>A in exon 14 and c.568G>A in exon 16) and the SNP in the splice donor site of intron 7(g.7197G>C). These selected SNPs were genotyped in DNA samples from the Egyptian buffalo. The genetic relationship among the *CSN1S2* CDS haplotypes in water buffalo was also investigated. Furthermore, a phylogenetic relationship between the two closely related animals, water buffalo and cattle (*Bos taurus*), both belonging to the subfamily *Bovinae*, has been constructed based on *CSN1S2* CDS haplotypes. The findings of this study could be beneficial to both animal scientists and livestock producers for efficient management and farming strategies aimed at improving the nutritional quality of Egyptian buffalo milk.

Material and Methods

Sequences retrieved from GenBank

The published CDS haplotypes of the *CSN1S2* gene in river buffalo, swamp buffalo, and cattle were retrieved from the GenBank databases at the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/nuccore/?term=buffalo+CSN1S2+CDS and https://www.ncbi.nlm.nih.gov/nuccore/?term=cattle+CSN1S2) and analyzed in this study.

Samples collection

The protocol applied during this study was reviewed and approved by the Medical Research Ethics Committee at the National Research Centre (reference number 01019102021). The permission for the blood samples collection was appropriately obtained from the director of the Animal Production Research Institute. As reported in the breeding farm records, the buffaloes enrolled for sampling were all healthy adults without direct kinship. They were derived from Mehalet Mousa Station for Animal Production, Kafr El-Sheikh Governorate, Egypt. Blood samples were collected from 38 lactating Egyptian buffaloes for DNA extraction.

DNA extraction, primer design, polymerase chain reaction (PCR), and sequencing

Genomic DNA was isolated from blood samples using the phenol-chloroform protocol described by Sambrook and Russell (2001) (23). The DNA quality was assessed using a 1.5% agarose gel and quantified using a NanoDrop LITE spectrophotometer (Thermo Fisher Scientific, USA). Subsequently, for the polymerase chain reactions, three primer pairs were

designed in the intronic regions flanking exons 7, 14, and 16 (Table 1) using the whole *CSN1S2* gene reference sequence in the Mediterranean river buffalo (accession no. MW159135.1) and Primer3 software (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) (24).

Each PCR (50 µl) contained 200 ng DNA template, 50X DreamTaq™ Green buffer, 50 µM of each primer, 25 mM of dNTPs, and 1 µl (1 unit) of DreamTaq™ Green DNA polymerase. The reaction mixture was run in Q-Cycler, HVD LifeSciences. The Thermocycling conditions consisted of an initial denaturation at 95°C for 4 minutes, followed by 35 cycles consisting of a denaturation step at 94°C for 45 seconds, a 45-second annealing step at the appropriate annealing temperature for each primer pair (**Table 1**), and an extension step at 72°C for 2 minutes. A final extension of 10 minutes at 72°C was accomplished to end the reaction. Successful PCR products were purified using the MEGAquick-spinTM Total Fragment DNA Purification Kit (iNtRON Biotechnology) following the manufacturer's recommended protocol. Purified products were further sequenced bidirectionally using the Sanger method.

Sequence data analysis

The *CSN1S2* CDS haplotypes retrieved from GenBank databases and *CSN1S2* DNA sequences from the present study were analyzed, aligned, and compared for nucleotide polymorphism detection using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) (25) and AliView (http://www.ormbunkar.se/aliview/downloads/) (26). The Coding DNA sequences were translated into proteins using the Expasy translation tool (https://web.expasy.org/translate/). Translated proteins were aligned and compared using the two previously mentioned programs (Clustal Omega and AliView). The allelic and genotypic frequencies of the SNPs determined in the Egyptian buffalo *CSN1S2* were calculated using PopGen32 (https://sites.ualberta.ca/~fyeh/popgene_download.html) (27). The potential impact of aa substitutions, due to the detected non-synonymous nucleotide variations, on protein function was predicted using PROVEAN (http://provean.jcvi.org/index.php) (28). PROVEAN considers substitution tolerated when the score is above -2.5 and intolerant if the score is below -2.5.

To investigate the probable genetic relationship among the haplotypes of *CSN1S2* CDS in river and swamp buffaloes, a median-joining network was constructed by the PopART program (https://popart.maths.otago.ac.nz/download/) (29). Furthermore, a maximum-likelihood (ML) phylogenetic tree based on the haplotypes of the *CSN1S2* CDS in buffalo and cattle was constructed using the function "build" of ETE3 v3.1.1(30) as implemented on the GenomeNet (https://www.genome.jp/tools/ete/). The tree was constructed using FastTree with slow NNI and MLACC=3 (to make the maximum-likelihood NNIs more exhaustive) (31). The *CSN1S2* CDS in the goat (*Capra hircus*) and sheep (*Ovis aries*) was used for rooting the tree.

Results

Polymorphism detection in the CSN1S2 coding sequence

Alignment of all the published sequences for the *CSN1S2* complete coding region in the river and swamp buffaloes from various countries (accession numbers MT316464.1, MT316465.1, MT316466.1, KY399458.2, FM865618.1, DW007983.1, DQ133467.1, MT316468.1, DQ173244.1, KX896650.1, FM865619.1, and NM_001290865.1 as a reference sequence) identified nine CDS haplotypes. These nine haplotypes resulted from nine SNPs: c.15T>C in exon 2, c.234A>C in exon 9, c.381T>A in exon 11, c.382A>G in exon 11, c.391G>A in exon11, c.459C>T in exon 13, c.484T>A in exon 14, c.568G>A in exon 16, c.642C>G in exon 16 and a deletion of exon 7 (27 nucleotides deletion from c.170 to c.196) (Fig. 1). An overview of the nine detected haplotypes is provided in Table 2. The river and swamp buffalo types shared haplotypes 1 and 3. Haplotypes 2, 4, 6, 7, and 9 were identified in river buffalo, while haplotypes 5 and 8 were detected in swamp buffalo. The determined length of *CSN1S2* CDS for the river and swamp buffalo was 669 nucleotides. The deletion of 27 nucleotides from the gene transcript was noticed in the swamp and river buffalo (haplotype 8 and haplotype 9, respectively).

Translation of the investigated CDS haplotypes to proteins indicated that the nucleotide polymorphisms c.234A>C, c.382A>G, c.391G>A, c.484T>A, c.568G>A, and c.642C>G were non-synonymous, leading to the amino acid substitutions p.Glu63Asp, p.Lys113Glu, p.Ala116Thr, p.Phe147Ile, p.Ala175Thr, and p.Asn199Lys in the mature peptide, respectively. The most frequent non-synonymous SNPs among the investigated CDS haplotypes were c.484T>A in exon 14 and c.568G>A in exon 16 (Fig. 1). They were only recognized in river buffalo haplotypes. Further on the DNA level, the two SNPs c.484T>A (MW159135.1: g.12803T>A) and c.568G>A (MW159135.1: g.14067G>A), along with the SNP in the splice donor site of intron 7 (MW159135.1: g.7197G>C), which is responsible for the deletion of 27 nucleotides (exon 7) from the *CSN1S2* transcript, were genotyped in the Egyptian buffalo.

Genotyping of the selected CSN1S2 SNPs in Egyptian buffalo DNA

The three designed primer pairs in the intronic regions flanking exons 7, 14, and 16 in *CSN1S2* successfully amplified DNA segments of the expected sizes: 541 bp, 213 bp, and 596 bp, respectively. Sequences of the three amplified segments in the investigated buffalo samples were determined and analyzed. The sequences of the three DNA segments are indicated in Fig. 2 A, B, and C, respectively. The positions of the three investigated SNPs in the amplified segments and their corresponding locations in the complete *CSN1S2* gene reference sequence MW159135.1 are illustrated in Table 3. The table also presents the genotypic and allelic frequencies of the three SNPs, aa substitutions or deletions resulting from the three SNPs, and the tolerance score for the aa substitutions.

In the present work, the nucleotide positions of the three investigated SNPs in the Egyptian buffalo samples are numbered according to their position in the *CSN1S2* whole gene sequence (accession number: MW159135.1). As indicated in Table 3, homozygous genotypes CC and TT at the nucleotides g.7197 and g.12803, respectively, were not detected in the Egyptian buffalo. The highest allele frequencies of 0.92, 0.92, and 0.74 were detected for g.7197G, g.12803A, and g.14067A, respectively. PROVEAN scores for the aa substitutions p.Phe147Ile and p.Ala175Thr were -2.350 and 1.936, indicating that the two aa substitutions are predicted to be tolerated.

The genetic relationship between the haplotypes of the CSN1S2 coding sequence

The possible genetic relationships among seven of the nine buffalo *CSN1S2* CDS haplotypes identified in the present study were investigated by constructing a median-joining network (Fig. 3). Although haplotypes 8 and 9 had the highest percentages of mutation among the investigated haplotypes (deletion of exon 7), we excluded these two haplotypes during the network construction because the PopART program does not consider the gaps in the DNA sequence during network construction.

The constructed network indicates that the *CSN1S2* CDS haplotypes 2, 3, 4, 5, 6, and 7 may originate from haplotype 1 (the reference sequence for the *CSN1S2* coding region) as follows: Haplotype 2 evolved from haplotype 1 through a transition c.568G>A, and haplotype 4 evolved from haplotype 2 through a transversion c.484T>A. Haplotype 6 and haplotype 7 evolved from haplotype 4 through a transversion c.381T>A and a transition c.382A>G in haplotype 6, and a transition c.15T>C and a transversion c.642C>G in haplotype 7. Haplotype 3 evolved from haplotype 1 through the transition c.391G>A, while haplotype 5 evolved from haplotype 3 through a transversion c.234A>C.

Phylogenetic relationship between water buffalo and cattle based on CSN1S2 CDS haplotypes

The phylogenetic tree (Fig. 4) based on the *CSN1S2* CDS haplotypes of buffalo, investigated in the present study, and cattle (*Bos taurus*) was rooted using the corresponding haplotype sequences of the two species, belonging to the subfamily *Caprinae*: sheep (*Ovis aries*), and goat (*Capra hircus*). The phylogenetic tree was clustered into three principal clades. The first clade included *Bos taurus* samples. The second clade consisted of the river and swamp buffalo. The third included only *Ovis aries and Capra hircus*.

Discussion

The latest advances in molecular knowledge and techniques have revealed a multitude of genetic polymorphisms at the DNA level, which are associated with animal performance. Single-nucleotide polymorphisms (SNPs) have become a popular marker of phenotypic variation among individuals in a population (32).

Milk proteins are essential for a healthy diet and provide numerous health benefits (33). The four caseins αS1-CN, β-CN, αS2-CN, and κ-CN play an essential role in improving the quality and quantity of milk-derived products. Genetic variants of milk proteins can significantly influence technological and nutritional characteristics of milk (2). Therefore, over the past few decades, numerous studies have linked genetic polymorphisms and protein structure to economically important milk traits in key ruminant species, including cows, sheep, and goats (34-35-36). Water buffalo has received less attention, as casein genes and their protein structure have not been thoroughly explored in this economically significant species.

The αS2-casein is the third most abundant casein fraction in buffalo milk. It was reported to be associated with the nutritive value of milk (5-20). In the present study, we investigated the different haplotypes of the complete coding region of the *CSN1S2* gene in river and swamp buffaloes from various countries, including Egypt, India, Italy for river buffalo, and China for swamp buffalo. Our goal was to identify the locations of the SNPs within the coding region.

We identified nine haplotypes for *CSN1S2* CDS which resulted from nine SNPs: c.15T>C in exon 2, c.234A>C in exon 9, c.381T>A in exon 11, c.382A>G in exon 11, c.391G>A in exon11, c.459C>T in exon 13, c.484T>A in exon 14, c.568G>A in exon 16, c.642C>G in exon 16 and a deletion of exon 7. The exon deletion was observed in both buffalo types. Most of the detected SNPs are haplotype-specific, except for the two non-synonymous SNPs c.484T>A in exon 14 and c.568G>A in exon 16, which were the most common

among the investigated river buffalo *CSN1S2* CDS haplotypes. These two SNPs, c.484T>A (g.12803T>A) and c.568G>A (g.14067G>A), were genotyped in DNA samples from Egyptian buffalo, along with SNP g.7197G>C, located in the splice donor site of *CSN1S2* intron 7.

Results indicated that the homozygous genotype CC at nucleotide position g.7197 was not detected in the Egyptian buffalo samples. The prevalent genotype at this position was GG. Previous studies (20,21) indicated that the transversion G>C at the splice donor site of intron 7 in buffalo *CSN1S2* gene led to the deletion of exon 7 (27 nucleotides) from the *CSN1S2* transcript and 7 amino acids from the translated protein. Therefore, increasing the frequency of the *CSN1S2* GG genotype at this position in river buffalo is desirable for selection in breeding programs (20). In this context, the high frequency of the GG genotype detected in the present study at the 5' end of the intron 7 (nucleotide position g.7197) in Egyptian buffalo *CSN1S2* strongly prevents exon 7 skipping.

Genotype analysis revealed that the homozygous genotype TT at nucleotide position g.12803 in *CSN1S2* was not detected in the Egyptian buffalo samples. The predominant genotype at this position was AA, followed by the TA genotype. The amino acid substitution p.Phe147Ile, resulting from g.12803T>A, was predicted to be tolerated. To the best of our knowledge, no previous studies have associated this SNP with the nutritional and technological quality of buffalo milk.

The three genotypes GG, GA, and AA, with frequencies of 0.21, 0.11, and 0.68, respectively, were determined for the SNP g.14067G>A at *CSN1S2* of Egyptian buffalo. The g.14067A and g.14067G allele frequencies detected were 0.74 and 0.26, respectively. The allele frequencies for the same SNP were determined in Indian and Italian river buffalo, where the highest frequencies were 0.59 and 0.51 for the G allele, respectively (3-20). Although the aa substitution p.Ala175Thr due to the SNP g.14067G>A was predicted to be tolerated in the present study, Cosenza and colleagues (20) reported that SNP g.14067A>G had a significant association with the palmitic acid content in buffalo milk. Their results indicated that homozygous GG and heterozygous GA buffaloes had lower amounts of palmitic acid, 34.13% and 34.71%, respectively, compared to the AA genotype (35.23%).

Palmitic acid (C16:0), a long-chain fatty acid, is found in all living organisms, from bacteria to humans. It naturally occurs in palm oil, meat, milk, butter, and cheese, serving as a key component of cell membranes and playing important roles in protein palmitoylation and signaling molecules (37-38). In buffalo raw milk, palmitic acid constitutes about 34.8% of total saturated fatty acids, the highest percentage among ruminants (39). High concentrations of palmitic acid in buffalo dairy products are notable (40). However, palmitic acid is associated with obesity and reduced insulin sensitivity, increasing the risk of type 2 diabetes and cardiovascular diseases due to higher cholesterol levels (41-42). Thus, diets rich in palmitic acid can adversely affect human health.

Our findings show that the frequency of the GG genotype at nucleotide position g.14067 in the *CSN1S2* gene of Egyptian river buffalo is relatively low, at 0.21. This is in comparison to the frequencies observed in riverine Indian and Italian buffalo, which are 0.339 and 0.27, respectively (3-20). Therefore, it is crucial to implement breeding programs that utilize marker-assisted selection technology to increase the frequency of the *CSN1S2* GG genotype at g.14067 in river buffalo, especially within the Egyptian breed.

The constructed network described the relationship between 7 of the 9 *CSN1S2* CDS haplotypes. This is because the PopART program (28) does not consider gaps in the DNA sequence; therefore, haplotypes 8 and 9 were excluded from the network due to 27 gaps in their sequences (resulting from exon 7 skipping). We suggest that haplotype 8 originated directly from haplotype 1 through deletion of 27 nucleotides, and haplotype 9 originated from haplotype 8 through a transition c.459C>T.

Water buffalo (*Bubalus bubalis*) and cattle (*Bos taurus*) are closely related genera in the subfamily *Bovinae*. The *CSN1S2* CDS length in the buffalo and cattle *CSN1S2* gene is equal (669 bp), while the nucleotide variation between the two species is approximately 2.24% (3). The phylogenetic tree, based on this percentage of nucleotide variation between the *CSN1S2* CDS haplotypes of buffalo and cattle, successfully differentiated between the buffalo and cattle samples, reflecting the real relationship between the two closely related animal species. This result could be a useful tool for detecting milk and cheese fraud. Similarly, *CSN2* and *CSN3* coding sequences were previously used for the differentiation between buffalo and cattle species (43-44).

Conclusion

In the course of elucidating the molecular characteristics of the *CSN1S2* gene in Egyptian buffalo, DNA samples were genotyped for two non-synonymous SNPs: g.12803T>A in exon 14, g.14067G>A in exon 16, and a SNP in the splice donor site of intron 7 (g.7197 G>C), known to cause exon 7 skipping. Results revealed that the Egyptian buffalo possesses a high frequency of the GG genotype at g.7197 in *CSN1S2*, suggesting a reduced likelihood of exon 7 skipping. Also, the prevalent AA genotype at g.14067 may be associated with increased palmitic acid in the milk, which could pose health risks. Therefore, it is crucial to implement

breeding programs that utilize marker-assisted selection technology to increase the frequency of the *CSN1S2* GG genotype at g.14067 in river buffalo, especially within the Egyptian breed. The phylogenetic tree analysis revealed that the *CSN1S2* CDS successfully differentiated between buffalo and cattle samples, suggesting that it could be used to detect milk and cheese fraud.

Acknowledgments

Not applicable.

Funding statement

This study didn't receive any funding support

Declaration of Conflict of Interest

The authors declare that there is no conflict of interest.

Ethical approval

The protocol applied during this study was reviewed and approved by the Medical Research Ethics Committee at the National Research Centre (reference number 01019102021).

TABLE 1. The sequences of primers for the *CSN1S2* genes exon 7, 14, and 16, including their PCR annealing temperatures and amplicon lengths.

Amplified region	Primer sequence (5' to 3')	Annealing temperature	Product length (bp)
<Exon 7>	F: ATTATGAGACTCGCTGGGGA R: CAGGCTTCTTGGCCGCTTGT	57°C	541
<Exon 14>	F: GCAGACAAAATTACTGGTGGGC R: AGGCGATTCTTATCTTCTTCAGTCA	62°C	213
<Exon 16>	F: CCCCTTTCCTTCTTACACAA R: CCAATGCAAGCACAATGTAT	60°C	596

TABLE 2. *CSN1S2* CDS haplotypes: Base composition, country of origin, and presence in the two buffalo types

Haplotype	Accession Numbers	Buffalo Type / Origin	Base composition of the haplotype
Haplotype 1	NM_001290865.1 MT316464.1	Reference seq. River buff. /India Swamp buff. /China	TATAGCTGC
Haplotype 2	MT316465.1	River buff. / India	TATAGCTAC
Haplotype 3	MT316466.1	River buff. / India Swamp buff. / China	TATAACTGC
Haplotype 4	KY399458.2 FM865618.1	River buff. / Egypt River buff. /Italy	TATAGCAAC
Haplotype 5	MT316468.1	Swamp buff. / China	TCTAACTGC
Haplotype 6	DW007983.1 DQ133467.1	River buff. / India River buff. / India	TAAGGCAAC
Haplotype 7	DQ173244.1	River buff. /India	CATAGCAAG
Haplotype 8	KX896650.1	Swamp buff. / China	TATAGCTGC C.170- C.196 deleted
Haplotype 9	FM865619.1	River buff. / Italy	TATAGTTGC C.170 -C.196 deleted

TABLE 3. Polymorphic loci in the Egyptian buffalo: position, genotypic and allelic frequencies, aa substitutions or deletions, and tolerance score.

SNP	Position in the DNA amplified segment	Position in the *CSN1S2* gene reference sequence	Genotype frequency	Allele frequency	aa substitution or deletion	PROVEAN score
G>C **Intron 7**	257	g.7197	GG (0.84) GC (0.16) CC (0.00)	G (0.92) C (0.08)	deletion 9 aas (EVIRNANEE)	
T>A **Exon 14**	79	g.12803	TT (0.00) TA (0.16) AA (0.84)	T (0.08) A (0.92)	P.Phe147Ile	-2.350
G>A **Exon 16**	285	g.14067	GG (0.21) GA (0.11) AA (0.68)	G (0.26) A (0.74)	P.Ala175Thr	1.936

PROVEAN score:>-2.5=tolerant, <-2.5 intolerant The *CSN1S2* gene reference sequence: MW159135.1

Accsession No.	Haplotype No.	c.15	c.170	c.171	c.172	c.173	c.174	c.175	c.176	c.177	c.178	c.179	c.180	c.181	c.182	c.183	c.184	c.185	c.186	c.187	c.188	c.189	c.190	c.191	c.192	c.193	c.194	c.195	c.196	c.234	c.381	c.382	c.391	c.459	c.484	c.568	c.642
NM_001290865.1	Hap.1	T	A	A	G	T	T	A	T	A	A	G	G	A	A	T	G	C	A	A	A	T	G	A	A	G	A	G	G	A	T	A	G	C	T	G	C
MT316464.1	Hap.1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MT316465.1	Hap.2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.
MT316466.1	Hap.3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
KY399458.2	Hap.4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.
FM865618.1	Hap.4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	.
MT316468.1	Hap.5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	A	.	.	.	.	.
DQ133467.1	Hap.6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	A	A	.
DW007983.1	Hap.6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	A	A	.
DQ173244.1	Hap.7	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	A	G
KX896650.1	Hap.8	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.	.	.	.	.	.	.
FM865619.1	Hap.9	.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.	.	.	.	T	.	.	.

Fig. 1. Sequence difference sites between the investigated buffalo *CSN1S2* coding sequence haplotypes. The number above each base represents the base position in the reference sequence NM_001290865.1. Dots (.) represent the identity with the NM_001290865.1. Nucleotide substitutions are denoted by different letters.

A:

ATTATGAGACTCGCTGGGGATACAACAATAAACACAGTAAACAAAAATAAATACCACTGTCATC
AAATAGGAAATTCTTTTATAACTTATACAAAACTCATACAATGAAACCGAATCCAGTCTAATGC
ATCTCTTAATGAGAGAGAAATTTTTCCTTTTATATTCAGGAGAAAATAGCATAGCAATGCTGTG
GGGAAATACATCTAATTATTTTTCTTCTCTCTTTAAG*GAAGTTATAAGGAATGCAAATGAAGAG*
GTGAGAAAATTTTATACATTTGAACATTTTTTGTTCATATTACCAATATGTAATATTCTAAAAT
TAGTAATGGCATATGGGAACATGTTTAAATTTCCTTTTTTAGTAAACTGGAAAATGATTTATAA
GTTGCTTCAAAATGACAGTTTTTCCCTGTCCACTTGAGTAATGTCCACCTAAATCACTGTGTAT
TCTCTTGAGCCACTTACATTTAAATATATGATCATTTTTATCTCTTCCACTTATTTGCTAAAGG
TGATTAAAAACAAGCGGCCAAGAAGCCTG

B:

GCAGACAAAATTACTGGTGGGCTATTCAAGTAAAGCTGATTTTCTTTTTTTTTTTTCCATAAG*G*
*AATCAACAGAAGTA***A***TCACTAAG*GTAAGTAATTTCTCTTCAAAAAAAATTAATTACATGCTATC
CTATAATATTTGCTTTTGTATTTATTGCAGTTTGTTTTTATTATTAACAGAAAACTAAACTGAC
TGAAGAAGATAAGAATCGCCT

C:

CCCCTTTCCTTCTTACACAAATATTTTTTTCTTTTTTTCCCACTTAAGCATTTTGGCACTAGAG
ATTGAAATAACCAAAGAATATTTTTAAACCTACCATTTATTTTATAAAAGTAAACATAATTTTA
TATGAATAAACTTTACTTTGCTTAAGTCTGTTTGGTATCATTTAGAATTTATCCAAGTAAAACT
TTATTTGGTTAGGTCCAGGTGTTCTGACACATGGGATAATGGAATAATTATAATTTTCCTAG*AA*
*AAAAATCAGCCAGCATTACCAGAAATTC***A***CCTGGCCCCAGTATCTCAAGACTGTTTATCAGTAT*
*CAGAAAGCTATGAAGCCATGGACTCAACCTAAGACAAACGTTATTCCCTATGTG*GTGAGTTCTC
CCTTTTATTTTTAATTTTAAAACTGAGTTGTCTTTTGTAATAAAAATAAGATAGGGAAATGAAA
TATGGAAATAAAATTCTAAGTTCTAAAGTAAAAGCTAAAACAAAATTAAATGAATGTTCTAGGC
TGGAAATTGAATTATAAATCATATAGGATGAATAGAGCTACAATAAGCAGAAATAATTAGCATG
ATACATTGTGCTTGCATTGG

Fig. 2. Sequences of three DNA segments from the *CSN1S2* gene in Egyptian river buffalo. A: 541 bp representing a partial sequence of intron 6, exon 7, and partial sequence of intron 7. B: 213 bp representing a partial sequence of intron 13, exon 14, and partial sequence of intron 14. C: 596 bp representing a partial sequence of intron 15, exon 16, and partial sequence of intron 16. Exons are italic red. The nucleotide in the SNP position is bold and represented by the abundant allele in the investigated samples.

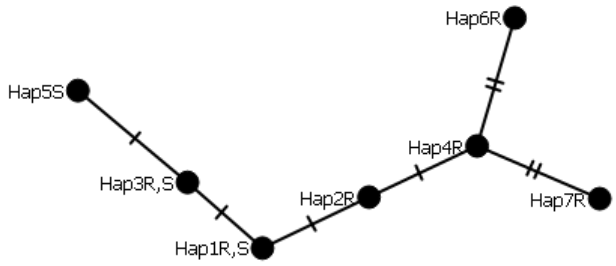

Fig. 3. Network of buffalo *CSN1S2* CDS haplotypes: 1-7. Each circle represents a haplotype. Dashes represent the number of mutations along the branch. Hap: haplotype, R: the haplotype presence in river buffalo, S: the haplotype presence in swamp buffalo.

Fig. 4. A maximum-likelihood (ML) phylogenetic tree based on the haplotypes of the *CSN1S2* CDS in buffalo and cattle. The tree is rooted with goat and sheep haplotypes

References

1. Smith, N.W., Fletcher, A.J., Hill, J.P. and McNabb, W.C. Modeling the contribution of milk to global nutrition. *Front. Nutr*. **8**, 716100 (2022). https://doi.org/10.3389 /fnut.2021.716100
2. Pauciullo, A., Versace, C., Gaspa, G., Letaief, N., Bedhiaf-Romdhani, S., Fulgione, A. and Cosenza, G. Sequencing and characterization of αs2-casein gene (*CSN1S2*) in the old-world camels have proven genetic variations useful for the understanding of species diversification. *Animals*, **13** (17), (2023). https://doi.org/10.3390/ani13172805
3. Fan, X., Gao, S., Fu, L., Qiu, L and Miao, Y. Polymorphism and molecular characteristics of the *CSN1S2* gene in river and swamp buffalo. *Archives Animal Breeding*, **63**(2), 345–54 (2020). https://doi.org/10.5194/aab-63-345-2020
4. Petrova, S.Y., Khlgatian, S.V., Emelyanova, O.Y., Pishulina, L.A. and Berzhets, V.M. Structure and biological functions of milk caseins. *Russian Open Medical Journal*, **11**, (2022). https://doi.org/10.15275/rusomj.2022. 0209
5. Wedholm, A., Hallén, E., Bach Larsen, L., Lindmark-Månsson, H., Hans Karlsson, A. and Allmere, T. Comparison of milk protein composition in a Swedish and a Danish dairy herd using reversed phase HPLC. *Acta Agriculturae Scand Section A*., **56**(1), 8–15 (2006). https://doi.org/10.1080/09064700600836745
6. Nilsen, H., Olsen, H.G., Hayes, B., Sehested, E., Svendsen, M., Nome, T., Meuwissen, T., and Lien, S. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. *Genetics Selection Evolution*, **41**, 1–12 (2009).https://doi.org/10. 118 6/1297-9686-41-24
7. Michelizzi, V.N., Dodson, M.V., Pan, Z., Amaral, M.E.J., Michal, J.J., McLean, D.J., Womack, J.E. and Jiang, Z. Water buffalo genome science comes of age. *Int. J. Biol. Sci*. **6**(4), 333–349 (2010). http://www.ncbi.nlm.nih.gov/pubmed/20582226
8. Minervino, A.H.H., Zava, M., Vecchio, D. and Borghese, A. *Bubalus bubalis*: A short story. *Front. Vet. Sci*., **7**, 570413 (2020). https://doi.org/10.3389/fvets.20 20.570413
9. Youssef, N.A., Curaudeau, M., El Nahas, S.M., Hassan, A.A.M. and Hassanin, A. Haplotype diversity in the mitochondrial genome of the Egyptian river buffalo (*Bubalus bubalis*). *Mitochondrial DNA Part B*., **6**(1), 145–147 (2021). https://doi.org/10.1080/23802359.2020.1852622
10. FAO. Food and Agriculture Organization of the United Nations. Statistics of Live Animals. 2021. http://www.fao.org/faos tat/en/ data/QA (accessed on 6 May 2024).
11. Pauciullo, A., Gaspa, G., Zhang, Y., Liu, Q. and Cosenza, G. CSN1S1, CSN3 and LPL: Three validated gene polymorphisms useful for more sustainable dairy production in the Mediterranean river buffalo. *Animals*, **14**(10), 1414 (2024). https://doi.org/10.3390/ani14101414
12. Pauciullo, A., Versace, C., Perucatti, A., Gaspa, G., Li, L.Y., Yang, C.Y., Zheng, H.Y., Liu, Q. and Shang, J.H. Oocyte aneuploidy rates in river and swamp buffalo types (*Bubalus bubalis*) determined by Multi-color Fluorescence In Situ Hybridization (M-FISH). *Scientific Reports*, **12**(1), 8440 (2022). https://doi.org/10.1038/s41598-022-12603-9
13. Rehman, S., Feng, T., Wu, S., Luo, X., Lei, A., Luobu, B., Hassan, F.U. and Liu. Q. Comparative genomics, evolutionary and gene regulatory regions analysis of casein gene family in *Bubalus bubalis*. *Frontiers in Genetics*, **12**, 1–11 (2021). https://doi.org/10.3389/fgene.2021.662609
14. Basilicata, M.G., Pepe, G., Sommella, E., Ostacolo, C., Manfra, M., Sosto, G., Pagano, G., Novellino, E. and Campiglia, P. Peptidome profiles and bioactivity elucidation of buffalo-milk dairy products after gastrointestinal digestion. *Food Research International*, **105**, 1003–1010 (2018). https://doi.org/10.1016/j.foodres.2017.12.038
15. Iannuzzi, L., Di Meo, G.P., Perucatti, A., Schibler, L., Incarnato, D., Gallagher, D., Eggen, A., Ferretti, L. and Cribiu, E.P., Womack, J. The river buffalo (Bubalus bubalis, 2n= 50) cytogenetic map: assignment of 64 loci by fluorescence in situ hybridization and R-banding. *Cytogenetic and Genome Research*, **102** (1–4), 65–75 (2003). https://doi.org/10.1159/00007 5727
16. Cosenza, G., Pauciullo, A., Coletta, A., Di Francia, A., Feligini, M., Gallo, D., Di Berardino, D. and Ramunno, L. Short communication: Translational efficiency of casein transcripts in Mediterranean river buffalo. *J. Dairy Sci.*, **94**(11), 5691–5694 (2011). https://doi.org/10.3168/jds.2010-4086
17. Kim, J.J., Yu, J., Bag, J., Bakovic, M. and Cant, J.P. Translation attenuation via 3′ terminal codon usage in bovine *CSN1S2* is responsible for the difference in αs2- and β-casein profile in milk. *RNA Biol*., **12**(3), 354–367 (2015). https://doi.org/10.1080/15476286.2015.101 7231
18. Lefèvre, C.M., Sharp, J.A. and Nicholas, K.R. Characterisation of monotreme caseins reveals lineage-specific expansion of an ancestral casein locus in mammals. *Reprod. Fertil. Dev*. **21**(8), 1015–1027 (2009). https://doi.org/10.1071/RD 09083
19. Corral, J., Padilla, J., Izquierdo, M. Martínez-Trancón, M., Parejo, J.C., Salazar, J. and Hernández-García, F.I., Detection and genetic characterization of ovine *CSN1S2**B polymorphisms and their associations with milk production traits. *Livestock Science*, **153**(1–3), 10–19 (2013). https://doi.org/ 10.1016/j.livsci.2013.01.008
20. Cosenza, G., Gallo, D., Auzino, B., Gaspa, G. and Pauciullo, A. Complete *CSN1S2* characterization, novel allele identification and association with milk fatty acid composition in river buffalo. *Front. Genet.*, **11**, 1–12 (2021). https://doi.org/10.3389/fgene.2020.622494
21. Cosenza, G., Pauciullo, A., Feligini, M., Coletta, A., Colimoro, L., Di Berardino, D. and Ramunno, L. A point mutation in the splice donor site of intron 7 in the as2-casein encoding gene of the Mediterranean River buffalo results in an allele-specific exon skipping.

Anim. Genet., **40**(5), 791 (2009). doi:10.1111/j.1365-2052.2009.01897.x

22. Berget, S.M. Exon recognition in vertebrate splicing. *Journal of Biological Chemistry*. **270**(6), 2411–2414 (1995). 10.1074/jbc.270.6.2411 External Link

23. Sambrook, J. and Russell, D.W. Molecular Cloning: Ch. 8. In Vitro amplification of DNA by the polymerase chain reaction. Vol. 2. Cold Spring Harbor Laboratory Press; 2001.

24. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. Primer3-new capabilities and interfaces. *Nucleic Acids Res*., **40**(15) (2012). https://doi.org/10.1093/nar/gks596

25. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. and Söding, J. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol*., **7**(1), 539 (2011). https://doi.org/10.1038/msb.2011.75

26. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, **30**(22), 3276–3268 (2014). https://doi.org/10.1093/bioinformatics/btu 531

27. Yeh, F.C. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J. Bot*., **129**, 157 (1997). https://sid.ir/paper/571184/en

28. Choi, M.K., Le, M., Nguyen, D., Choi, H., Kim, W., Kim, J.H., Chun, J., Hyeon, J., Seo, K. and Park, C. Genome-level identification, gene expression, and comparative analysis of porcine SZ-defensin genes. *BMC Genet*., **13**(1), 98 (2012). http://www.biomedcentral.com/1471-2156/13/98.

29. Leigh, J.W., Bryant, D. and Nakagawa, S. POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol*., **6**(9) (2015). doi: 10.1111/2041-210X.12410

30. Huerta-Cepas, J., Serra, F. and Bork, P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol*., **33**(6), 1635–1638 (2016). https://doi. org/10.1093/molbev/msw046.

31. Price, M.N., Dehal, P.S. and Arkin, A.P. Fast Tree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol*., **26**(7), 1641–1650 (2009). https://doi.org/10.1093/molbev/msp077

32. Younis, A., Hussain, I., Ahmad, S.N., Shah, A., Inayat, I., Kanwal, M.A., Suleman, S., Kamran, M.A., Matloob, S. and Ahmad, K.R. Validation of *Bos taurus* SNPs for milk productivity of sahiwal breed (*Bos indicus*), Pakistan. *Animals*, **14**(9), 1306 (2024). https://doi.org/10.3390/ani14091306

33. Arslan, A., Kaplan, M., Duman, H., Bayraktar, A., Ertürk, M., Henrick, B.M., Frese, S.A. and Karav, S. Bovine colostrum and its potential for human health and nutrition. *Front. Nutr*. **8**, 651721 (2021). https://doi.org/10.3389/fnut.2021.651721

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

35. Selvaggi, M., Laudadio, V., Dario, C. and Tufarelli, V. Major proteins in goat milk: an updated overview on genetic variability. *Mol. Biol. Rep*., **41**, 1035–1048 (2014). https://doi.org /10.1007/s11033-013-2949-9

36. Ozdemir, M., Kopuzlu, S., Topal, M. and Bilgin, O.C. Relationships between milk protein polymorphisms and production traits in cattle: a systematic review and meta-analysis. *Arch. Anim. Breed*., **61**(2), 197–206 (2018). https://doi.org/10.5194/aab-61-197-2018

37. German, J.B. Dietary lipids from an evolutionary perspective: sources, structures and functions. *Maternal and Child Nutr*., **7**, 2–16 (2011). https://doi.org/10.1111/j.1740-8709.2011.00300.x

38. Burch, A.M., Pineda, A. and Lock, A.L. Effect of palmitic acid-enriched supplements containing stearic or oleic acid on nutrient digestibility and milk production of low- and high-producing dairy cows. *J. Dairy Sci*., **104**(8), 8673–8684 (2021). https://doi.org/10.3168/jds.2020-19913

39. Cosenza, G., Macciotta, N.P.P., Nudda, A., Coletta, A., Ramunno, L. and Pauciullo, A. A novel polymorphism in the oxytocin receptor encoding gene (OXTR) affects milk fatty acid composition in Italian Mediterranean river buffalo. *Journal of Dairy Research*, **84**(2), 170–180 (2017). https://doi.org/10.1017/ S0022029917000127

40. Pena-Serna, C., Gómez-Ramirez, B. and Zapata-López, N. Nutritional aspects of ghee based on lipid composition. *Pak. J. Nutr*., **18**, 1107–1114 (2019).

41. Praagman, J., de Jonge, E.A.L., Kiefte-de Jong, J.C., Beulens, J.W.J., Sluijs, I., Schoufour, J.D., Hofman, A., van der Schouw, Y.T. and Franco, O.H. Dietary saturated fatty acids and coronary heart disease risk in a Dutch middle-aged and elderly population. *Arterioscler Thromb. Vasc. Biol*., **36**(9), 2011–2018 (2016). https://doi.org/10.1161/ATVBAHA.116.307578

42. Imamura, F., Fretts, A., Marklund, M., Ardisson Korat, A.V., Yang, W.S., Lankinen, M., Qureshi, W., Helmer, C., Chen, T.A. and Wong, K. Fatty acid biomarkers of dairy fat consumption and incidence of type 2 diabetes: a pooled analysis of prospective cohort studies. *PLoS Med*., **15**(10), e1002670 (2018). https://doi.org/10.1371/journal.pmed.1002670

43. Fan, X., Zhang, Z., Qiu, L., Zhang, Y. and Miao, Y. Polymorphisms of the kappa casein (CSN3) gene and inference of its variants in water buffalo (*Bubalus bubalis*). *Arch. Anim. Breed.*, **62**(2), 585–596 (2019). https://doi.org/10.5194/aab-62-585-2019

44. Fan, X.Y., Xue, J., Qiu, L.H., Wang, R.P. and Miao, Y.W. Polymorphisms of beta casein (Csn2) cds and inference of its variants in river and swamp water buffalo (*Bubalus bubalis*). *J. Anim. Plant Sci*., **31**(5), 1240–1251 (2021). https://doi.org/10.36 899/JAPS.2021.5.0324

التوصيف الجزيئي لجين *CSN1S2* في الجاموس النهري المصري

الحسن شهوان[1]، نادية أبو المعاطي[1]، أمل حسن[1]، إيمان صالح[2]، هناء عرابي[1] و أسامة أ. إبراهيم[2]

[1] قسم بيولوجيا الخلية، معهد بحوث التكنولوجيا الحيوية، المركز القومي للبحوث، القاهرة، مصر.

[2] قسم الكيمياء الحيوية، كلية العلوم، جامعة عين شمس، القاهرة، مصر.

الملخص

خلال عملية توضيح الخصائص الجزيئية لجين CSN1S2 في الجاموس المصري قمنا بتحليل عدة أنماط وراثية مختلفة لتسلسل الترميز الكامل (CDS) لجين CSN1S2 في الجاموس النهري وجاموس المستنقعات من بنك الجينات. تم تحديد اثنين من تعدد الأشكال المتبادلة للنوكليوتيدات غير المترادف (SNPs) (g.12803T>A) c.484T>A في اكسون 14 و (g.14067G>A) c.568G>A في اكسون 16, كأكثر التغيرات شيوعًا بين الأنماط الوراثية المدروسة، وقد وُجدت حصريًا في الجاموس النهري. لذلك، تم دراسة الأنماط الوراثية لعينات الحمض النووي التي تم جمعها من ثمانٍ وثلاثين رأسًا من الجاموس المصري النهري للكشف عن هذه التعددات النوكليوتيدية الأحادية . (SNPs) كما تم أيضًا تحليل تعدد طفرة أخرى في موقع الربط في انترون7 (g.7197 G>C) من جين *CSN1S2* , والمعروفة بأنها تسبب حذف اكسون 7. كانت الأنماط الجينية السائدة في الموقعين g.7197 و g.12803 هي GG و AA على التوالي بنسبة تكرار 0.84 لكل منهما. ولم يتم الكشف عن النمطين الجينيين CC و TT في هذين الموقعين، على الترتيب. أما ترددات الأنماط الجينية في الموقع g.14067 فقد بلغت 0.21 للنمط GG و0.11 للنمط GA و 0.68 للنمط AA. هذا و أظهرت نتائج التحليل أن الجاموس المصري يمتلك تكرارًا عاليًا للنمط الجيني GG في الموقع g.7197 من جين CSN1S2، مما يشير إلى انخفاض احتمالية حذف الإكسون 7 .في المقابل، فإن النمط الجيني السائد AA في الموقع g.14067 قد يكون مرتبطًا بزيادة محتوى حمض البالمتيك في الحليب، مما يشير إلى احتمال وجود مخاطر صحية. و للحد من هذه المخاطر، ينبغي أن تركز برامج التربية المعتمدة على الاختيار بمساعدة العلامات الوراثية على زيادة تكرار النمط الجيني GG لجين CSN1S2 عند الموقع g.14067. بالإضافة إلى ذلك، فحصت هذه الدراسة العلاقات الوراثية بين الأنماط الوراثية للمنطقة المشفرة لجين CSN1S2 في الجاموس، و تم إنشاء (phylogenetic tree) شجرة تطورية تميز بين عينات الجاموس والأبقار.

الكلمات الدالة: كازين αS2، الجاموس، تسلسل الحمض النووي، الأنماط الوراثية، تعدد الأشكال.