CHARACTERIZATION OF THE α $_{S1}$-CASEIN I^{RV} ALLELE PROVIDES EVIDENCE FOR PHYLOGENY OF THE ANCIENT ROMANIAN GREY STEPPE CATTLE, MOLDAVIAN STRAIN

V.A. Balteanu[1], F.D. Pop[1], A. Vlaic[1], T.C. Carsai[1], Şt. Creangă[2], A.R. Rusu[3]

[1]U.S.A.M.V. Cluj-Napoca, Faculty of Animal Husbandry and Biotechnology, Cluj County, Romania
[2]U.S.A.M.V. Iasi, Faculty of Animal Husbandry, Iasi County, Romania
[3]U.S.A.M.V. Cluj-Napoca, Quality Control Department, Cluj County, Romania
e-mail: avbalteanu@yahoo.com

Abstract

The objective of this work was the complete characterization of the α_{S1} casein (α_{S1}-CN) I^{RV} allele, previously identified in the Moldavian strain of Romanian Grey Steppe cattle. After α_{S1}-CN cDNA sequencing several substitutions were observed in I^{RV} allele (GenBank EU908730), as compared with common B and C alleles. These results suggest that the I^{RV} allele arose phylogenetically from the C variant. Based on A-T substitution (exon 11), a PCR-RFLP test was developed for identification of I^{RV} allele. Blood samples of 63 living individuals of this breed an representing different ages and preserved semen samples of 6 deceased bulls were genotyped, resulting in a 0.123 frequency of I^{RV} allele. The discovery of an ancestral casein allele in such a small endangered population is critical in the context of genetic resources preservation, highlighting the unique phylogeny of Romanian Grey Steppe cattle.

Key words: Grey Steppe cattle, α_{S1} casein I^{RV}, milk, polymorphism, cattle phylogeny

INTRODUCTION

The genetic polymorphisms of major milk proteins were studied in cattle relative to milk quality and it's manufacturing properties [8], [17], either to study phylogenetic relationships [6], [13] or as Zebu introgression markers in *Bos taurus* breeds [12]. From the first detected polymorphism in the cattle α_{S1}-casein (**α_{S1}-CN**) locus [16] through 2004, nine genetic variants were identified at this locus: A, B, C, D, Eyak, Ebali, F, G, and H [9]. The alpha S1 B variant is the most common in European breeds, with the highest frequency (90-100%) in Holstein [10]. In Yak and Zebu, the C allele is predominant [15]. The B variant has its phylogenetic origin from C, which was likely the ancestral allele; all the other known variants in *Bos taurus* were derived from B, except for E variant found only in *Bos grunniens* [11], [14]. In 1991, Mahé et al. [15] observed 2 new possible α_{S1}-CN variants by isoelectric focusing (IEF) in African Kuri and Sudanese Fulani. One variant was characterized as H and the other, migrating between the B and C variants, but closer to C, was not characterized because the milk quantity was insufficient [15]. In 1992, Kawamoto et al. [14] observed two unknown α_{S1}-CN IEF profiles. The first profile, named X, was identified in Nepalese *Bos taurus*. The second one, named Y, occurred only in *B. taurus* x *B. grunniens* crosses. The unknown X variant had a similar IEF migrating profile [14] to the one observed earlier in 1991 (Mahé, unpublished data) and Mahé et al. [15]. The existence of these new possible genetic variants has not been confirmed by molecular analysis, so they are not yet recorded as new genetic variants [9], [10].

In 2005 Balteanu et al. observed the same α_{S1}-CN IEF pattern in the Romanian Grey Steppe cattle, Moldavian strain, a local *Bos taurus primigenius* breed belonging to the Grey Cattle group of breeds [3], [4]. The genetic nature of this new IEF pattern was confirmed by pedigree studies and the 2 carrier cows of this new α_{S1}-CN allele had a common bull ancestor. It was concluded that

this is a new α_{S1}-CN allele, named α_{S1}-CN I^{RV} [3], [4], which was never observed in other European cattle breeds [9] or in Romanian breeds [4]. This allele was partially characterized at the protein level [1] and fully by cDNA sequencing [2] (GenBank accession number EU908730). More recently, a similar IEF profile was observed by Caroli et al., in a study on Carora cattle, named also I [7].

Considering this background and novelty of the α_{S1}-CN I^{RV} allele belonging to Romanian Grey Steppe cattle breed, our objective was to describe its complete characterization and verify its frequency in this breed.

MATERIAL AND METHOD

I. RNA extraction from milk somatic cells

Milk samples collection was from 2 Romanian Grey Steppe cows previously identified by IEF as carriers of the new α_{S1}-CN I^{RV} allele in the heterozygous condition (BI^{RV}, CI^{RV}) and 2 reference samples (BB, CC). Somatic cells from the collected milk samples (50 ml) were pelleted by centrifugation at 2,000 g for 15 min at 4^0C. The cell pellet was resuspended in 1 ml PureZOL reagent (BioRad Laboratories, Hercules, CA, USA) and the following extraction steps were performed according to the protocol recommended by the producer.

II. DNA extraction from blood and semen

Blood samples were collected from the jugular vein into containers with K-EDTA anticoagulant from the 2 carrier cows (BI^{RV}, CI^{RV}) and 3 reference individuals (BB, CC, BC). In order to study the frequency of this allele in this breed, blood samples were also collected from the entire known existing population, consisting of 63 individuals of different ages: 57 from S.C.D.C.B. Dancu, Iasi County, Romania and 6 from the U.S.A.M.V. Cluj-Napoca, Cluj County, Romania). Six straws with conserved semen, belonging to the bulls used in the national conservation program of this breed, were also recovered from S.C.D.C.B. Dancu. The DNA extraction was performed as described before [2].

III. cDNA obtaining and sequencing was done as described before [2]

IV. BseGI restriction map of cDNA belonging to BI^{RV} sample and the 2 references (BB, CC), was performed with BseGI enzyme, in order to confirm one of the mutations characterizing the α_{S1}-CN I^{RV} variant observed following comparative sequencing. The restriction was performed on 15 µl PCR product as recommended by manufacturer (Fermentas, Vilnius, Lithuania). The restriction products were analyzed in 2.5% agarose gel containing 1X Sybr Safe (Invitrogen, Eugene, OR, USA).

V. PCR-RFLP of exon 11 and whole population genotyping was done from all collected samples as described before [2].

RESULTS AND DISCUSSION

I. Sequencing results

Comparative analysis of chromatograms obtained following sequencing (with the forward and reverse primer, respectively) of the 2 samples carriers of α_{S1}-CN I^{RV} (BI^{RV}, CI^{RV}) and 2 reference samples (BB, CC), revealed the mutations characterizing this new genetic variant. Alpha S1 B allele (the most common variant in European cattle breeds) differs from I^{RV} allele by 2 substitutions: substitution of an adenine from the B allele (exon 11, nucleotide 297, codon 99 gaA coding for Glu) with a thymine in the I^{RV} allele (gaT coding for Asp); substitution of an adenine from the B allele (exon 17, nucleotide 620, codon 207 gAa coding for Glu) with guanine in the I^{RV} allele (gGa coding for Gly). Alpha S1 C allele differs from I^{RV} allele by 1 substitution: an adenine from C allele (exon 11, nucleotide 297, codon 99 gaA coding for Glu) is substituted with a thymine in I^{RV} allele (gaT coding for Asp). In the position 620, C and I^{RV} alleles are similar. There are 2 additional substitutions observed in the B, C and I^{RV} alleles (sequenced from Romanian Grey Steppe cattle breed and mentioned in GenBank EU908730) that differ from those published in GenBank, with no effect on protein sequence.

II. BseGI restriction map of cDNA

The α_{S1}-CN cDNA restriction map of the 695 bp product was performed with BseGI enzyme, in order to confirm the A to T

substitution at position 297. Five restriction products were observed in B and C variants because of the presence of 4 restriction sites for this enzyme (460bp, 97bp, 75bp, 42bp, and 10bp). The A to T substitution present in I^{RV} allele creates a new restriction site for BseGI in the 297 position, and therefore the 460 bp product is cut into 2 additional products, evidencing the presence of this mutation (Figure 1).

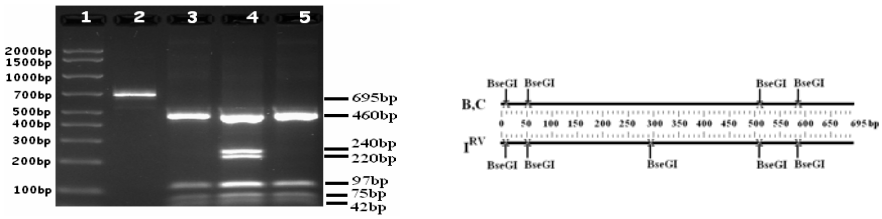

Figure 1. Comparison among α_{S1}-CN B, C and I^{RV} cDNA's restriction maps with BseGI enzyme. Lane 1 has Ampli Size™ Molecular Ruler (BioRad, Hercules, CA, USA)

III. Alpha S1 B, C, I^{RV} proteins sequences comparison and phylogenetic origin of I^{RV} allele

Protein sequence of the I^{RV} variant (as compared with B and C variants), was deduced from the mutations found following cDNA sequencing and from our previous sequencing work at the protein level [1]. Accordingly, the B protein variant differs from the I^{RV} variant by a Glu (E) to Asp (D) substitution at position 99 (position 84 in mature protein) and a Glu (E) to Gly (G) at position 207 (position 192 in mature protein). Further, the C protein variant differs from I^{RV} variant by a Glu (E) to Asp (D) substitution at position 84. At position 192 the C and I^{RV} protein variants are similar (Figure 2).

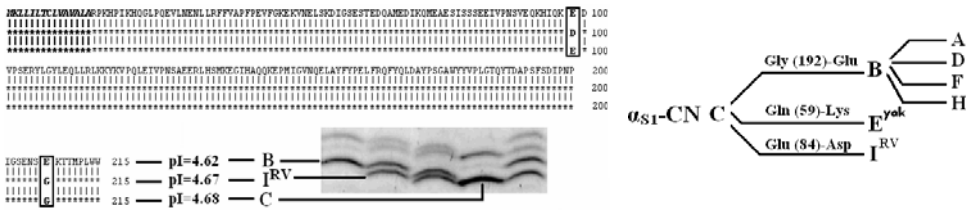

Figure 2. Left figure: Comparison among protein sequence features of B, C and I^{RV} variants (Romanian Grey Steppe cattle breed). The mutations, which comprise the difference among these 3 protein variants, are marked with rectangles. Note the differences among isoelectric points (pI) of the 3 proteins, which explain the observed IEF profiles (gel image from the right side). Signal peptide is highlighted with bold italic letters; Right figure: Phylogenetic origin of the α_{S1} I^{RV} allele

The calculation of α_{S1}-CN B, C and I^{RV} mature proteins' isoelectric points (4.62, 4.68, 4.67, respectively) explained the mobility differences previously observed in IEF profiles [3], [4], indicating that the I^{RV} variant migrates closer to the C variant, rather than B (Figure 2). This is perfectly in agreement with the protein sequence deduced from cDNA sequencing of the B, C and I^{RV} alleles. Comparing the results obtained at the DNA (present work) and protein levels (performed in our previous work), we confirmed the Gly-Glu substitution from position 207 (position 192 in mature protein) which comprises the difference between B and C or I^{RV} variants. The other mutation characterizing α_{S1}-CN I^{RV}, presumed in our previous work at the protein level involving the glutamine (Q) residues from exons 8, 9 or 11, was not confirmed at DNA level, although the mutation detected at DNA level was located in one of the presumed exons, namely exon 11.

The phylogenetic origin of alpha S1 I^{RV}allele can be deduced from the available molecular data. The mutation event which occurred in the 11th exon, concerning an A to T substitution from position 21, responsible for the Glu to Asp substitution at position 84 of the mature protein, distinguishes I^{RV} from the B and C alleles. These data together with the differences observed between the B and I^{RV} alleles at position 31 of exon 17, responsible for the Gly to Glu substitution at position 192 of the mature protein, and the similarity between the C and I^{RV} alleles at this position, suggest that I^{RV} allele arose phylogenetically from C variant (Figure 2).

Gathering all existing literature information and the molecular data obtained in this study, we can conclude that the I^{RV} variant is the only known α_{S1}-CN allele in *Bos taurus* issued directly from the ancestral C allele and not from the B allele (Figure 2). A similar mutation event led to the appearance of the E allele in yak, which was also proposed as having the C allele as its origin [11]. This theory is also sustained by the fact that the C variant has a high frequency in cattle breeds from Africa and Asia. However, this is the first molecular evidence that brings a useful clue about phylogenic relationships among the Romanian Grey Steppe cattle and *Bos genus* representative breeds from Africa and Asia.

IV. PCR-RFLP genotyping and frequency of I^{RV} allele in this breed

The identification of alpha S1 I^{RV} from the B and C alleles was possible by amplification of a 422bp fragment, containing a part of intron 9 and intron 11 and including entire exons 10 and 11. The 422bp DNA fragment amplified from B and C allele has 1 restriction site for the BseGI enzyme, with 2 fragments resulting from digestion. Those fragments are 156 bp and 266 bp, corresponding to the BB, BC or CC genotypes. The substitution of an adenine from B and C alleles with a thymine in I^{RV} at position 21 of exon 11 creates a new BseGI restriction site for the amplified DNA fragment belonging to I^{RV}. Therefore the 266 bp fragment is further cut twice at lengths of 196 and 70bp (Figure 3).

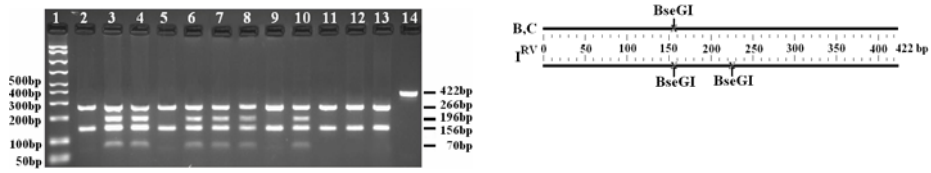

Figure 3. Identification of α_{S1}-CN I^{RV} allele by means of PCR-RFLP. Lane 1 has Ampli Size™ Molecular Ruler (BioRad, Hercules, CA, USA).

The whole population genotyping revealed that 4 deceased bulls and 13 living females of different ages carry this α_{S1}-CN I^{RV} new allele, making the calculated frequency of this allele 0.123. This frequency is relatively high for a new genetic variant. Of course some could attribute this to a low number of individuals left, which possibly caused inbreeding. This theory can be easily refuted by the fact that the conserved semen belonging to 4 unrelated bulls (from the 1980's) carries this new variant, as shown in this study (Figure 3). This suggests that I^{RV} variant had an even higher frequency at the time when this breed and its progenitors were the only ones existing in Romania. However, the declining number of individuals has also likely allowed this breed to remain relatively isolated from introgression of other alleles from other breeds. The relatively high frequency of I^{RV} variant is interesting considering that, based on the similar IEF profiles, this variant is likely similar to uncharacterized alpha S1 variants observed at a low frequency in *Bos taurus* - African Kuri breed (0.01) and *Bos indicus:* Sudanese Fulani (0.01) [15], Nepalese *Bos taurus* from the Dhaulaghiri zone (0.07) [14] and Carora cattle (0.016) [7]. However it is lower than that observed by Kawamoto et al. [14] in

Nepalese *Bos taurus* from the Pokhara and Sauraha regions (0.25 and 0.16, respectively). The results obtained in this study are in contrast with those formulated by Caroli et al. [7], who concluded that their I „variant'' is a *Bos indicus* specific allele and can be used as a marker to prove *Bos indicus* introgression in *Bos taurus* breeds.

CONCLUSIONS

A new α_{S1}-CN allele, named I^{RV}, previously discovered in Romanian Grey Steppe cattle breed, Moldavian variety was completely characterized and used to consider the phylogenetic evolution of this breed and its ancient progenitors. Using the combined approach of DNA sequencing and restriction map analysis, we identified and confirmed the hypothesis of the existence of this new additional variant previously detected [3], [4].There are 2 remaining questions linked to X „variants'' observed in Africa in Kuri (*Bos taurus*) and Sudanese Fulani (*Bos indicus*) cattle breeds in 1991 [15], in Asia in Nepalese *Bos taurus* and *B. taurus x B. grunniens* crosses [14] and in Carrora cattle from west-central Venezuela, for the last one the authors proposing the name I [7]. Are this similar with α_{S1}-CN I^{RV} observed and characterized in Romanian Grey Steppe cattle [1], [2], [3]? Because the detection method was the same and based on IEF profiles observed in our study as compared with the other three, we can say that it is probably the same α_{S1}-CN genetic variant. Further studies should be performed to characterize this X „variant" at it's origin in Africa and Asia, in order to have a complete picture of cattle phylogenetic relationships on the 3 continents. If the mutations will be similar with those observed in our work, it is highly probable that mutation events which leaded to the appearance of this new variant took place in a common ancestor of the 3 cattle groups from Europe, Africa and Asia, sustaining the monophyletic origin of cattle breeds.

Our theory is sustained by the fact that molecular data showed that α_{S1}-CN I^{RV} allele is the only genetic variant described in *Bos taurus* so far in this locus, which originates directly from C variant (the most common variant in cattle breeds from Asia and Africa), as E variant described in *Bos grunniens*. Another interesting point is that α_{S1}-CN I^{RV} allele was never observed in other European cattle breeds [9], not even in other breeds belonging to Grey Cattle Group, including the close related Hungarian Grey Steppe cattle [5] or other Romanian cattle breeds [4]. Because the presence of this new α_{S1}-CN allele was noticed only in pure breed individuals, its origin in this breed is unquestionable. In Italy, Hungary, France breeds from the Grey Cattle Group were continuously improved for meat production, so they lost a big part of their ancestor's characteristics. Among many varieties existing in Romania until 1940, belonging to Grey Cattle Group, in Romanian Grey Steppe cattle, Moldavian variety little breeding work has been done, so the breed has many common characteristics with their wild ancestors making it the closest relative of aurochs. Therefore its conservation should represent a national priority. Unfortunately Heck cattle is considered today as being reconstructed auroch, but we hope that further studies which are in progress in our laboratory will sustain our theory.

The discovery of an ancestral casein allele in such a small population, so endangered, is very important in the context of necessity of world genetic resources preservation. We should take a closer look to this almost extinct breed, which could give us many answers about cattle origin and their evolutionary history. Being a part of our national history we think that saving it from extinction is in our benefit and in the benefit of international genetic heritage.

ACKNOWLEDGMENTS

This research was supported by Romanian Ministry of Education, Research and Innovation trough CNCSIS Project No. 1122/2008-2011. The authors are grateful to Marie Françoise Mahé from INRA, Jouy en Josas, France, for her support.

BIBLIOGRAPHY

[1] Balteanu V.A., Vlaic A., Pop F.D., Rusu A.R., Martin P., Miranda G., Creanga S.: Characterization at protein level of the new α_{S1}-casein allele I^{RV} discovered in Romanian Grey Steppe cattle breed Moldavian variety, Scientific Papers Anim. Husbandry and Bitech. Timisoara, 2008, 41: 1- 10.

[2] Balteanu V.A., Pop F.D., Vlaic A., Rusu A.R., Creanga S.: Molecular characterization of α_{S1}-CN I^{RV} allele discovered in Romanian Grey Steppe cattle and it's frequency in this breed, Bulletin of USAMV-CN, Anim. Sci. and Biotech., 2008, 65: 477.

[3] Balteanu V.A., Vlaic A., Rusu A.R., Creanga S., Pop F.D., Odagiu A., Pântea M.L., Hâncu V.: Milk proteins polymorphism in Romanian Grey Steppe cattle studied by isoelectric focusing technique (IEF). Identification of a new allele alpha S1-casein allele: alpha S1 I^{RV}, Bulletin of USAMV-CN, Anim. Sci. and Biotech., 2007, 63/64: 304 - 310.

[4] Balteanu V.A.: Milk proteins polymorphism in Romanian cattle, buffalo and goat breeds studied by isoelectric focusing technique. Identification of 2 new casein allele: α_{S1}-casein I^{RV} in Romanian Grey Steppe cattle and α_{S1}-casein B^{RV} in Romanian Buffalo. MSc Thesis, University of Agriculture Sciences and Veterinary Medicine, Cluj-Napoca, Romania, 2005.

[5] Baranyi M., Bösze Z.S., Buchberger J., Krause I.: Genetic polymorphism of milk proteins in Hungarian Spotted and Hungarian Grey cattle: a possible new genetic variant of β-lactoglobulin, J. Dairy Sci., 1993, 76: 630 – 636.

[6] Beja-Pereira A., Erhardt G., Matos C., Gama L., Ferrand N.: Evidence for a geographical cline of casein haplotypes in Portuguese cattle breeds, Anim. Genet., 2002, 33: 295 – 300.

[7] Caroli A., Chessa S., Chiatti F., Rignanese D., Melendez B., Rizzi R.: Short Communication: Carrora cattle show high variability in α_S1- casein, J. Dairy Sci, 2008, 91:354 – 359.

[8] Delacroix-Buchet A., Marie C.: Comparaison des variants A et C de la caséine k des laits de vaches Tarentaises en modèle fromagerde type beaufort, Lait, 1994, 74: 343-360.

[9] Farrell H.M., Jimenez-Flores Jr.R., Bleck G.T., Brown E.M., Butler J.E., Creamer L.K.: Nomenclature of the Proteins of Cows' Milk-Sixth Revision, J. Dairy Sci., 2004, 87: 1641-1674.

[10] Formaggioni P., Summer A., Malacarne M., Mariani P.: Milk protein polymorphism: detection and diffusion of the genetic variants in Bos genus, Ann. Fac. Med. Vet. Univ. Parma 1999, 127 – 165.

[11]Grosclaude F., Mahé M.F., Mercier J.C., Bonnemaire J., Teissier J.:Polymorphisme des lactoprotéines de Bovinés Népalais. I. Mise en évidence, chez le yak, et caractérisation biochimique de deux nouveaux variants: b-Lactoglobuline Dyak et caséine aS1-E, Ann. Génét. Sél. Anim., 1976, 8: 461- 479.

[12] Ibeagha-Awemu E.M., Prinzenberg E.M., Jann O.C., Luhken G., Ibeagha A.E., Zhao X.: Molecular characterization of bovine CSN1S2 B and extensive distribution of zebu specific milk protein alleles in European cattle, J, Dairy Sci, 2007, 90: 3522 – 3529.

[13] Jann O.C., Ibeagha-A.E.M., Ozbeyaz C., Zaragoza P., Williams J.L., Ajmone-Marsan P.: Geographic distribution of haplotype diversity at the bovine casein locus, Gen. Sel. Evol., 2004, 36: 243 -257.

[14] Kawamoto Y., Namikawa T.A., Adachi T., Amano T., Shotake T., Nishida T.: A population genetic study on yaks, cattle and their hybrids in Nepal using milk protein variations, Anim. Sci. Tech., 1992, 63: 563-575.

[15] Mahé M.F., Miranda G., Queval R., Bado A., Zafindrajaona P.S., Grosclaude F.: Genetic polymorphism of milk proteins in African Bos taurus and Bos indicus populations. Characterization of variants aS1-Cn H and k-Cn, J. Genet. Sel. Evol., 1999, 31: 239 – 253.

[16] Thompson M.P., Kiddy C.A., Pepper L., Zittle C.A.: Variations in the α_S1-casein fraction of individual cow's milk, Nature, 1962, 195: 1001 – 1002.

[17] Vlaic A., Oroian T., Pascalau S.: Genotypisation of bulls used for artificial insemination at k-casein and β-lactoglobulin locus, Bulletin of USAMV-CN, Anim. Sci. and Biotech., 2001, 55/56: 112 – 117.