Genetic Polymorphism of Kappa-Casein Gene in Senegalese Local Cattle Breeds

Ndiaye B [1, 2*], Ndiaye N.P[1] and Sembène M[1]
1. Department of Animal Biology, Faculty of Sciences and Techniques, Cheikh Anta Diop University of Dakar, P.O. Box 5005, Dakar-Fann, Senegal
2. National Livestock and Veterinary Research Laboratory, BP 2057 Route du Front de Terre Dakar-Hann, Senegal.
Corresponding author: Bakary NDIAYE, Phone: +221 77 327 38 28, E-mail: morbakaryndiaye@hotmail.fr / bakary2.ndiaye@ucad.edu.sn
Key words: kappa-casein, genetic variants, dairy production, cattle breeds, Senegal

Submitted 18/01/2023, Published online on 31/12/2023 in the *Journal of Animal and Plant Sciences (J. Anim. Plant Sci.) ISSN 2071 – 7024*

1 ABSTRACT

In Senegal, the dairy production improvement of the local cattle breeds, is one of the most political and economic issues challenged by public authorities and stakeholders. To address such challenges, the study of genes related to milk production is essential to accelerate the genetic improvement of these cattle breeds. Among these lactoprotein genes, kappa-casein (*CSN3*) is one of the most targeted because of its effects on milk composition and cheese making properties. The objective of this study was to identify the kappa-casein (*CSN3*) genetic variants and to estimate its polymorphism in the Senegalese local cattle breeds. The animal sampling consisted of 48 cattle from the four breeds namely Gobra zebu, Maure zebu, Djakore and N'Dama taurine. The characterization of the *CSN3* gene exon IV genetic polymorphism was performed using PCR-Sequencing method. The identification revealed four genetic variants *A*, *B*, *A^I* and *H*, from which six haplotypes were resulted: *AA*, *AA^I*, *AH*, *A^{I}H*, *BB* and *BH*. The highest genetic variability of the *CSN3* gene was observed in Maure zebu and Djakore with four haplotypes for each breed. Genetic variants *CSN3*A* and *CSN3*B* associated with high milk protein content, cheese making properties, and milk coagulation properties are the most represented in the Senegalese local cattle breeds. These preliminary results should encourage the inclusion of these genetic variants among the selection criteria when designing breeding programs for improving dairy production in Senegal.

2 INTRODUCTION

Senegalese local cattle consisted of three indicine breeds, namely Gobra zebu, Maure zebu and Djakore and one taurine breed, the N'Dama. Due to precarious breeding conditions with few inputs, their production remains marginal (Ndiaye *et al.,* 2015; Sambe *et al.,* 2019; Badji *et al.,* 2020). Their average daily milk yield does not exceed 3 litres for the best producing cows (Bertrand, 2006; Ndiaye, 2015). Although their low milk production, these breeds contributed to over a half of the national production ~ 243.5 million litres, representing 54% of the country's milk production (MEPA, 2018). In 2017, milk and dairy products imports reached 46.46% of the national consumption (ANSD, 2020). The dairy bill is very expensive and is close to 60 billion CFA francs in 2018. The introduction of highly performant dairy breeds through artificial insemination programs have been established to boost the milk production of the local cattle breeds (Ndiaye *et al.,* 2015). However, the lack of an insightful national plan for the genetic improvement of the local cattle breeds reduces the chances of successful programs. The growth

of dairy farms on the outskirts of large cities such as Dakar, Thies and Touba has contributed to improve the national production, and the milk quantity consumed per capita/per year rose from 27 litre/inhabitant/year from 1994 to 29.41 in 2016 (ANSD, 2020). This increasing of individual consumption remains low and does not directly affect imports of dairy products, of which quantities continue to rise every year. Those imports increased from 29,773 tonnes to 30,618 tonnes between 2016 and 2017, representing an increase of 2.8% (MEPA, 2018). Bovine lactoproteins have been considerably investigated during the past years due to their possible association with milk yield and some reproductive performance in dairy cattle (Grosclaude *et al.,* 1988; Dadhich *et al.,* 2006; Karimi *et al.,* 2009). Several studies showed that certain lactoproteins may be associated with milk protein content and cheese making properties (Bovenhuis *et al.,* 1992; Caroli *et al.,* 2009). Therefore, milk protein genes could be useful as genetic markers for additional selection criteria in the dairy industry (Otavianio *et al.,* 2005; Anggraeni *et al.,* 2010). Bovine milk contains six main milk protein sub-groups classified into soluble and insoluble fractions (Grosclaude, 1988, Hamza *et al.,* 2011). The soluble fraction, called whey protein, consists of α-lactalbumin and β-lactoglobulin. The insoluble fraction known as whole casein is the main component of milk proteins (78%-82%) and consists of$\alpha_{s1} - CN$, $\alpha_{s2} - CN$, $\beta - CN$ and $\kappa - CN$ (Anggraeni *et al.,* 2010). Milk proteins polymorphisms are due to its extensive genetic variation and their variants are transmitted by simple and non-dominant Mendelian inheritance (Otaviano *et al.,* 2005). Kappa-casein ($\kappa - CN$) differs from other caseins in structure and other intrinsic characteristics (Hamza *et al.,* 2011). The primary structure of this protein was determined by Dumas *et al.* (1972) who described it as consisting of 169 amino acids of mature protein. The *CSN3* ($\kappa - CN$) is subdivided into five exons and four introns with a high number of genetic variants in cattle populations (Alipanah *et al.,* 2005; Getachew, 2010; Hamza *et al.,* 2011). So far, thirteen milk protein variants (*A*, *B*, B^2, *C*, *D*, *E*, F^1, F^2, G^1, G^2, *H*, *I*, and *J*) and one synonymous variant (A^1) of *CSN3* gene have been identified in cattle breeds with the most common as alleles *A* and *B* (Farrell *et al.,* 2004; Caroli *et al.,* 2009). Owing to the importance of bovine lactoproteins in deciphering the dairy abilities of cows, thus Senegalese cattle breeds need to be characterized at molecular level as alternative to traditional selection so far used in breeding programs. Given their low milk yield, the study of the main milk protein variants could be useful to assess genetically the dairy production performance of these breeds. The objective of this research was to identify the kappa-casein (*CSN3*) genetic variants and to estimate its polymorphism in the Senegalese local cattle breeds.

3 MATERIALS AND METHODS

3.1 Animal Sampling: The study was conducted from October to December 2013 in three (3) agro-pastoral regions in Senegal, namely Saint-Louis (16°02'00"N and 16°30'00"W), Kaolack (14°08'35"N and 16°05'45"W), Kaffrine (Latitude: 14° 06' 21.38" N Longitude: -15° 33' 2.88" W) and Kolda (13°01'60"N and 14°52'00"W) (Figure 1). These regions represent the main livestock breeding areas of Gobra zebu, Maure zebu, Djakore and N'Dama breeds. Samples were collected in 15 localities within these regions and in the Zootechnical Research Center (ZRC) of Kolda. The choice of the locations in each region was based on the presence of the target breed. The animals were selected by random sampling according to the age and the physiological state of cows (lactating). The animal sample was composed of 48 unrelated cows with 12 from each breed. Blood samples were collected from the animal jugular vein into 4 mL vacutainer tubes containing the disodium salt of ethylene diamine tetra-acetate (EDTA) as anti-coagulant and stored ± 4ºC until DNA extraction. During

the blood sampling, adequate measures were taken to minimize animals' pain and discomfort. This study was approved by the Ethics Committee of the Cheikh Anta Diop University of Dakar. Participants' approval was obtained after fully explanations of the study.

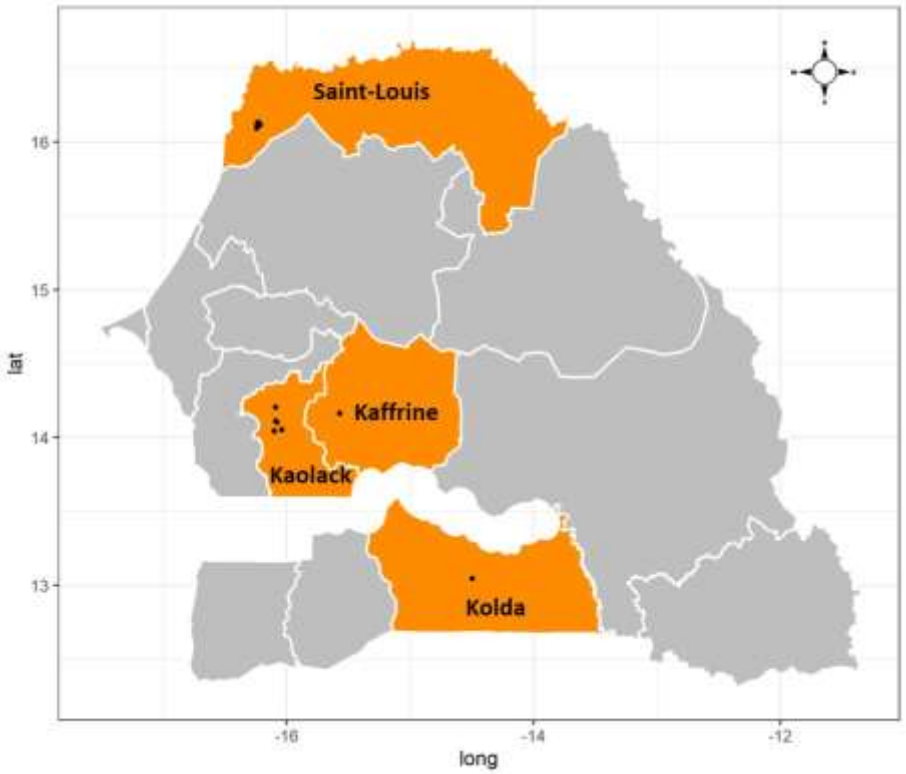

Figure 1: Mapping of study sites across three agro-pastoral regions in Senegal (Saint-Louis, Kaolack, Kaffrine and Kolda). lat: latitude, long: longitude. The map was built using the R programming language (R Core Team, 2020).

3.2 DNA extraction and PCR-Sequencing: DNA was isolated from whole blood sample using Gentra Puregene Blood kit standard protocol developed by QIAGEN® group. The amplification of the exon IV of the *CSN3* gene was carried out in an Eppendorf thermocycler (Master cycler Gradient 5331 version 2.30.31.09) according to the conventional PCR protocols with forward primer CN-F : 5'- CAG CGC TGT GAG AAA GAT GA-3' and reverse primer CN-R: 5'- CCC ATT TCG CCT TCT CTG TA-3'. Polymerase chain reaction (PCR) amplifications were performed using cycling conditions as: initial denaturation at 94°C for 3 min, followed by 34 cycles of denaturation at 94°C for 1 min, hybridization at 60°C for 1 min and elongation of the complementary DNA strand at 72°C for 1 min, and final elongation at 72°C for 10 min ended the PCR reactions. A hold at 10°C was set to store the PCR products until their removing from the thermocycler. The DNA sequencing was performed in South Korea from 30 µl of PCR product.

3.3 Genetic Analysis: The raw sequences were aligned and corrected manually using BioEdit v7.2.0 (Hall, 1999) to determine site homology, to identify the different variants of the *CSN3* gene exon IV and define haplotypes. To find similarity regions between the query sequences and the reference sequence *Bos taurus* kappa-casein (*CSN3*A* allele), the Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov) available in GenBank was used. The *CSN3* gene exon IV sample sequences were aligned and compared to

the reference sequence to identify the different kappa-casein variants present in the Senegalese local cattle breeds.

3.4 Identification of genetic variants of the *CSN3* gene exon IV: The identification of *CSN3* ($\kappa - CN$) variants in the Senegalese local cattle breeds was carried out using the *Bos taurus* reference sequence *CSN3*A* retrieved from GenBank with the accession number AY380228 (https://www.ncbi.nlm.nih.gov/nuccore/AY38 02280). This reference sequence is a fragment of 633 bp and consists of part of intron III (4 bp), the whole exon IV (516 bp) and part of intron IV (113 bp). The reference sequence variant *CSN3*A* showed different variants previously described by Caroli et al. (2009) at 14 positions (12690/10, 12940/93, 12950/97, 12951/97, 12971/104, 13065/135, 13068/136, 13096/145, 13104/148, 13111/150, 13119/153, 13124/155, 13162/167, 13165/168).

3.5 Estimation of the kappa-casein genetic polymorphism within breeds: The sample sequences of the *CSN3* gene exon IV enabled to emphasized punctual mutations such as substitutions and indels (insertions and deletions) within breeds through the SNPs (Single Nucleotide Polymorphisms) using the MEGA software v6.05 (Tamura *et al.,* 2013). The basic genetic parameters such as the number of variable sites, the number of parsimony informative sites, the nature of mutations and the substitution rate (ratio transitions/transversions) were also calculated using MEGA software. In addition, haplotypic and nucleotide diversities (Hd and Pi) (Nei, 1978) were estimated using the DnaSP software v5.10.1 (Rozas *et al.,* 2010).

3.6 Genetic distances and differentiation of the kappa-casein gene between breeds: The genetic distances between individuals from the same breed and between breeds were determined using the MEGA software with the Kimura-2Parameters model (Kimura, 1980). The pairwise F_{ST} values were estimated using the Arlequin software v3.11 (Excoffier *et al.,* 2005). The F_{ST} index provides information on population differentiation and subdivision. According to Wright (1978), more the F_{ST} is closer to 1, more the populations are genetically structured. However, if the F_{ST} is equal to zero, there is no genetic differences between populations. The F_{ST} values are considered statistically significant at the significance level of $P < 0.05$.

3.7 Haplotype network: The Minimum Spanning Tree (MST) or haplotype network was built using the Network software v4.6.1.3 (Polzin and Daneshmand, 2003) based on a matrix of pairwise distances calculated between all haplotype pairs using a modified algorithm. The haplotype network is a tree that highlights clusters of individuals sharing common nucleotide sequences and the existence of possible associations between haplotypes. In the minimum haplotype network, each circle stands for a haplotype, each link represents a mutation and the size of each circle is proportional to its frequency.

4 RESULTS

4.1 Nucleotide sequences of the *CSN3* gene exon IV: A total of 48 nucleotide sample sequences resulted from the PCR-Sequencing. After data editing, 31 sequences with a length of 418 base pairs (bp) were left and used to perform downstream analyses. The BLAST in GenBank output a similarity rate of 95%. The individuals removed showed a rather important nucleotide differences compared to those kept for this study during the sequence alignment. Among the individuals kept for the analyses, 9 were from the Djakore breed, 7 from the Gobra zebu, 6 from the zebu Maure and 9 from the N'Dama taurine.

***4.2* *CSN3* gene exon IV genetic variants and haplotypes:** A total of 5 substitutions including 3 non-synonymous and 2 synonymous allowed to identify 4 variants out of 5 from the 12 sites identified (Table 1). No substitution was found at the positions 12940, 12950, 12951, 12971 and 13096 for all the 31 sample sequences of the dataset (Tables 1 and 2). Only 6

individuals showed no nucleotide changes at the *CSN3*A* allele reference positions and therefore have the variant A and are considered as *AA* haplotype. At the positions 13068/136 and 13104/148, two non-synonymous substitutions (C > U) and (A > C) were observed in the 2nd position of the codon, respectively and from which resulted a change in amino acid (Thr > Ile and Asp > Ala) in both cases. In addition, a synonymous substitution in the 3rd position of the codon (A > G) without a change in amino acid was detected at the position 13165/168. The mutations at these 3 positions corresponded to the variant B (*CSN3*B* allele) and therefore the individuals having those changes in amino acid are from *BB* haplotypes. The variant A^I was also detected and is due to a synonymous substitution in the 3rd position of the codon (A > G) at the position 13111/150 without change in amino acid. This mutation was observed in 4 individuals of which 3 are AA^I haplotypes because they only showed this transition among the reference positions of the *CSN3*A* variant. The remaining individual presents in addition to the substitution at this position (13111/150), a non-synonymous mutation of type transition (C > U) in the 2nd position of the codon at the position 13065/135 with a change in amino acid (Thr > Ile) and therefore is a A^IH haplotype. Individuals presenting only this substitution at the position 13065/135 with respect to the *CSN3*A* variant are *AH* haplotypes. The substitutions at the positions 13065/135 (variant *H*), 13068/136 and 13104/148 (variant *B*) are found in one individual and constitute the *BH* haplotype. However, the substitutions found at the positions 13119/153 and 13124/155 did not correspond to any variant relative to the reference sequence. The positions 12690/10 and 13162/167 were excluded during the identification process because having deletions in the 1st position of the codon compared to the reference sequence.

Table 1. CSN3 genetic variants identified in the Senegalese local cattle breeds.

	CSN3 variants[1]			
SNP, AA positions[2]	A	B	A^I	H
12690	**CGC**			
10	**Arg**			
12940	ACU			
93	Thr			
12950	CGU			
97	Arg			
12951	CGU			
97	Arg			
12971	UCA			
104	Ser			
13065	A**C**C			A**U**C
135	Thr			Ile
13068	A**C**C	A**U**C		
136	Thr	Ile		
13096	ACU			
145	Thr			
13104	G**A**U	G**C**U		
148	Asp	Ala		
13111	CC*A*		CC*G*	
150	Pro		Pro	
13119	AUU			

153	Ile	
13124	AGC	
155	Ser	
13162	**ACT**	
167	**Thr**	
13165	GC*A*	GC*G*
168	Ala	Ala

[1]Different CSN3 variants identified in *Bos taurus* from the CSN3*A allele corresponding to the reference sequence found in GenBank No. AY380228. [2]SNP: single nucleotide polymorphism; AA: amino acid.

The identification of *CSN3* genetic variants resulted in a total of 4 which are *A*, *B*, A^I and *H* identified in 16 individuals out of all the 31 sample sequences. These variants allowed to classify the individuals into 6 haplotypes: *AA*, *BB*, *AH*, AA^I, A^IH and *BH* (Table 3). Among 16 individuals identified, 4 were from the Djakore breed, 3 from Gobra zebu, 5 from Maure zebu and 4 from N'Dama taurine. The *AA*, *BB* and AA^I haplotypes were predominant in the local cattle breeds in Senegal. The *AA* haplotype was identified in 6 individuals, *BB* haplotype in 3 individuals and AA^I haplotype in 3 individuals. The A^IH and *BH* haplotypes were found in two breeds, respectively in Djakore and Maure zebu. The *AH* haplotype was found in two individuals of which one is belonged to the Gobra breed and the other one from the Djakore (Table 3). The Djakore and Maure zebu were the most diversified breeds with four different haplotypes for each. In contrast, only 3 haplotypes have been recorded in Gobra zebu and 2 haplotypes in N'Dama.

Table 3. Distribution of CSN3 gene haplotypes according to the four cattle breeds.

	Djakore	**Gobra zebu**	**Maure zebu**	**Ndama taurine**
AA	0	1	2	3
BB	1	0	1	1
AH	1	1	0	0
AA^I	1	1	1	0
A^IH	1	0	0	0
BH	0	0	1	0

4.3 Genetic variability of the kappa-casein (*CSN3*) within breeds: After correction and alignment, all the 31 sample sequences with a length of 418 sites were analysed. From the total sites, 76.31% were monomorphic and 23.68% were variable sites, among which 11.49% were singletons and the remaining 12.20% were parsimony informative sites. The mutations of type transversion were predominant and accounted for 52.79% of all substitutions while transitions represented 47.21% (Table 4). In Djakore and Gobra zebu breeds, transitions are found to be more important than transversions whereas in Maure zebu and N'Dama, the transversions are more numerous than transitions. The greatest genetic variability of the *CSN3* gene was found in Gobra and N'Dama breeds (14.11% of variable sites) and the lowest in the Djakore breed with 2.63% of variable sites. The Djakore breed had the highest substitution rate (2.58) among the four local cattle breeds. The *CSN3* gene variability showed a high haplotypic diversity and a low nucleotide diversity in all the four bovine breeds

(Hd = 0.983 ± 0.014 and Pi = 0.037 ± 0.007) and by breed separately (Table 4). The haplotypic diversity is maximum (Hd = 1) in the two zebu breeds and the Djakore while the nucleotide diversities are very low (0.009 < Pi < 0.049). The N'Dama breed also showed when considered the whole population a high haplotypic diversity and a low nucleotide diversity (Hd = 0.917 ± 0.092 and Pi = 0.048 ± 0.012).

Table 4. Summary of the basic genetic diversity parameters.

	Djakore	Gobra zebu	Maure zebu	Ndama taurine	Whole population
Transitions	73 .18	50.13	47.17	42.99	47.21
Transversions	26.82	49.87	52.83	57.01	52.79
SV	2.63	14.11	11	14.11	23.68
R	2.576	0.930	0.849	0.704	0.838
Hd	1.000 ± 0.052	1.000 ± 0.076	1.000 ± 0.096	0.917 ± 0.092	0.983 ± 0.014
Pi	0.009 ± 0.002	0.049 ± 0.015	0.040 ± 0.013	0.048 ± 0.012	0.037 ± 0.007

SV: variable sites; R: substitution rate; Hd: haplotypic diversity; Pi: nucleotide diversity.

4.4 Genetic differentiation of the Kappa-casein (*CSN3*) between breeds: Values of the within and between breeds' genetic distances were very low in general and summarized in table 5. The lowest intra-population genetic distance was observed within Djakore breed (0.010) and the largest in Gobra zebu (0.052). Regarding the inter-population genetic distances, the largest value was recorded between Djakore and Gobra zebu (0.054) and the lowest value between Maure zebu and the N'Dama breed (0.026). Considering the genetic differentiation between breeds, the highest pairwise F_{ST} value was observed between the N'Dama and the Gobra zebu (0.324, P = 0.01) while the lowest was obtained between N'Dama and Maure zebu (0.171, P = 0.02) (Table 6). The two non-significant pairwise F_{ST} values were observed between Maure zebu and Djakore, and between Maure zebu and Gobra zebu. The N'Dama breed remains the most genetically differentiated among the cattle populations.

Table 5. Genetic distance estimates within and between breeds.

	Intrapopulation distances	**Interpopulation distances**			
		Djakore	**Gobra zebu**	**Maure zebu**	**Ndama taurine**
Djakore	0.010 ± 0.003	—			
Gobra zebu	0.052 ± 0.007	0.054	—		
Maure zebu	0.042 ± 0.006	0.046	0.047	—	
Ndama taurine	0.050 ± 0.007	0.034	0.033	0.026	—

Table 6. Genetic differentiation estimates between breeds.

	F_{ST} values			
	Djakore	**Gobra zebu**	**Maure zebu**	**Ndama taurine**
Djakore	—	**0.175** (0.009)	0.129 (0.135)	**0.198** (0.009)
Gobra zebu		—	0.107 (0.108)	**0.324** (0.009)
Maure zebu			—	**0.171** (0.018)
Ndama taurine				—

Significant F_{ST} values are in bold with a p-value of 5%. Values in parentheses are the p-values.

4.5 Haplotype network: The Minimum Spanning Tree revealed two main haplogroups, two major haplotypes represented by individuals from N'Dama and Maure zebu breeds, and the remained are single haplotypes (Figure 2). The first haplogroup is constituted by *BB* and *BH* haplotypes while the second haplogroup composed of *AA*, *AH*, AA^I and A^IH. The major haplotype from the N'Dama breed appeared as the transitional haplotype (*AA* haplotype).

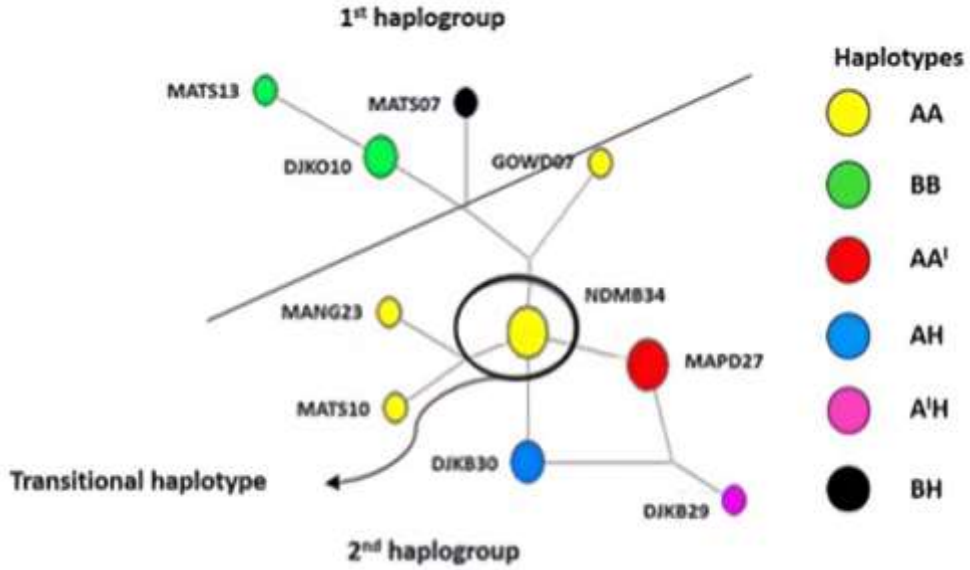

Figure 2. Minimum Spanning Tree of haplotypes of the Senegalese local cattle breeds.

5 DISCUSSION

The objective of this study was to identify the genetic variants of the kappa-casein gene and to estimate its polymorphism in the Senegalese local cattle breeds. A total of 31 sample sequences were compared to the *Bos taurus* reference sequence *CSN3*A* variant to identify the main genetic variants of the kappa-casein gene through single nucleotide polymorphism (SNP) substitutions. The sequencing of the *CSN3* gene exon IV resulted in four variants (*A*, *B*, A^I and *H*) and six haplotypes (*AA*, *BB*, AA^I, *AH*, *BH*, A^IH) distributed among the local cattle breeds in Senegal. A total of sixteen (16) sample sequence out of the thirty-one (31) were identified. The remaining fifteen (15) were discarded from the dataset because showing mutations that are different from those present in the reference sequence *CSN3* allele A. These mutations could be source of novel variants of the kappa-casein gene not yet identified. The predominance of variants *A* and *B* of the *CSN3* gene reported in this study is consistent with the

results found in previous work done on this gene around the world (Farrell *et al.,* 2004; Anggraeni *et al.,* 2010; El Nahas *et al.,* 2013). The haplotypes identified are unevenly distributed among the four studied breeds. Maure zebu and Djakore breeds had all the *CSN3* genetic variants identified and therefore the most diversified haplotypes. *AA*, *BB*, AA^I and *BH* were identified in Maure zebu while the haplotypes *BB*, *AH*, AA^I and A^IH were identified in Djakore breed. Regarding the Gobra zebu, three variants (*A*, A^I and *H*) that constituted three haplotypes *AA*, AA^I and *AH* have been identified. In the N'Dama breed, only two variants *A* and *B* that formed haplotypes *AA* and *BB* were observed. The presence of *CSN3* variant *B* in N'Dama taurine populations from Ivory Coast and Senegal was reported in the late 20th century by Mahé *et al.* (1999). Interestingly, the variants A^I and *H* have been only identified in the local zebu breeds (Gobra and Maure) and Djakore in this study. These variants are considered specific to zebus and therefore are not found in pure taurine breeds according to Caroli *et al.* (2009) and Gallinat *et al.* (2013). When the variants A^I and *H* are found in taurine breeds as the case of those from the southern Europe, it is the result of zebu introgression into taurine populations (Jann *et al.,* 2004). In addition, when these variants are combined to form a single haplotype A^IH, it is found only in Djakore breed. Since the *CSN3* gene is non-dominant and transmitted through Mendelian inheritance (Otaviano *et al.,* 2005; Anggraeni *et al.,* 2010), it is therefore easy to understand that this haplotype could be found in a population resulting from a natural crossing between the Gobra zebu and the N'Dama taurine. Djakore cattle should be considered as a hybrid population as its haplotype A^IH is resulted from a genetic introgression. These results agree with those of Ndiaye *et al.* (2019) who confirmed the hybrid genotype of the Djakore breed from 11 microsatellite loci markers. Characterizing lactoprotein variants is crucial. The kappa-casein protein plays an important role in the formation, stabilization, and aggregation of casein protein micelles and has effects on cheese making properties (Chessa *et al.,* 2003). *CSN3* variant *A* is recognized as the most common with the effects of increasing milk yield but lowering protein content (Neubauerova, 2001; Kucerova *et al.,* 2005; Matejicek *et al.,* 2008); while the variant *B* is considered favorable for the quality and quantity of milk-derived cheese products (Anggraeni *et al.,* 2010). Denisenko (2004) reported that cheese abilities may be improved by 5% in cows with variant *B* compared to those with variant *A*. The presence of variant *B* of the *CSN3* gene in the Senegalese local cattle, in particular in N'Dama breed would indicate that its milk might provide high protein content and good cheese abilities. Bovine kappa-casein variant *B* is known to be conducive to the production of high-quality milk and high cheese yield (Grosclaude, 1988; Otaviano *et al.,* 2005; Matejicek *et al.,* 2008). Such variant is included in dairy cattle selection strategies to improve productivity (Alipanah *et al.,* 2005; Dadhich *et al.,* 2006; Hamza *et al.,* 2011). Despite the milk quality produced by cows with variant *B*, cows with variant *A* would produce a higher milk yield as described by Bovenhuis *et al.* (1992). These authors reported a lower milk yield of 173 kg in cows with variant *B* compared to those with variant *A*. The variability of the *CSN3* gene (23.68% of variable sites) observed in the Senegalese local cattle breeds confirmed the results obtained by Ndiaye (2014), who showed using the mitochondrial gene cytochrome B (22% of variable sites) that these cattle are subjected to genetic admixture. Genetic admixture in animals not belonging to the same production systems is attributable to breeding practices such as animal mobility for feeding, uncontrolled reproduction when herds meet during grazing and at watering points, and the introduction of exotic breeds to improve dairy production (Ndiaye, 2014; Ndiaye *et al.,* 2014). The polymorphism of the kappa-casein gene in the Senegalese local bovine breeds in general is characterized by high haplotypic diversity and low nucleotide diversity. This presence of many haplotypes that differ by breed might be related to the multiplicity of maternal lineages as reported by Bradley *et al.* (1996) and Karpinski *et*

al. (2006). In addition, such diversity in domesticated ruminants suggests a certain stability of breeding populations. The genetic differentiation revealed by the kappa-casein gene among breeds showed that the N'Dama breed is genetically distant from the other cattle populations by high and significant F_{ST} values. The greatest genetic differentiation was observed between N'Dama and Gobra zebu. The genetic differentiation between the N'Dama and other local cattle populations confirms the differences between taurines and zebus from two different subspecies of *Bos taurus*. This might be explained by the fact that the N'Dama cattle are breed in an environment which is not suitable for zebus because of trypanosomiasis disease presence. The production environment and the type of subspecies (zebu or taurine) would explain this high genetic differentiation indicating therefore the presence of well-differentiated subpopulations. However, the high F_{ST} values characterizing the genetic differentiation among the Senegalese local cattle breeds do not evolve in the same direction as the pairwise genetic distances which are very low. These low values could be pinpointed to the low nucleotide diversity levels observed in these cattle breeds. The haplotypes identified through the *CSN3* genetic variants allowed to build the haplotype network divided into two haplogroups and showing a clear separation between *AA* and *BB* haplotypes. The transitional haplotype was found to be from *AA* haplotypes and therefore confirms that the *CSN3* allele *A* is indeed the reference variant. In other words, it is from this so-called ancestral haplotype, following a series of local variations of which the DNA sequence has been subjected, that the recently discovered haplotypes in the Senegalese local cattle breeds were derived.

6 CONCLUSION

The genetic polymorphism estimation of the kappa-casein gene (*CSN3*) revealed the existence of four milk protein variants *A*, *B*, A^I, *H* and six haplotypes (*AA*, *BB*, *AH*, AA^I, A^IH and *BH*) which were characterized by high haplotypic diversity and low nucleotide diversity in the Senegalese local cattle breeds. Among the haplotypes identified, *AA*, *BB*, and AA^I were the most spread, and are clues of good dairy performance in these breeds. As the variant *B* is the most desirable allele and found in almost three breeds, so the *CSN3 BB* haplotypes should be included among the selection criteria in breeding programs, by selecting animals carrying the variant *B* to better optimize production yields of such resilient breeds rearing in various breeding systems. These results give strong evidence that selection criteria focused on milk protein variants is a promising breeding strategy at coping the economic dairy production challenges.

7 REFERENCES

Alipanah M, Kalashnikova L. and Rodionov G: 2005. Kappa-casein genotypic frequencies in Russian breeds Black and Red Pied cattle. Iran Journal of Biotechnology. 3: 191-194.

Anggraeni A, Sumantri C, Farajallah A. and Andreas E: 2010. Kappa-Casein Genotypic Frequencies in Holstein-Friesian Dairy Cattle in West Java Province. Media Peternakan. 33:61-67. http://repository.ipb.ac.id/handle/123456789/43269.

ANSD: 2020. Situation Economique et Sociale du Sénégal en 2017 - 2018. Agence Nationale de la statistique et de la Démographie. Point E-Dakar, Sénégal.

Badji MN, Diouf MN, Ciss M, Baldé AT, Sambe BS, Diop M. and M. Sembène: 2020. Etude phénotypique des taurins Ndama du Centre de Recherches Zootechniques Techniques de Kolda (Senegal). European Scientific Journal. 16: 514 - 534. http://dx.doi.org/10.19044/esj.2020.v16n3p514.

Bertrand B: 2006. Bilan et analyse de l'utilisation de l'insémination artificielle dans les programmes d'amélioration génétique des races laitières en Afrique soudano-sahélienne. DVM Thesis. Claude Bernard Université Lyon I.

Bovenhuis H, Van Arendonk JAM. and Korver S: 1992. Associations between milk protein polymorphisms and milk production traits. Journal of Dairy Science. 75: 2549-2559. https://doi.org/10.3168/jds.S0022-0302(92)78017-5.

Bradley DG, MacHugh DE, Cunningham P. and Loftus RT: 1996. Mitochondrial DNA diversity and the origins of African and European cattle. Proceedings of the National Academy of Sciences. 93: 5131-5135. https://dx.doi.org/10.1073%2Fpnas.93.10.5131.

Caroli AM, Chessa A. and Erhardt GJ: 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. Journal of Dairy Science. 92: 5335-5352. https://doi.org/10.3168/jds.2009-2461.

Chessa S, Budelli E, Gutscher K, Caroli A. and Erhardt G: 2003. Short communication: simultaneous identification of five κ-casein (CSN3) alleles in domestic goat by polymerase chain reaction-single strand conformation polymorphism. Journal of Dairy Science. 86: 3726-3729. https://doi.org/10.3168/jds.S0022-0302(03)73978-2.

Dadhich S, Patel RK, Soni KJ, Singh KM. and Chauhan JB: 2006. Estimation of allelic frequency of κ-casein and β-lactoglobulin genes in *Bos indicus* cattle breeds. International Journal of Cow Science 2:48-51.

Denisenko EA: 2004. Milk production and technological property milk of Black Pied breeds with difference genotypes of kappa casein in region Siberia. PhD Thesis. Lesni poliani, Moscow.

Dumas BR, Brignon G, Grosclaude F. and Mercier JC: 1972. Structure primaire de la caséine ß bovine. Séquence complète. European Journal of Biochemistry. 25:505-514. https://doi.org/10.1111/j.1432-1033.1972.tb01722.x.

Excoffier L, Laval G. and Schneider S: 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evolutionary Bioinformatics. Online. 1: 47-50. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2658868/pdf/ebo-01-47.pdf.

Farrell HM, Jimenez-Flores R, Bleck GT, Brown EM, Butler JE, Creamer LK, Hicks CL, Hollar CM, Ng-Kwai-Hang KF. and Swaisgood HE: 2004. Nomenclature of the proteins of cows' milk—sixth revision. Journal of Dairy Science. 87:1641-1674. https://doi.org/10.3168/jds.S0022-0302(04)73319-6.

Gallinat, JL, Qanbari S, Drogemuller C, Pimentel ECG, Thaller G. and Tetens J: 2013. DNA-based identification of novel bovine casein gene variants. Journal of Dairy Science. 96: 699-709. https://doi.org/10.3168/jds.2012-5908.

Getachew R: 2010. Molecular characterization of the bovine gene CSN3 and LGB milk protein genes using sequencing and PCR-RFLP markers in Ethiopian Indigenous cattle populations. MS Thesis. Science Faculty Biotechnology Program, Addis Ababa Univ., Ethiopia.

Grosclaude F: 1988. Le polymorphisme génétique des principales lactoprotéines bovines. Relations avec la quantité, la composition et les aptitudes fromagères du lait. INRA Production Animale. 1: 5-17. https://doi.org/10.20870/productions-animales.1988.1.1.4430.

Hall TA: 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows

95/98/NT. Nucleic Acids Symp. Ser. 41:95-98.

Hamza AE, Yang ZP, Wang XL, Chen RJ, Wu HT. and Ibrahim AI: 2011. The impact of kappa casein gene polymorphism on milk components and other productive performance traits of Chinese Holstein cattle. Pakistan Veterinary Journal. 31: 153-156.

Jann OC, Ibeagha-Awemu EM, Ozbeyaz C, Zaragova P, Williams JL, Ajmone-Marsan P, Lenstra JA, Moazami-Goudarzi K. and Erhardt G: 2004. Geographic distribution of haplotype diversity at the bovine casein locus. Genetics, Selection, Evolution 36: 243-257. https://doi.org/10.1186/1297-9686-36-2-243.

Karimi K, Nassiri MTB, Mirzadeh K, Ashayerizadeh A, Roushanfekr H. and Fayyazi J: 2009. Polymorphism of the β-lactoglobulin gene and its association with milk production traits in Iranian Najdi cattle. Iranian Journal of Biotechnology 7: 82-85.

Karpinski M, Junkuszew A, Drozd A. and Gruszecki TM: 2006. A phylogenetic comparison of wild sheep (*Ovis musimon*) and domestic sheep (*Ovis aries*) represented by BCP synthetic line using mitochondrial cytochrome b gene sequence analysis. Archiv Fur Tierzucht. 49: 310-316.

Kimura M: 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 16: 111-120. https://doi.org/10.1007/BF01731581.

Kucerova J, Nemcova E, Stipkova M, Jandurova O, Matejicek A. and Bouska J: 2005. The association between CSN3 genotypes and milk production parameters in Czech Pied Cattle. Pages 5-8 in 56th Ann. Meet. EAAP. Uppsala, Sweden.

Mahé M-F, Miranda G, Queval R, Bado A, Zcafindrajaona PS. And Grosclaude F: 1999. Genetic polymorphism of milk proteins in African *Bos taurus* and *Bos indicus* populations. Characterization of variants α_{s1}-Cn H and $\varkappa$-Cn J. Genetics, Selection, Evolution 31: 239-253. https://hal.archives-ouvertes.fr/hal-00199729.

Matejicek A, Matejickova J, Stipkova M, Hanus O, Gencurova V, Kyselova J, Nemcova E, Kott T, Sefrova J, Krejcova M, Melcova S, Holzelova I, Bouska J. and Frelich J: 2008. Joint effects of CSN3 and LGB genes on milk quality and coagulation properties in Czech Fleckvieh. Czech Journal of Animal Science 53: 246-252.

MEPA: 2018. Rapport de revue annuelle du secteur de l'élevage au Sénégal. Ministère de l'Elevage et des Productions Animales. Accessed Oct. 8, 2018. https://elevage.sec.gouv.sn/sites/default/files/Revue%20annuelle%202018.pdf .

El Nahas SM, Bibars MA. and Taha DA: 2013. Genetic characterization of Egyptian buffalo CSN3 gene. Journal of Genetic Engineering and Biotechnology. 11:123-127. https://doi.org/10.1016/j.jgeb.2013.08.003.

Ndiaye M: 2014. Diversité génétique et évolution démographique des races locales et métisses de bovins du Sénégal, inférées des séquences partielles du cytochrome b de l'ADN mitochondrial. MS Thesis. Department of Animal Biology, Cheikh Anta Diop University, Dakar.

Ndiaye NP: 2015. Evaluation de la diversité et de la structure génétique des races bovines locales (Artiodactyla : Bovidae) du Sénégal. PhD. Department of Animal Biology, Cheikh Anta Diop University, Dakar.

Ndiaye NP, Sow A, Ndiaye S, Sembène M. and Sawadogo GJ: 2014. Phenotypical characterization of Senegalese local cattle breeds using multivariate analysis.

Journal of Animal Veterinary Advanced. 13:1150-1159. https://medwelljournals.com/abstract/?doi=javaa.2014.1150.1159.

Ndiaye NP, Sow A, Dayo GK, Ndiaye S, Sawadogo GJ. and Sembène M: 2015. Genetic diversity and phylogenetic relationships in local cattle breeds of Senegal based on autosomal microsatellite markers. Veterinary World. 8:994-1005. https://doi:10.14202/vetworld.2015.994-1005.

Ndiaye NP, Sow A, Sawadogo GJ. and M. Sembène: 2019. Diversity and genetic structure of Senegalese local cattle breeds. Asian Journal of Applied Science. 12: 149-157. https://scialert.net/abstract/?doi=ajaps.2019.149.157.

Nei M: 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 89: 583-590. https://www.genetics.org/content/genetics/89/3/583.full.pdf.

Neubauerová V: 2001. Detection of genetic markers and possibility of their use in cattle and other subungulates. PhD Thesis. University of South Bohemia, Ceske Budejovice.

Otaviano AR, Tonhati H, Sena JAD. and Munoz MFC: 2005. Kappa casein gene study with molecular markers in female buffaloes (*Bubalus bubalis*). Genetics and Molecular Biology. 28: 237-241. https://doi.org/10.1590/S1415-47572005000200010.

Polzin T. and Daneshmand SV: 2003. On Steiner trees and minimum spanning trees in hypergraphs. Oper. Res. Lett. 31:12-20. https://doi.org/10.1016/S0167-6377(02)00185-2.

R Core Team: 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rozas J, Librado P, Sánchez-Del Barrio JC, Messeguer X. and Rozas R: 2010. DnaSP version 5.10.1. Accessed Mar. 25, 2015. http://www.ub.edu/dnasp/index_v5.html.

Sambe BS, Diouf MN, Ciss M, Baldé AT, Badji MN, Diop M. and Sembène M: 2019. Phenotypic characterization of Gobra zebu cattle of Centre de Recherches Zootechniques of Dahra. International Journal of Advanced Research 7: 26-34. http://dx.doi.org/10.21474/IJAR01/9190.

Tamura K, Stecher G, Peterson D, Filipski A. and Kumar S: 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution. 30: 2725-2729. https://doi.org/10.1093/molbev/mst197.

Wright S: 1978. Evolution and the genetics of populations. Volume 4: Variability within and among Natural Populations. University of Chicago Press, Chicago.

Table 2. Distribution of CSN3 gene haplotypes identified

	Haplotypes		Position des substitutions nucléotidiques/ acides aminés											
			12940/ 93	12950/ 97	12951 /97	12971/ 104	13065/ 135	13068/ 136	13096/ 145	13104/ 148	13111/ 150	13119/ 153	13124/ 155	13165/ 168
GoWd07	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
GoKO25	AH	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr					
GoKO30	AA^{I}	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCG/Pro			
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCA/Pro			
MaTS07	BH	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile	AUC/Ile	ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr	ACC/Thr	ACU/Thr	GAU/Asp				GCA/Ala
MaTS13	BB	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		AUC/Ile	ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		ACC/Thr	ACU/Thr	GAU/Asp				GCA/Ala
DjKb29	$A^{I}H$	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr		CCG/Pro			
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr		CCA/Pro			
DjKO10	BB	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		AUC/Ile	ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		ACC/Thr	ACU/Thr	GAU/Asp				GCA/Ala
DjKb28	AA^{I}	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCG/Pro			
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCA/Pro			

Q: query sequence; S: reference sequence.

Tableau 2: Distribution of CSN3 gene haplotypes identified (continuation and end)

	Haplotypes		Position des substitutions nucléotidiques/ acides aminés											
			12940/ 93	12950/ 97	12951 /97	12971/ 104	13065/ 135	13068/ 136	13096/ 145	13104/ 148	13111/ 150	13119/ 153	13124/ 155	13165/ 168
DjKb30	AH	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr					
MaTS16	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
MaNg23	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
MaPD27	AAI	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCG/Pro			
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCA/Pro			
NdCK29	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
NdMB34	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
NdMB35	AA	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					
NdMB37	BB	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		AUC/Ile	ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		ACC/Thr	ACU/Thr	GAU/Asp				GCA/Ala

Q: query sequence; S: reference sequence

Appendix 1 : Distribution of *CSN3* gene unidentified haplotypes

	Haplotypes		Position of nucleotidic substitutions / amino acid											
			12940/ 93	12950/ 97	12951 /97	12971/ 104	13065/ 135	13068/ 136	13096/ 145	13104/ 148	13111/ 150	13119/ 153	13124/ 155	13165/ 168
DjKO04	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr	GAU/Asp				GCA/Ala
DjKO16	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCG/Pro			GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr		CCA/Pro			GCA/Ala
DjKO21	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					GCA/Ala
DjNf23	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		AUC/Ile	ACU/Thr					
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser		ACC/Thr	ACU/Thr					
DjKb27	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr	GCU/Ala				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr	GAU/Asp				GCA/Ala
GoWd06	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr	ACU/Thr	CCC/Pro			GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr	ACU/Thr	CCA/Pro			GCA/Ala
GoPD35	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	AUC/Ile		ACU/Thr	ACU/Thr				GCG/Ala
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser	ACC/Thr		ACU/Thr	ACU/Thr				GCA/Ala
GoPD37	?	Q	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					GUU/Val
		S	ACU/Thr	CGU/Arg	CGU/Arg	UCA/Ser			ACU/Thr					GCA/Ala

Q: query sequence; S: reference sequence; ? : unidentified individuals